Results and lessons from implementing the Water Assets Accounts in the EEA area

From concept to production

Contents

Table of annexes

Table of figures

Table of maps

Table of tables

Acknowledgment

These are the first EU-level water accounts that display water balances at monthly and sub-basin levels. The accounts were developed in the hope that the many data gaps and methodological imperfections will be ironed out in the future.
These accounts are the result of extensive collaboration over many years with a number of pioneers in the field of environmental accounting. Without their expertise, commitment and generosity this first work would never have been completed.
Among these pioneers, there are three people whose help was especially important. They are: the renowned hydrologist Jean Margat, who many decades ago advanced the then-controversial idea that water resources could be accounted for; Professor Michel Meybeck, who during his research and teaching career taught that applied science supported by rigorous methodology was a natural progression from fundamental research; and Jean-Louis Weber, who has worked tirelessly to promote the use of environmental accounting as a key element of environmental assessment and policy support.

This report has been prepared by Philippe Crouzet (EEA project manager) with contribution from Guillaume le Gall (Pöyry), Paul Campling (Vito), Miriam Basso (SCM) and the EEA staff members: Jean-Louis Weber, Oscar Gomez and Blaz Kurnik.

1 Executive summary

1.1 Scope of the project

The United Nations **System of National Accounting** (**SNA**) framework provides an internationally agreed methodology for national economic accounts. SNA accounts are the main source of information for the internationally comparable economic aggregates and indicators which are used to assess the economic performance of countries. Examples are gross domestic product (GDP), value added, income, consumption, economic growth rate and government deficit,.

GDP is hence the best-known measure of macroeconomic activity. It has also come to be regarded as a proxy indicator for overall societal development and progress in general. However, GDP does not measure environmental sustainability or social inclusion, and these limitations need to be taken into account when using it in policy analysis and debates. The need to strengthen the data and indicators that complement GDP is increasingly recognised, and several international initiatives have been launched to address these issues. Taking stock of these, in August 2009, the European Commission adopted a communication 'GDP and beyond - measuring progress in a changing world' (EC, 2009). This communication explicitly addresses the need for environmental accounting (Section 3.5) and recalls that since 2006, the Commission had called on the European Union (EU) and its Member States to 'extend the national accounts to key aspects of Sustainable Development. The national accounts will therefore be complemented with integrated environmental-economic accounting that provides data that are fully consistent'. The development of the accounts is eagerly anticipated, since 'in the longer term it is expected that more integrated environmental, social and economic accounting will provide the basis for new top-level indicators'.

From 2000, the EEA has experimented with the computation of water accounting (EEA, 2001a, 2001b and 2001c) to test river quality accounting and analyse highly significant indicators. These developments were based on principles, established in the mid 1980s (Weber, J.-L., 1986); the hydrologically based improvements were tested in a couple of countries only, with France being one of these (Babillot, 1995).

Building hydrologically consistent water accounting to usefully address the balance between resource and uses is a very complex task. Here, the resource is the water that can be exploited by the economy at a certain place in the catchment at a certain moment in time and uses the actual abstractions, evaporation and returns in the same place at the same time. However, and even if the needs for maintaining ecosystem functions are set aside for simplifying the approach, it is not possible to estimate the resource as the sum of volumes of water in the different compartments because the intrinsic specificities of the water pathways (water flows through rivers, exchanges between soil and underground systems, multiple uses of water along a river, etc.) on the one hand and the uses as the simple sum of abstracted volumes on the other hand. At the end, there can be 'competition' between resource and uses which identification requires appropriate methodology and data to mitigate uncertainties if information and gaps in knowledge.

Following the fundamentals developed from the mid 80s and supported by different policies related to biodiversity (e.g. the EU 2010 strategy and the Millennium Ecosystem Assessment), the physical accounts were developed by the EEA with the intention of addressing new challenges and their computation carried out to check the effectiveness of the approach and the appropriateness of the existing data collection systems. .

The development of the economic analysis of the relationships between ecosystems and biodiversity (The Economics of Ecosystems and Biodiversity (TEEB)) increased ambitions of contributing to the preservation of ecosystem and natural services in the long term, by including them in the economic framework: 'Being spatially explicit is important in order to take into account the spatial heterogeneity of service flows and of the economic values that can be assigned to them … It also allows the identification of mismatches of scales as well as analysing the distributional implications of decisions that affect ecosystems and exploring trade-offs.' (de Groot et al., 2010).

These two complementary views of the Commission and TEEB reinforce the approach used by the European Environment Agency (EEA), with the active support of the Directorate-General for the Environment (DG Environment). This approach aims at being spatially explicit, so as to accurately cover the reality of systems with their physical constraints, as well as appropriately timed, so that policy-relevant information can make use of seasonal effects and time-trends. These are also the requirements for building useful indicators; since the EEA is not exclusively focused on the production of the formal accounting tables, their accounting approach targets integrated assessment capable of supporting other important environmental issues as well.

The 2012 Water Blueprint (COM/2012/0673 final) ([1]) served as an opportunity for DG Environment and the EEA to fully implement the water resource assets accounting: DG Environment hired a consultant (Poÿry) after public tendering, and the EEA provided data and information and provided technical support to the DG Environment. This report details the rationales and methodological developments that resulted, and presents two types of outcome: results proper on the one hand, and lessons in developing methodology, reference systems and data flows, on the other. The lessons point to improvements needed if water asset accounting is to form the basis for a set of 'new top-level indicators' (among other outcomes), as required by the communication mentioned above.

1.2 Main results and ancillary outcomes

Factual results and more general outcomes must be analysed under the very definition of accounting. Water accounting ([2]) is one of two ways of calculating water balances over large areas; the other is modelling. There is a fundamental difference between water accounting (and accounting for any other component of the environment as well) and modelling. Modelling is an attempt to reproduce the causal processes between different 'compartments'; accounting is placing the observations of these compartments side by side (acknowledging that the causal relationship is established), and analysing the degree to which they match.

Gaps in data sets are not expected to be reconstructed by using data from another compartment: this would breach the fundamental principle of independence of data sets in the accounting process. Hence, accounting is quite effective in identifying gaps in data sets and inconsistencies in relationships across data sets.

Consequently, the expected result is the consistency of data sets. This is a very important result since the data sets at stake are the benchmarks of policy implementation and effectiveness; the water balances, with their associated indicators, reveal the spatial and temporal structure of resources and scarcities.

The main lessons are as follows.

([1]) COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS A Blueprint to Safeguard Europe's Water Resources. http://ec.europa.eu/environment/water/blueprint/index_en.htm

([2]) In this report, the terms 'water accounts' or 'water accounting', when used without supplementary adjectives, refer to the SEEAW methodology as upgraded by the EEA in the spatial (sub-basin instead of country) and time (month instead of civil year) dimensions, and not to the simplified I/O tables derived from annual statistics.

- Making time (month) and space (sub-basin) disaggregated water balances under the System of Environmental and Economic Accounting for Water (SEEAW) enhanced methodology is technically feasible, affordable and informative. The quality of the balance has been demonstrated (Section 5.1) to hold a direct relationship to the relevance of meteorological inputs and river discharge, that are the pillars of the accounts.
- Information resulting from the assessments clearly demonstrates that water resource issues (for uses and ecological support) are extremely diversified and significant in many EU areas, not just in structural water scarce areas; hence, they call for finely tuned policies.
- The current data flows, as collected in the EEA European Environment Information and Observation Network (Eionet) flows, were not envisaged to serve the needs of water accounting. Their restructuring requires revisions both of the networking (under the Shared Environmental Information System (SEIS)) and of internal management, to address the responsibility of data collection by universe (e.g. all relevant aspects of 'urban', of which urban water issues), instead of by topic (e.g. all water uses of which urban uses). The approach by topic omits certain parts of knowledge which cannot be categorized easily.

1.2.1 Result no 1: feasibility of the asset accounts confirmed, but some data questionable

The exercise confirmed that making assets accounts at monthly and sub-basin resolutions was feasible. This may appear to be stating the obvious, but in fact, no such exercise had ever been attempted at EU level over the past 8 years (the initial 10-year target could not be achieved).

Indeed, for this first exercise, some resources had to be mobilised with a significant share of investment in making the systematic update, as a follow-up of affordable EU policy in the current economic context.

However, this systematic update demands a rather radical revision of the data collection schemes (if it is to be affordable and effective as support to other policies); in parallel, it would significantly contribute to all EEA and Commission work (particularly by offering better data for Joint Resource Centre (JRC) modelling and forecasts).

Accounts production does not allow for delivery of figures with uncertainties; in physical accounting, it is necessary to flag results based on questionable or insufficient information. The approach taken in the reported exercise is to score the essential data sets, and compare the data scoring per sub-basin to a standard reference, indicating the median data quality that may be accepted as a short-term target for data collection.

All maps are presented with the result overlaid with a special pattern that blurs the results of the areas which quality is lower than the median quality target. For reasons detailed in methodological sections, it is not possible to monitor and calculate uncertainty; this presentation of results tells however the reader on the degree of likelihood of the results presented. By contrast, summary statistics cannot take into account such quality limitations from scoring. This is summarised in Section 1.3 .

1.2.2 Result no 2: time- and space-disaggregated indicators

Robust, relevant and timely indicators are at the heart of high-level policy assessments and communication. However, the simpler the indicator is, the larger the precautions called for in its construction. Attempts to set up a revised Water Exploitation Index Plus (the WEI+) were less successful than expected, because of inconsistencies in the definition that resulted from the political process of setting it up, and the inappropriateness of data provided by the Member States.

Fortunately, it can be demonstrated that a wide set of hydrologically consistent indicators (the different avatars of WEI) can be directly produced from the accounts. A normalised WEI (nWEI) has been calculated, by assessing the actual water exploitation in the most comparable way. It represents the possibility for the economy to actually obtain the required water volumes, irrespective of whether they are returned. The indicators can be presented in two ways:

1. as statistical aggregates (e.g. annual averages) preserving the seasonal differences;
2. as statistical events (e.g. percentile X %), whose analysis explains the characteristics of water scarcity in structural, recurrent or episodic terms, hence opening the way to use the results for policy purposes.

Combining these indicators provides a spatially defined and statistically representative assessment of water exploitation at the European level. The results are presented from Chapter 5 onwards; fundamental findings are reported below.

Of 411 sub-basins, one half are in the interannual WEI average of less than 10 %; 57 (14 %) could not be computed owing to lack of essential data, in this case only outlet information. This means that at least half of the sub-basins are not under systematic water scarcity threat.

By contrast, 87 sub-basins are in the 10 % to 25 % range, meaning that (on average) 16 % of resources are at any given time incorporated into the economy, possibly reaching 15 % to 50 % of resources, with a return time of one month per year. This rate suggests possible harm to the ecosystem, without, however, suggesting significant risk of water provisioning. But since the uses are rather underestimated, this class and the basins involved are to be further examined after data revision.

The two last classes, 46 and 17 sub-basins, totalling 63, make up a percentage in number in sub-basins of between 15 % to 18 % of the total number of computed catchments, on the unlikely assumption that the non-documented basins are all equally apportioned across the classes or unproblematic.

In these basins, the average quadratic mean of monthly WEIs ranges between 36 % and 54 %, meaning resources are under a great deal of pressure. In the scarcest group, the 10 % nWEIs (those reflecting the high water period) are also very high, suggesting a structural scarcity for at least 17 % and up to 20 % of sub-basins.

The last group probably covers two categories and is likely to also comprise sub-basins, in which the scarcity is more a recurrent than a structural issue; this is suggested by the mapping of the nWEI in the next sections where geographical distribution is discussed.

1.2.3 Result no 3: information on scarcity and water use

Similarly, an indicator of net consumption has been computed (called ‘pseudo WEI+’, because it is not produced under the WEI+ process), and shows that two ([3]) groups of sub-basins present both a high interannual average (in practice ~10 % and ~20 % of resources totally consumed), and 90 % values close to 50 %, indicating structural overuse of water. On average, 16 % to 19 % of sub-basins are likely in significant overconsumption of resources, whereas 6 % to 7 % are in sharp overuse of resources.

Risk of scarcity is clearly driven both by low resources and by irregularity in resources. This factor is recognised as highly relevant, and can be addressed effectively only if reasonably long time-series, disaggregated below the season, become available. In the current exercise, the time disaggregation

([3]) The detailed analysis produces three categories, but the most consuming have been grouped together in the synthesis.

is satisfactory; the duration of the computed period had to be limited to 8 years (96 months) because of insufficient data.

During the validation phase, the representatives of countries where interannual variability is exacerbated pinpointed that 8- or 10-year periods were too brief. This is accurate and relevant especially when long-term reservoirs (underground or surface) are found only once over scores or decades.

In this instance, variability has been assessed by considering the ratio of the nWEI percentiles; higher variability patterns are evident in areas showing Mediterranean and Atlantic regimes . More detailed analyses should be carried out, considering the geological background and characteristics of groundwater systems that were not taken into account because these data could not be delivered in time for the exercise.

1.2.4 Result no 4: ancillary information and ecological flows

The ecological flow is an important driver for meeting the EU Water Framework Directive (WFD) ([4]) objectives of restoring or keeping the best-suited ecological status. Proposing any indication would lie outside the scope of the report. However, two categories of outcomes must be mentioned. Progress is possible thanks to the development of a consistent river reference system to back the water balances and the enlargement of the assessments based on reported data, thanks to the combination of this reference system with the reported data and the water accounts side-results.

First, a global result has been computed as a test indicator that represents the share of resources that remain downstream of any catchment. This can cover any return period and has reasonable frequencies of 2 % to 10 % trespassing (a share which is larger 98 % to 90 % of the time) — roughly one month per ~5 years to once a year on average. This poses a significant threat to ecological resources that are more sensitive to extreme events (in restored areas) and that are more deeply impacted in their restoration by frequent adverse events than by water supply: an ecosystem that 'dies' every five years disappears, another threatened every year cannot recover, whereas a water shortage with the same return time is compensated by exceptional measures (e.g. banning private car-washing, limiting irrigation of golf courses, etc.).

Over the period computed, one tenth of sub-basins are likely to be submitted to systematic stress; whereas ~30 more (close to 20 % in number of sub-basins) should be explored under this issue. More accurate results can be achieved with two simple supplementary actions.

1. Having better data and a longer period explored, to prepare the assessment of the appropriate 'ecological flows'.
2. Deepening the analysis with a comparison at river segment level, between the hydrological conditions (by reference to the catchment's conditions) and waterbody status. This is a very simple undertaking since all data are in the same reference system (European Catchments and Rivers Network System (ECRINS)).

1.3 Data issues: lessons learnt

Environmental accounting is possibly the most effective means to quality assess data sets. This is due to the methodological obligation to process data sets independently (to avoid any circularity)

([4]) Directive 2000/60/EC of the European Parliament and of the Council establishing a framework for the Community action in the field of water policy.

on the one hand, and to rigorously confront independent data while closely mimicking the natural cycle, on the other.

Innumerable data issues were encountered; these could only be partly addressed during the water accounting process as presented here. These issues, along with the proposed solutions, constitute one of three categories of issues calling for targeted solutions, with the central one linking all three.

1.3.1 The reference systems

Appropriate reference systems have a key role. At the moment, environmental accounting methodological principles should form a central framework of data processing for all environmental assessments related to spatial distribution.

However, only ECRINS has been developed to a point where its use is feasible, as an EEA-wide reference for surface hydrological systems. Gaps and errors remain, and conceptual developments are needed concerning canals and defluences that are essential in water conveying. These changes should form part of version 1.5, and in a few years, version 2, with geometrical accuracy closer to 1:100 K rather than 1:250 K.

Despite this, the attachment of point objects (monitoring networks, dams, pumping, etc.) is not yet a routine maintenance step. Moreover, it is clear from recent ancillary productions and despite the INSPIRE Directive (Directive 2007/2/EC) ([5]) recommendations, that the central role of the reference system to attach all these categories has not yet 'copied' in the intellectual schemes of some experts.

The acknowledged relevance of ECRINS should not conceal the critical gap represented by the insufficient development of the other irreplaceable reference systems required for environmental (not only water) accounting:

- for groundwater systems, the good example of the French BD Lisa (Base de données des limites de systèmes aquifers: aquifer's systems delineation database) should foster comparable developments and integration, hopefully with the support of EuroGeoSurveys[6] for example, with the current developments by the European Topic Centre for Spatial information and Analysis (ETC/SIA) being an intermediate step;
- bedrock and soil systems integration;
- major artefacts on land, namely the cities and their relations as spatial objects.

The interrelationships between these objects, to outreach the geographic information system (GIS)-based correspondence and achieve correspondence between identifiers ([7]), is the way to dramatically increase the productivity of assessments, as anticipated by ETC/SIA work plans in the past years.

([5]) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

[6] EuroGeoSurveys is an organisation of 33 European Geological Surveys. Our statutory aims are to address the European issues, to promote contribution of geosciences to EU affairs, to assist EU to obtain technical advice and to provide a network between the geological surveys http://www.eurogeosurveys.org/

([7]) Finding that A relates to B by GIS is long, resource-consuming and better done once, verified, and then processed as ID of A relates to ID of B. This is simple in principle, but calls for planning, organising and maintenance. Once done, the processing is increasingly speedier.

1.3.2 Improving the conceptual model of data organisation

Environmental accounting is not processing one data set; rather, it is processing numerous data sets in their spatial context and aiming to 'blend' them together. The experience from water accounting, applicable for all categories of environmental categories, suggests that improvement calls for envisaging a radical change in the data organisation paradigm: collecting data in their spatial systems, and not integrating the collected data later in their spatial containers.

This has many practical and organisational impacts. For instance, data are collected per data category (per topic) and are not driven by the universe in which data are relevant. Data uses are collected as one of the many 'water data', and not using a 'user-comprehensive approach': water used for human consumption is hence not collected from the 'city' perspective in the topic approach, and eventually, data collected in this way cannot be used for the accounting exercise. Considering water in the city, for example, the driver is the water cycle in the specific city, not water use in cities in general; collecting domestic water data as part of the water process does not provide information about cities, and water data are insufficient as well. As a result, none of the data sets collected from a topic perspective are complete, accurate and correctly usable.

Similarly, there is little information on industrial or energy production water uses, because this is not embedded into an industrial activity or energy production activity in which water is a component.

It may be considered self-evident that river-discharge data collection follows the appropriate process. This is not the case: in the reported exercise, 2 000 of 9 000 documented (with discharge values) gauging stations could not be used, since they could not be properly attached to river drains owing to insufficient placement information. Moreover, many discharge data were considered of poor quality due to not meeting the expected range of values for the basin they drain.

This highlights the need to embed all spatially related information (city, industry, gauging stations, sewage plants, etc.) in a hierarchical spatial context of time-event, spatial 'superstructure' (the location) and the 'infrastructure 'the global context of the point located': this hierarchy ensures the soundest way to quality assure the information. This assurance, again for performance in using resources and accuracy of reporting reasons, should be carried out in three steps. Exemplifying (simplified) with river discharge is self evident:

1. time series are validated by time irregularities (and when documented, with historical data);
2. flow values at stations are validated by reference to productivity at stations (catchment needed);
3. stations are validated in the basin context (forest, other stations, etc.) by reference to the spatial infrastructure.

These findings are detailed in Section 3.5

1.3.3 Data storage and management

Water accounting cannot be the outcome of processing two sets of data, i.e. of time and space variability. The very fact that rivers are individually significant and the necessity to balance results of many classes demand large data sets. These data sets are not collected just for the sake of water accounting; they have to be fully consistent with other applications, and cover a very large area (the order of magnitude is in the range of 10 million kilometres squared).

As demonstrated in the report, many data sets have to be processed from the daily resolution, to provide accurate monthly aggregates. All these data need storage space: tables and databases require up to several terabytes (TB), in contrast to MS Access® desktop databases (limited to 2 GB).

This structure has been developed as a prototype for the accounts (for example, the climate monthly data are ~36 GB ([8]) and the source discharge is ~20 GB), with the management tools allowing the operators to manipulate data.

The architecture of Water Information System for Europe (WISE)/Waterbase, used within this project, is not tailored to these developments, and is understood to serve as summary data for the general public, with all time-dependent information ranging from meteorological to uses being stored in a single MS Access® database. Currently the database is undergoing enlargement and development towards a common data structure, which captures the complex needs of the efficient integration between spatial and tabular data. This will provide a system, allowing bringing the results of water accounts not only to internal use between EEA and the Commission, but also to share it with a wider audience as part of the EEA environmental assessments.

Some developments and integration are needed to render this summary database the outcome of the aggregation process from the professional database — that must itself be completed for systematic running of the accounts.

1.3.4 Practical brakes on data flows

Improving the conceptual model of data organisation is irrelevant if no data are eventually collected. Data collection, with prior data identification and location, is an underestimated task, managed alongside 'orphan data', those essential data that are not part of any data collection process.

There are three major issues of data collection for environmental purposes, addressed in the next three sections.

Inaccurately identified data

In these data sets, data are supposedly present, but actually are missing or are not suited to the context. Most water usage data fall into this category (with the supplementary jeopardy of access restrictions). In most cases, inaccurately identified data are a result of incorrect reporting processes: the most prominent example is the European Pollutant Release and Transfer Register (E-PRTR), which provides information on industrial emissions. In fact, it contains no information on water volume, a key vector of liquid pollution.

Inaccurately identified data could be mitigated by two synergistic processes:

- since water uses have a very asymmetric distribution, identify the reference population and address the values and spatialisation using a stratified statistical approach;
- since information access is split between 'political actors' that may provide it (but cannot), and technical actors that can deliver data (but may not), create the conditions for political bodies to allow technical associations so that they provide or track information, under the conditions of the previous process.

As an Eionet main node, the EEA could foster such a development, fully in line with the already highlighted concepts of processing information by universe and not by topic.

Known data with restricted access

([8]) For other purposes, daily data have to be stored, rendering the size for 10 years in the range of 1.4 TB.

Accessibility to data in Europe and even in the EU varies. For example, for data as essential as that of river discharge (used for all environmental accounts and many assessments beyond accounts), the status ranges from fully and freely available online, to absolute restriction, in some countries even extending to restricting knowledge of where data are stored.

Another significant restriction in data access stems from privatisation of many former public services. For example, reservoir changes in volumes were publicly accessible before privatisation in meeting EU directive targets: these data are now considered 'industrial secrets' and must be reconstructed.

There are three ways, to be explored in parallel, to make essential data available for environmental accounting, and more widely for environmental assessment and support of EU objectives of sustainable development in the context of climate change and the best use of natural resources.

1. Continue the processes started for the accounts; and organise (within SEIS and Copernicus ([9])) and maintain the inclusion of essential data from those countries open to provision, while trying to convince others.
2. Use a stepwise process, under the aegis of international organisations (e.g. the World Meteorological Organization (WMO)) towards centralised data collection. An initiative to use the Global Run-off Data Centre (GRDC) for river discharge data is under way. However, as demonstrated in this report, this pathway cannot substitute direct data collection if no substantial revision of the data collection scheme is first set up by these organisations.
3. Jointly with the Commission, elaborate upgrades of the EU legislation, so that some data become part of compulsory exchanges; however, this method will not cover the EEA, whose mandate extends beyond the EU.

Orphan data

The category of orphan data clusters those data that exist and are accessible (even if lacking sufficient density) but whose use for the process requires deep and consistent specific processing. The most significant is meteorological data: rainfall, actual evapotranspiration, temperature, etc. are data essential for all environmental processes (water accounts, carbon accounts, ecosystem services, etc.). Despite this, there is no defined process to draw up these data.

The case of meteorological data serves as a good example: the development of water accounts is founded successively on three different sources.

1. Advanced Terrestrial Ecosystem Analysis and Modelling (ATEAM) data: fine spatial density but insufficient time density; discontinued in 2000, and hence no longer suited.
2. Monitoring Agricultural Resources (MARS) (JRC-sourced) data: fine time density but insufficient spatial density, with restricted accessibility; odd quality for the accounts (oriented to agriculture in plains); no longer used by the EEA from 2010.
3. The ENSEMBLES E-OBS[10] data set, obtained via the European Climate Assessment and Data (ECA&D) database: fine time density and acceptable spatial density (with some noticeable exceptions, which could be improved); odd quality (depending on the

([9]) Copernicus (not Kopernicus) is the new denomination of GMES, from December 2012 onwards.

[10] The ENSEMBLES project (contract number GOCE-CT-2003-505539) is supported by the European Commission's 6th Framework Programme as a 5 year Integrated Project from 2004-2009 under the Thematic Sub-Priority "Global Change and Ecosystems". http://www.ensembles-eu.org/

Ensembles data set); in-house modelling not planned for the next years and no alternate solution envisaged. Time series are updated regularly every six months.

Without stable and consistent meteorological data sources, the accounting cannot be continued.

River discharge data is to some extent orphan data as well, since its current organisation, as

supported by the collection in EIONET Member Countries left open issues in terms of meta data description regarding spatial integration and possible time series which gives a certain limitation to the use in the Accounts calculations.

1.3.5 Orientations

New sources become available, especially from space and global climate reanalysis. Two promising new sources of information must be mentioned: although they have not been used yet (NVDI has been checked in another context), they should probably become validation sources for the water balances, and be further integrated with carbon and ecosystem accounting:

- NVDI, resulting from red/infrared processing, has been analysed for forest assessment; it seems very promising following its processing and integration into a database server for validating soil humidity ([11]);
- The Gravity Recovery and Climate Experiment (GRACE) (GRACE, 2013) project on microgravity changes seems a reasonable framework for assessing groundwater reserve changes as well as ice caps changes — both stocks that call for more data and that should be tested (after the aquifers have been inserted as reference systems, of course).
- The Soil Moisture and Ocean Salinity (SMOS) (ESA, 2013) project. As its name suggests, the SMOS satellite was designed to measure how much moisture is held in soil and how much salt is held in the surface waters of the oceans. Data series have been available since early 2010 and may potentially be used in future.

1.4 Organisation of the report

This report aims at being as comprehensive as possible: it describes the outcomes of the full-scale realisation of water assets accounting across Europe. Water accounting is a combination of methodology and exploitation of heterogeneous data sets, and it seemed important to cover all issues related to methodological adjustments, data processing, data flows, organisation, and results, as these may contribute to policy support.

To achieve these different goals, non-essential technical insights were excluded from the main text. These insights are instead reported in the appendix section that covers methodology, reference systems, and data issues. The relevant appendices are indicated in the main text and can be read independently if required.

([11]) NVDI / NDVI, the Normalized Difference Vegetation Index (NDVI) is a simple graphical indicator that can be used to analyse remote sensing measurements and assess the photosynthetic activity. A report has been prepared by consultant (SCM: Société de Calcul Mathématique), under framework contract with the EEA.

2 Water resources assessment and water accounts

2.1 Assessment of water resources and water accounts

The EEA conducts many activities related to water resources. Obtaining the expected outputs requires a versatile system in which the data collected could be used to reach all the targets, meeting the member countries' demand to 'collect once, use many'.

The core objective of the EEA 2009–13 strategy, which drives EEA activities, is to continue to produce European, pan-European and regional environment-related data and indicator sets, integrated environmental assessments, and thematic analyses, in order to provide a sound decision basis for the EU and member countries' environmental policies. Water issues are a key component of this strategy, both as an important component of the environment and as a key EU environmental policy (the WFD, the Floods Directive (Directive 2007/60/EC) ([12]), etc.).

During 2011, in addition to its regular, ongoing production, the EEA placed extra emphasis on three major areas: the first two, 'resource efficiency, the green economy and physical ecosystem accounting' and 'implementation of new ICT to support environmental observation, monitoring, reporting and assessment', are directly connected to environmental accounting.

Resource efficiency indicators are an important issue: in 2011, detailed consideration was given to extending the indicator base beyond material flows accounting, to include energy, water, land, biodiversity and ecosystems, and economic sectors as baseline information of the environment related policies.

The need for comprehensive and targeted water resource–related information encompasses several key objectives of the EEA (text reflects the EEA Annual Management Plan 2011, but is likely to inform on the EEA objectives for the years to come as well):

- studying biodiversity-specific indicators (river fragmentation), biodiversity-related aspects of land and water accounts, and other relevant spatial assessments;
- assessment of the post-2010 biodiversity policy in relation to other policy areas, namely the WFD;
- supporting the 2012 Water Blueprint and the integration of WFD country-reporting into a common water information system (WISE); assessment of state and pressure information from the first River Basin Management Plans under the WFD (including water accounts and water economics); and evaluating results in the context and perspective of the EEA's State of the Environment (SoE)–related information;
- supporting DG Environment on European policies related to water quality and quantity, and providing regular updates on the EEA priority data flows and core set indicators;
- assessing the vulnerability and integrity of water ecosystems, groundwater and water management, and – potentially – Member States flood-mapping systems;

([12]) Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks (Text with EEA relevance).

- providing an indicator framework for the Water Scarcity and Droughts policy, including intersectoral dependencies, water pricing and costs of services, and vulnerability to climate change and adaptation; the continuation of water assets accounts will be a central methodological element underpinning such a framework, relevant to the objectives and needs of the WFD.

In all cases, there is a strong need to elaborate a sound assessment of water resources that would depict accurately European diversity without breaching the subsidiarity principle. This rather complex equilibrium between opposing constraints demanded a feedback approach that was eventually finalised by implementing adjusted principles and methods taken from the SEEAW accounting methodology, designed to be a consistent framework for using the best achievable water statistics.

EEA work focused on using the SEEAW conceptual framework to make it the overarching approach to the production of versatile water balances, so they would meet EEA needs for supporting the Commission, facilitate its mandatory assessments, and also match the SEEAW requirements related to water accounts (they comprise physical assets and uses but their structuring must relate to Eurostat developments of the economics side of the accounts). From this perspective, water statistics become one of the possible sources of information; the main sources are all the water-related data. Similarly, the water assets accounts (understood as under the SEEAW) are one of the possible outcomes from a comprehensive implementation of water and land data sets and calculation models.

The current implementation at EU level thus aims to accurately encompass the relevant aspects of river continuity and catchments occupation, and to mimic the environmental water cycle, at the most relevant time and space resolution for all reasonable assessments (water accounting and ecosystem accounting). An analytical resolution close to 1:250K, and aggregates at a resolution in the range of 1:1M or less detailed, were considered to adequately meet both constraints mentioned above. Their relation to the main sectors of the economy is considered at the level of their relationships with the environment; the hybrid accounting of water uses into the economy is not implemented, since it is expected that insufficient data will be available in future.

This widened approach is in direct follow-up to the 2003 SEEAW that expanded the 1993 SNA asset boundary to include all water assets and their quality, and explicitly identified produced assets used for mobilising water resources. This approach aims at both providing the water assets accounts and helping the contextual quality assurance of data.

In late 2011, DG Environment decided to launch an ambitious project of making water balances across the EU, using the SEEAW methodology and based on the EEA methodological developments (ECRINS) and data collection processes engaged to develop the methodology at sub-basin and monthly level. The target of DG Environment was to strongly back the 2012 Water Blueprint; its invaluable unintended effect was to foster the full-scale implementation of water balances and allow the building of a comprehensive database that was used in turn to reinforce the methodology and the production of indicators.

This exercise also provided a unique opportunity to assess the appropriateness of data flows and propose improvements. This report describes the results obtained, how the SEEAW was implemented, and which methodological, modelling and data issues had to be tackled to make this implementation happen. It sets out the requirements for making this implementation the basis for regular assessment supporting all related EEA tasks in the near future, as well as for integration into the overarching 'ecosystem services account' that may be the foundation for the future strategy of the EEA.

2.2 *Improving the capacity for assessing resource efficiency*

As with the main national accounts, the SEEA accounting framework provides a score-keeping function from which key indicators can be derived, and a management function for use in the analysis of policy options. The accounts provide a sound basis for the calculation of measures that may already be included in sets of sustainable development indicators, but may also be used to develop new indicators, such as environmentally adjusted macro-aggregates that would not otherwise be available. As with most information systems, the potential uses of the environmental accounts are greatly enhanced once a consistent and coherent time series is established. This would call for the accounts to be considered part of wider national accounts and produced on a routine basis.

2.2.1 Supporting WFD goals of sustainable water use

The water balances under the water accounts can be applied to support the implementation of sustainable water resource management across Europe in a number of ways, including the following:

1. The accounts provide the basis for estimating a Europe-wide quantification of water availability, not only at river basin scale, but also on a monthly basis. In doing so, 'hotspots' of water stress will be readily identified. By processing longer time series including the one published in hydrological annual books, the return time of water scarcity events can be assessed.
2. As a major side outcome, the actual resource can be assessed, whereas annual aggregates tend to consider all renewable water as a resource, which is far from being the case; it helps define the 'ecosystem base flow'.
3. Sectoral water use will be quantified, including cases of overexploitation. In this way, measures to address unsustainable water use can be targeted in an optimum and cost-effective manner.
4. The linkage of water availability and use information will enable the role of drought (a natural phenomenon) to be discriminated from water abstraction by economic sectors, with respect to the impacts upon availability.
5. The accounts enable assessment of scenarios of water resource management to be examined, including the impact of a range of measures. The likelihood of such scenarios depends largely of the lengths of the analysed period (see first listed item).
6. The accounts will enable water scarcity and drought indicators to be refined and improved, for example, through an improvement in their temporal and spatial scale. The same applies to other indicators like the reuse index, for instance.
7. The 'hybrid' water and economic accounts will improve understanding of the role of economics, for example, with respect to the impact of water pricing upon water abstraction and use across different sectors. However, this last input requires still more efforts in data collection compared to the present situation.
8. Last but not least, the accounting methodology helps pinpoint data gaps and inconsistencies.

2.2.2 Prerequisites for better quantifying water scarcity and drought indicators

Water resources are irregularly distributed in space and time ([13]). They are under pressure due to major population change and increased demand. Europe has a very diverse hydrological background, reflecting its varied climate and topography. In the south, there is very significant variation in flow through the year, with long and dry summers. To the west, there is less extreme variation, and in catchments underlain by absorbent aquifers, flows remain reasonably substantial throughout the year (but with noticeable exceptions). In the north and east, much precipitation falls as snow, so much flow occurs during the spring snowmelt period. Major rivers (such as the Rhine, Rhone, Po, and Danube) distribute water from the 'water tower' of the Alps. Superimposed on this varied hydrological base are a wide variety of water uses, pressures, and management approaches that include man-made reservoirs, natural lakes and transfer canals.

These varied conditions of resources are diversely used for supply: huge problems occur where and when a strong demand is exerted on limited resources. This is analysed through 'water scarcity and droughts': this covers two distinct phenomena that are ultimately combined in the water accounts framework.

- Drought is a natural hazard that cannot be prevented and occurs with a different return time. Indicators for drought are quite complex, and their description is beyond the scope of this report. Drought is defined as a sustained and regionally extensive occurrence of below-average ([14]) natural water availability. Drought affects all components of the water cycle with a deficit in soil moisture, through reduced groundwater recharge and levels, and up to low river flows or dried-up rivers. It is a reoccurring and worldwide phenomenon, with spatial and temporal characteristics that vary significantly from one region to another. Drought has wide-ranging social, environmental and economic impacts. Drought should not be confused with aridity, which is a permanent feature of a dry climate.
- Water scarcity implies a long-term imbalance of available water resources and demand. Severe water scarcity is observed when there is demand almost equal to and even larger than available resources during an intense drought period. It can also occur when the quality of resources is depleted below reasonable requirements. The most severe social consequences of scarcity are, however, found in arid or semi-arid regions where the availability of water is already low under normal conditions. However, water scarcity events are observed in wetter areas as well, especially if the demand has been set considering close-to-average resource conditions; the return time of scarcity events depends therefore on demand, the return time of the drought event and inter-seasonal storage capacities.

For the time being, there is little combined information for water resources and water demand. Hydrological and meteorological analyses provide patterns of water resources at different scales. Water scarcity events are primarily made known after they have occurred, and depending on many factors.

Drought events have occurred regularly across Europe over the last 30 years. The duration of each event, as well as the area and population affected have been variable throughout this period.

Information provided by Member States (EC, 2007) made it possible to identify **severe events** that yearly affected more than 800 000 km² of the EU territory (37 %) and 100 million inhabitants (20 %) in 1989, 1990, 1991, and more recently in 2003.

([13]) Much information has been taken from the 2009 EEA report on water resources.

([14]) The averaging is historical and is meaningful only if observations are more or less Gaussian. This is not the case in southern rivers. Using averages for calculating volume of resources may be tricky and misleading if rains are very irregular.

The EEA report on water availability issued in 2009 (EEA, 2009) mentions that one relatively straightforward indicator of the pressure or stress on freshwater resources is the Water Exploitation Index (WEI), which is calculated annually as the ratio of total freshwater abstraction to the total renewable resource, possibly at a river basin scale and from water statistics. It is suggested that some thresholds identify the degree of sustainability of water resources. This index is discussed in further sections, under the new information made available thanks to the computed EU water balances.

The main issue with the use of this readily understandable indicator is that such analysis still struggles to reflect fully the level of stress upon local water resources. This is primarily because the WEI is based on annual data and cannot, therefore, account for seasonal variations in water availability and abstraction. As a result, totally opposite situations can be reported with the same numeric value, making the information misleading. During the summer months in southern Europe, for example, agricultural and touristic water demands peak at a time when the natural water resources reach a minimum. The annual average approach of the WEI is unable to capture this, and cannot therefore fully reflect the potential threat to both human uses and the freshwater ecosystem.

Moreover, the WEI can overestimate water stress, because it does not account for the consumptive use of water. Where abstraction is dominated by power generation, for instance, nearly all the abstracted water is returned to the source. Such issues were detected by Margat (1993), who suggested a wider set of indicators: the 'water wearing' index was to take into account non-consumptive uses that nevertheless involve introduction of water into the economy, before returning it to the environment.

Water accounting, provided it is carried out at the sub-basin level and at a monthly resolution, can help answer these questions and provide a fully comparable, sound and reliable indicator of stress; if carried out over a reasonable time period, it can also indicate number of times a certain even can occur. The observed frequency can then be presented as 'probability of occurrence of event X'. The probability of occurrence of a certain event (e.g. tension for watering crops) is not just the drought return period, because of the negative feedback that is likely to occur in such case. Agricultural water use is all the more intense that the lack of rainfall drives greater abstraction in order to fulfil crop water requirements, hence lowering the available resources for other uses.

2.3 Quick summary recall of SNA and SEEA

2.3.1 Historical milestones

The System of National Accounts (SNA) deals exclusively with macroeconomics, whereas the System of Environmental and Economic Accounts (SEEA) aims at intertwining the economy and environmental components. There are a number of aggregate measures in the national accounts, most notably GDP ([15]) (the most widely used measure of aggregate economic activity in a period), disposable income, savings and investment. The very clear book edited by the Organisation for Economic Cooperation and Development (OECD) (Lequiller and Blades, 2006), which has been used as a source for this section of the report, is proposed as further reading for those wishing to

([15]) GDP is the market value of all final goods and services produced within the borders of a country in a year. Its capacity of measuring the standard of living has come under increasing criticism and many countries are actively exploring alternative means of doing so. A further criticism is that the consumption of fixed capital (assimilated to amortisation) is an incorrect estimate of the true consumption of natural capital that results in depletion of ecosystem services, including marketable services.

learn more. The development of the SNA and its international standardisation is indeed an important historical process that was fostered by the economic crisis triggered by World War II.

The original motivation for the development of national accounts and the systematic measurement of employment was the need for accurate measures of aggregate economic activity to assess the state of the economy, and hopefully undertake development of corrective measures. This task was made all the more pressing after the Great Depression began in 1929. The author of the first formal national accounting system (in 1941) is the famous economist J. M. Keynes. Alongside the British economist Richard Stone, he published the first national income statistics for the United Kingdom in the same year. Similar approaches were developed in other countries, e.g. in France during the Vichy regime. These and other events led to the implementation of national systems after the end of World War II.

A strong driver for making these accounts was the reconstruction effort, supported in western European countries by the Marshall plan, which had to be followed up by estimates of effectiveness. As a result, in 1952 and under the aegis of the Organisation for European Economic Co-operation (OEEC) (precursor of the OECD), Richard Stone published *A Standardised System of National Accounts*, endorsed and slightly adjusted by the United Nations Statistics Division (UNSD) in the 1953 SNA. The role of the United Nations is very important because international standardisation of national accounting is crucial for comparison. However, this was not yet the case at the time, since a large part of the world was ruled by communist economies with differing standards and paradigms.

A major development was the synthesis of the national accounts as input/output (I/O) tables, formalised by Wassily Leontief. The I/O tables have hence become the cornerstone of all national accounting work, including the environmental accounts.

After some time, national accountants reached a consensus that it was time to revise the 1953 SNA, to take stock of the improvement of the I/O tables and of all the suggestions proposed by the countries that implemented the SNA (France, the Netherlands, the United Kingdom and United States being leaders). In addition, the political objective shifted from *ex post* assessment to the idea of actively planning the future developments of the economy. As a result, many countries implemented different forms of 'indicative planning'. The outcome of this was the 1968 SNA, an upgraded release of the 1953 SNA that improved it without altering anything fundamental.

As a result of discussions at the annual meetings organised by the OECD for national accountants from member countries, the decision to revise the 1968 SNA was made in the early 1980s. The revision and drafting process involved more than 50 statisticians and economists. As a result, and having consulted several international agencies, the 1993 SNA was a joint publication of the OECD, Eurostat, the World Bank, the International Monetary Fund, and the United Nations.

During the revision process of the SNA, the fall of the Berlin Wall (November 1989), gave it a further boost: the countries of the former Soviet bloc that had their own system of national accounts (material products system) switched to the SNA. At the present time ([16]), only two countries have not formally adopted the 1993 version as the basis for their official national accounts: Cuba and North Korea. The United States produces accounts that are conceptually consistent with the 1993 SNA, but does not publish the same tables and groupings.

GDP is a single indicator, and its development over time is generally of strongest interest to economic policymakers, although the detailed national accounts contain a rich source of information for economic analysis, for example in the I/O tables, which show how industries

([16]) July 2012.

interact with each other in the production process. GDP is sometimes confused with GNI (gross national income), which measures the total income (excluding capital gains and losses) of all economic agents residing within the territory (households, firms and government institutions), disregarding the source of income.

2.3.2 Counting environmental resources — patrimonial accounts

When the current SNA was released in 1993, the economic accounts and environmental statistics were each being developed as independent areas subject to their own conventions and classifications. This situation has not markedly changed since that time. Economic accounting is carried out almost entirely in monetary terms, and although the economy operates within the natural environment, the inputs from the environment have until recently been seen to be 'free'. As a result, the impact of the environment on the economy has not been readily identifiable within the economic accounts. Sets of environmental data are often compiled with specific regulatory or administrative purposes in mind and, therefore, use a variety of concepts, methods, classifications and units of measurement according to the need they serve. For this reason, disparate sets of environmental statistics are generally not integrated with one another, or with sets of data relating to the economy or society.

Since 1970 and the 'Club de Rome' (The Club of Rome, 2013) publications, the feeling that economic growth could be jeopardised by the consumption of natural resources (whose stocks, rate of renewal and quantity used were not well known) has become a growing concern. Indeed, the effect of mankind's activity upon the environment has become an important policy issue throughout the last part of the twentieth century, and is now understood to be a major driver of human development in the decades to come. On the one hand, there has been growing concern about the impact of each country's economic activity upon the global and local environment. On the other hand, there has been increasing recognition that continuing economic growth and human welfare are dependent upon services provided by the environment.

Growing consumption of fixed resources cannot be sustainable: this self-evident message has been rephrased to target decision-makers. Currently, the parallel issue of monetary debt has resulted in presenting the excessive use of natural and ecosystem services as an 'ecosystem capital debt'. This presentation is all the more accurate in that the consumption weighs on stock-based resources rather than on short-term flow-based resources. The services include the provision of raw materials and energy used to produce goods and services, the provision and recycling of water, the provision of suitable conditions for food production, the absorption of waste from human activities, and the basic roles in life support and the provision of other amenities such as landscapes. The building blocks of ecosystem services have been set for the 'Millennium ecosystem assessment' (2005).

Since the late 1970s ([17]), many national accountants and environmental statisticians have become increasingly committed to appropriately assessing 'common goods'. This concern is not exclusively driven by environmental protection: market failures had been recognised by economists as a gap in economic theories, requiring that the externalities be addressed. The concept of externalities is based on the simple idea that an externality exists when a person makes a choice that affects other people who are not accounted for in the market price. For instance, a firm discharging pollution (into the atmosphere or in water) will typically not take into account the costs that this pollution imposes on others. The 'polluter pays' principle first aimed at retailing identical concurrence between economic actors, for example one plant that pollutes vs. another that purifies its wastes,

([17]) In France, for example, the Prime Minister Raymond Barre created in 1978 the Commission of accounts of the national patrimony. This commission, chaired by A. Vanoli, with J .L Weber as its secretary, published in 1986 a fundamental report on environmental accounting; its many findings inspire much of the current work at the EEA.

and not necessarily at protecting the environment; French water agencies created as a result of the 1964 Act worked for decades under this conceptual framework, and based their estimate of external cost to the averaged correction cost (the fee for pollution equilibrating the subsidies required for accelerating the construction of the wastewater treatment plant (WWTP), during a certain period of time).

However, it is difficult to compare different sets of environmental data and statistics, and equally difficult to swiftly and accurately estimate the environmental pressures of economic activity (e.g. irrigation versus water resource over time, renewable energy production versus river ecosystems, etc.). The response proposed by national accountants and statisticians was to develop a physical accounting system using the same conceptual framework as the SNA, so that the economy and the physical environment could be compared. The mirroring concept in the SEEA is that natural capital produces services, as monetary capital does in the SNA. This introduces a considerable change in paradigm: the objective is no longer to restore true concurrence but to compensate for losses of natural capital, in order to ensure sustainable development. The difference is not academic; all the environmental issues can be encapsulated in the new concept, whereas only limited monetary compensation could be ensured under the former one. This approach had been anticipated when considering the possible roles of the French water agencies, whose approach to charging for pollution and water use has significantly evolved during the five decades of their enforcement (Kaczmarek, 2006).

2.3.3 The building blocks of the SEEA

The SEEA is largely inspired by national accounts and statistical offices habits: its key vocabulary and concepts inherited the specific language and concepts that based the SNA. The SNA is older, and for more than 60 years it has driven the key economic indicators; the need for an internationally agreed accounting system for monetary flows in the economy has long been accepted, and the SNA has been widely used in most countries of the world for many years.

Figures such as GDP and national income derive from the SNA. The fact that they are derived from an internationally recognised standard helps to ensure not only their comparability across countries, but also their credibility. One goal of the designing of the SEEA is precisely to start to establish the same international acceptance of environmental accounting.

The 2003 SEEA was a rather close transposition of the SNA to the environmental system: it defined the building blocks of the accounts from a narrow economic perspective. This attitude may have changed since that time; regarding water, the acceptation of 'assets' in the implementation is indeed wider, as is shown in the following sections.

The fundamentals of the SEEA are the following four categories of accounts.

a) **Physical flow accounts:** the flows expressed in physical terms (e.g. m^3/year) are between and within the environment and economy. They can belong to four categories: products, natural resources, ecosystem inputs, and residuals. Physical flow accounts consist in aggregating accounts for products, natural resources, ecosystem inputs and residuals, each account being expressed in terms of supply to the economy and use by the economy. The accounts in this category also show how flows of data in physical and monetary terms can be combined to produce so-called 'hybrid' flow accounts. Emissions accounts for greenhouse gases are an example of the type included in this category.

b) **SNA flow accounts, *sensu lato*:** this category is purely economic and takes those elements of the existing SNA that are relevant to the good management of the environment, and shows how the environment-related transactions can be made more explicit. For this, it is necessary to go beyond the supply and use tables, and examine the whole system of the SNA flow accounts, including income distribution and redistribution. This includes environmental protection and resource management, along with environmental taxes and permits to use environmental resources.

c) **Environmental assets:** these are measured in physical and monetary terms. Timber stock accounts showing opening and closing timber balances and the related changes over the course of an accounting period are an example. These accounts include how natural resources contribute to income and the measure of wealth in the national balance sheet. The current developments deal with the water resource assets that are a subset of this specific category and take a part of the physical flow accounts, since there is no practical possibility of decoupling stocks and flows in water accounting.

d) **Valuation and environmental adjustments:** this final category of SEEA considers how the existing SNA might be adjusted to account for the impact of the economy on the environment. Three sorts of adjustments are considered: those relating to depletion, those concerning so-called defensive expenditures, and those relating to degradation. This category, despite many efforts, is still conceptually and practically more provisional than the rest of the system.

2.3.4 Scope of physical flows accounts in the SEEA

In the 2003 SEEA, it is clearly stated that to a large extent, physical flow accounts can be disconnected from the economic part of the accounts, provided the classifications used do not contradict future integration. Working in physical terms does not in most cases require in-depth knowledge of economic accounting. For this reason, both the production and the use of physical flow accounts may be more accessible to those who are orientated more towards the natural sciences than towards economics. While it is usually possible to compile data in physical terms without the corresponding economic values, compiling monetary accounts is facilitated by a foundation in physical accounts.

However complex the process is, it is generally possible (in principle), to establish a balance, because the first law of thermodynamics states that matter (mass or energy) is neither created nor destroyed by any physical transformation process, whether of production or of consumption. This law provides a proven basic principle for a physical bookkeeping system for the consistent and comprehensive recording of inputs, outputs and material accumulation.

Following the material balance principle, the physical flow accounts are constructed in such a way that net material accumulation is equal to the excess of total inputs over total outputs. This identity may be expressed in terms of inputs and outputs, or in terms of supply and use, and eventually combined to opening and closing stocks. It also presupposes a direction of flows: any outflow from one area (and to a period) is a candidate inflow to another area (and period) and vice versa. A particular point of interest here is the flows between the economic and environmental spheres, to fuel the other accounting categories. However, a fine-tuning of the flows within the environment and between the environment and the economy is a prerequisite for addressing accurately all physical flows, and not only those that are understood at one moment to capture those related to a substance supplied to the economy.

A key issue in the implementation of the accounts is to overcome the formal limitations first included in the 2003 SEEA. It is stated that 'the development of environmental accounts is closely linked to the concepts embodied in the SNA. As such, environmental accounts are most suitably compiled for national areas on an annual basis. Environmental issues that are seasonal (such as shortages of water in the summer) or local (such as a reduction in air quality in a particular location) do not lend themselves easily to analysis in the accounts. Although quarterly and regional accounts are feasible in theory, in practice few countries have the data from which to compile such accounts.' (UNSD, 2003).

This last sentence has long been used to justify the production of physical accounts at the country-level at annual resolution. The outcome of many discussions and tests is that physical accounting should be carried out at the time and space resolutions that best reflect the environmental mechanisms of the system to be accounted. For example, water cycles are deeply driven by seasonality and catchment characteristics; water asset accounting must consider these features. This is why the EEA decided to carry out these accounts at the monthly and sub-basin levels. By contrast, land take is not marked by seasonality; hence land take accounts are based on inter-annual comparisons.

However, final aggregation at country and year levels to meet monetary standards must be feasible. Indeed, seasonal and regional issues are not easily related to monetary accounting, but this is because money and water don't follow the same algebraic system,.

The major challenge is to stick to the spirit and targets of the SEEA and produce water-resource asset accounts that provide relevant and non-misleading outputs that also meet SEEA requirements (see item (b) in Subsection 2.3.3 .

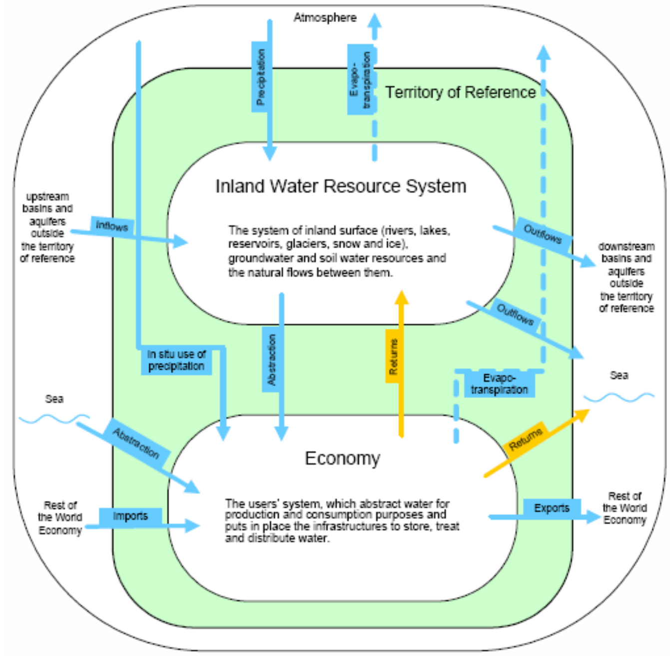
2.4 Mimicking the water cycle and water balance constraints

2.4.1 The SEEAW conceptual model

The summary results presented in the next sections in this report meet SEEAW (EEA, 2009) requirements. Before displaying the results, these requirements have to be presented for the reader to contextualise the information reported.

The SEEAW summarises its framework as displayed in Figure 2.1: the economy, the system of water resources, and their interactions. In the SEEAW, the economy and the inland water resource system of the reporting area — named the 'territory of reference' — are represented in the figure as two separate boxes. The inland water resource system of a territory is composed of all water resources in the territory (surface water, groundwater and soil water), and the natural flows between them. The economy of a territory consists of resident water users who abstract water for production and consumption purposes, and put in place the infrastructure to store, treat, distribute and discharge water. The inland water system and the economy are further elaborated in Figure 2.2 in order to describe the main flows within each system and the interactions between the two systems that drive the final tables' structure.

Figure 2.1 Flows between the economy and the environment



Source: SEEAW (EEA, 2009), Figure 2.1.

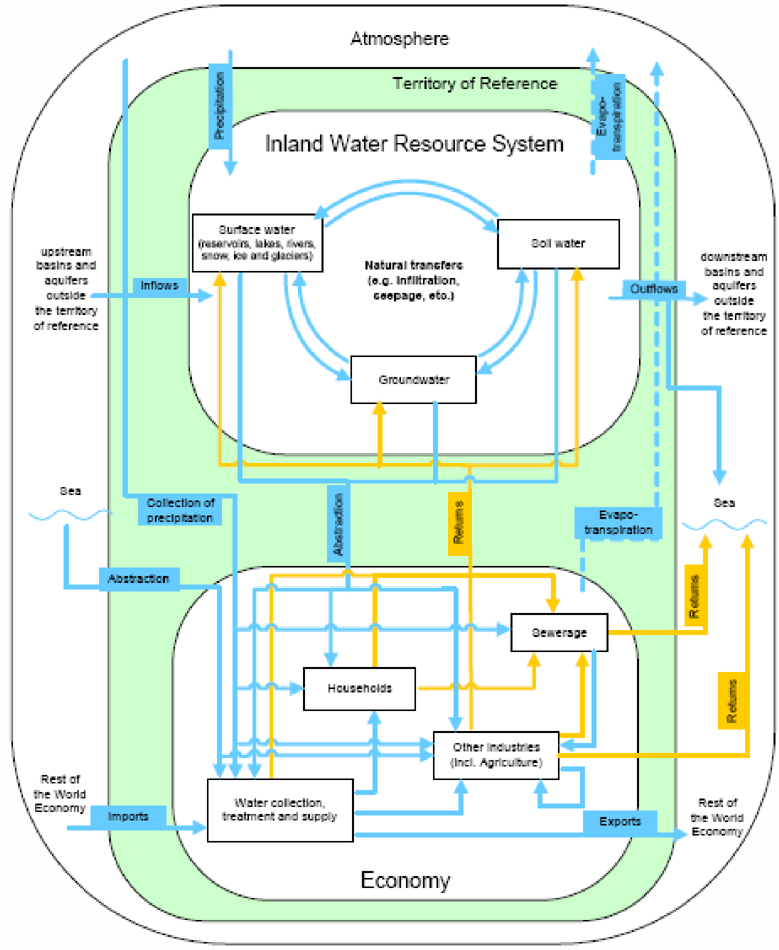
Following the SEEAW, the inland water resource system and the economy of a given territory – which can be a country, an administrative region or river basin – can exchange water with any component of other territories through imports/exports of water (exchanges of water between economies), and through inflows from upstream territories and outflows to downstream territories (exchanges of water between inland water systems). Figure 2.1 also shows exchanges with the sea and the atmosphere, which are considered outside the inland water resource system. These flows are also captured in the SEEAW accounting framework.

The atmosphere and the sea are both source and sinks for water resources that are considered outside the accounting balance; hence there are no opening and closing stocks for either atmosphere or sea.

The second figure proposed by the SEEAW and shown here in Figure 2.2 indicates the different compartments of each subsystem and their interrelationships as they should be reflected in the I/O tables. The three main compartments of the inland water resource system ([18]) are surface waters (rivers, lakes, artificial reservoirs, snow, ice and glaciers), soil water, and groundwater within the territory of reference. These resources are hence considered at the resolution of the territory of reference. The question of whether the territory of reference is an analytical unit or a reporting unit remains unclear; the response to this question, however, is crucial, and has practical consequences that are discussed in later sections. For implementation at EU level, it is considered a reporting rather than an analytical unit.

([18]) The acronym IWRS is intentionally omitted here: it is confused with the *International Recommendations for Water Statistics* (IRWS), the brochure produced under the direction of the United Nations Statistic Division (UNSD), and to guide the compilation of statistics and the production of the water accounts to IWRM (integrated water resource management). IRWS is hence mentioned as IR-Wat. Stats.

Figure 2.2 Main flows within the inland water resource system and the economy



Source: SEEAW (EEA, 2009), Figure 2.1.

The SEEAW accounting framework cannot be implemented as such. Dealing with flows of water, it has to define the analytical resolution in time and space. There are important questions behind the implementation: Can this volume be abstracted from this territory and be supplied to that segment of the economy? Can this volume be abstracted at this moment in time and be provided?

The rationales behind this are implicitly indicated in the SEEAW: 'The inland water resource system is composed of: (a) all inland water resources from which water is, or **can be ([19])**, abstracted; (b) water exchanges between water resources within the territory of reference (e.g. infiltration, runoff, percolation); and (c) water exchanges with water resources of other territories (i.e. inflows, outflows). Exchanges of water between the water resources are also referred to as natural transfers.' (EEA, 2009). The important wording of 'can be' calls for the physical and temporal possibility of water abstraction to be considered in the implementation.

There is no simple solution to this important question (further discussed later, in Subsection 2.4.2 . The response to this question requires a sound definition of the inland water resource system, the spatial allocation of the economy within the territory of reference, and eventually, a definition of the operational territory of reference.

([19]) Emphasis added to the original text, to open the possibility of addressing only those exploitable resources rather than total resources. The question of what resource can be exploited raises very difficult technical and cost issues. Simple example is: is a flood a resource? The fraction of flood volume that is stored is a resource; the volume that cannot be stored is not.

In the EEA implementation, inflows and outflows are considered only between other territories of reference or between the sea and atmosphere. The atmosphere is considered one indiscriminate compartment from which rainfall is provided to the appropriate statistical unit of the territory of reference and from which evaporation subtracts resources.

The focus of water accounting is on the interactions between water resources and the economy, where the economy is considered the system that abstracts water for consumption and production activities, and puts in place the infrastructure to mobilise, store, treat, distribute and return water into the environment.

The economy compartment is at the same time extremely detailed (SEEA designers are first and foremost statisticians) and a bit ambiguous. Considering Figure 2.2 alone, the economy exchanges water volumes through imports and exports from the same territory of reference, the sea or the 'rest of the world'; should this be understood as a lumped source or recipient? This ambiguity is reinforced by the fact that the inland water resource system is explicitly indicated as importing or exporting resources from upstream or downstream systems respectively, hence explicitly understating a geographical location, topologically related to the current inland water resource system.

In conclusion, the asset accounts module of the SEEAW describes the inland water resource system in terms of stocks and flows: it provides information on the stocks of water resources at the beginning and end of the accounting period and the changes therein. These changes are described in terms of flows brought about by the economy and by natural processes. Asset accounts can therefore be thought of as a formalised description, in accounting terms, of the hydrological water balance. Hence, the assets accounts should be produced as a standardised representation of hydrological balance, computed as close as possible to reality, and displayed in the accounting matrix. In the implementation scheme, this second definition does not cover exactly the same concepts, even though the result is practically identical.

There is, however, an important issue, not discussed by the SEEAW approach, that does not explicitly take into account the major difference in the physical possibilities of counting monetary items and physical items with the same spatial and time resolutions.

This difference is that while money can be transferred virtually, water cannot. Similarly, money can be loaned, while water cannot. These apparently simple differences have huge consequences in practically implementing the water balances under the SEEAW methodology. Water accounting is probably the system where the algebraic differences between monetary accounts and physical accounts are the greatest, including the hierarchical systems of flows in the catchment systems and the limited meaning of stocks in water accounting (flows represent in most water systems several times the stocks, contrasting, for example, with land or carbon accounting[20]).

At the moment, stocks are just a fraction of annual flows; seasonality matters.

Water resources are hence also described in the SEEAW in terms of their quality. Quality accounts describe the quality of the stocks of water at the beginning and end of the accounting period and, if effectively implemented, can trace the path of changes; this is how they were tested at the EEA in 2000 ([21]) (EEA, 2001b and 2001c), and also how they have been produced from 2010 onwards. Quality can be defined in terms of one pollutant, a combination of them, or in terms of physical characteristics (e.g. salinity level) of water. In new developments, quality is defined as 'exergy' that

[20] There are always exceptions: a small endorheic lake, a confined aquifer, etc.
([21]) Only internal working reports were drafted and summarised in different presentations.

combines in a single indicator both quantitative resource and some fundamentals in water quality (Carrasquer, 2012; Valero et al., 2007).

2.4.2 Compromising the appropriate time resolution

The evidence provided in the above sections (indicators for water scarcity and droughts, water resource accounting proper, water quality accounting, etc.) strongly suggested that the elements of water balance should be captured at a monthly time resolution, and that computations should distinguish the analytical level of the statistical units and their restitution level (the territory of reference). This later distinction is not addressed in the IR-WAT. STATS because the IR-WAT. STATS, as its name indicates, is based on existing statistics. The use of existing statistics to carry out water-resource assets accounting is not, according to the experiments carried out at the EEA, the best way to obtain the expected results.

The reasons for this are as follows.

- Water resource statistics are necessarily based on monitored data that are collected over a largely varying time resolution; moreover, the resource is expected to vary more intensively over seasons than needs.
- Computed statistics related to uses are (by definition of public statistics) aggregated over legal time (e.g. year) and space (e.g. country) – levels of resolution that are not necessarily those suited to the purpose of water accounting implementation at the EEA. Moreover, if the population considered by the statistics is very asymmetrical or oddly distributed in space, the legal aggregates provide figures that are not easily usable.

There is a more fundamental rationale behind this. When accounts address large stocks that are modified by relatively small flows, the annual periodicity is not an issue. By contrast, water accounts are characterised by a wide range of relative size of flows and stocks: rivers have small stocks, renewed over days or weeks; lakes may be renewed over weeks or years (in exceptional cases, centuries); groundwater is renewed every couple of months (sandy aquifers in granite, karstic aquifers, for example) to centuries (large chalk aquifers), and in exceptional cases, several millennia ([22]). Ice and snow have residence times of a couple of months to years and centuries respectively.

When considering the largest volumes involved in accounts, the flows widely predominate over the stocks, and this is a second argument for assessing the changes in stocks over a time congruent with flows. The month is the soundest affordable compromise.

The question of usage data will be discussed later. For water resources, it is sounder and more effective to build the resource figures needed to populate the I/O tables of the assets accounts, from monitored information and with appropriate extending of these data on the relevant reference system.

This difficulty is mentioned in the SEEAW, although for erroneous reasons: 'Yearly accounts often hide potential seasonal variability of water use and supply as well as of availability of water resources in the environment. Ideally, quarterly water accounts would be useful in the analysis of intra-annual variations. They are, however, very data-demanding and thus are often not considered a feasible option'. (UNSD, 2007).

([22]) The Albian green-sands aquifer beneath Paris has an estimated residence time of between 30 000 and 100 000 years. By contrast, the average residence time of water in oceans is only 2 500 years (de Marsily, 1995).

The position of the EEA is this: preferring a scientifically questionable time resolution because it is possibly easier to achieve or calls for less data is not an acceptable option, especially when this preference is not backed by any evidence or testing

When implementing the water assets accounts at the EEA, the issues related to the most appropriate compromise of time and space resolution were considered in detail, and scientifically acceptable and practically implementable responses were found. The conclusion of the previous section is that the spatial (state) and time (year) resolutions that are implicitly recommended for water accounting analytical levels are not sound and do not meet the hydrological requirements. By contrast, the compilation of the accounts should allow such restitution, thanks to the appropriate aggregation method.

The issues related to time resolution can be addressed relatively easily. The selected **time resolution is calendar month,** the best achievable compromise between expected outcomes and data provision. Monthly time resolution has the following advantages.

- It allows any aggregate (quarters, calendar years or hydrological years being the most likely to be demanded). Months are suited to quality accounting and to exergy analysis.
- It allows building the statistical distribution of the findings, thus permitting detailed assessment of pressures and threats and in-depth understanding of the 'exploitable resource'.
- It is both short enough to allow seasonal analysis, and long enough to require only hydrological (balance) models to compute essential variables (shorter time resolutions would require hydraulic (transient/process) models to take into account transient situations).
- The most sensitive data ([23]) are river run-off and meteorological data, whose availability stretches from days to months. Other data, for example usage data, can be rather simply apportioned from yearly (or model-estimated) and monthly values.

The single disadvantage of using the calendar month is that computations are 12 times longer and voluminous than when using the year as the time resolution, although this is now of secondary importance, considering the low price of powerful computers and storage options ([24]). The appropriate organisation of applications also cuts costs in terms of not requiring personnel. Moreover, the data collection issue must be resolved in a wider scope, considering all other applications and the benefit of consistent accounts. Summarising, the issue of annual time pace for the water resource accounts is no longer an issue. In contrast, there are still some difficulties in addressing the usage data; not only regarding the time resolution of existing data, but in the data's existence or availability.

2.4.3 Space resolution: adjusting the 'statistical units' concept for the environment

Space resolution poses a more difficult problem that raises questions related to the central notion of statistical units as defined in the SEEAW, and analytical units versus restitution units ('territories of reference') as completed for implementation at the EEA level. The technical developments were developed and eventually published in 2012 after many improvements in the 'international

([23]) See the data discussion in later sections.
([24]) Now below EUR 10/GB/year.

recommendations for water statistics' ([25]) (UNSD, 2012); hence, further citations may refer to non-final reports.

'Statistical units' were formally introduced in the IR-WAT. STATS; in contrast, the term 'statistical unit' is not precisely defined or even mentioned in the overarching documents ([26]). It is extremely important: '… to address the definition and classification of **statistical units** as they relate to the collection, compilation, analysis and dissemination of water statistics. A **statistical unit** is the entity about which information is sought and for which statistics are ultimately compiled. It is the unit at the base of statistical aggregates and to which tabulated data refer'. (Margat, 1993).

This definition is not directly operative for implementation and potentially presents a contradiction ('statistics ultimately compiled' on the one hand, and 'base of statistical aggregates' on the other hand). The level at which data is collected varies from one class of data to another, and the final 'compilation' is very unlikely to be at the same level as that at which data have been collected. Based on this, and on the practical consideration that a statistical unit should be defined as any element of the system for which **a single value would be representative of the unit at a certain moment in time**, the EEA proposed mitigating the IR-WAT. STATS to allow sounder implementation of the water assets resource accounts, while still permitting simplified deliveries so as to remain in line with the recommendations.

In conclusion, the working definition of the statistical unit is 'the elementary piece of the system being represented by a single value at a moment in time (a volume for example), the aggregation of all statistical units making the compilation of the accounts at the chosen reporting level, called the territory of reference'. This in turn calls for a definition of the appropriate delineation of statistical units, as discussed below.

The need for a workable and accurate definition of natural statistical units and their delineation thanks to an accurate and comprehensive reference hydrographical system is implicitly stated in the definition of stocks for rivers*:* 'To keep consistency with the other water resources, the stock level of a river should be measured as the volume of the active riverbed determined on the basis of the geographic profile of the riverbed and the water level. This quantity is usually very small compared to the total stocks of water resources and the annual flows of rivers. However, the river profile and the water depth are important indicators for environmental and economic considerations. There might be cases, however, in which the stocks of river may not be meaningful, either because the rate of the flow is very high or because the profile of the riverbed changes constantly due to topographic conditions. In these circumstances, computing the stock of rivers is not realistic and can be omitted from the accounts'. (UNSD, 2007).

However, for any hydrologist, the 'stock of [water in a] river' is meaningless: what could be the stock on the Danube as a whole, for example? The stock must be either computed at the statistical unit level, or compiled at the territory of reference level (by aggregation of all statistical units in this territory of reference).

The latest release of the IR-WAT. STATS has indeed taken into consideration the remarks issued by the EEA. The statistical units of the environment are now understood as those that 'may be observation units or analytical units, but not reporting units. For example, a lake can be a statistical unit, but any information about the lake will have to be reported by a unit of the economy that owns, manages or monitors the lake or part thereof.' (Margat, 1993). This definition is easily exemplified, and the operational statistical units are explained and shown per category in Table 2.1.

([25]) IRWS is an acronym that may lead to ambiguities; it is mentioned as IR-WAT. Stats in this report.
([26]) The term is found once, in Section 5.156 in the SEEA 2003 document, but it has neither definition nor scope.

Table 2.1 Components of surface waters and corresponding statistical units

Category	IR-WAT. STATS definition	EEA statistical unit
Elementary catchment	Omitted	The functional elementary catchment (FEC) is the elementary area of land, participating to a river catchment, for which rainfall and evaporation are computed; by extension, both values are broken down per land-use category if required. The functional elementary catchment is the host for soil water.
Lakes	Depressions in the earth's surface occupied by bodies of standing water. Lakes generally contain large bodies of standing water, but also include small and shallow waterbodies such as ponds and lagoons.	Standing volume of the lake at month M, monthly inflow and outflow, inputs from rain and evaporation. From the accounting perspective, a natural lake is as well a storage where water is regularly available. Lakes and reservoirs are considered equally, which might be not the case under strict understanding of the SEEAW.
Artificial reservoirs	Man-made reservoirs used for storage, regulation, and control of water resources.	Standing volume of the reservoir lake at month M, monthly inflow and outflow, inputs from rain and evaporation. A reservoir is storage where water is regularly available.
Rivers and streams	Consist of channels where water flows continuously or periodically.	River segment within any **elementary catchment.** A river segment can be characterised by a discharge averaged per month and has a volume, and hence a possible stock.
Wetlands	Areas of marsh, fen, peat land, swamp or shallow water, permanently, intermittently or seasonally saturated with water.	Applies only to those waterbodies not recorded as lakes. In this case, the set of wetlands is grouped per elementary catchment, and attached values are the standing volume of the lumped waterbody, monthly inflow and outflow, and inputs from rain and evaporation.
Glaciers	Accumulation of ice of atmospheric origin, generally moving slowly on land over a long period. These include ice sheets, ice caps, ice fields, mountain glaciers, valley glaciers and cirque glaciers.	Ice stock per elementary catchment (or lumped catchments at the glacier unit), standing volume, inputs and outputs. The melting period is taken into consideration (a glacier is storage where water availability is postponed and lasts years).
Snow and ice	Areas where seasonal or permanent layers of snow and ice form on the grounds surface.	Snow stock per elementary catchment (or lumped catchments at the glacier unit), standing volume, inputs and outputs. The melting period is taken into consideration (snow is storage where water availability is postponed and does not last years).

The SEEAW manual makes it explicitly and precisely clear what data should populate the tables. The SEEAW asset classification of water resources consists of the following categories, under the EA.13 'Water Resources (measured in cubic metres)' key entry:

1. EA.131 Surface water , of which:
    a. EA.1311 Artificial reservoirs
    b. EA.1312 Lakes
    c. EA.1313 Rivers and streams
    d. EA.1314 Glaciers, snow and ice
2. EA.132 Groundwater
3. EA.133 Soil water.

The SEEAW asset classification expands the 2003 SEEA classification by including the categories EA.1314 Glaciers, snow and ice, and EA.133 Soil water. While the 2003 SEEA acknowledges the importance of these resources in terms of flows, it does not include them in the asset classification, because they represent only a temporary storage of water. This argument is correct if accounts are analysed at the yearly level, but void if the data are analysed monthly: snow and ice constitute temporary natural storage of water that makes up a volume becoming a resource with a time shift.

It is important to note that a change in time resolution is not just a multiplication of values; it impacts the meaning and relevance of elements considered ([27]). The explicit inclusion of glaciers, snow, ice and soil water in the SEEAW asset classification reflects the increasing importance of these resources in terms of temporary stocks (in particular soil water), and also allows for a clearer representation of water exchanges between water resources, while making it possible to draw an accurate picture of seasonal effects on assets.

Glaciers are included in the asset classification even though their stock levels are not significantly affected by human abstraction. The melt derived from glaciers often sustains river flow in dry months and contributes to water peaks. Moreover, monitoring glacier stocks is also important for monitoring climate change. The main difference (from the accounting perspective) is that snow is water storage over months, whilst glaciers are water storage over decades and possibly centuries.

The issue of artificial reservoirs has been intensely debated; these objects were originally considered as economic objects and not considered to be part of the environment. The EEA position was that this distinction between economic objects and parts of the environment was problematic (considering, for example, the huge number of semi-natural lakes/semi-artificial reservoirs) and would complicate calculations for no reason. Quite surprisingly, this issue has not yet been finally resolved. In the meantime, the ambiguous conclusion is that 'the discussion on whether to consider water in a reservoir as a produced asset has not yet concluded. For this reason, the SEEAW has retained the classification of the SEEA-2003' ([28]), leading to the non-operational recommendation that 'the present situation is that while the wall of the reservoir (or dam wall) is part of the economy, the water behind it is not. Until the matter is resolved, it is recommended to separately identify artificial reservoirs from other surface water resources, and countries may choose to adopt a presentation of data items that does not show artificial reservoirs as part of the environment.'

This recommendation is not fully operational because not all lakes can be identified as artificial or natural; the question also remains of how to determine the relative percentage of natural volume and artificial volume, with some very old artificial lakes no longer being operative. There is hence a much scope for uncertainty in this. The only information that can be assumed with some degree of confidence is the volume abstracted or diverted, and the volumes released in the environment for ensuring abstraction capability along river courses. In this later case, however, the notions of import/export from other territories (between river statistical units) is extremely uncertain, and possibly outside the possibility of being accurately addressed. In cases where the reservoir empties into a long canal that in turn supplies many users, the proper assignment of economy and natural assets is contingent upon the availability of appropriate data, which are often missing or not accessible.

Soil is the primary recipient for rainfall, and the water content of soil is an important compartment in the water cycle. Soil water is defined as water suspended in the uppermost belt of soil, or in the zone of aeration near the ground surface, that can be discharged into the atmosphere by

([27]) The impact of the change in scale (time, space) is much wider that the apparent gain/loss in precision; it makes a change in nature. Incidentally, the argument that snow is a temporary resource should have also excluded temporary rivers and lakes that are dry a part of the year or over years, and are therefore temporary assets as well.
([28]) SEEAW, Section 6.24.

evaporation of soil water and transpiration from plants that take up soil water. An important use of this water is agricultural production (i.e. rain-fed agriculture).

From the accounting perspective, the IR-WAT. STATS (Section 3.13) states: 'The soil containing water and the area it occupies could be considered a statistical unit of the environment, but this is not necessary for these recommendations (considering that soils can be a relevant statistical unit for the environment where monthly data are available, for example for agricultural areas. Such data are often unavailable or difficult to collect and compile over large areas.).'

The recommendations are ultimately quite contradictory, and reflect the difficulties of considering this issue. As a statistical unit, soil and its area cannot be by any means a calculable statistical unit: the area is fuzzy and the relationship with the soil's statistical units with other statistical units beyond reasonable assessment. In the current implementation, this difficulty has been mitigated by creating a lumped soil compartment by elementary catchment that makes as many statistical units, congruent with river segments inside each elementary catchment. Important information that could be derived from soil functional maps is the potential soil reserve, to derive indicators.

Aquifers are underground zones that contain sufficient saturated permeable material to yield significant quantities of water to wells and springs. It is important to note that aquifers receive water from surface waterbodies (e.g. rivers, reservoirs and snow fields) and precipitation that infiltrates into the ground (from soil water), and from other parts of the groundwater system such as aquitards ([29]). For the purposes of water statistics, only the groundwater in aquifers is considered, as only this water can be used. An aquifer is a porous or fractured geological formation capable of storing water in quantities large enough to allow economic quantities of water. The water table level (top of the saturated zone) may be hundreds of metres below the historical water table level in the same aquifer in the case of overexploitation. The water table level depth is a good indicator of changes in stocks in the aquifer, but can hardly be related to accurate volumes ([30]).

Since the notion of aquifers is to some extent dependent on the economic value of the water, the actual stock depends on abstraction equipment and cannot be measured; however, in some cases, a rough estimate may be computed. Aquifers are candidate statistical units of the environment and may be classified according to depth (e.g. shallow or deep) or as being unconfined or confined.

The distinction between unconfined and confined aquifers is more relevant when attempting to apportion them into elementary statistical units.

Unconfined aquifers are bounded below by an aquitard and have no overlying confining layer. Their replenishing area is the ground above and they are used thanks to wells drilled in the same ground area. Depending on the size of the aquifer, statistical units can be created quite arbitrarily, depending, for example, on the flowing directions if the aquifer contributes to different surface catchments.

A confined aquifer is bounded above and below by an aquitard. Being bounded above by an aquitard, their main replenishing area is not the ground above and may be very distant. By contrast, the usage area is more likely around the area of highest pressure. Defining the appropriate

---

([29]) 'Aquitard' (etymologically, that which 'slows water') is defined in the IR-WAT. STATS, reporting the FAO definition that contains a typing error making the definition erroneous. The FAO definition of an aquitard is 'a geological unit that is relatively impermeable over a short time-frame. The unit may be permeable enough to transmit water in significant quantities when viewed over large areas and long periods, but the hydraulic conductivity of an **aquifer** (erroneous: should read aquitard) is low enough to typically act as a 'floor' for the groundwater table (FAO, 2013)

([30]) For two different reasons: first because porosity (and the effective porosity) is poorly documented at the EEA level, and second because the monitored level may be much lower than the actual table level if intensely exploited.

apportionment as statistical units and their relationships with the other compartments of the water cycle is not currently feasible at the EEA area level.

Hence, for the time being, the water asset accounts implemented do not apportion aquifers into statistical units; this is postponed until a reasonably accurate map of aquifers is related to the surface water system (see Subsection 4.2.4).

2.4.4 Nomenclature of statistical units related to economic entities

The statistical units as they relate to economic activities are clearly addressed in the IR-WAT. STATS and need no further development.

In terms of practical computation, the situation is quite different. As discussed in Section 4.3 the currently, envisaged affordable data collection processes cannot extract the economic information at the resolution of the statistical units (as defined in the IR-WAT. STATS) over the EEA area for the different details of the IR-WAT. STATS nomenclature of economic activities. This impossibility is more structural than economic: a larger allocation of resources would not in itself resolve the issue.

In the implementation of the accounts, the nomenclature of activities used and populated is presented in Annex 1 6.1

2.4.5 Supply and uses tables

The 'water supply' is an economic activity that requires infrastructures. Hence, it is not to be considered as the ecosystem services that make this supply possible by transforming the effective rainfall into water suited to human use. In the SEEA, water supply is hence considered as the **active extraction of water from resources and its further inclusion into the economic cycle.** Water supply requires water availability in the exploited resource; this reciprocal element is not defined in the SEEA, where the availability is largely the ecosystem service.

The supply and use tables are built on the same pattern as the assets tables. They comprise two distinct tables: the supply and use proper, making explicit, at the selected space and time resolution, the source used to provide water to an economic sector, sorted by International Standard Industrial Classification (ISIC) category.

The information presented in SEEAW Table 3.1 (physical use from the environment, in Figure 2.3) is indeed part of the resource assets accounts under the EEA understanding, because water use by economic sectors and the return flows have direct impacts on the resource. The appropriate computation of data from Tables 6.1, 6.2 and 3.1 is the way to build the WEI ([31]) indicators, scarcity and drought indicators, etc.

Water in oceans, seas ([32]) and the atmosphere is not part of the nomenclature, because the stocks of these resources are enormous compared to the abstraction. These assets, in general, do not incur depletion. Water in oceans, seas and the atmosphere is recorded in the accounts only in terms of abstracted water.

([31]) The relevance of WEI indicators as classically computed is discussed in another section.

([32]) In the EEA implementation that follows international nomenclatures of seas and lakes, a sea is a subdivision of a global ocean, notwithstanding the given name or the salt contents. For example, the 'Caspian Sea' is a lake, and hence should be part of the assets under EA.1312.

However, the physical supply and use tables record the following: (a) water abstracted from and returned into the sea (in the case, for example, of abstraction of sea water for cooling purposes or for desalination); (b) the precipitation directly used by the economy (in the case, for example, of rainwater harvesting); and (c) evaporation and evapotranspiration, which occur within the economic sphere as part of water consumption.

The asset accounts record the following: (a) water flowing into oceans and sea (outflows from rivers and, in some circumstances from groundwater); (b) water vaporised and evapotranspired ([33]) from water resources; and (c) precipitation into water resources (flow from the atmosphere into the inland water resources ([34])).

Figure 2.3 Facsimile of the physical supply and use table for water, as demanded by the SEEAW 2007

Table 3.1: Standard physical supply and use tables for water

Physical use table

Physical units

		Industries (by ISIC categories)							Households	Rest of the world	Total
		1	2-33, 41-43	35	36	37	38,39, 45-99	Total			
From the environment	U1 - total abstraction (=a.1+a.2=b.1+b.2):										
	a.1- Abstraction for own use										
	a.2- Abstraction for distribution										
	b.1- From water resources:										
	Surface water										
	Groundwater										
	Soil water										
	b.2- From other sources										
	Collection of precipitation										
	Abstraction from the sea										
Within the economy	U2 - Use of water received from other economic units										
U=U1+U2 - Total use of water											

Physical supply table

Physical units

		Industries (by ISIC categories)							Households	Rest of the world	Total
		1	2-33, 41-43	35	36	37	38,39, 45-99	Total			
Within the economy	S1 - Supply of water to other economic units										
	of which: Reused water										
	Wastewater to sewerage										
To the environment	S2 - Total returns (= d.1+d.2)										
	d.1- To water resources										
	Surface water										
	Groundwater										
	Soil water										
	d.2- To other sources (e.g. sea water)										
S - Total supply of water (= S1+S2)											
Consumption (U - S)											

Note: Grey cells indicate zero entries by definition.

Source: SEEAW manual.

The two latter points pose specific difficulties, which are addressed thanks to the appropriate reference hydrographical system used for computing the accounts (see Subsection 3.2.2 , page 44).

([33]) Evapotranspiration (ET) is the sum of evaporation and plant transpiration from the Earth's land surface to the atmosphere.

([34]) This may pose some problems of accuracy. In many cases, the achievable rainfall data are calculated after removal of the leaf interception (because of water not reaching the ground). When used to populate cells related to water surfaces, there is a significant underestimation, since no interception occurs in this case.

Figure 2.4 Facsimile of the matrix of flows of water within the economy: exchange between agents, as demanded by the 2007 SEEAW

Table 3.2: Matrix of transfers of water within the economy

Physical units

User ⇒ / Supplier ⇓		Industries (by ISIC categories)							Households	Rest of the world	Total
		1	2-33, 41-43	35	36	37	38,39, 45-99	Total			
Industries (by ISIC categories)	1										
	2-33, 41-43										
	35										
	36										
	37										
	38,39, 45-99										
	Total										
Households											
Rest of the world											
Total											

Source: SEEAW manual.

By contrast, the use of water within the economy poses the problem of data sources and of the capacity of inserting 'the exchanges between agents' that monitors the share of water received from other economic units as part of the usage in the receiving sector.[35] For the time being, this share is unknown, and hence this is a source of potential bias in the results of any WEI computed without the reused data figures.

The information requested in the part of the table named physical supply table is indeed found purely in the economic sector, and is not addressed in this report when the supply within the economy is involved. By contrast, there is a strong likelihood that the bottom part of the table can be populated, at least for some cells (e.g. cooling water from open systems' return to the surface of sea; sewage returns to the surface, most non-evaporated irrigation water returns to soil, then to groundwater/surface, etc.).

2.4.6 Environmental assets tables

The standard table for asset accounts for water resources is presented in Figure 2.5. The columns refer to the water resources as specified in the asset classification (Table 2.1, Subsection 2.4.3), and the rows describe in detail the level of the stocks and the changes therein due to economic activities and natural processes.

[35] For example: cooling water from thermal plant used to warm greenhouses; tap water from city system used in slaughterhouses, etc. the volumes, in practice can be estimated only from monetary flows between sectors, and in most cases the statistical units the establishment) is not known.

Figure 2.5 Facsimile of the assets accounts table, as demanded by the SEEAW 2007

Table 6.1: Asset accounts

physical units

		EA.131 Surface water						
		EA.1311 Reservoirs	EA.1312 Lakes	EA.1313 Rivers	EA.1314 Snow, Ice and Glaciers	EA.132 Groundwater	EA.133 Soil water	Total
Opening Stocks								
Changes due to human activities	Abstraction							
	of which Sustainable use							
	Returns from the economy							
Changes due to natural processes	Precipitation							
	Inflows							
	from upstream territories							
	from other resources in the territory							
	Evaporation/Actual evapotranspiration							
	Outflows							
	to downstream territories							
	to the sea							
	to other resources in the territory							
Other changes in volume								
Closing Stocks								

There is an underestimated difficulty in comparing water uses and natural assets that relates to the way assets can be measured from monitoring. When the water use represents a substantial share of the natural resource, the monitored river discharge for example reflects this use (the observed figures are roughly equal to natural resources minus use). In most cases, the final resource is estimated by (or at least calibrated against) the river outflow at catchment's outlet, otherwise double accounting of resource could be calculated. Because the many uncertainties in the respective values of resource and uses volumes, some possible mismatches may result from the calculation process. This possible mismatch becomes very sensitive when computing at a monthly level. In such cases, the abstractions may become significantly larger than the current apparent resource, just because the actual abstraction is reflected in the measurement of the resource and because the large uncertainties in data.

This apparent gap is directly reflected in the computation of the WEI that is demand/resource. In such cases, the calculated WEI may tend to the infinite (division by ~0). The consequences are discussed in another chapter, alongside an exploration of the meaning and interpretation of different WEI computation methods.

Exchanges of water between water resources are also described in more detail in a separate table (Table 6.2 in the 2007 SEEAW), displayed here in Figure 2.6. This table, which expands the information in rows 4.b and 7.c of Table 6.1, displayed in Figure 2.5, provides information on the origin and destination of flows between the water resources of a territory of reference, allowing for a better understanding of the exchanges of water between resources.

This table is also useful for the calculation of internal renewable water resources and for reducing the risk of double counting when assessing this indicator separately for surface water and groundwater due to the water exchanges between these resources (as taken from the FAO-AQUASTAT database (FAO, 2013a), for example). The exchanges between water resources assist in identifying the contribution of groundwater to the surface flow as well as the recharge of aquifers by surface run-off, or at the very least, in verifying the likelihood of reported volumes.

Figure 2.6 Facsimile of the matrix of flows between water resources, as demanded by the SEEAW 2007

Table 6.1: Asset accounts

physical units

		EA 131 Surface water				EA 132 Groundwater	EA 133 Soil water	Total
		EA 1311 Reservoirs	EA 1312 Lakes	EA 1313 Rivers	EA 1314 Snow, Ice and Glaciers			
Opening Stocks								
Changes due to human activities	Abstraction							
	of which Sustainable use							
	Returns from the economy							
Changes due to natural processes	Precipitation							
	Inflows							
	from upstream territories							
	from other resources in the territory							
	Evaporation/Actual evapotranspiration							
	Outflows							
	to downstream territories							
	to the sea							
	to other resources in the territory							
Other changes in volume								
Closing Stocks								

Building the accounts tables is a step-by-step process that can be improved by assimilating more data with better accuracy. There are two different paradigms in making the accounts.

- The historical method, based on comprehensive statistics, with each cell eventually being filled with compiled and verified statistics whose precision and accuracy are likely to be good, but whose relevance, timeliness and spatial resolution are likely to be insufficient; most (if not all) national accounts published in past years were carried out under this paradigm.
- The method proposed in the EEA implementation. Here, the data assimilation aims at providing a data set that is as comprehensive as possible, fully relevant, and time adjusted at the expense of reconstructing missing data where, when and for those sectors where the data collection scheme is not operative. Some significant uncertainties are expected, especially for those essential data for which no affordable (or even existing) source has been found.

This second paradigm is more likely to fulfil, after a breaking-in and improvement period, the full objectives of the accounts. Implementation carried out under this paradigm calls for the essential data (that cannot be substituted) and the data that can be surrogated by modelled data (under the strict rules that apply to model data for accounting purposes, mentioned in the introduction) to be considered separately. In this section, only data that are time dependent or related to works or measures are addressed. Structural information, such as soil field retention capacity is not targeted, having already been collected once with the reference system.

3 Key features of the Water Asset Accounts in the EU/EEA area

3.1 Area and period of computations

3.1.1 Purpose of computations

DG Environment (unit D1:‘ Protection of Water Resources) has set up a tendering procedure for the computation of water balances under the SEEAW, using the ECRINS reference layer for the EU-27 to the consortium led by Pöyry[36] (with Vito and SCM). Their role is to carry out the calculation of water resource assets accounts (WRAA), once they have homogenised all natural data, and domestic and agricultural data. The target is hence to produce balances, including abstraction and returns at the exception of exchanges between economic agents. The role of Vito is to reconstruct and provide uses in relation to energy and industry. The role of SCM is to bridge data gaps (time series and spatial gaps). The work was performed in close interaction with and with the permanent support of the EEA, whose role was to collect, prepare and provide all relevant data sets, including climatic and referential data.

These computations were intended to support the BP 2012 assessments, and to provide homogeneous ancillary data to those consultants working on the desertification issues and on case studies related to improvement of water management in test areas.

This full-scale implementation is a very important step in implementing the water assets accounts computation procedure, and in analysing the methodological lessons, data flows, gaps and improvements, along with the information resulting on water resource issues at EU-27 level.

3.1.2 Computation areas and time period covered

The computation area covers the catchments included in the EU-27, plus the areas necessary for connecting these catchments. For example, Serbia and Croatia are not part of the EU-27; however, since the Danube flows in these countries from and to EU-27 countries, relevant areas of these non-EU countries have been considered in the computations as well. By contrast, because of the lack of resources and difficulties in obtaining data, the Turkish catchments have not been considered, despite falling inside the EEA area. Hence, data reconstruction in this very important area has not been fully developed either.

It was jointly agreed to use the Ecrins Functional Elementary Catchments (FECs) (details in Section 6.2 page 117 and thereafter) as statistical units, and the natural sub-basins (designated as ‘SB’ within ECRINS) as territories of reference. The functional river basin districts (FRBDs) that are aggregates of SB are used as well.

The monthly time step has been systematically used over the period from 2000 to 2010. Possible data gaps have restricted the set of produced data in some areas. In the end, only eight years could be fully computed, because some data gaps were too large to be adequately bridged.

---

[36] Following restructuration in capital, the Pöyry component that carried out the project is now named Naldeo.

3.1.3 Units used in the water accounting

For reasons of consistency, the same units are used systematically in all the water accounting processes. These units are selected both for the sake of keeping with convention and for practical considerations of readability.

- Water volumes: specific coefficients are expressed in m^3/t (metric ton of produced material) or litre/person, to match standard knowledge; river discharge is in m^3/s (or per period); all volumes are in million m^3, noted in hm^3 to meet international MKS standards (1 hm^3= cube of 100 m edge). Very large volumes can be expressed in km^3 (= 1 billion m^3 = 1 000 hm^3).
- Rainfall and evaporation: mm (millimetre), one mm/m^2= 1 litre and 1 000 mm/km^2= 1 hm^3.
- Areas are generally expressed in km^2 (= 1 million m^2); however, to match some usual coefficients, for example irrigation, ha (=10 000 m^2 = 1/100 km^2) can be used instead.
- Specific coefficients are expressed in their usual unit and in this case are explicitly informed.

3.2 Water accounts application

3.2.1 Calculation requirements and organisation

The production of WRAA at the European level requires a combination and aggregation of data produced by radically different processes, from local to regional level, and from natural processes to water uses, so that the final information is spatially comprehensive, statistically representative, thematically consistent and politically relevant, which means it also covers socioeconomic aspects.

To be politically relevant, the water accounts must meet short-notice deadlines created by the political agenda. This time pressure prevents fully fledged data collection and data validation to be implemented. To that end, existing data flows were used, that were developed within other process then WRAA.

Many existing data required for accounting are not stored in databases and sometimes are just not accessible. Nevertheless, the challenge is to produce a full EU-wide set of water volumes over a period covering several years. The way to prepare fully fledged accounts is to construct the full population of statistical units and to populate each of the individuals with the best achievable data for the accounting categories that are present in the statistical unit. Once this task has been completed, compilation of accounts at the different required aggregation level can be carried out.

This process may seem exceedingly complicated. In practice, this is the only reproducible (and hence refutable) way to integrate all data (collected, reconstructed or just estimated) and eventually process the data sets through systematic algorithms that can be checked, validated and possibly corrected.

This challenge can only be met thanks to the use of a structured application system that is as operational as possible and is capable of providing spatially comprehensive results, despite many time and space gaps in data sets, with acceptable statistical representativeness. A secondary objective in the development of the application was to avoid excessive data collection where redundancy is not a guarantee of data quality. For example, ideally, changes in lake and reservoir volumes should be recorded and entered to assess storages and deliveries. However, the effort involved in getting these data was excessive and likely to provide very odd data sets ([37]) compared to the alternative method of computing the change in reserve from differences in river discharge.

These considerations led to separating data sets that contain 'irreplaceable data' (i.e. data that cannot be surrogated), from those that contain 'replaceable data'. The biggest efforts were hence devoted to minimising the quantity of missing irreplaceable data, and where gaps persisted, to reconstruct these data with the most affordable appropriate techniques.

The application used for water accounting is Nopolu System2, developed by Pöyry consultants. This application was developed in MS Access®, so as to be highly flexible, and was improved and tested in a similar context across the EEA area (in 2000-2001 and since 2006). It has been precisely designed to handle large data sets (thanks to the twin possibility of accessing local databases or tables stored as remote server data sets) and capable of handling data with gaps, thanks to modules of specific data processing and reconstructing.

Because the ECRINS reference system was developed and produced as personal geodatabases that are MS Access® databases and that it had been previously designed to process topologically consistent river systems, it was immediately compatible with the ECRINS river reference system.

From an application perspective, the water accounts production consists in arranging volumes from different management sub-applications and creating fully attributed tables, so that the final I/O tables can be compiled. The development of Nopolu System2 as an organiser of modular applications made it possible to focus on separate tasks just having to manage the appropriate identifiers in each sub-task to reconcile data at the end. More than 15 different applications handle the different data clusters plus the reference system. They are set out in greater detail in Annex 1 (page 101).

The 'water accounts application' is conceptually a 'query handler and manager' that handles data stored in tables (that may belong to other applications: hydrologic modelling, industrial water uses, etc.) and gets the homogenised data through queries coded in SQL. Each query makes explicit the relationship between one line (source) and one column (recipient) of any matrix to be populated as standard SEEA tables by a structured coding of the name of the query (or queries) that harvests the data.

In practice, this concept has been implemented in a more end-user–friendly approach and is currently managed by explicit menus (see Annex 1 Section 4).

Making the procedures for the tables to be queried is the major task at hand, and these tasks are detailed in the next sections and in the annexes. Systematising and automating the accounts has naturally imposed the use of the FEC as the elementary catchment statistical unit; without this, neither water transfer nor upper aggregation could be calculated. This requirement emphasises the importance of the hydrographical reference system, which must be topologically consistent. Of course, the aggregation at non-hydrographical levels is subject to uncertainty (sharing of catchments between countries).

In practice, several procedures complete the process, since populating the final tables is not merely a retrieval process; it requires some degree of scenario-building to mitigate data defects and uncertainties. However, the final process of making the accounts can be carried out rather independently, per category of data; it eventually consolidates all different data sets into the final

---

([37]) Many countries deny access to changes in reservoir volumes, for different reasons (Électricité de France (EdF), for example). Where data can be accessed, this is generally at the dam's management level rather than being centralised, with the exception of Portugal.

requested tables. Some ancillary post-processes have been developed as such until they are finally accepted and can be incorporated as end-user procedures.

3.2.2 The hydrographical reference system ECRINS

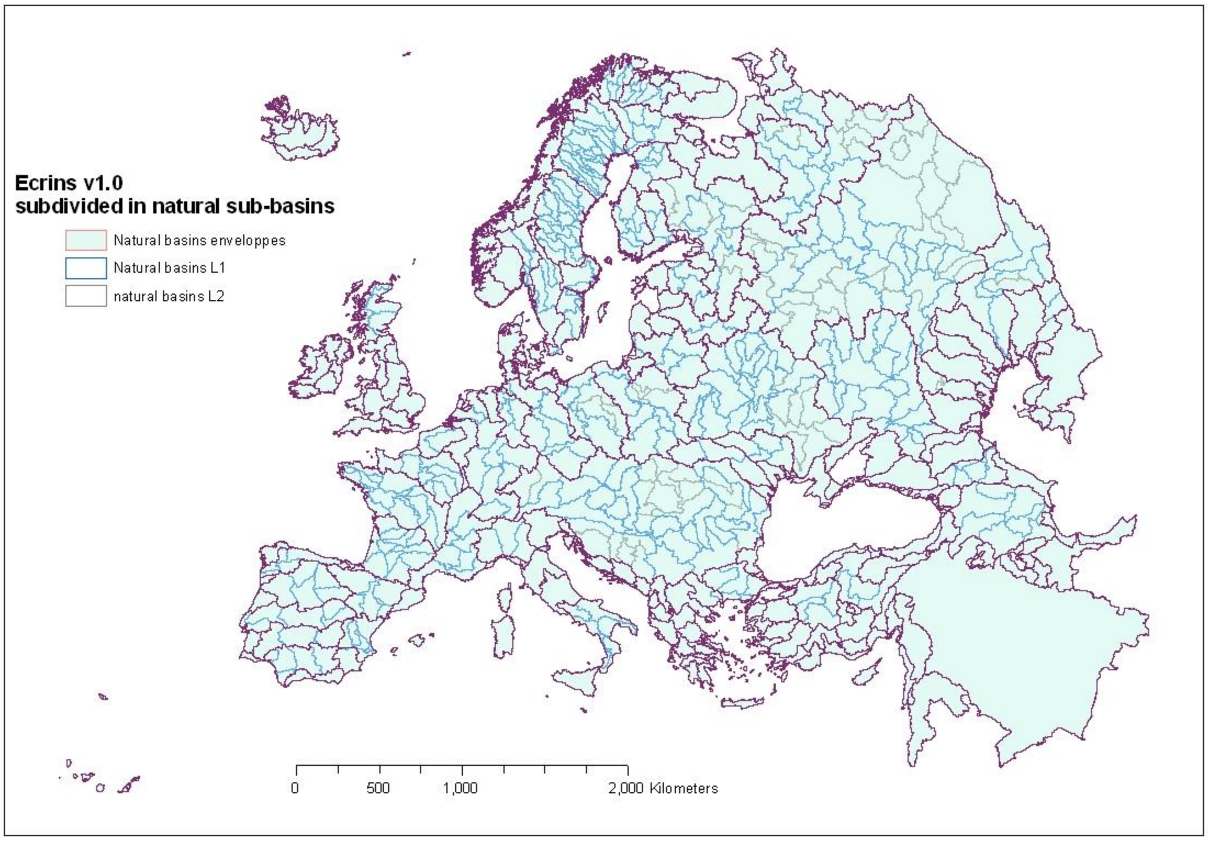
An appropriate hydrographical reference system is a prerequisite for implementing the resource assets accounts in such a way that firstly, regular production at the relevant time and scale resolutions can be carried out, and secondly, water assets accounts are made part of a wider continental water integrated assessment system.

In 2008, the EEA took the initiative of developing a Europe-wide system, ECRINS, from the Catchment Characterisation and Modelling (CCM) (Vogt, 2003) produced by the JRC and complementary sources of information. The first version, released in mid-2009 (v0/Beta) was used for the first trial implementation of the accounts. A new version (V1.0), improved in some detail identifiers although not y changed in its objectives, was produced and used for the EU-wide accounts exercise. This system is now ready for general use on the EEA data service (EEA, 2012). It will be upgraded to version 2 (eliminating most remaining topological errors and enhancing geometrical accuracy) in the coming years, based on the RDA project outcomes ([38]).

The major input (with relevance to the accounts) of the ECRINS v2 is the drawing and connecting of defluences and canals that are not processed in ECRINS v1. ECRINS v1 is composed of catchments and drains (the name 'river' is used when a set of drains is given a river identifier (ID) and name), lakes (natural and artificial) and dams whose topology is fully implemented. It is organised from a layer of 181 071 'functional elementary catchments (FECs)' whose average size is ~62 km^2, fully connected with explicit ID relationships and upstream areas. Catchments are grouped into different categories of entities, including the sub-basins and the river basin districts (actual and functional to meet hydrographical continuity) that are relevant for the purpose of the accounts.

The catchments are also built according to their sea shore of emptying to meet Marine Strategy delineations. The algorithm for making of catchments ensures that no FEC is split between different shores, is organisation. A 'shore' is the edge of catchments between sea limits. Since se delineations and hydrography result from independent mechanisms, there is some possible slight differences in the official shores and the limit of FECs.

([38]) RDA refers to the 'Initial GMES service for Geospatial Reference Data Access'. It comprises an access system and the production of data elevation model at a 30 m resolution, and revision of the 'CCM2-like' data sets used to build ECRINS.



Map 3.1 Sub-basins delineation in ECRINS v1, embedded into the river basin districts

Notes: the sub-basin delineation is the version 0 used for calculations. Italy and some others are not perfect and were revised; for reasons of consistency and recomputing costs, this was not justified where no data is present, corrections have not been entered.

Source: EEA processing.

Catchments are drained by 1 348 163 river segments, sorted as 'main drains' (connecting the FECs together) and secondary drains (internal to a FEC). River segments mimic the natural drainage, however, fulfilling the topological constraint of '0, 1 or 2 upstream, single or 0 downstream'. Each segment is populated with distance to the sea, to ease further processing. They are connected to elementary catchments and nodes documented with altitude. The topological constraints inhibit both defluences ([39]) and deltas. Featuring for the management of flows through defluences and deltas is not yet implemented in ECRINS. This featuring is planned for the next version, along with canals insertion. Segments are also documented with a fully populated 'dummy river code' that earmarks each segment with the most distant to the outlet in each drainage basin and, wherever possible, with a 'true river' ID based on river naming.

The consultant (Pöyry) has added an attribute indicating the catchment area at the segment level, which was made possible, but not implemented. This attribute shall form part of the next version.

([39]) A defluent is the reciprocal of a confluent: if a river flow splits into two or more branches for a long distance, the de-branching reaches are defluents. The same rule will probably be kept in ECRINS v2, but defluences will be identified as such and the data model revised to allow their processing, namely to address the unsolved issue of defining a catchment when defluences occur. A delta is a special case of defluence, where river empties to the final recipient, making it possible to have land-locked deltas.

A layer of lakes and dams has been elaborated. Lakes polygons (70 755) are taken primarily from Corine Land Cover (CLC) checked against the EuroGeographics lakes (ERM ([40]) v3), the WFD's Article 13, and in some cases from the source CCM 'water layer'. River segment constituting inlets (can be many) and outlet (must be unique) to each lake are set with the segment ID, and where relevant, the dam making the lake is attached and documented. All lakes whose depths and volume were found have been updated (see Annex 5 , page 150).

FECs constitute the elementary analytical area and the elementary statistical units for rainfall, evaporation and modelling areal abstractions and returns.

The sub-basin is the 'territory of reference'. Since sub-basins are fully apportioning FRBDs ([41]), it is also possible to provide account aggregates at the equivalent of the river basin district.

ECRINS v1 data sets are fully and publicly available and downloadable from the EEA website (EEA, 2012). Specific featuring and processing required by the water accounts procedure are reported in ECRINS processing in Annex 2 pages 125 and thereafter.

3.2.3 Groundwater systems

Groundwater is one of the major inter-seasonal storage compartments dealt with by the accounts, with lakes and reservoirs on the one hand, and ice and snow on the other hand. Groundwater systems are probably the biggest natural storage systems of water (inter-seasonal and inter-annual in many cases). The major difference is that groundwater management ideally requires a reference system of aquifers, whereas lakes are assimilated to individual statistical units in topological relationship with rivers, and that ice and snow are avatars of rainfall over the FECs, catchment statistical units.

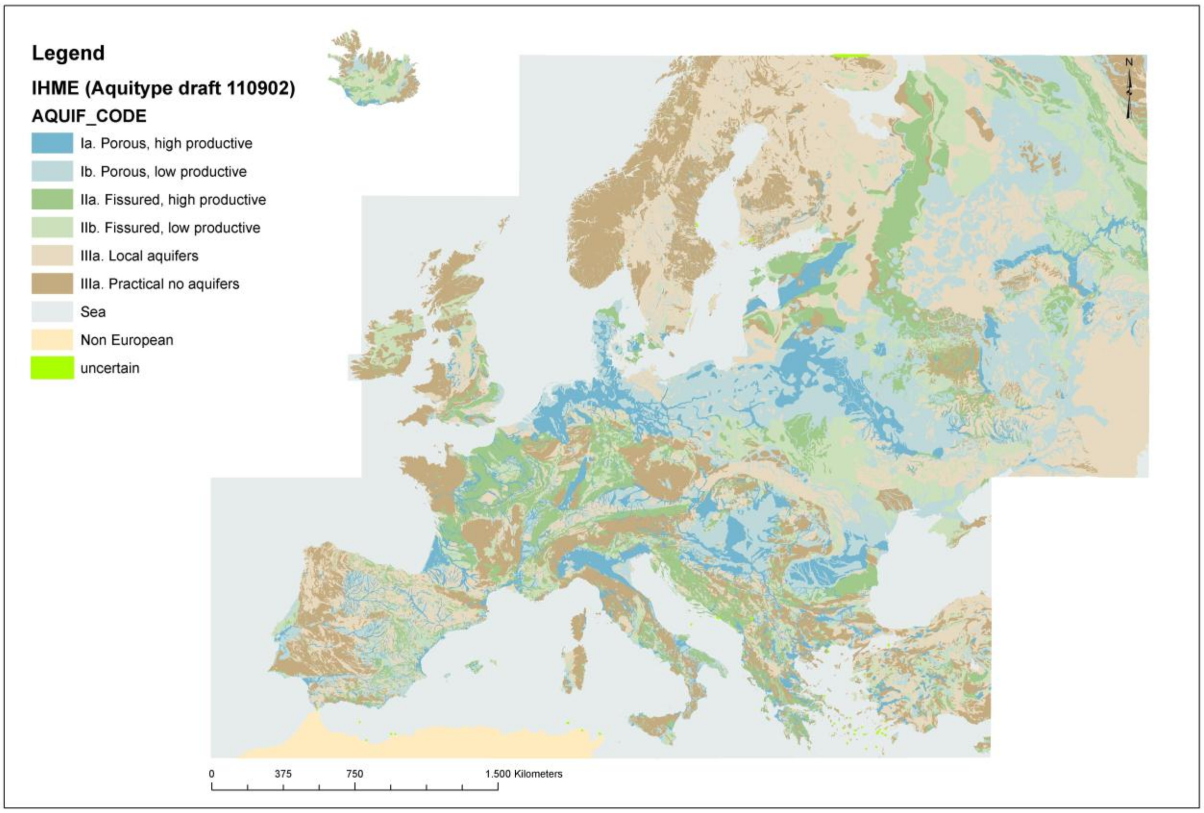
The SEEAW prescribes analysing the exchanges of groundwater with soil and rivers. In the real hydrology, the relationships are quite complex and exchanges occur in all directions. In the simplified water cycle, two major exchanges should be considered, at least: soil to groundwater, and groundwater to rivers.

Similarly, groundwater is in some countries the major (or even the exclusive) source of water supply. Groundwater is hence a subsystem of the water balance, having an opening stock and a closing stock tuned by exchange flows. The convention is, however, that no precipitation directly reaches groundwater, which makes soil water a compulsory intermediate in this case.

From the water balance tables' point of view, groundwater systems are fed by soil water and exchange water to and from rivers and lakes, and can be a source of abstractions.

([40]) Following licensing constraints, the EEA is not permitted to provide licensed polygons from ERM to the public. However, it is possible to check a lake, if from CLC it is recorded as a lake in a geographical data set.
([41]) Where river basin districts exist, of course. FRBDs are the cluster of FECs making a river basin district, corrected for hydrological inconsistencies resulting from the political delineation of RBDs.

Map 3.2 Sketch display of the used map of aquifers (IHME draft version)



From the water accounts calculation procedure, this means there are several prerequisites that cannot be adequately met or can only be only partly met.

- Groundwater systems as geometrically defined entities do not exist at the European level; there are country systems, far too detailed for the accounts, that are not accessible in most cases. The EEA had discussed this with EuroGeoSurveys some years ago, but it was not possible to finalise the production of a simplified map.
- The EEA developed a map from the German Federal Institute for Geosciences and Natural Resources (BGR) 'Transboundary aquifers in Europe (IHME)' and thanks to support from the BGR, it was made into a vector map with no mismatching information ([42]).
- With the support of the ETC/SIA and the geological map of Europe, the EEA made a GIS map of simplified aquifers (just as containers). However, the large gaps in documentation concerning bedrock prevented the production of a groundwater GIS with the minimum featuring for the water accounts: porosity and depth, making it possible to compute a possible storage volume. These gaps cannot be bridged for the time being, and the source, the International Hydrogeological Map of Europe (IHME), is not public ([43]).

([42]) Source map is a BMP scanning of the manually assembled maps, comprising for example roads, city names, etc. that are as many 'pollutants'.

([43]) The BGR expert commented: 'The restrictions on the draft version, which you received last year, are still in force. The draft version comprises errors and false reproduction. Therefore it is not for publishing or to be shared with any third parties. You may depict the draft version as a small-scale picture though.'

However, the restrictions apply to dissemination, not to aggregated use, as was the case for the accounts.

- The transfer between soil and groundwater was carried out thanks to previous EU research (the SUrface water/GroundwAter contRibution index (SUGAR) programme); however, the absence of capacities makes this transfer a bit 'blind'.
- The transfer between aquifers and rivers (using depletion curves in receiving rivers, for example) has been postponed because allocated resources had to be diverted to the reconstruction of missing discharge values.
- Aquifers as a source of water are unknown in all cases, either from volumes of place or of abstraction. Allocation between surface and groundwater has hence been carried out from a statistical source, usually from data at the country or river basin-district levels, neither of which are relevant for the purpose, hence generating much uncertainty.

This data source has been deeply processed and adjusted to the catchment limits. The final map was not finalised at the time of drafting this report. Details can be found in Annex 6 , page 153.

3.3 *Outputs and indicators*

The SEEAW standard tables are the normal output of the process. They have been presented in the appropriate figures (Figure 2.3, Figure 2.4, Figure 2.5 and Figure 2.6). However, much side information and possible indicators can be derived from the databases. These specific 'ancillary' outcomes from the process have been at the heart of the assessments carried out at the opportunity of the BP 2012 process.

Moreover, the very process of making the accounts at a monthly resolution raised awareness of the new possibility to deeply revise the very scope and definition of commonly agreed indicators.

This section deals with these new possibilities that might become as important as the classic I/O tables.

3.3.1 Standard outputs from the disaggregated accounting process

First and foremost, the calculations have produced a systematic value of extrapolated monthly discharge for all the river segments of the calculable areas and abstractions and returns computed, at least at the resolution of the territory of reference. Almost all historically defined indicators can be derived from the calculations with enhanced computations. No hypothesis is made about the accurateness of results, and this section deals only with the relevance of indicators, how they can be improved, and their values derived from the data computed by the application.

3.3.2 The historically defined indicators

Hydrology is a science with a long history and water is a vital fluid; hence several indices related to water uses and availability have been proposed and tested. One of the father-developers of the water accounts, J. Margat, had previously collected, analysed and proposed 11 different indicators, of significant relevance for water-scarce countries. These indicators were first recalled in an unpublished Bureau de Recherches Géologiques et Minières (BRGM) internal report (that can nevertheless be read online (Margat, 1993). These indicators were further published with special regard for the clients: Plan Bleu and FAO (Margat, 1996). However, being analysed first for water-scarce countries where long renewal time groundwater constitutes the main resource, these indicators are set for annual estimates and reporting.

The main interest of these indicators lies in the fact that they cover very fundamental aspects of water availability and use, and that all could be easily computed from the database of the water accounts and address many questions. Their summary list is set out below.

Table 3.1 Summary of questions and indicator story-telling

Question posed	Proposed response and indicator proposed
Quality of the knowledge used	1. Validity of hydrological basis: is the source of data for making further indicators reliable?
Characteristics of the natural water resource	2. Density of internal resource **(resource per unit of area)**
	3. Concentration index of the resource **(is the resource concentrated or distributed over a territory?)**
	4. Regularity index of the resource **(is the resource available over time or at certain moments only?)**
	5. Independence of the reference territory **(is the resource more or less dependent on external (= upstream) territories?)**
	6. Freedom of action index **(can the territory freely exploit its resource or should it share with downstream neighbours?)**
	7. Resource per capita (or capita per unit of resource)
Human pressure on the resource	8. Exploitation index (ratio: **all abstracted resources/renewable resources)**: currently targeted by the WEI
	9. Consumption index (ratio: **consumed resource/renewable resource)**
Human pressure on the quality of the resource	10. Water resource wearing **(how many times is a single volume being entered and returned from activities?)**
	11. Water sanitation and purification index

Source: summarised from Margat (1996).

Amongst these indices, the most famous is the 'exploitation index', now known as the 'Water Exploitation Index', which poses many unsolved questions. It is analysed in greater detail in the next section.

However, most of the suggestions made by Margat are very important: for example, the indicators of independence and freedom for action are essential in scarcity situations and in assessing upstream versus downstream dependency (as in most international catchments).

3.3.3 Critical analysis of the way of computing the many 'water exploitation index'

The EEA and EU's currently published indicator is the **Water Exploitation Index** (WEI) (c.f. bullet 8 in the Table 3.1). It is defined as the sum of all abstracted resources divided by the sum of renewable resources (over a long period and for a certain area). Under the source definition of this index, evaporation from reservoirs is part of the abstraction, since this loss holds a direct relationship to the usage of water.

The computation of this seemingly simple indicator poses important questions: 'what is a resource?' and 'what is a demand or an exploitation?' in the different observed hydrological situations, and 'how can agreed definitions can be computed with the data that can be observed and collected?' The second class of questions is about the way seasonally contrasted situations could be aggregated at the annual level to surrogate inadequate calculations based on annual averages. Seasonality is an important issue that is now recognised by the EEA (EEA, 2012). The absence of consideration of seasonality in indicators jeopardises the accurate differentiation between scarcity and drought events.

As an integrated indicator, it is essential that contrasted situations display different values. Unfortunately, the WEI is very sensitive to seasonal differences. This is exemplified by four opposed situations displayed as theoretical examples in Figure 3.1 (best and worst) and Figure 3.2 (favourable and less favourable) below. The four cases, having the same annual average for resource and demand, were sketched to present a gradient of favourable to unfavourable patterns, with the worst case displaying a scarcity episode because of larger irregularity of resources.

The expected figures are that the bigger the pressure on the resource, the larger the value of the indicator is. The displayed examples, by contrast, provide the same figures for the four cases, since their annual averages are identical. Several attempts were made to find aggregation methods that would provide visible differences between the different cases.

Assuming first that abstraction and resource are both accurately estimated, the approach is to consider the monthly ratios and propose an appropriate aggregation method based on the analysis of the averaging on the one hand, and on the nature of the information to provide on the other hand, compared to the discriminating capacity of the final indicator. In principle, WEI is a ratio, and the truest average of a set of ratios (assuming that all have the same weight because each is valid for one month) is the harmonic mean. However, the message being delivered seems both erratic and in contrast to what is expected: the annual WEI is less in the scarcity situation than in more abundant situation. This is because computing the harmonic average of a series of WEI eventually gives the maximum importance to the smallest WEI. In this case, the worst individual situation and the annual aggregate vary in opposite directions, which is not manageable.

In fact, what is visually considered as the worst situation is the one with the maximum number of small WEIs, not compensated by some very high ones; this just suggests that if the message to deliver is the risk of temporary tension on the resource, this aggregation method is not relevant.

Table 3.2 Comparison of WEI aggregation methods for the four theoretical situations presented

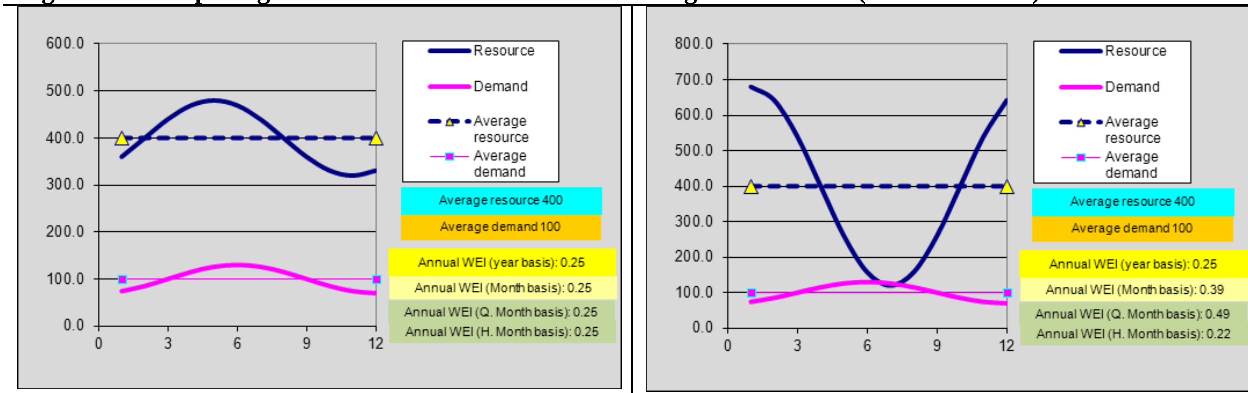
Method of aggregation	Best (1)	Second best (2)	Acceptable (3)	Worst (4)
Ratio on annual averages	0.25	0.25	0.25	0.25
Arithmetic average of monthly WEI	0.25	0.26	0.31	0.39
Quadratic average of monthly WEI	0.25	0.28	0.34	0.49
Harmonic average of monthly WEI	0.25	0.23	0.26	0.22

Source: data taken from Figure 3.1 and Figure 3.2 below.

The aggregation should indicate if the considered period has suffered a scarcity event or not (this is the target of the WEI). To measure if such target is met using a single central aggregate; two candidates are possible: arithmetic average or quadratic average.

The final analysis hence suggests that either the arithmetic mean or quadratic mean of monthly average can provide a discriminating central yearly aggregate of the WEI, with a larger range in the case of quadratic mean that indeed reinforces the weight of large individual WEIs that measure the significant scarcity event.

Figure 3.1 Comparing two different situations and resulting WEI indexes (best and worst)

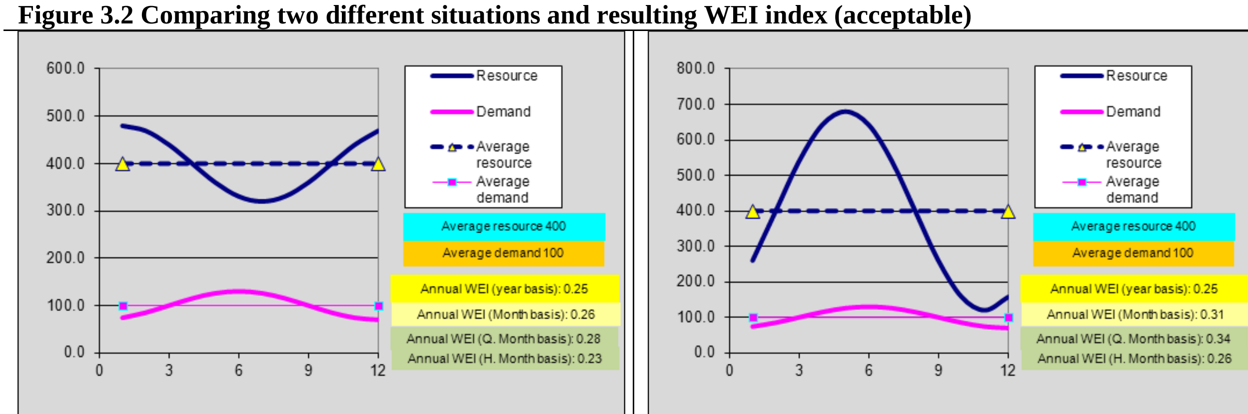


Best situation (1): resource and demand develop in parallel

Worst situation (4): demand crosses resource

In the worst situation displayed, the question of deciding whether the highest discharges are resources or not is raised as well, which is not the case for the best situation.

Figure 3.2 Comparing two different situations and resulting WEI index (acceptable)



Second best situation (2): demand and resource differ

Acceptable situation (3): demand approaches resource

However, it is too optimistic to assume that both demand and resource are correctly measured, and this is not the case when the resource is scarce and demand is high: precisely where the indicator is needed. The categories of problems encompass what to include, and how to overcome the difficulty in collecting representative resource information.

First of all, there are many discussions about which volumes to incorporate or put aside from the demand component (numerator). According to some recommendations, all abstractions that are immediately returned into the resource (turbining, once-through cooling) should be excluded. This poses methodological and practical difficulties, since the data sources do not clearly distinguish such uses.

- If any volume actually abstracted is excluded from the calculations, what is the scope of 'water exploitation'? Exploitation requires assessing if the demanded volumes are actually present and can be exploited. Arbitrarily excluding some volumes calls into question the relevance of the indicator.

- Moreover, if, for example, the once-through cooling or turbining is performed thanks to a transfer from a resource to another, what decision should be taken? Some practical examples illustrate the point. Water diverted into the Provence Canal (from the Durance River, affluent of the Rhone) is turbined at the St Chamas plant and emptied into the brackish Berre lagoon. This water is indeed removed from the Rhone river system over more than 100 km of main drain, and cannot be exploited for any other purpose, including the ecological requirements of the Durance River.

- Groundwater is intensely used in Baden-Wurtemberg to cool lignite-fuelled electricity plants, then emptied into surface water. Indeed, groundwater is depleted from a substantial resource (that perhaps has use otherwise, but from an accounting perspective, this is a transfer).
- The last example was from the upper Loire system to the Rhone system (the transfer has been severely cut now). Turbining water from the Montpezat storage lakes system to the Ardèche river, to take stock of considerable difference in altitude is consumption, since it is transferring water from the spring area of a system to the medium reaches of another ([44]).

Taking into account all volumes is not a secondary issue. The WEI at a certain place expresses the current demands and the way they are satisfied. Consequently, the permissions for exploiting water upstream (of which diversions) are informed of these demands. This can be supplementarily documented with a 'freedom of action' indicator (see bullet 6 in Table 3.2 above). If substantial volumes are not taken into account, the possible availability of the resource may be radically overestimated, resulting in inappropriate decisions.

The issue is the accurate assessment of what constitutes a resource, and it is reported in the denominator. Two distinct questions are still to be addressed. The first is the annual share of volume that is actually a resource: during the high water period, should all floods (including devastating volumes that cannot be used) be considered as resource or not? The response is significant, and largely depends on the regime of floods and storage capacities. Large floods in permanent regimes are indeed not a resource, whereas in arid regions, a flood once over 10 years may fill a reservoir (for example, in southern Morocco).

From a practical perspective, the provision of an accurate figure of resources is a difficult question. From a theoretical perspective, the resource in the denominator is the resource before abstraction (the total resource from which uses are satisfied). This assessment poses limited problems if the resource is considered as long-term annual over large territories, or where the resource widely exceeds the demand. In both cases, and with the exception of the use of non-renewable fossil resources, the computed WEI is much smaller than one, and the error in relation to the inaccuracy in data is negligible.

By contrast, when computing the WEI at sub-basin and monthly level, it may happen that the observed resource is smaller than the demand: WEIs much larger than 1 are computed. In the water accounts exercise, some sub-basins yield uncorrected WEIs larger than 1 000. This is because the resource is assessed as the discharge at the outlet (which is in many cases the only obtainable information). Indeed, many catchments have a limited outlet discharge value, since most usable water has been used for economic purposes, possibly beyond the sustainable resources ([45]).

([44]) A similar system, albeit larger, is used to feed the Rio de Janeiro conurbation with potable water, making a substantial transfer from the Paraiba do Sul Basin to the Atlantic Ocean

([45]) The most illustrative example is the Colorado River, United States, that empties into the Gulf of California, Mexico. Based on overestimation of resources of at least 20 %, excess water allocations, uses and diversions have been permitted and carried out on this catchment. As a result, the river discharge to the sea is now close to nil, instead of more than 15 km^3/year, causing many ecological and production problems in the sea, but also on land. Flow regulation by damming and abstractions had devastating consequences on ecosystems (no floods, uniform cold temperature, sediment trapping and downscaling flowing volumes, riparian areas shrinking), on agriculture (salinisation of land) and are even dangerous for human health.

3.3.4 Critical analysis of the WEI+

Identifying the fact that the different avatars of the original WEI presented some limitations (analysed above), the EEA worked both in the development of the water accounts (refined at monthly basis and sub-basin disaggregation) and with the WFD CIS Expert Group on Water Scarcity & Drought towards an improved formulation of this indicator (the so called WEI+) with the purpose of better capturing the balance and critical thresholds between natural renewable water resources and abstraction, in order to assess the prevailing water stress conditions in a catchment.

Together with the water quantity experts in the Eionet and the WFD Working Group on Water Scarcity and Drought, further data collections have been initiated to calculate a modified water exploitation index called the WEI+, which aimed at mitigating some of the shortcomings mentioned above. The new calculation method was agreed by water directors in June 2012 (CIRCABC, 2012), and is subject to further testing ([46]). The WEI+ calculation method does not address temporal and spatial aggregation areas, which may pose practical problems of effective computation.

The new definition agreed is: 'The proposed WEI+ aims mainly at redefining the actual potential water to be exploited (i.e. availability), since it incorporates returns and accounts for changes in storage, tackling as well issues of temporal and spatial scaling and proposing the use of environmental requirements for the formulation of adequate thresholds'. The WEI+ is formulated as follows:

$$WEI{+}= \frac{(abstractions - returns)}{renewable\ water\ resource}$$

The calculation formula contradicts the definition because the eventually inserted exclusion of returned waters from the numerator as a result of negotiations in the working group. This change in formulation, however, transforms the assessment of **exploitation** to an assessment of **net consumption**, which is a radical change not sufficiently pinpointed and hence not capable of assessing the availability of water (indeed before making a net consumption, a raw abstraction has to be carried out!).

The decision has proposed two ways of addressing the denominator (renewable resource).

Option 1 refers to the calculation of renewable water resources (RWR) based on the hydrological balance equation, using precipitation, external inflow, actual evapotranspiration and change in natural storage as components:

$$RWR = ExIn + P - Eta - \Delta S$$

Where ExIn are the inflows to the computation area, P the rainfall on the computation area, Eta the actual evapotranspiration and ΔS the changes in storage of stored resources. The term 'P-Eta' is not accurate as a resource assessment beyond a couple of days, since the Eta is driven by the availability of water to evaporate from soil (and is hence P-dependent). This is why the computation of evapotranspiration is carried out at daily levels in the accounts. This option is unlikely to be operational in calculations.

Option 2 refers to the calculation of renewable resources based on the naturalisation of stream flow, using outflow, abstraction, return and change in artificial storage as components:

([46]) '…endorsed the WEI+ indicator as part of the overall indicator set for water scarcity and drought, with the understanding that thresholds still need to be tested and agreed…'. Extract from the minutes of the decisions from the Water Director's meeting mentioned in the text

$$RWR = Outflow + (abstraction - returns) - \Delta Sart$$

Where Δsart (variable taken from the water Director's document) is change in storage of artificial reservoirs. This formula is close to the ones mentioned in the previous section, except for the exclusion of the storage that makes the results unstable at monthly level, which is precisely the time resolution for which its inclusion is demanded in the decision.

Option 2 has the advantage of being based on rather easily observable information, the discharge at the outlet, the most reliable and cheapest to obtain amongst all possible options. However, the current monitoring systems (as used in water accounting) provide the residual of natural flow (including the artificial recharge and groundwater inputs, which cannot be known directly) once water consumption has been completed, as well as returns in the same calculation area. The calculation formula requires subtracting the changes in inputs to reconstruct the natural flow. The proposed computation aims at reconstructing the current resource, which is in true agreement with the water accounts outputs.

However, since ΔSart volumes (that represent the quantity input into the system by artificial recharge) are both part of the formula and are often not known (and their delivery dismissed in several countries for commercial reasons), they can be estimated only by the change in river discharge upstream and downstream reservoirs, which may be many and nested in the same aggregation sub-basin, consequently leading to increased uncertainty in the estimates. An indirect method of assessing such volumes is highly subject to double counting or gaps.

Introducing the ΔSart in the computation formula poses a theoretical issue, making the computation unlikely to be carried out at monthly level. In most (if not all) water-scarce catchments, reservoirs were made and are operated to enhance low-water discharges, and are the only source of water. This single source makes the demand below the available resource; hence,most of the observed discharge results from reservoir withdrawal. Considering all the uncertainties related to abstractions, returns, enhancement, etc., the risk of getting zero or negative resources is high, and would make their computation highly erratic and their interpretation questionable ([47]).

The attempts to compute WEI+ with the Eionet-delivered data failed in practice: the obtained results were not discriminating and provoked criticism from Member States that could not recognise in the produced maps the water scarcity problems encountered in their countries. Hence these results could not be published. This makes it necessary to analyse the reasons why officially provided data computed with officially agreed formulae eventually yield unacceptable results. This strongly suggests that both computation agreement and flows beneath the production should be radically revisited.

3.3.5 Calculation of indices for the accounts

Considering the difficulties attached to different indicators, practical ways of addressing indicators relevant for the presentation of water accounts results have been undertaken and checked, based on data obtained from the water accounting exercise.

First, since WEI is a widely recognised indicator in principle, its calculation has been 'normalised', in order to mimic the overall resource before abstractions while fully endorsing the notion of 'exploitation of resource'. In practice, the normalised WEI is computed as:

([47]) The source equation was like x/(Q+x-y) that became (x-y)/(Q+x-y). It is unstable if y->x; this is the case if in a catchment either returns tend to abstractions OR if returns are made just close to downstream sub-catchment. In this case, small changes in catchment delineation or placement of return point may, in practice, make the index quite erratic.

$$WEIn = \frac{abstraction}{outflow + abstraction - return}$$

all being elements of the accounts. Under this definition, WEI values are all in the [0 – 1] range, except for large irregularity in data that suggests errors.

This indicator nWEI is computed monthly and by sub-basin.

The nWEI addresses only the abstraction of water and the share of water entering into economic activities. The ecological needs are an important issue and are not covered by this indicator. The normal way of addressing the ecological needs is the 'ecological flow'. In the absence of agreement on what an 'ecological flow' is, it is only possible to consider the share of resources remaining after uses, which also helps assess the actual resource for downstream uses. The ecological flow is understood as being defined differentially per waterbody and considering the natural variability (EEA, 2012).

The acronym ESIr has been created as a potential indicator of ecological stress for rivers. It could be computed as:

$$ESIr = \frac{outflow}{outflow + abstraction - return}$$

This makes it a symmetrical ratio compared to the WEIn.

However, this calculation method poses two problems, one theoretical and the other practical. The practical issue is uncertainty concerning abstraction and return values, which, when placed in the denominator, makes the calculation unstable (many cases tend to 0 if outflow is scarce).

The theoretical problem is deciding what is important from an ecological point of view: is the stress to aquatic ecosystems dependent on the final balance or impacted by a series of local withdrawals? Even if the series of local withdrawals are returned at the end, they make cumulative lengths of rivers with depleted flow and possible peaks of bad quality.

Balancing both issues, it was eventually preferred to modify the calculation and set this indicator, on a testing development for this report as:

$$ESItr = \frac{outflow}{(outflow + abstraction)}$$

These are the values that are displayed and processed in Map 5.6. As suggested in the conclusions, the equation is not rigorous and should be reworked and possibly replaced by a radically different process, for example a process that would use river discharge and water body status reporting since they are both in the same reference system.

The WEI+, a consumption index, is a very interesting indicator as well, and was systematically computed under a simplified formula, dropping the commented correcting term Δsart . It becomes hence a consumption index, noted WEI+c and computed as:

$$WEI_{+c} = \frac{(abstraction - return)}{outflow + abstraction - return}$$

3.3.6 Presentation and aggregation

Presenting the values of indices requires the selection firstly of a scale of thresholds and secondly of what to represent.

The thresholds are far from neutral: according to their chosen values, the visual impression of the reader is changed: red, green and blue colours have strong visual effects. It is possible that a key reason for the rejection of the WEI+ maps, (apart from the fact they were based on annual averages), is that the selected thresholds were taken from exploitation, whereas the indicator is indeed a consumption index: being too high, the class values appeared as 'no problem' everywhere.

There is no international consensus on what should be the optimum ranging of thresholds for exploitation. Based on Plan Bleu maps, values of 0–10 % (blue, no problem), 10–25 % (green, warning), 25–50 % (yellow, risky), 50–75 % (orange, excessive) and 75–100 % (not sustainable) were used.

By contrast, when addressing the consumption index, for the fraction of water not returned, thresholds must be lower. Using the same source of inspiration, the following classes were taken, with a different progression. Retained values are 0–1 % (blue, no problem), 1–5 % (green, warning), 5–15 % (yellow, risky), 15–25 % (orange, excessive) and 25–50 % (not sustainable). Consumption over 50 % was not considered and would be represented in dark violet or black as 'dangerous'.

To map the share left for ecological purposes (ESIr), the following classes were used: 100–90 % (green, no problem), 90–65 % (yellow, warning), 65–50 % (orange, risky), 50–25 % (red, excessive) and 25–15 % (violet, not sustainable) and 15–0 % (destructive).

To be fully sound, the threshold should be differentiated according to the frequency of return; this option has not been followed since it could be misinterpreted by preventing comparison of maps and is not supported by sufficient evidence.

Once the threshold setting is populated, and since WEIn, WEI+c and ESIr are all computed at the monthly level, their presentation has to be carried out after some aggregation (otherwise, there would be a minimum of 12 × 8× no of sub-basins values).

Different categories of aggregation were carried out:

1. a classical averaging across the years, using the method described above;
2. frequency analysis of the distribution that opens to new assessments;
3. ratios of indices, that follow other classes.

Once many monthly values are available, a percentile distribution becomes a very powerful tool to aggregate the indices all across the time of computation. Three frequencies were chosen: 10 %, 50 % and 90 %. These rounded figures represent the index respectively ([48]), so that roughly only 1 month per year presents smaller index, 6 per year larger indexes and 1 per year a larger index.

The mapping of the indices at 50 % suggests structural water availability issues if the indices fall into scarcity classes; by contrast, the 90 % indices show where a recurrent (albeit non-permanent) water supply issue is endured. The ratio of 90 % to 50 % indices tells about the variability of the resource.

The aggregation by averaging hence becomes only one of the possible exploitation means of the results. Moreover, since all components are present, the other indices can be provided, at sub-basin or district level ([49]), similarly to the WEIs.

([48]) The accurate percentages should be 8.3 %, 50 % and 91.7 %.
([49]) Assuming that an appropriate aggregation method is defined, for the time being, simple summing is being carried out, despite being considered an unsatisfactory method.

By contrast, the latest tested indicator, the ESItr, cannot be presented with same frequencies. The smaller the indicator, the larger the threat for the ecosystem. A reasonable return time for considering a severe threat of once every two or five years, roughly translates as 5 % or 2 % at monthly level.

The threshold values for displaying were taken quite arbitrarily, 5 %, 10 %, 25 %, 50 %, 75 % and beyond.

3.3.7 Uncertainties and uncertainties assessment

The implicit paradigm backing the accounts is the total consistency of figures. This paradigm results from the transposition of currency book-keeping to environmental issues: full consistency is indeed the central objective of accountants that aim at balancing to the nearest cent. However, if full consistency of figures is theoretically (albeit practically not achievable) possible in economics (because all processes are human-defined), there are too many sources of uncertainty and lack in knowledge in environmental issues, because of the following.

1. The process is beyond our control and cannot be monitored for practical reasons; for example, rain cannot be monitored for each and every square metre; it is monitored locally and the rain field is reconstructed.
2. The process cannot be monitored for fundamental reasons; for example, actual evapotranspiration cannot be monitored (monitoring would modify the phenomenon). (This occurs in economics as well with errors, robberies, cheating, etc., reasons why there are accepted unbalances in monetary accounts).
3. Data related to processes that can be monitored result from radically different systems and eventually do not address the same information content. As a consequence, these data are not structurally consistent. For example, using water counters of many small distributions never matches with the values recorded by central counters, for error summing reasons.

When addressing environmental balances and relationships between environmental resources and their economic use, there are huge potential gaps in data. One of the trickiest issues of the environmental accounts is that none of the values is recorded with known accuracy (by contrast, the precision ([50]) of each measurement can be estimated, which is paradoxical!).

As a consequence, perfect line and column matching is the result of adjustments whose justification should be carefully recorded.

A major source of inaccuracy is the incompleteness of the accounts. This is the result of different factors that are listed below (the list is possibly not comprehensive).

- Intrinsic uncertainties on values derived from monitoring. This is the case, for example, of discharge in rivers. The precision and accuracy of each gauging is rather high; by contrast, the accuracy of the recorded chronicle ([51]) depends on the frequency of gauging, on the observation of rare events, on morphological factors (e.g. under bed flow, not recorded), on maintenance of equipment, and lastly on the density of stations. Moreover, intense

([50]) It is quite simple to estimate the precision of a rain gauge (for liquid precipitation); by contrast the accuracy of the areas monitored by this rain gauge is unknown.

([51]) A discharge time series is modelled from the monitored elevation of water at the monitoring point calculated with the calibration curve made from gauging at certain dates. The assumption that the h/Q relationship is unambiguous is only globally true, even in normal situations.

flood events may completely reshape the river bed at gauging sections, making the h/Q relationship obsolete ([52]).

- Uncertainties related to spatial extension of point measurements. Most rainfall is accurately recorded at meteorological stations. Errors may occur when rainfall intensity exceeds the capacity of the sensor, but the major source of inaccuracy is the spatial extension. This is carried out by reanalysis and different modelling; outputs are gridded data whose precision of reconstruction is known, but whose accuracy is not, because this depends on the density of stations provided (see the analysis of the scoring, Annex 3 , Section 4).
- The need to use data that cannot be monitored. For example, the accounts include actual evaporation over lakes and evapotranspiration over land. None of these variables can be monitored for the areas to which they apply. For example, evaporation is monitored in Weather Bureau Class A evaporimeter tanks at meteorological stations. Such monitoring is subject to huge uncertainties, even in developed countries ([53]), and is merely a value used for calibrating modelling evaporation on the field.
- Use of monitored values to populate variables partly outside the scope of this monitoring. This is the general case for water uses. Most water uses that are monitored (with counters) have a specific role: generally checking permits, establishing the asset of payment and taxes, etc. There is no certainty that such a survey may provide direct water uses (abstraction, consumption, return) for the different categories of activities.
- Relative importance of events, e.g. transfer to canals, snow melting, etc. whose location and intensity are widely unknown and not monitored.
- Unknown values of the opening stocks for the first calculated period.

Uncertainties in the statistical sense could not be estimated in the water accounts production because there are no supporting data for such production. By contrast, simple techniques were developed to assess if data provision was sufficient or not and to provide a scoring category. Data scoring, when established, is produced in the annexes dealing with data and when relevant, is indicated in the maps of results to inform the reader of the degree of confidence he or she can have in the results. That is used to overlay mapped results and determine if the colour class is more or less unlikely.

Data scoring has been applied to all data sets that could be assessed in this way. The rationale of data scoring is comparing a representative assessment of data obtained versus expected data computed from accepted references or statistically enabled references. The obtained data is computed for each target unit using density of data provisioning (as documented sites, quantity of time series, etc.). this approach aims at preventing spurious scores resulting from more than expected documentation in some areas that cannot compensate for poor documentation in other areas.

([52]) The Var River (catchment 2 820 km^2) for example, experienced a big flood on 5 November 1994. The riverbed level at one of the gauging sections had been changed by 2 metres due to flood and the downstream gauging was wiped out when flow reached 800 $m^3.s^{-1}$, whilst the estimated peak was 3 700 $m^3.s^{-1}$ (Gourbesville, 2009).

([53]) Personal observation in Senegal in a critical meteorological station used to manage dam releases (tank upside down because maintenance personnel fired owing to lack of resources), and in Singapore in 2010 in a critical station for managing urban water reserves (tanks used by monkeys as bathroom).

3.4 *Data collection sources, surrogates and data extension summary*

3.4.1 Rationales

Making the accounts table is a stepwise process that can be improved by assimilating more data with better accuracy. There are two different paradigms in making the accounts, discussed below.

- The one historically used. This is based on comprehensive statistics, with each cell eventually being filled with compiled and verified statistics whose precision and accuracy are likely to be good, but whose relevance, timeliness and spatial resolution are likely to be incorrect. Most (if not all) published national accounts are carried out under this paradigm.
- The way proposed in the EEA implementation. The data assimilation aims at providing a data set that is as comprehensive as possible, fully relevant and time adjusted at the expense of reconstructing missing data where, when and for those sectors where the data collection scheme is not operative. Some significant uncertainties are expected, especially for those essential data for which no affordable source has been found.

This second paradigm is more likely to fulfil, after a breaking-in and improvement period, the full objectives of the accounts. Implementation carried out under this paradigm calls for the essential data (that cannot be substituted) and the data that can be surrogated by modelled data to be considered separately. In this section, only data that are time dependent or related to works or measures are addressed.

It is important to be precise concerning the acceptations of terms. Modelling is the creation of information for a certain variable using information on another variable, by applying a process of simulation that implies a causal relationship between the source data and the created data.

Water accounts (and environmental accounts at large) are not based on models: there is no process simulation whatsoever in the comparison of data. Accounts observe if sets of data (that are necessarily related by some causal relationship) match under this understanding.

For example, a hydrological model of rain–discharge is based on a process, simulated in computed code, that computes the river discharge thanks to rainfall information. By contrast, accounts observe rainfall and river discharge separately, and analyse if rain and discharge develop in comparable values (having considered evaporation, storage, etc.). This section only deals with the fundamental differences.

The common confusion between modelling and accounting lies in the fact that many data used for the accounts are incomplete and must be reconstructed using statistics, extrapolation and sometimes modelling. When modelling is used to complement data sets under the accounting process, this modelling does not address processes between categories, but rather exclusively intra-category processes.

For example, there is an essential difference in modelling water uses by thermal plants from their CO_2 emissions (intra-category process) and modelling discharge from rainfall (between-categories processes). Hence, when the term ‘modelling’ is used in the information attached to data provision, it refers only to the intra-category processes, unless otherwise contextually obvious or indicated.

3.4.2 Essential data

Essential data are those data that are the heart of the accounting process and that cannot be substituted, except by breaching the principle on non-modelling that is set as the fundamental feature of the accounts. Essential data is divided into two very different groups.

1. Environmental data: climatic essential variables (rainfall, (snowfall ([54])), potential evaporation, and – in practice – evapotranspiration from crops and vegetation), river discharge at a sufficient number of gauging stations (e.g. x/1 000 km of main drain length ([55])), groundwater levels. Reservoir and lake variations in reserve would preferably be collected, but since such data do not exist, they are estimated from river discharge.
2. Usage volumes and abstraction places that are related to point activities for which no distributed variable is available or for which location and characteristics are unknown. For example, it is impossible to model cooling water needed by a nuclear power plant using information on the plant's cooling technology and its location. All energy abstractions, industrial uses, large aqueducts for irrigated systems and urban wastewater plants size and discharge points fall into this category. Population numbers and land cover are also non-environmental data that cannot be substituted or reconstructed with acceptable accuracy.

3.4.3 Data that can be surrogated by internal modelling

The fact some data can be replaced by estimates does not mean that data collection efforts for important data classes can be slackened, while some information from modelling is fully acceptable. Hence, two categories of data are considered here, those that can be modelled as a surrogate of data collection and those that would better be modelled on the long term.

1. Values that can be modelled as a surrogate to data collection are those data which depend on spatially distributed variables and whose distribution is known. For example, domestic water volumes can be modelled from population numbers (cities) or density (scattered populations), even though the points and source of abstraction remain unknown. In this example, the source of water (surface versus ground) can be estimated from regional statistics.
2. Values that are candidates for modelling on the long-term are all those data that are related to a very large number of units, and whose relation with their position is strong, and for which individual data collection would be at the same time tedious and inaccurate. For example, all domestic consumption from small settlements, irrigation from pumping on the field, etc. are indeed better spatialised with modelling (using CLC for example) than from data collection: summing large numbers of items, each with great uncertainty, produces resource exhaustion and large errors. In this class of data, statistics are extremely useful if they can provide technical coefficients and framing values, or if national and regional systems can be used to substitute.

Discharge data in main drains segments, evaporation from lakes, etc. fall into this second category as well: drain discharge is fuelled by observations at gauging stations and then extrapolated with hydrological modelling.

However, surrogating data demands that the value to model has a well-defined surrogate; this is not systematically the case, and ways to proxy such categories of data are described in the dedicated chapters.

([54]) Snowfall is not obtained from monitoring by intra-category modelling from rain and temperature, hence it does not breach the principles set.
([55]) In the scoring, the minimum target is set in the range of 3.5 to 4.5 stations /1 000 km of main drain, at country and sub-basin levels respectively.

3.5 *Developments and improvements in data collection schemes*

3.5.1 Insufficient yield of data in the current scheme

The reasons for proposing a data conceptual model is because many data issues and much non-optimal use of collected data is owing to this absence of conceptualisation and hence appropriate collection levels, primary quality assurance and contextual quality assurance.

Lessons taken from the implementation of the water accounts and the development of other environmental accounting processes strongly suggest the need to reanalyse the complex relationships between the data and the EEA strategy targets. 'Putting data to work' has been the EEA's main role from the start. However, implementing this in practise is not yet feasible, as demonstrated by the many and systematic data gaps found during the development of the water accounts, which used simple data collected for decades. Literally speaking, 'putting data to work' means that same data should produce many outcomes from their work, hence a great deal of different information.

However, experience has shown that that many data could not work owing to them not being correctly located or insufficiently 'trained'. Why is it, after so many years of water-related data collection, that most of these data cannot be used, resulting in surrogate data collection? To a large extent, this incapacity of data to produce the desired information strongly challenges the data collection organisation, and queries whether this failure is really more a conceptual understanding issue than organisational insufficiency. Many processes, workshops and working groups have discussed data collection, metadata have been developed and populated, and at the end of the day, despite efforts, a large share of data cannot be accurately processed and does not contribute to the final results.

This waste of data working strength is not acceptable and must be remediated. The major lesson of the water accounts is that data that cannot be used at their maximum informative capacity were insufficiently embedded in their context. For example, it is clear for river discharge whose gauging station is attached to the wrong river, for water abstraction that is not related to a resource, for a wastewater treatment plant that is not related to the source of pollution, for a mapped city that cannot be related to administrative units: these are all cases of intense efforts for collection with poor yield.

3.5.2 Common features of all data in the environmental context

When analysing data from the point of view of their informative potential, and not from the metadata issue (what could be called the 'meta-information' attached to data), a very simple chain of relationships becomes obvious.

Any piece of data is an event (very short time: an hour to a week) that occurs in a spatially defined superstructure (duration: months to decades) that is related to a spatial infrastructure (permanent at human lifespan: decades to centuries).

If this simple scheme is endorsed, it explicitly requires processing of the quality assurance and the use of data in its full context of information and spatial integration. Such integration had not been made explicit in this simple way but started being implemented in Work Package 3 over the past years' implementation plans of the ETC/SIA. Together, the two spatially defined categories constitute a reference system.

- The spatially defined infrastructure comprises the geographical systems (with their topology) defined and built to cover data uses. In this category are catchments (virtually no change), rivers (some changes), lakes (changes observed), mountains (insignificant changes), geological background (insignificant changes, except locally (mines)), hydrogeological systems (insignificant changes), and soils (slow changes, except sealing);

- The spatially defined superstructures are the spatially defined features, related to the infrastructure (position, topological relationships). At this level, events occur that are the categories of data (= time series data, but 'event' is more appropriate because neutral form the time perspective). In some specific cases, the event has no superstructure as such.

  Most elements of superstructure change with time. For example, cities sprawl and modify their extent (events attached, e.g. population and energy demand. change faster). Dams are made and can be decommissioned, even destroyed. All monitoring networks are superstructures, even if virtual (there is no physical device to witness their implementation on the ground). The same definition applies to wells, roads, etc. This concept is not new; the important element is that an event cannot be considered outside its attachment superstructure, which must itself be connected to the relevant infrastructure(s).

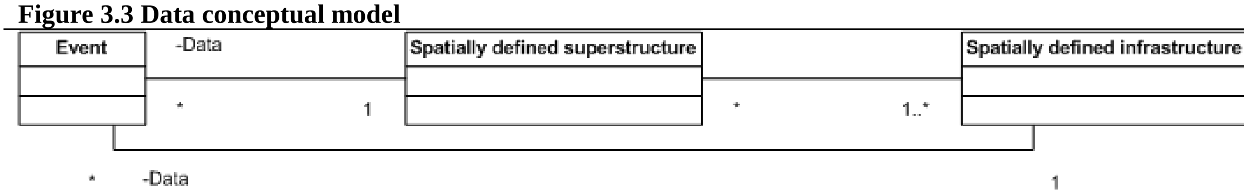
The last category deals with what is commonly understood as data and is named 'events', because time-series data is too limiting in its acceptation.

3.5.3 Conceptual model

Based on the rationales indicated above, a simple conceptual model can be drawn, to express the cardinalities between its components.

The conceptual model expresses that many categories of data can attach to a single superstructure (but cannot be defined out of this attachment!), and that many different superstructures can attach to one or many infrastructures (but cannot be defined out of it). This expresses simply that water discharge is related to a gauging station that is related to a river segment, and that a bridge or a pumping can be attached to the same river segment. Since the model reads from left to right, superstructure can exist without an event being attached.

Figure 3.3 Data conceptual model



An exception is seen in some events that attach directly to the infrastructure (this could be the case of land cover classes that are not making a superstructure). In this case, multiple attachments are not possible.

3.5.4 Use of conceptual modelling

The definition of such a simple model may appear to be stating the obvious, but the principles that the model sketches are not systematically applied: river discharges ([56]) are collected without clear knowledge of their monitoring stations, which in turn cannot be attached to the river (infrastructure system). The conceptual model plays a dual role.

([56]) The model is not for processing river discharge: this event is taken as example because a non-specialist can understand the consequences of the implementation of the concepts.

1. It is the systematic guideline in the data integration (discharge cannot be uploaded if the stations do not exist and the station cannot exist if not referred to the relevant river segment).
2. It defines the appropriate steps in Q/A that are threefold in this model. Using the same example, Q/A steps takes place:
    a. at the event level (new data set), Q/A might consider the self-defining outliers and suspects, namely the same values in contiguous dates (which almost never happens) and single peak changes, e.g. by one order of magnitude (suspicion of decimal point error);
    b. at the superstructure level, outliers are detected by classical statistics (does value exceed k standard deviation?) (historical context);
    c. at the infrastructure level, is the specific discharge in line with the given area, snapping area and specific discharges of stations in the same sub-basin within the same range of areas?

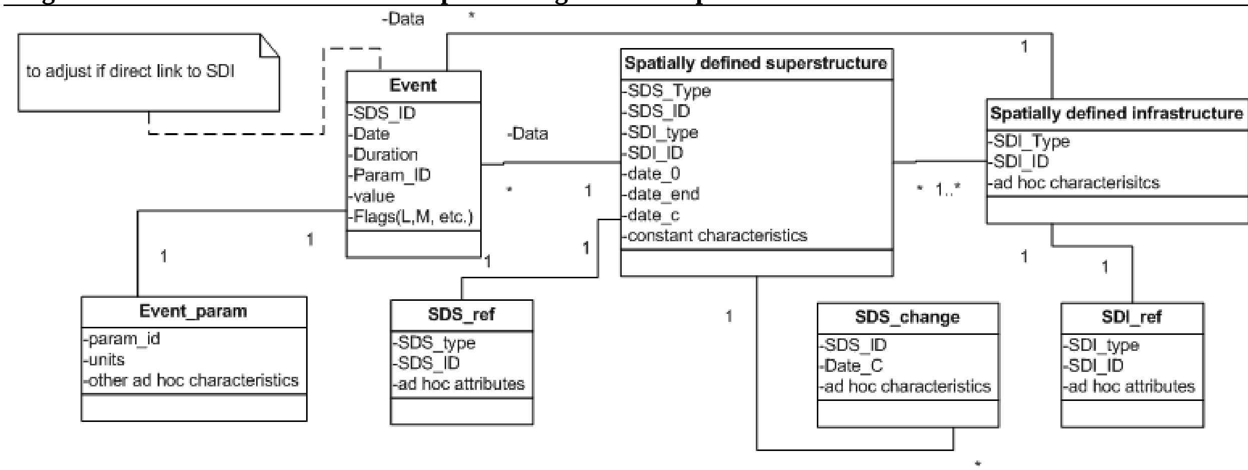
Such a rigorous definition of the different steps of validation can be applied for all the events contributing to the production of accounts. No such procedure was implemented in the production of the accounts, and this resulted in the rejection of a large share (~2 000 over 9 000 gauging stations) because, eventually, the time-series attached did not pass the specific discharge testing carried out at the production level and not at the data validation level.

The application of this conceptual approach can be elaborated for all data sets used for water accounts and could also be set up for other components of the environmental accounts.

3.5.5 Consequences as a data model

The data model shows the way the tables containing the data are designed and their relationships coded in the processing for their storage and validation. In many cases, the data sets implemented in Nopolu System2 implicitly follow the principle that no time-series data (in V_ tables) can exist if their characteristics table and the entry in the table does not exist; this (structure is vary analogous (albeit not explicitly mentioned as such and hence not comprehensively processed) to the 'superstructure'. However, there are gaps in the information and the relationships of superstructure item to infrastructure may be insufficient (for example, which version of the data set represents infrastructure?).

Figure 3.4 Generic data model for implementing data conceptual model



The generic data model is inspired by the European Lakes, Dams and Reservoirs Database (Eldred2) data model. It considers that the kernel of information is at the superstructure level: this entity is the most important vehicle of environmental information: it changes slowly but may present different avatars, recorded in the SDS_change generic table. The dam database exemplifies this featuring. A dam is commissioned (date_0) and decommissioned (possibly destroyed at date_end). In between, despite being the same dam, it may undergo some important changes (e.g. the height is changed and reservoir capacity is increased) at date_C. Similar avatars are possible for any superstructure and can be stored effectively this way.

4 EU-wide implementation

4.1 *Implementation historical steps*

The water accounts implementation supporting this report has been carried out by a consortium led by Pöyry, with Vito and SCM as consultants, under a service contract set in place by DG Environment, under the 'development of prevention activities to halt desertification in Europe'. The service contract aimed at 'building the water and ecosystem accounts at the EU level'. The contract proper terminated in late June 2012 and complements were provided until autumn 2012.

This last implementation stage is the first full-scale development to take stock from the past experience gained by the EEA in adapting the SEEAW principles and the simplified accounts based on global statistics (rooted in the examples in the foundations) (Weber, 1986). This experience was supported by diverse works and the implementation proper was carried out during the past five years, thanks to framework contracts led by the EEA.

The main outcomes of the previous implementations are related to computation procedures, reference systems and data collection. All categories of improvements were investigated, and to a large extent, these improvements were implemented. They are all related in the relevant chapters and sections and need not be discussed again

4.2 *Summary of data collection for EU implementation: natural assets*

4.2.1 Climatic data

Climatic data collection is described in detail in Annex 3 The principles backing this data collection were to achieve the widest possible coverage of all areas and to make it possible to reuse this data for many purposes. The application of these principles was realised in a three-stage data preparation process, carried out at the EEA([57]).

1. Selecting the appropriate source of information, sharing the twin constraint of being technically most sound (amongst the achievable ones) and free for use. This led to the selecting of ENSEMBLES project–derived data set previously mentioned (see as well Annex 3 and Haylock (2008)).
2. Modelling from this selected source the required variables at the daily step, and aggregates at the monthly step (critical for evaporation that must be computed daily). The model generates kilometric grid rasters, one per parameter and per month.
3. Post-processing the rasters to upload them as SQL-S tables (one table per period or month, but it could be days, as well) and sets of parameters, then aggregating the kilometric grid data into each FEC, and compiling the dates into a single output table to deliver to the consultant.

Data pre- and post-processed in this way are fully compatible with the requirements of the Nopolu System2 to compute the accounts. Snow is for the time being simulated as one avatar of rain within Nopolu System2. Since 10 years were computed (2000–10), there are 120 months, and the spatial distribution of data has often been presented in the form of videos, in different forums (e.g. the Water Forum, in Marseilles (France), March 2012).

([57]) Blaz Kurnik selected the data set and modelling; Philippe Crouzet exploited results, stored and aggregated as FECs.

Even though data processed in this way is the best currently achievable, and to a large extent improves the data sources used when preparing the EEA-wide water accounting (ATEAM ([58]) data re-gridded to the FECs level and MARS ([59]) data sets computed by the JRC for agricultural forecasting), and the soil moisture validation is the best that can be expected (Kurnik et al.) for the time being, the data provided to the consultants suffers from two categories of problems, detailed in Annex 3 and summarised here.

1. The source meteorological data is extremely odd in many countries (source data delivery being the country's responsibility, the quality scoring at sub-basin level matches country borders). The final weighted data quality scoring is only 19.8 %, with well-covered countries/areas being Ireland, Germany, the Netherlands and northern Italy with Slovenia as well as southern Scandinavia (Denmark excepted). See the score map in Annex 3, (**Error! Reference source not found.**, page **Error! Bookmark not defined.**).
2. The final grid omits many coastal catchments; these are not documented (see Map A3.3, page **Error! Bookmark not defined.**).

The source gridded data has then to undergo a complex process of modelling and post-processing to become a calculable source for the accounts. The improvements are detailed in Annex 3 but the most critical issue is that there is no secured process and resources to prepare these data. If this is not changed, **the main source for water accounting risks being neither maintained nor updated**.

4.2.2 River discharge data collection

River discharge is the touchstone of the water balance and a key component of the I/O tables, as well. This twin role of reference of the validity of the balances on the one hand, and as central resource on the other hand, make this group of information especially important.

The main priority has been given to collecting daily average discharge values at all possible stations situated on ECRINS main drains or in as many places as possible in all upstream catchments (i.e. coastal and true upstream FECs of continental catchments) where a 'main drain' has no hydrological meaning.

An important milestone in the data collection process was the Eionet meeting held in Copenhagen (Denmark) on 18 May 2011, when all countries were explicitly requested to provide:

- a comprehensive map of gauging (hydrometrical) stations;
- daily averages for at least the 10 past years on those stations meeting criteria above.

Data collection (to supplement previous data collection of which countries had been informed) was requested of the ETC/ICM, under a special task. The ETC was also expected to ensure the mapping of gauging stations to ECRINS and to quality assure the delivery of data, to ensure systematic provision as standard text files to the EEA.

([58]) Advanced Terrestrial Ecosystem Analysis and Modelling (ATEAM, 2013) is a Fifth Framework Programme (FP5) project that produced, among others, a monthly set of data gridded across Europe of key climatic variables at the monthly resolution.
([59]) Monitoring Agricultural Resources (MARS) (JRC, 2013) is a unit mission of the JRC supporting agricultural resources at large (http://mars.jrc.ec.europa.eu/mars/About-us/AGRI4CAST/Data-distribution MARS, under certain conditions, allows access and downloading of daily reconstructed climatic data over coarse grids, across Europe and its neighbours (not checked out of the EEA widest area of concern).

As was feared, the process was delayed, to a large extent because of poor responsiveness of some countries. The EEA and the consultants had to carry out many supplementary tasks, including mining many data into official websites and performing many snapping and data checks.

Final table C_quan (gauging stations) was eventually completed by the consultants and final V_quan table (daily values) was completed in the same way. The averaging daily data (including minor gaps bridging) was carried out with Nopolu System2 facilities. As a result, the following data statistics were produced; they illustrate the volume of basic information prepared for carrying out the water balances.

Table 4.1 Summary statistics on the river discharge data incorporated into the accounts

Country code	Country	Number of stations	Beginning period of time series	End period of time series	No of daily flows	No of missing flows in period in all time-series
AT	Österreich	502	01-Jan-51	31-Dec-08	6 256 564	30 502
BE	Belgique/België	46	31-Dec-68	31-Dec-10	397 619	22 668
CH	Schweiz/Suisse/Svizzera	206	01-Jan-60	31-Dec-09	3 054 039	27 577
CY	Kypros/Kibris	14	01-Jan-98	30-Sep-09	52 858	2 558
CZ	Česká republika	4	01-Jan-95	31-Dec-09	20 516	-
DE	Deutschland	321	02-Jan-00	31-Dec-10	7 123 856	12 872
DK	Danmark	140	01-Jan-89	31-Dec-06	872 595	1 959
EE	Eesti	38	01-Jan-02	31-Dec-10	884 293	15 989
ES	España	1 111	01-Jan-12	30-Sep-08	12 767 122	5 308 733
FI	Suomi/Finland	81	22-Jun-09	31-Dec-10	1 524 772	43 366
FR	France	2 119	01-Jan-13	07-Dec-09	18 188 335	1 154 812
GR	Ellada	14	01-Mar-61	30-Sep-09	103 992	12 811
HR	Croatia	19	01-Jan-96	31-Dec-10	104 040	30
HU	Magyarország	18	01-Jan-24	31-Aug-11	274 524	4 094
IE	Irish Republic	169	01-Jul-40	28-Jun-11	1 805 918	186 635
IS	Iceland	8	01-Sep-32	01-Sep-10	149 981	7 369
LI	Liechtenstein	1	01-Jan-75	31-Dec-09	12 784	-
LT	Lietuva	13	01-Jan-96	31-Dec-09	63 953	702
LV	Latvia	13	01-Jan-96	31-Dec-10	71 227	-
MK	Macedonia	8	01-Jan-95	31-Dec-05	31 410	3
NL	Nederland	27	02-Jan-00	31-Dec-10	164 753	103 833
NO	Norge	99	01-Jan-96	31-Dec-10	534 254	1 073
PT	Portugal	133	01-Oct-17	31-Aug-11	1 289 565	330 951
RS	Serbia	12	01-Jan-96	31-Dec-10	65 748	-
SE	Sverige	325	02-Jan-00	05-Sep-11	7 276 432	221 803
SI	Slovenija	35	01-Jan-61	31-Dec-09	496 044	2 555

Country code	Country	Number of stations	Beginning period of time series	End period of time series	No of daily flows	No of missing flows in period in all time-series
SK	Slovenská republika	10	01-Jan-96	31-Dec-10	54 790	-
TR	Türkiye	13	01-Oct-80	31-Dec-06	74 816	15 792
UK	United Kingdom	206	02-Jan-00	31-Dec-09	3 440 455	44 593

Source: Summary from the SQL server database prepared by Pöyry (latest inclusions of reconstructed data).

NB: country developed names as provided by countries, and have not been reset to standard name purposely. Country code is ISO 2 standard code.

4.2.3 Surface river discharge data reconstruction

River discharge data present and will present gaps: time gaps when data have not been recorded for a while (stations damaged by floods, recording interrupted for maintenance/works in the river, etc.); and spatial gaps. Spatial gaps have two main causes: the major cause in the 2012 calculations is due to the fact that some countries would not or could not provide enough data; this cause will be mitigated with time. The second cause is that some sub-basins are not sufficiently instrumented (no measurements) for any number of reasons. Such cases must be mitigated and data reconstructed so that usable values are produced at minimum cost, to avoid complex process modelling.

Time and monthly averages reconstructions are carried out with two different methods:

- time gaps (some months missing in time series) reconstructed by probabilistic complements, based on closely related stations;
- spatial gaps (no data at all) reconstructed by the Experimental Probabilistic Hypersurface (EPH) method, which was used for the first time.

Both methods are presented, discussed and exemplified in Annex 8 . As a summary conclusion, the time reconstruction provides gap-bridging, in a quick, reliable and economic way without requiring hydrological expertise. The method, being independent of any hypothesis of statistical distribution laws, is very robust.

By contrast, the EPH method offered rather erratic results in the production carried out in 2012. This was not totally unexpected: reconstructing river discharge where no data at all exist is quite a challenge, especially if the constraint is to exclude any external information (due to the consistency constraints of the accounting procedure).

There are two reasons for this: first, in the areas requiring spatial reconstruction, reference data were too scarce to reconstruct catchment productivities, and were oriented (due to time constraints) towards direct discharge reconstruction, and second, because reconstruction data (discharges at neighbouring stations) were also questionable.

In conclusion, a radical revision, leading to complements in ECRINS itself, was analysed, tested and made ready for further spatial reconstructions, from the perspective of repeating the water accounting perspectives. These conclusions that recommend reconstructing river productivity to recomputed monthly discharge are in Annex 8. The advantage of the method (once reasons for some erratic results are handled) is that it is robust and effective for the purpose.

The consultant partner (SCM) carried out a deeper analysis with the EEA and made calculations with an improved approach, detailed in Annex 8, based on:

- reconstructing only specific discharge (to avoid the unpredictability effect);
- using as seeds a limited subset of stations, precisely selected according to criteria of altitude, distance, sea outlet, and biogeographical region to separate out the 'most resembling stations';
- restricting the dimensions of the probabilistic space to three: catchment area, distance and altitude (other important factors having been used at selection).

The test was carried out with a single selection of stations, whose attributes attached to stations were left untouched (namely their drainage area, which may have been very inaccurate in many cases, had not been updated). The test nevertheless substantially improved the reconstructed information in three quarters of the tested stations, despite providing values that do not reach the expected quality for accounting purposes.

The provisional conclusions are the following.

- With discharge at stations being influenced by many factors, the reconstruction in the absence of data cannot mitigate direct information with high accuracy.
- The stations draining large areas are in some cases reconstructed with large uncertainty (this might be in relation to the limited number of 'seed' stations, and would lead to a differentiated selection process to exclude irrelevant stations). The hypothesis that they could not be reconstructed by any method meeting the accounting requirements should be considered for the largest ones.
- Stations with median drainage areas (up to some 1 000 km^2 up to some 10 000 km^2) are best candidates for the improved method, provided the documentation on stations is absolutely accurate.

The improved method is therefore very promising, and since it is much less time-consuming, it allows systematic testing of filtering hypotheses, making it a tool potentially suited for the systematisation of the water accounting at sub-basin level.

There is a very interesting domain of investigation to make possible the accounting development in those areas with high likelihood of similarity that are poorly instrumented; the large systems, on the other hand, must be documented by monitored data.

4.2.4 Groundwater data

In the absence of reference systems, no groundwater data were mobilised for this water accounts exercise. The assumptions were that groundwater is an indefinite recipient and source of water. In other words, this means that there is no processing of limits for groundwater recharge, and no limit for abstraction.

4.2.5 Quality scoring of natural assets data

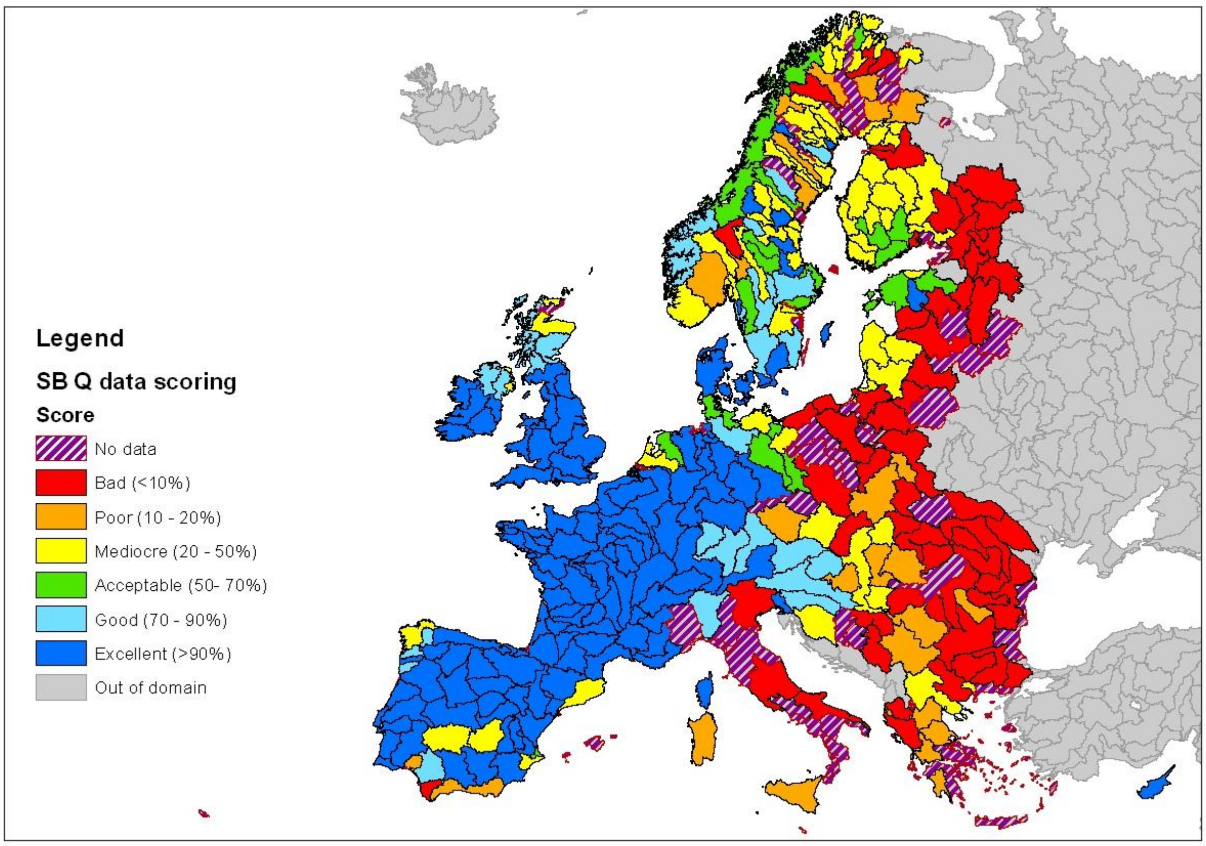
For the sake of presenting results with some information on the accuracy of these results, as well as to pave the way for better information collection and procurement, data scoring has been systematically carried out.

The scoring principle is comparing the data obtained to the reference volume of data and appropriate coverage. Since the water accounts are presented at the sub-basin level, scoring has also been computed at sub-basin level and, where relevant, at country level. The results at country level tend to be systematically better if data provision is unevenly distributed across the country.

Climatic data scoring (for details, see Annex 3 , Section 4) is based on the share of sub-basin area documented by climatic station and reported to a target density of climatic stations of 1/25 km × 25 km grid cell, with figures supported by statistical evidence from the Ensembles researchers. The stations are those used by the ENSEMBLES project to reanalyse climatic variables across Europe from a long-term perspective. The weighted average of scoring over the water accounting area is a bit less than 20 %, whereas it is larger than 90 % in significant areas mentioned in Subsection 4.2.1 above.

River discharge data scoring has been computed in a comparable way, also taking into account the type of data reconstruction. The target density has been established as the number of stations/1 000 km main drain in the sub-basin, considering the median density across all countries (including those countries that delivered no map) and counting the number of data per station.

Map 4.1 River discharge data scoring per sub-basin



Source: EEA computations (the map presents the primary delineation of sub-basins).

When the close to 1 million km main drains are considered, the weighted scoring is 49 %. Results are presented with a pattern shading corresponding to the result colour, so that sub-basins having above or below average discharge scores appear more or less shaded.

The most worrying issue is that, despite reconstruction efforts, several sub-basins have no data at all.

4.3 Summary of data collection: usage volumes

4.3.1 Water uses as a subset of the management of uses and pollution loads

The assessment of abstractions and returns deals with an item of the SEEAW called 'flows from the environment to the economy'. These flows then encapsulate the next component of the accounts, the 'exchange of flows between the economy'. While this exchange of flows within the economy is not part of this development, it is important to bear it in mind in order to understand the structure of the application and the milestones crossed, so that this component can be executed in line with the assessment of the assets and the main exchange between the economy and the environment.

Abstractions on the one side, and return flows from the economy back into the environment on the other side are the dual components of the use of resources; the calculation system has been developed on the Nopolu System2 platform so that all these item components can be assessed in the same process.

Water uses at large (that do not address volumes only, but pollution and purification as well) are a key issue of the environmental assessment. To assess these flows related to the economy from a volume perspective, the 'Nopolu Water Uses' has been developed. Built on the basis of the existing Nopolu Integrated Emission Inventory (IEI) application, it can deal with complex relations: where water is abstracted and by whom, who uses it, and where water is returned (partly directly to the environment and partly via sewage systems).

The enormous amount of conceptual work already carried out to implement the Nopolu IEI application in the past years (which is dedicated to the assessment of pollutant loads with linked driving force, pressure and impact) has hence been recycled for the sake of quantitative water accounts. This was all the more relevant since a characteristic of Nopolu IEI application is that the water flow is the vector of pollution, transported as a load but purified based on concentration. It was then a natural extension to develop the framework of water use in the economy on the basis of the IEI application structure.

The Nopolu IEI information system has been developed to handle the complexity of water uses topology between FECs, domestic entities, water supply features, sewage systems, etc. Another important and valuable characteristic is that it either processes values or computes values from activity volumes and activity coefficients. This characteristic is well adapted to water uses in an incomplete information system, where some items are documented and others are not.

4.3.2 Common features for all water uses categories

Total water abstraction in the EU-27 amounts to about 247 000 hm^3/year. On average, 44 % of total water abstraction in the EU is used for energy production, 24 % for agriculture, 17 % for public water supply and 15 % for industry. The major uses of water in the manufacturing industry sector are for heating and cooling, chemical processing, and as a solvent; nine manufacturing industry branches use 90 % of the total of the volumes in this class.

Compared to natural resources, water uses present radical differences. Water resources are quite evenly distributed compared to the range of uses. The range between highest rainfall and smallest rainfall between comparable statistical units (FECs) is in a 30-fold factor), whereas the range between a small user within a FEC and the largest ones can be up to a million-fold when considering abstracted volumes. At the statistical unit level, the distribution of resources is rather regular (if not cumulated), whereas the distribution of uses is extremely asymmetrical and hence calls for a specific approach to be populated.

As a result, the data collection and processing methods must be radically different: in the case of natural resources, all elements must be populated, with a rather uniform accuracy. However, when addressing water uses, a differential effort should be set up to address with the best possible accuracy the limited number of largest users, while possibly considering the smallest with a simplistic approach, based on best use of working resources. Natural resources are addressed by inventory process whereas uses are populated through stratified statistics approach.

There are a considerable number of water usage issues (and innumerable numbers of small entities) that cannot even be identified as elementary statistical units, making up a small share of the total activity, and a small number (some hundreds) of very big entities, that together make up more than half of the total volume.

Based on standard statistical concepts, all the usage entities could be categorised into three classes, generating as many strata, called by the appropriate method to populate the attached values.

1. Largest: 'individual survey' — the items should be individually populated, both from the point of view of the water source and return; identified as L (large).
2. Medium size: 'individual modelling' — the items in this category are individually located but the values attached can be modelled for each item, based on surrogate and technical coefficients; identified as M (medium). This category is created to simplify data collection in the cases where the source data don't comprise the detailed values.
3. Smallest size: 'areal modelling' — items belonging to this category are grouped and expressed as density at the statistical unit level or, by default, at the territory of reference level. They are identified as S (small). The items in this category do not have any good reason for being individually identified and having their data collected.

The implementation of this approach is a long-lasting process. It requires an assessment of the statistical population to sample. Amongst the four categories of uses at stake, only human population is fully known and can be distributed into subcategories. The manufacturing industry has limited sources of information (the largest are to some extent known). Energy production is also partly known, with the biggest being documented to some extent. The last category, agriculture, is better known when dealing with spatialised irrigation, and poorly known if considering large irrigation areas.

Details of data organisation and status of collection are reported in Annex 9 , page 197 for domestic and urban; Annex 10 , page 204 for cooling water uses and for manufacturing industry; and last but not least, Annex 12 page 252 for irrigation water, respectively.

In all cases, the purpose of the data collection is:

- to organise the usage category data sets so that each item is assigned its proper statistical unit and relationships with the natural system;
- to allocate the abstraction volume per source (with their location, with respect to the natural systems and type of resource), hopefully disaggregated per time unit;
- to find, estimate and store the amount of water consumed (not to be returned to natural systems);
- to allocate the return flows at their place of delivery with respect to the natural recipients;
- to populate the different items with their ISIC code, so as to meet SEEAW standards.

For the time being, the volumes exchanges between users are not considered, being outside the scope of the assessment of natural assets accounting. However, data organisation is capable of storing and processing such data, when such information becomes available, and if required.

4.3.3 Summary of findings for urban and domestic waters

Urban and domestic water uses follow a generic transfer scheme that schematically comprises abstraction, transport and preparation of tap water, uses (domestic and urban uses), disposal in sewerage system with (in general) wastewater purification, and then final disposal into the recipient.

Urban and domestic use are characterised by the fact that large cities may abstract water, and sometimes dispose of it, at significant distances from the city itself.

This category of uses best follows the stratified data collection scheme, since the reference statistical population is well known from population numbers from different sources; unfortunately, the EEA has to use the American source since the EU ones don't cover the EEA area, only the EU area.

Despite this favourable situation, much progress remains to be made. First of all, the delineation and composition of 'cities' as members of the stratum of the largest items comes from different and non-fully streamlined data sets. The Urban Audit is the most relevant source and should be developed into a database of the largest cities, with systematic information updates (the city being a spatial superstructure populated with time-series data and related to the spatial infrastructure, where relationships not fully established yet, e.g. what administrative items are included in the city?).

Similarly, the intermediate stratum is still imperfectly defined as a spatial superstructure related to the infrastructure (the same question as above may be posed). In the stratification conceptual model, this stratum is populated by 'individual modelling', which calls for relevant and highly specific technical coefficients for rates of abstraction, water production, losses, domestic and urban uses, etc.

In both cases, the collection of technical information that is required by the accounting methodology is not sufficiently structured and happens to be split between different services in the EEA ([60]), resulting in incomplete information eventually being collected. The same remark will have to be formulated for the other categories of water use. This is largely why the water uses data have been so complicated to produce and why the results have limited accuracy.

At the Eionet level, similar difficulties are observed and the future data collection should include more cooperation between the data owners on the one hand (often private companies and specific organisations) and administrative levels (the National Focal Points, that handles the possibility of allowing data dissemination, but who disposes of aggregated data only). In many cases, who owns rights on data has not the data and who has the data does not owns the rights to provide data, making data collection incomplete and resource consuming.

Taken together, the volumes abstracted and supplied are respectively 46 976 hm^3 and 36 441 hm^3, reflecting a loss between abstraction and supply of 22 %. The values indicated do not suggest a precision of 1/500 00 and are reported for consistency with sums in Table A9.4, page 211 of Annex 9

([60]) The share depends on the current organisation, and may change quite rapidly. In the current situation, one group collects aggregated volumes and another deals with the urban items. This organisation, based on 'topics' is orthogonal to the appropriate unity of data collection by 'universe' and categories.

4.3.4 Summary of findings for cooling water uses

Water for cooling is mainly, but not exclusively, related to energy production. The analysis carried out shows that in cooling of large combustion facilities (LCPs) (1 041 entities) and nuclear power plants (NPPs), a total of 10 775 hm^3 are abstracted from freshwater with a return of 3 232 hm^3 (all LCPs together, Table A10.8) and 5 704 hm^3 with a return of 1 711 $hm^3$for NPPs.

Abstraction from brackish and saline waters is estimated to be significantly larger (because a one-through cooling process is assumed), but in both cases, uncertainty concerning the volumes is very high, as with the location of facilities. In particular, data accessible at EU level do not indicate if freshwater abstractions are used by once-through cooling, which would significantly change the figures above. According to different sources, and without likelihood information, the uncertainty range is likely 0.5-fold to 8-fold, and the probability that volumes should be multiplied by 2 to 5 very high.

To summarise, the figures used and supported by evidence are as follows: 16 479 hm^3abstracted and 4 943 hm^3 returned (70 % consumed). This could be closer to 60 000 hm^3 for abstractions only, in further reassessments.

The improvement of data requires a radical revision of the data collection schemes; this would be effectively supported by shifting some items of the European Pollutant Release and Transfer Register (EPRTR) reporting from 'optional' to 'compulsory'.

Regarding the single issue of energy production, all those falling in the 'L' stratum should be addressed individually, possibly following a specific agreement with EURELECTRIC[61] within Eionet.

4.3.5 Summary of findings for the manufacturing industry

The absence of a statistically representative population of industries allowed only key sectors to be explored, and their related volumes estimated from different sources. The compilation of values might provide misleading information and is hence not carried out. All other calculation details are provided and volumes per country (before disaggregation by sites) are reported in Annex 10 , Section 6 and thereafter.

The improvement of data also requires a radical revision of the data collection schemes; this is likely more complicated than the one proposed for cooling. The cooling water volumes in manufacturing processes are, however, related to large plants, most of which fall into the 'L' stratum' and are likely to be processed as such.

The manufacturing sector has shrunk significantly in the past decades, and is mostly under the 'M' (and secondarily, the 'S') strata. The first difficulty posed is the absence of statistical references preventing an effective and affordable approach for data collection.

The most reasonable development involves creating a cooperative synergy with Eionet — possibly revisiting the concept of the National Reference Centres (NRCs) in this issue — and building a positive relationship with the relevant professional associations, in order to set up an effective data collection design within SEIS, so as to ensure the production of water accounts in the next EEA strategy.

---

[61] Electric power industry: Eurelectric (Union of the Electricity Industry, Boulevard de l'Impératrice, 66, B-1000 Brussels, Belgium, Tel.: +32 25151000) http://www.eurelectric.org/

4.3.6 Summary of findings for irrigation water uses

Water uses by agriculture are composed of two elements.

1. The soil water evapotranspired by rainfed agriculture. According to estimates, the cumulated area of rainfed agriculture is 1 356 330- 93 451= 1 262 879 km^2 (126 million ha), which, assuming a net use of ~200 mm/year, makes a total volume of 1 262 879 *.2=252 575.8 hm^3 of water a year.
2. The abstractions for irrigation. After correction (Table A12.5 in Annex 12), these are 13 536 hm^3 from the JRC source plus at least 7 260.1 hm^3 (identified correction, except Spain, where the correction should be ~ 20 500 hm^3 according to Spain's comments on the study, lacking, however, sufficient information for being incorporated). The current range of ~20 - 800 hm^3 to ~41 300 hm^3 would probably not exceed 50 000 hm^3 by much, since most significantly irrigating countries (except Turkey, which had no relevant data) are included in this assessment.

Irrigation water is likely to represent ~1/6 of total volumes involved in agriculture; this is a huge volume, especially since volumes are mobilised only in certain regions and during drought periods. Consequently, better assessment of volumes and location is necessary. The currently available source at the European scale seems reporting only about one quarter of the estimated volumes, which is insufficient indeed.

It is suggested that a stratified approach be developed jointly with the JRC, in which:

- large systems are identified, with the help of Eionet, and based on a common specification;
- infrastructures linked to agriculture are inserted into ECRINS, and related to resources (to better document the change in reserves);
- specific information on weather-dependent technical coefficients (if data on volumes are not available) is collected with countries' support, at sub-unit level, for example, to correct current estimates, in close cooperation with the Eionet NRCs.

5 Obtained results

5.1 Likelihood assessment of results

5.1.1 Rationales for analysing the quality of water balances

The water balances computed at monthly level are combined data from different sources, with very different levels of accuracy, and they reflect water uses. This is the role of accounts: to deal with storage, uses, transfers to groundwater, etc.

At the scale of the study, it is conceptually and practically impossible to balance precisely the flows of water. There are, however, some simple analyses that can be carried out.

The first systematic control is based on the recognised fact that, generally speaking, there is a strong deterministic relationship between the upstream flow and the flow at the outlet of any catchment. The analysis will help assess whether this is correct and should help detect significant inconsistencies. By definition, this analysis can be carried out only if the sub-basin has an upstream sub-basin. When this condition is met, the outflow should normally be larger than the inflow. The following events can disrupt the relationship:

1. significant errors in discharge values; this may result from errors in data, errors in placement of gauging stations, and incorrectly reconstructed data;
2. significant abstractions and diversions without restitution.

This case shall be addressed by simple linear regression between inflows and outflows, filtered by the discharge data scoring.

The second systematic control aims at assessing if the share of the outflow that should result from inputs inside the sub-basin is explained by the volumes that are produced and available for this production of flow. River discharge is principally fuelled by effective rainfall (hydrological acceptation: rain minus actual evapotranspiration). In the long term, this is the only source of water in rivers and the volume discharged at the outlet should match the volume resulting from the cumulated effective rainfall.

The short-term assessment is the cornerstone of the accounts. In this case, a full recovery of the output volumes is not expected because storage, abstractions and other factors that are not synchronised with the monthly data sets. Unlike the first case, all catchments can be investigated, since the sub-basins with no upstream just have a zero external import. Many factors, more than in the first case, tune the response.

1. Since all catchments are considered, those with no upstream and no downstream may have several outlets, which are not all monitored.
2. The share of water produced is based on the simple balance formula, close to the formula used for computing the indices:

   $$local\ outflow = outflow - inflow + rain - ETR + abstraction - return$$

   This simple formula does not take into account storage and lags between rain and ETR. Hence, at a certain month, the component rain –ETR can be strongly negative. To mitigate this, further statistical analysis is carried out with the explanatory variable of up to four months before.

Moreover, the result depends a great deal on the quality of usage data, especially if the volumes are large compared to natural flows.

3. It is expected that the relationships firmly depend on the quality of data, and are hence filtered by both meteorological scores and river discharge scores.

A third systematic control is based on the general hypothesis that over some years, the stock should equilibrate. This is a bit uncertain, since the starting stock is unknown and estimated from closing stock by a feedback process. However, in many catchments and without calibration, the processed sources match well.

5.1.2 Data preparation

All data have been recomputed from the water accounts outcomes (as presented in the dynamic Excel workbooks, see Annex 1 , Section 9 by a series of MS Access® queries embedded in a macro (for repeatability, and in the possible further implementations, as a service in the application).

The variables computed are the following.

- Identification: sub-basin, year, month and rank (rank= 1 for first month of first year, 13 for the first month of second year, and so on).
- Raw inflows and outflows into the sub-basin, for the date (RawInpQ, RawOutQ), used for the first assessment.
- Local reconstructed outflow (= rawOutQ – RawInpQ + abstraction – return). This is the explained variable Local0.
- Reconstructed inputs rain – ETR (as computed, this is the most likely run-off from the accounting tables), placed in variables prodM0 to ProdM4 (same month to the fourth month before). These are the independent variables.

Of course, missing variables are replaced by appropriate jokers (-1 if filters, -32 768 if values, so that they can be accurately processed by the statistical package).

The table is exported as text with column headers and imported into Statgraphics® Centurion XV, and processed by the EEA. The data sets prepared contain 39 552 records (sub-basin * year *month) and covers eight full years. Unfortunately not all records are fully populated with quality data.

5.1.3 Results of the I/O analysis per catchment

Linear regressions were carried out, filtering inputs and outputs < 10 hm^3/month ([62]) and making a series of regressions with filtering from no to any scoring, and then by 10 % scores until 90 % and beyond. The characteristics of regressions are the following.

Table 5.1 Synthesis of the upstream–downstream river discharge correlations

([62]) This is a rule of thumb and choice to take round value that corresponds to discharges less than 3.76 4.13 m^3/s, (month 31 to 28 days) making the excluded cases extremely insignificants sub-basins at the EEA level.

Regression indicators	Score Q (%)									
	0	10	20	30	40	50	60	70	80	90
Nb couples	11161	8155	7575	6715	5947	5275	4891	4315	4027	3451
R^2	0.709	0.838	0.826	0.865	0.876	0.874	0.874	0.915	0.916	0.92
r	0.594	0.916	0.909	0.93	0.936	0.935	0.956	0.957	0.959	0.93
slope	0.7	1.07	0.94	1.09	1.103	1.115	1.115	1.16	1.13	1.112
initial	467	232	318	196	225.8	226	232	194	177	157

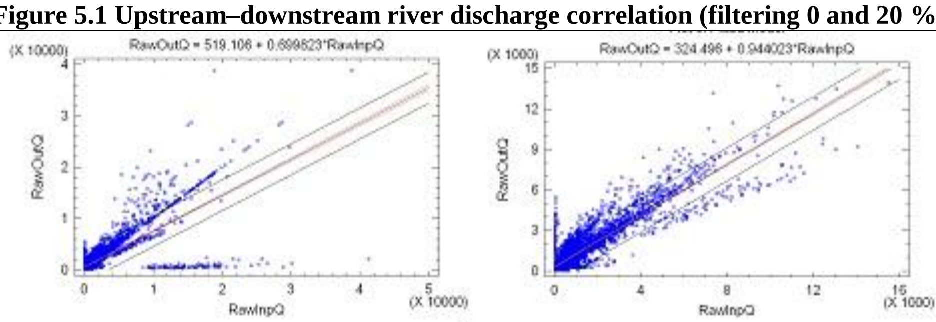
Source: Accounts v5, November 2012, Pöyry data preparation and computation by EEA.

This table suggest rather important conclusions. The first conclusion is that the number of sub-basins having upstream inflows is less than 1/3 of all sub-basins. Indeed many sub-basins are coastal that cannot have upstream catchment. This might however pose questions regarding the delineation of sub-basins and could require delineating smaller entities.

The second conclusions relates to data quality, (as computed by scoring, even though scoring is an imperfect indicator) as driver of the relationship between upstream and downstream discharge. Analysis suggest a strong increase in relationship certainty as demonstrated by both the R^2 and the r coefficients in relation with the data scoring. The numeric value of r is more sensitive to the number of data couples involved, and may suggest a good relationship just because of the number of data. In all circumstances, data scoring below 20 % should signal that caution is called for in considering the balances. Policy decisions may only be taken based on scoring above 50 %. This second proposal is backed by the instability of both value at origin and slope of the relationships.

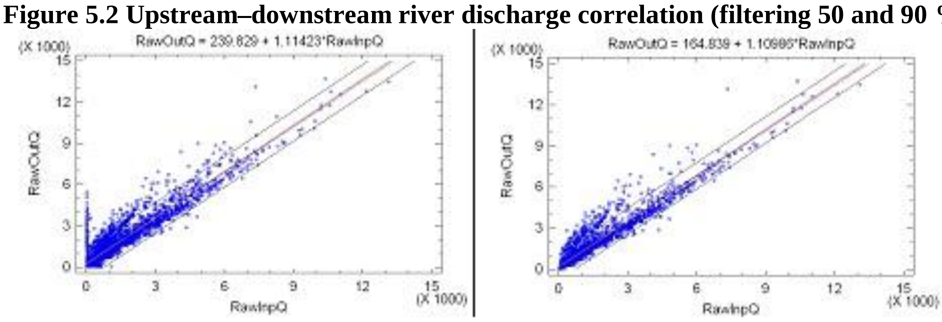
This instability is further understood by examining data distribution: it seems that four categories of data are addressed and stepwise eliminated when increasing the filtering constraint. This is displayed in Figure 5.1 and **Error! Reference source not found.**.

Figure 5.1 Upstream–downstream river discharge correlation (filtering 0 and 20 %)



Source: accounts v5, November 2012, Pöyry data preparation and computation by EEA.

Figure 5.2 Upstream–downstream river discharge correlation (filtering 50 and 90 %)



Source: accounts v5, November 2012, Pöyry data preparation and computation by EEA.

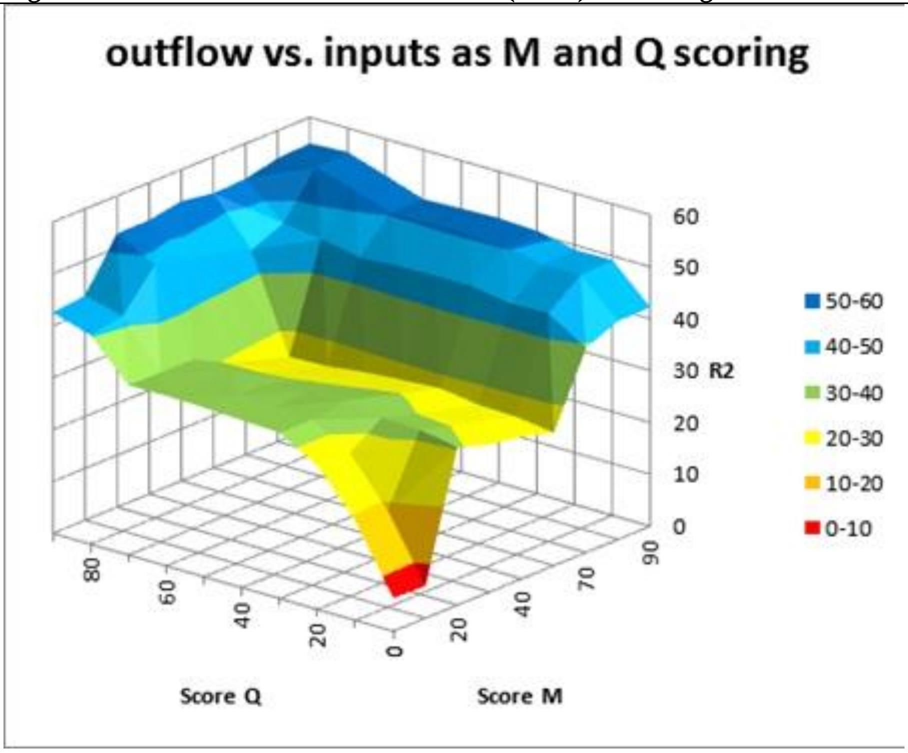
The cases in Figure 5.1 and **Error! Reference source not found.** are very instructive.

- No filtering (Figure 5.1, left) displays a series of data in which a large range of upstream (variable RawInpQ) values refer to virtually nil outflows (variable RawOutQ). Such cases are possible (see Footnote 450), but seem related to questionable data, since the points are no longer present after 20 % score filtering (same figure).
- In all cases, there is seemingly a twin population of relationships (very obvious in **Error! Reference source not found.**, right): one in which the outflow grows more than the inflows (slope >1), and another in which outflows decrease (slope <1). The analysis of these cases is beyond the scope of this report, but is an important question to solve and filter with appropriate flags in further production of the accounts.
- There is, in the three first filtering levels, up to 50 % included, a set of couples with very limited input and a quite large range of outflows (the points close to the RawOutQ axis) that are possibly the result of data insufficiency, despite certain filtering. This is the symmetrical of the case mentioned in first bullet (large input, no output, in this case the large output with no input are stranger) but is unlikely to result from reservoir management or transfers until fully documented). If management reasons are suspected, they should be documented as well, to flag cases.

5.1.4 Check of the balance effective rainfall – outlet discharge

In the case of the accounts, it was considered that the reconstructed outlet discharge (variable Local0, as described in Subsection 5.1.2) should correlate with the current month's effective rainfall and the four months before, to take into consideration the way river flows are affected by rain and transfers across the catchment. One hundred regressions were carried out for all filtering combinations, and the results are displayed in Figure 5.3. This calculation addresses the possibility to finalise the I/O tables by comparing effective rainfall to river discharge, disregarding if there is an upstream catchment or not. Appropriate cumulative data preparing have been done to this end.

Figure 5.3 Distribution of the R2 coefficient (in %) vs. scoring for rainfall–outlet discharge



Source: accounts v5, November 2012, Pöyry data preparation and computation by EEA.

The figure uses a gradient of colours to indicate the distribution of R^2 coefficients (that express the percentage of variance of the dependent variable by the independent variables) as a rather regular function of data quality.

This simplistic approach does not aim at evaluating a rainfall – discharge model but only checking the likelihood of the relationship between effective rainfall (primary source of water, highly estimated) and the result of the process (river discharge, primarily resulting from monitored data) and if they match well disregarding the scaling: the R^2coefficient tell nothing about the factors, which analysis would be beyond the scope of this paper, all the more they address in almost all cases 5 independent variables. The one to one balance between volumes is the point addressed by the I/O tables proper.

The information from the figure can be summarised as follows.

1. When both scores are low (below 30 %), the relationship is poor and likely non-existent. In such cases (i.e. for a large share of Europe), achieving a full balance is not possible.
2. When both scores are excellent, the relationship is good, considering the many factors of imbalance; the full accounts can be carried out with a high likelihood of accuracy.
3. When one of the scores is very good to excellent and the other poor, the relationship lies in a domain where the change in quality is very abrupt: a change of 10 points in the worst of scores means a drop up to 30 % in R^2. However, resilience to bad meteorological scoring is stronger than that related to discharge. This is probably because the meteorological data, even when of poor quality, are spatially distributed and hence less sensitive to lower data quality than river discharge, which depends totally upon the river and is hence less substitutable.

4. When both lie in a median range of quality, the relationship is possible, but difficult to qualify.

From a practical point of view, the most critical factor is river discharge data; this is why this information is overlaid to the maps presenting the results.

For further assessments, it seems clear that significantly better scoring should be targeted: scores over 60 % do not demand much data and are achieved in many countries for the time being.

In conclusion, there is no guarantee that good scoring will secure an accurate relationship between effective rainfall and river discharge; only the existence of a representative relationship is assessed. By contrast, the poor scoring is a high likelihood of poor relationship. This is partly explored in the next section.

5.1.5 Global assessment based on trend of stock closing

A simple method based on cumulative volume of water balance over a long period called a 'shift' can be used to examine the accuracy of results on basins. In fact, this shift should be equal to 0 in a long 'standard conditions' period and without global warming impact, meaning that equilibrium is reached between inflows and outflows. This approach is based on the verified assumption that on the long term inputs and outputs are balanced at any catchment level.

In order to assess the accuracy of water balance on such a calculation period, this shift can be weighted with the only and external inflow going to the basin, and precipitation, which on most basins represents more than 95 % of total inflows (excepting 'Total 4b': 'Inflows from resources in the territory', which is equal to 'Total 7c': 'Outflows to other resources in the territory') because there no upstream territories for a whole basin.

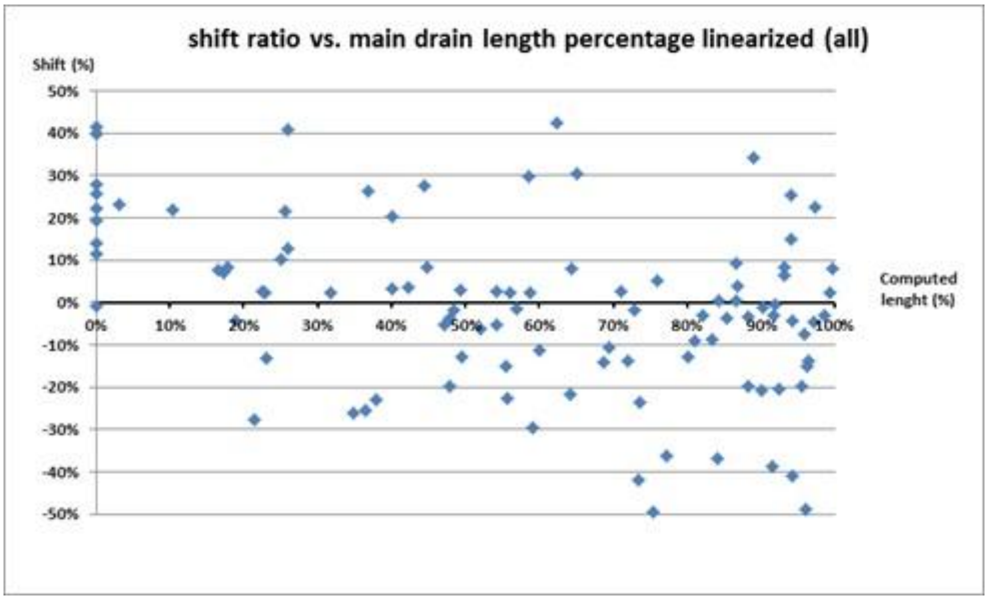
For all basins, such a shift ratio can be calculated as:

Annual shift ratio = total (shift over period / (average rainfall x number of years for period)

The results fall into two categories: all basins together, and only those having more than 80 % of their main drain linearised, i.e. basins with enough flow data to be calculated (independently of flow data quality).

The assessment quality probably relies on two components: the proportion of linearised rivers in the catchment (quantity of river covered) on the one hand, and the quality of river discharge data availability. This is what is shown in the next figures.

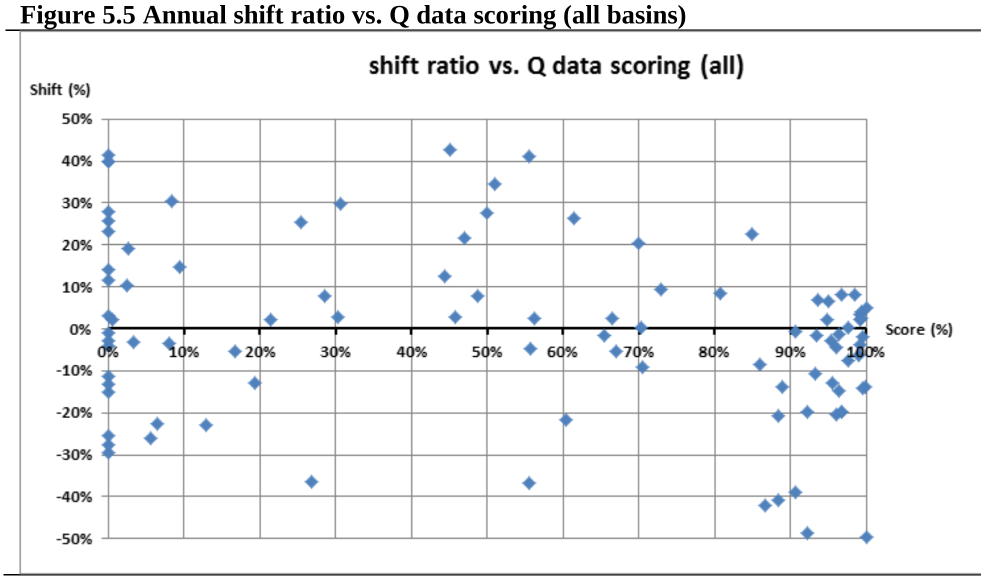
Figure 5.4 Annual shift ratio following main drain length



Source: Pöyry computations, post reporting.

Figure 5.5 shows a larger proportion of small shifts when the computed length increases.

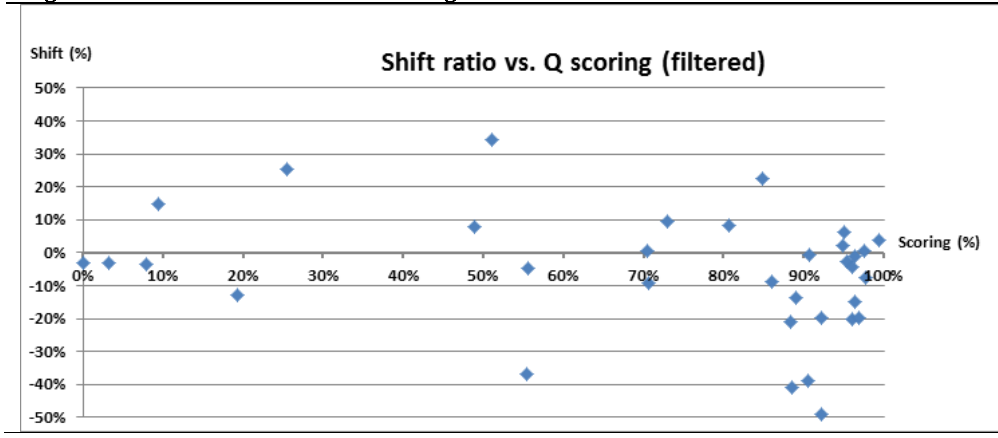
Figure 5.5 Annual shift ratio vs. Q data scoring (all basins)



Source: Pöyry computations, post reporting.

Combining both filters suggests first a global improvement of the relationship (as this had been already suggested from the existence of a relationship in previous section) and a grouping around in the ranges -20% / +20% and another narrower grouping in the range -10% / + 10% with some outliers, that indicate the need for better inclusion of rainfall data and groundwater exchanges.

Figure 5.6 Annual shift ratio following score 0 for all basins



Source: Pöyry computations, post reporting.

Comment: filtering refers to the selection of only those basins whose percentage of computed river length is over 80 %.

Most of the basins with annual shift ratio worse that -10% have underestimated rain supply, or more rarely, have overestimated outlets flows. Several basins were analysed individually, while checking the basin reference values in order to determine the type of data to generate the shift. An example is given for the Dee Basin (ZG00000030) in the following table.

Table 5.2 Sample balance: the Dee Basin illustration of the 'shift analysis' approach

Item	Value	Unit	Source, then balance	
Area	1 816	km^2	Up to Chester Weir	
Area modelled total	1 928		ECRINS	
Rain UK met basin (Chester Weir)	1 110	mm/year	1 999	
Rain modelled	994	mm/year	1 806	hm^3/year
Actual evapotranspiration	500	mm/year		
Actual evapotranspiration modelled	459	mm/year	884	hm^3/year
Module	520	mm/year	937	hm^3/year
Module modelled (avg.)	996	mm/year	1 794	hm^3/year
Abstraction			22	hm^3/year
Abstraction river modelled (avg.)			56	hm^3/year
Shift per year			**- 895**	hm^3/year
Explanation shift	on flow	857	857	hm^3/year
	on rain	- 193	193	hm^3/year
	on ETP	- 41	- 41	hm^3/year
	Total		**1 009**	hm^3/year

Source: Pöyry computation, post-delivery of report to DG Environment.

Table 5.2 illustrates how large flow outlet overestimation and small rain underestimation can generate very large shift values and ratios (see the 'Explanation shift' row).

5.2 *Expected results from the accounting methodology*

The usual presentation of the accounts follows the I/O tables that are specified by the SEEAW (copies accompany this report). These tables are produced in Excel workbooks, under a very similar presentation to that of the Nopolu System2 application used to manage all the databases. This is detailed in Annex 1 .

As mentioned in the introduction, the number of tables is considerable: as an order of magnitude, 411 sub-basins × 8 years × 12 months mean potentially 39 456 tables, plus the aggregation tables at year and RBD levels, for example.

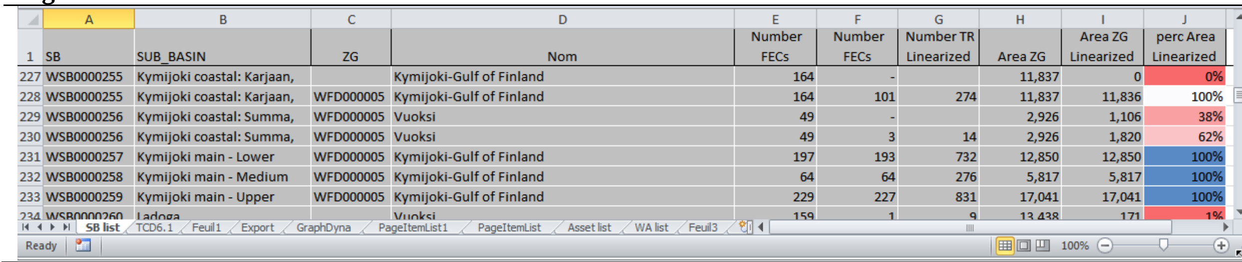
This is why carrying out the water accounting at a detailed level makes classical exploiting of results impossible : the I/O standard tables as unique output of results are obsolete approach. By contrast, these tables should now be considered dynamic elements that allow the lay user to examine the synthesis for a certain basin and a certain period, and not just the single outcome of accounting.

The results are entered into dynamic Excel workbooks, one for sub-basins and one for districts, and are automatically updated by the water accounting application. The following figures present selected views.

5.2.1 Summary information on the information produced

The Excel workbooks are built on the same structure and contain several sheets to make all documentation simple and accessible.

Figure 5.7 Excel workbook of results: documentation of entities



Source: Pöyry production for DG Environment

Comments: the screen has been reduced in size to display the name of sheets present.

This sheet displays all the information about the way each and every sub-basin (district) has been computed: its code (ECRINS), its district (coded ZG), the number of FECs that compose it, the number of FECs actually computed, the number of river segments in the main drainage system, and the number of these segments for which discharge has been computed, along with the area drained by these segments.

This information is all the more important for detailed analysis, because it expresses the elements of 'likelihood': a large drainage can be computed from few segments if only a couple of stations are involved.

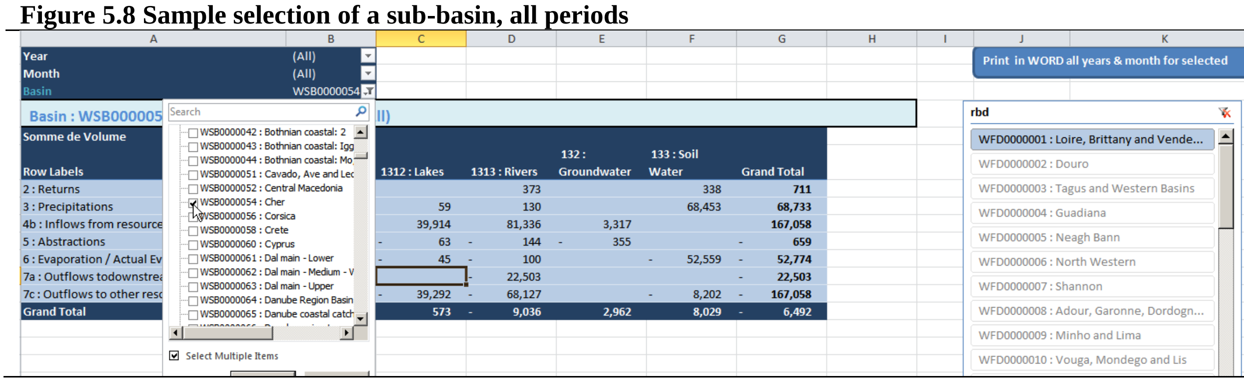
5.2.2 Ancillary information

Nomenclature is set out in two sheets: the WA_list and asset list display and define the water accounting elements (natural systems with their codes) and the assets elements (abstraction, outflows, etc.).

5.2.3 Tabular and graphic presentation facilities

All information (the compiled data sets) has been copied by Nopolu System2 into the sheet Export. This sheet is exploited by two sets of dynamic pivot procedures; one is for making tables, the other for making graphs. Sheet TCD6.1 makes the tables and sheets, and GraphDyna makes the graphical presentation that is useful for synthesising the seasonal development of the balances.

Figure 5.8 Sample selection of a sub-basin, all periods



Source: Pöyry delivery to DG Environment.

Selecting a sub-basin dynamically lightens the list of the RBDs of attachments, as shown in the figure. The sub-basin selected is displayed in the upper left box.

In the display, the Cher (affluent of the Loire river) is aggregated over all the years for all months: any selection can be made and displayed. Figure 5.9 displays the same basin for year 2001 and the month of June.

Figure 5.9 Sample selection of a sub-basin, one month and one year

Year	2,001
Month	6
Basin	WSB0000054

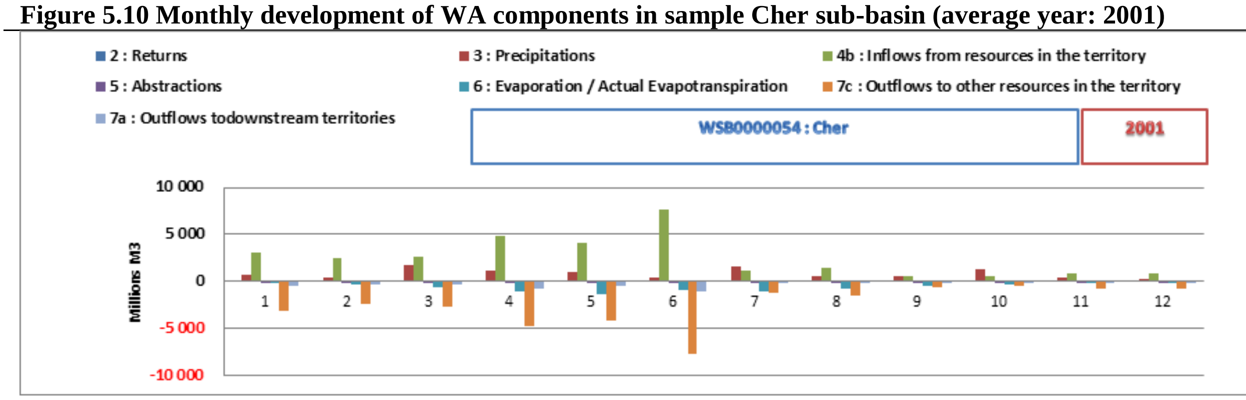
Basin : WSB0000054 : Cher Year : 2001 Month : 6

Somme de Volume / Row Labels	1311 : Reservoirs	1312 : Lakes	1313 : Rivers	132 : Groundwater	133 : Soil Water	Grand Total
2 : Returns			4		3	7
3 : Precipitations	1	0	1		414	416
4b : Inflows from resources in the territory	2,100	1,919	3,663			7,682
5 : Abstractions	- 1	- 1	- 1	- 3		- 6
6 : Evaporation / Actual Evapotranspiration	- 1	- 1	- 2		- 925	- 929
7a : Outflows todownstream territories			- 1,067			- 1,067
7c : Outflows to other resources in the territory	- 2,669	- 1,703	- 3,310			- 7,682
Grand Total	- 570	- 215	- 712	- 3	- 508	- 1,579

Source: Pöyry delivery to DG Environment.

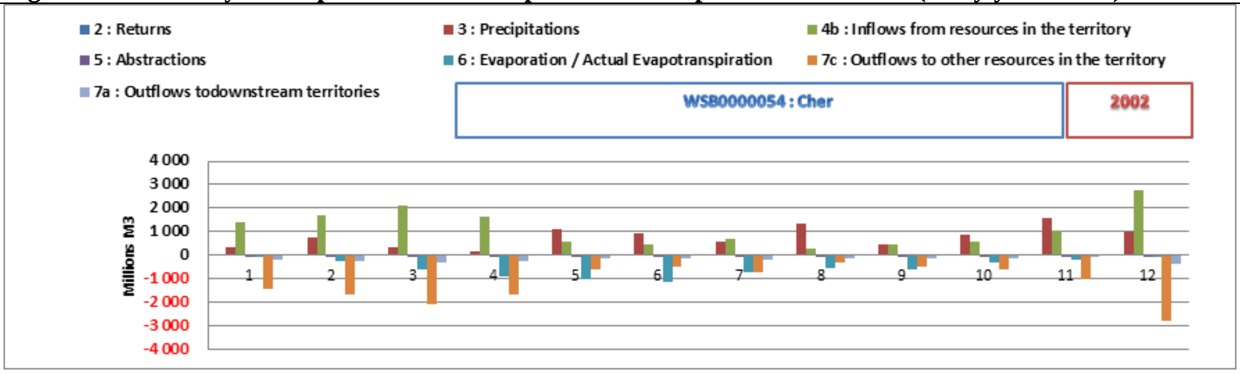
Such a display reveals the figures and, as previously mentioned, is of interest only at regional level. Since the detailed data have become available, the trend analysis (seasonal or interannual) is also possible. The Excel workbooks have hence been fitted with facilities to make this graphical analysis, from the same data set imported into the Excel workbook.

Figure 5.10 Monthly development of WA components in sample Cher sub-basin (average year: 2001)



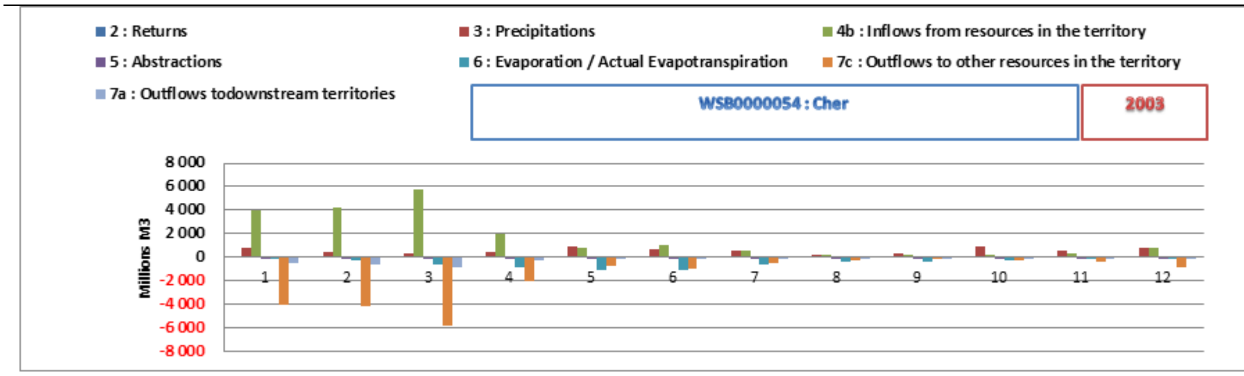
Source: Pöyry delivery to DG Environment.

Figure 5.11 Monthly development of WA components in sample Cher sub-basin (rainy year: 2002)



Source: Pöyry delivery to DG Environment.

Figure 5.12 Monthly development of WA components in sample Cher sub-basin (dry and hot year: 2003)



Source: Pöyry delivery to DG Environment.

These three figures exemplify the importance of seasonal analysis and interannual trends analysis, especially where the hydrological features of the basin are limited groundwater storage capacities, average irrigation demand, and insignificant surface water usable reservoirs.

5.3 Results presented as spatial indicators based on the WEI

5.3.1 Data preparation

Previous discussions strongly suggested that a classical analysis based on tabular exploitation would be unnecessarily tedious, and would not achieve the overarching goal of providing to policymakers a clear view of the water situation and its developments. Results from the water accounting consolidation are imported and processed by a set of functions in the ancillary database W_Wany4Map.mdb, which has been specially designed by the EEA to process the results. The functions in this database compute the scores, import the results, and process them under three tables, preformatted to be used by ArcGIS®[63] map documents.

The functions create three different tables sharing the same format; they contain the aggregates at the sub-basin level over the period: indicator percentiles; number of data; and the arithmetic, quadratic and harmonic averages of monthly indices, as resulting from the rationales developed in Subsection 3.3.6 .

[63] http://www.arcgis.com/about/

The three tables, WEIa4ForMap, WEIb4ForMap and WEIPF4ForMap, manage respectively the 'ecological share', the normalised water exploitation index, and the quasi WEI+ values (see Subsection 3.3.5 for definitions).

5.3.2 Overall results for the normalised WEI

The overall results are summarised as statistics of nWEI, based on sorting out the quadratic average over the period, as this average was set the most relevant for aggregating the monthly WEIs. The quadratic means are sorted into classes 0–.1, .1–.25, .5–.5; .5–.75 and .75–.9, as suggested in Subsection 3.3.6 . The results are reported in the next Table 5.3

The results are reported in Table 5.3.

Table 5.3 Synthesis of the nWEI, based on averaged Quadratic means

Class	no entities	average Q mean	Min. (10 %)	Max. (10 %)	Min. (50 %)	Max. (50 %)	Min. (90 %)	Max. (90 %)
ND	57	NC	NC	NC	NC	NC	NC	NC
0.0–0.1	204	0.03	0	0.04	0	0.1	0	0.17
0.1–0.25	87	0.16	0	0.16	0.02	0.26	0.15	0.5
0.25–0.50	46	0.36	0.02	0.5	0.15	0.5	0.3	0.69
0.50–0.75	17	0.54	0.34	0.68	0.46	0.68	0.51	0.71
	411							

Source: accounts v5, November 2012, Pöyry data preparation and computation by EEA.

Of more than 411 sub-basins, one half are in the interannual WEI average of less than 10 %, and 57 (14 %) could not be computed, because of missing essential data, in this case, only the outlet information. This means that at least half of the sub-basins are not under systematic water scarcity threat.

By contrast, 87 sub-basins are in the 10 % to 25 % range, meaning that (on average) 16 % of resources are at any one time incorporated in the economy, possibly reaching 15 % to 50 % of resources with a return time of one month per year. This rate suggests possible harm to the ecosystem, while it does not suggest a significant risk for water provisioning. However, since the uses are rather underestimated, this class and the basins involved are to be kept under scrutiny after data revision.

The two last classes group 46 and 17 sub-basins, totalling 63, with a percentage in number in sub-basins of between 15 % and 18 % of the total number of computed catchments, on the extreme hypothesis that the non-documented basins are all without problems or are equally apportioned across the classes.

In these basins, the average quadratic mean of monthly WEIs ranges between 36 % and 54 %, meaning a huge pressure on resources. In the scarcest group, the 10 % nWEI (those reflecting the high water period) is also very high, suggesting a structural scarcity for at least 17 % and up to 20 % of sub-basins.

The last group probably presents two different categories and is likely to also comprise sub-basins in which the scarcity is more a recurrent than structural issue; this is clearly suggested by the mapping of the nWEI in the next sections, where geographical distribution is the focus.

5.3.3 Overall results of the pseudo WEI+

The 'pseudo WEI+' emulates the WEI+, that is, as demonstrated in Subsection 3.3.4 , the net consumption. Since consumed volumes are necessarily much below the abstracted ones and their impact is different, a different set of thresholds is suggested in Subsection 3.3.6 . The distribution of results, with the same variables as presented for the nWEI, are presented in The same number of 57 sub-basins could not be computed, because at least one key category of data was missing.

Table 5.4.

The same number of 57 sub-basins could not be computed, because at least one key category of data was missing.

Table 5.4 Synthesis of the pseudo WEI+, based on averaged Quadratic mean

Class	no entities	Mean Q mean	Min. (10 %)	Max. (10 %)	Min. (50 %)	Max. (50 %)	Min. (90 %)	Max. (90 %)
ND	57	NC	NC	NC	NC	NC	NC	NC
0.00–0.01	164	0	0	0.01	0	0.01	0	0.02
0.01–0.05	97	0.03	0	0.02	0	0.05	0	0.1
0.05–0.15	67	0.09	0	0.12	0	0.14	0.05	0.3
0.15–0.25	24	0.19	0	0.19	0	0.24	0.17	0.43
0.2–0.50	2	0.31	0	0.36	0.17	0.36	0.36	0.43
Together	**411**							

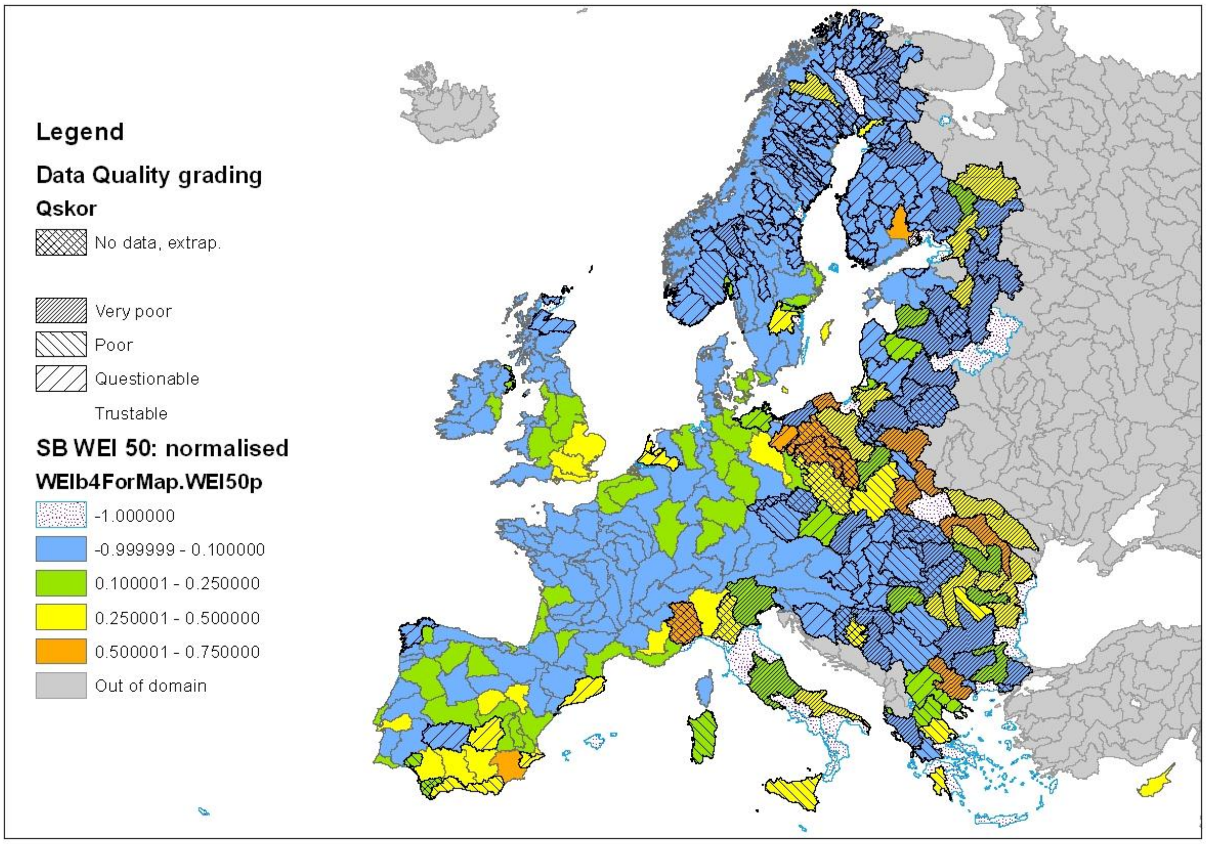
Source: accounts v5, November 2012, Pöyry data preparation and computation by EEA.

Similarly to the true exploitation, a large number of catchments have a reasonable proportion of water consumption: 164 + 97: 261 over 411 (354) basins are below 10 % of consumption, in crest periods.

However, three groups (the most consuming being limited to only two sub-basins (hence lumped together with the before last in next comments)) present both a high interannual average (in practice, ~10 % and ~20 % of resources are totally consumed) and 90 % values close to 50 %, meaning a structural overuse of water. The maximum values are not computed despite being mapped, because of uncertainty over data. On average, 16 % to 19 % of sub-basins are likely in significant overconsumption of resources, whereas 6 % to 7 % are in sharp overuse of resources. Their geographical distribution is presented in next sections.

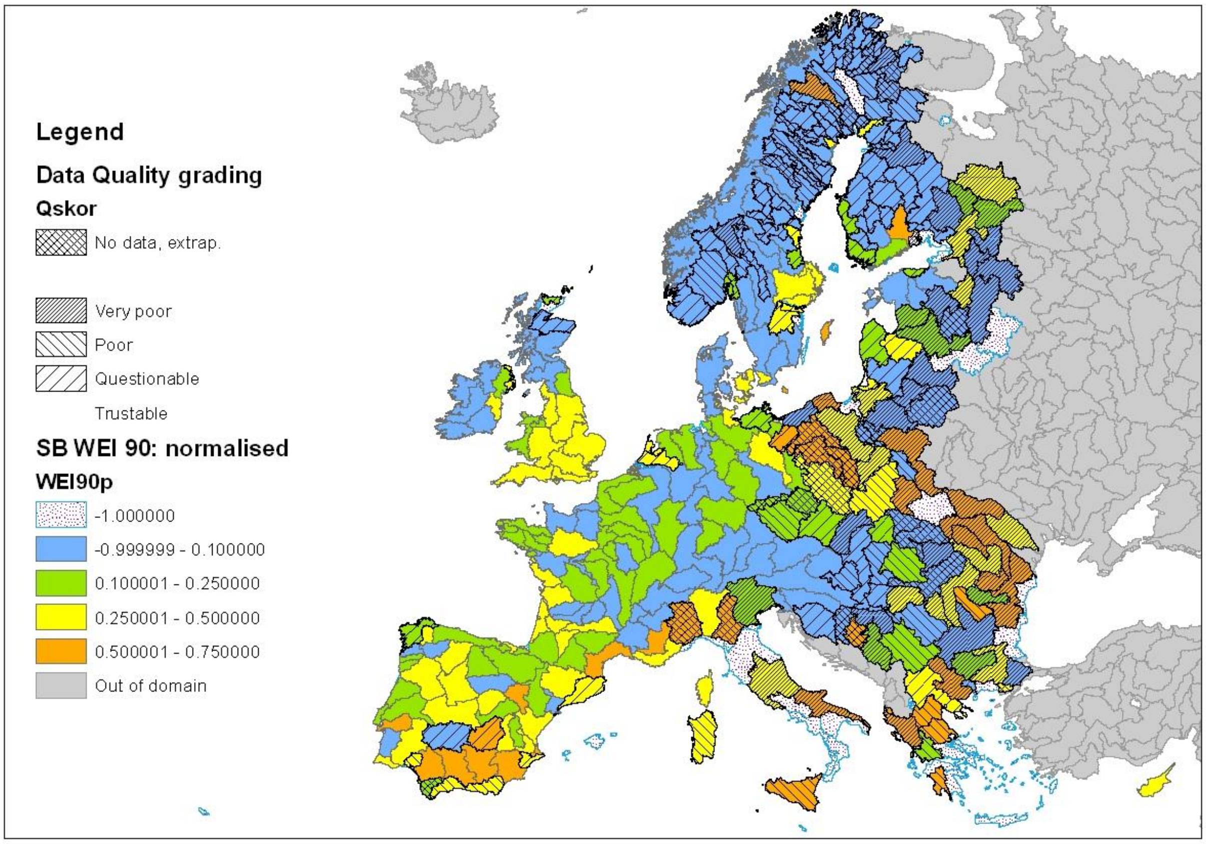
5.3.4 Mapping of the normalised WEI

Map 5.1 The 50 % normalised WEI based on water accounting



The distribution of the 50 % (median) monthly nWEI is not limited to traditionally dry areas: high demands are the fact of intense human activities in small areas with a reasonable supply; the factor of relative scarcity is hence more a matter of population density than of meteorology. The statistics in Table 5.3 should be considered with caution, since many of the areas with greatest scarcity are in badly scored catchments.

Map 5.2 The 90 % normalised WEI based on water accounting



Source: WA calculations, August 2012 EEA from Pöyry data sets for DG Environment.

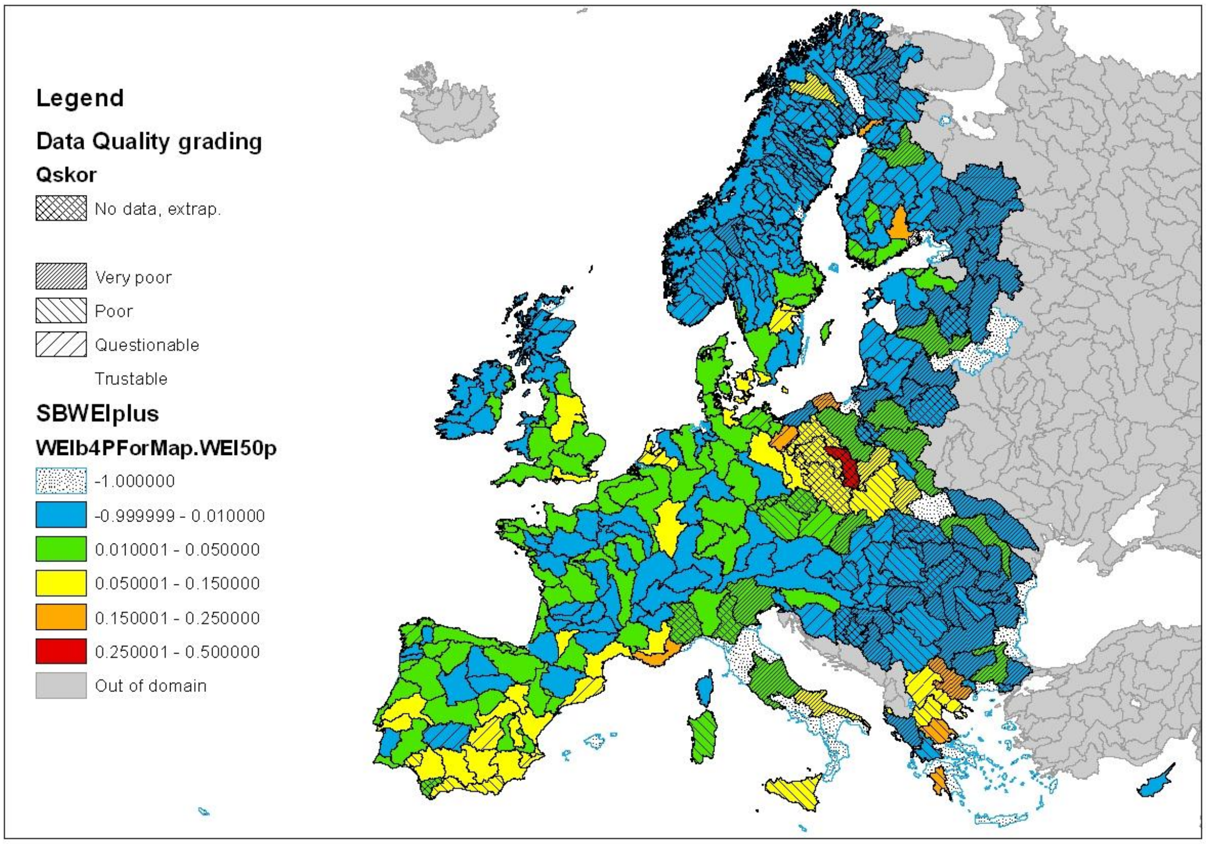
Comparing the maps offers important insights that reflect the rather spatially extended issue of potential water issues across Europe. Only dark green sub-basins have sufficient water; many sub-catchments are either in light green (hence in a range of 20 % demand versus resource, generally considered as the problematic ratio), and shift to much larger ratios in non permanent but not that exceptional situations (a 90 % percentile on a monthly basis means 12 months worse every 10 years on the average, which can be a sign of a chronic tension situation).

The map of the 90 % nWEI confirms and makes more precise the former analysis: many sub-basins with acceptable median nWEI drop one or two grades in the dry period of the year and face systematic droughts, forming a median scarcity level. This is, for example, the case in the Atlantic sub-basins of France. This mapping corresponds with findings by the French authorities, which suggests accuracy of data.

There is also a substantial share of nWEI falling in poorly scored areas, making the statistics somewhat fragile.

5.3.5 Mapping of the pseudo WEI+

Map 5.3 The 50 % pseudo WEI+ based on water accounting

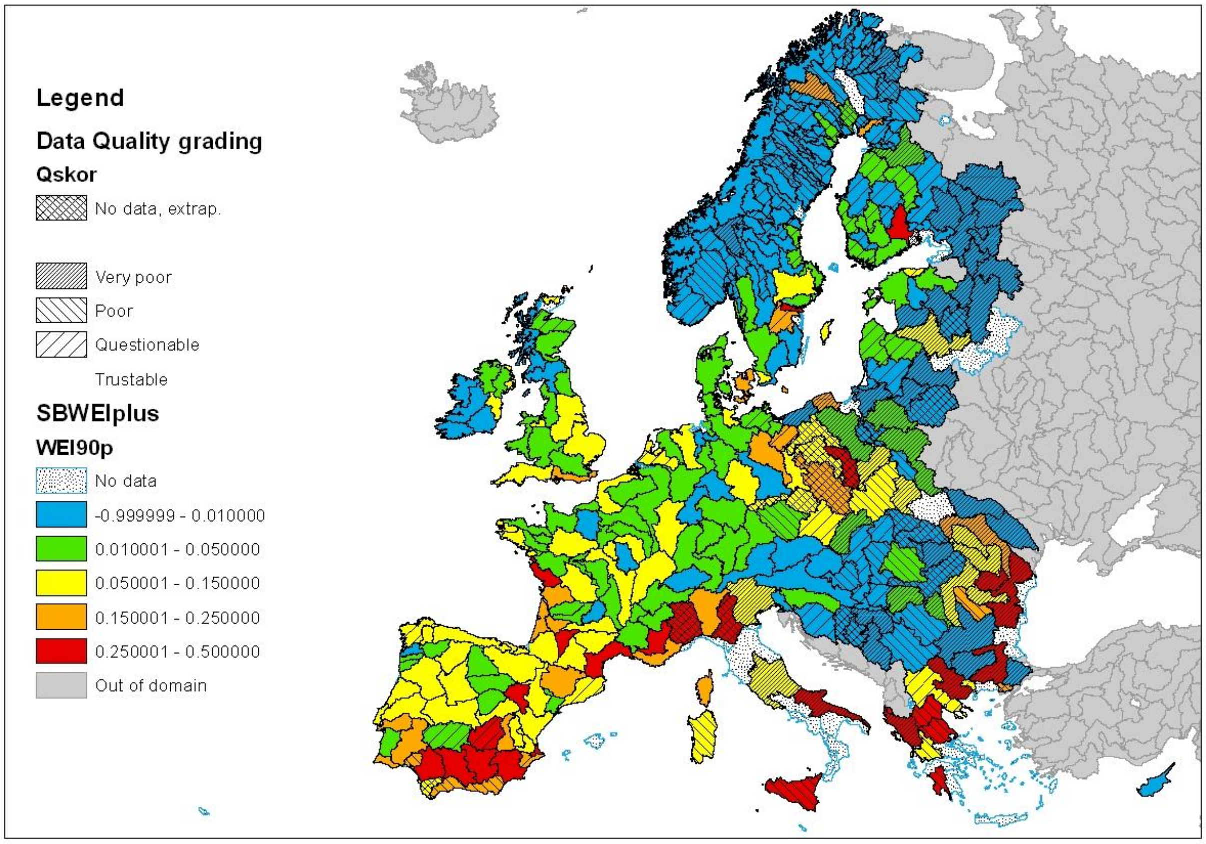


Source: WA calculations, August 2012, EEA from Pöyry data sets for DG Environment.

The spatial distribution of net consumption of water shows a rather regular pattern, with the exception of an extreme index in a Polish catchment, likely in relation to the quasi-absence of data in these catchments, likely to be erroneous. The high consumptive indices in Finland and other areas with potentially sufficient water availability may be the outcome of erroneous consumption values. In the complete absence of comments on the data submitted, the values used in the accounts are understood as 'not incorrect'.

Regarding the peak consumption index (pseudo WEI+ 90 %), the spatial distribution of values seems rather accurate (until it is demonstrated otherwise by better data). The low consumption index on the central Guadiana basin in Spain possibly reflects either the mapped insufficiency in river discharge, or the significant difference in the irrigation volumes between the JRC source used and the proposed overall volumes provided by Spain (albeit lacking spatial information, hence making the value not yet usable).

Map 5.4 The 90 % pseudo WEI+ based on water accounting



Source: WA calculations, August 2012, EEA from Pöyry data sets for DG Environment.

5.3.6 Variability

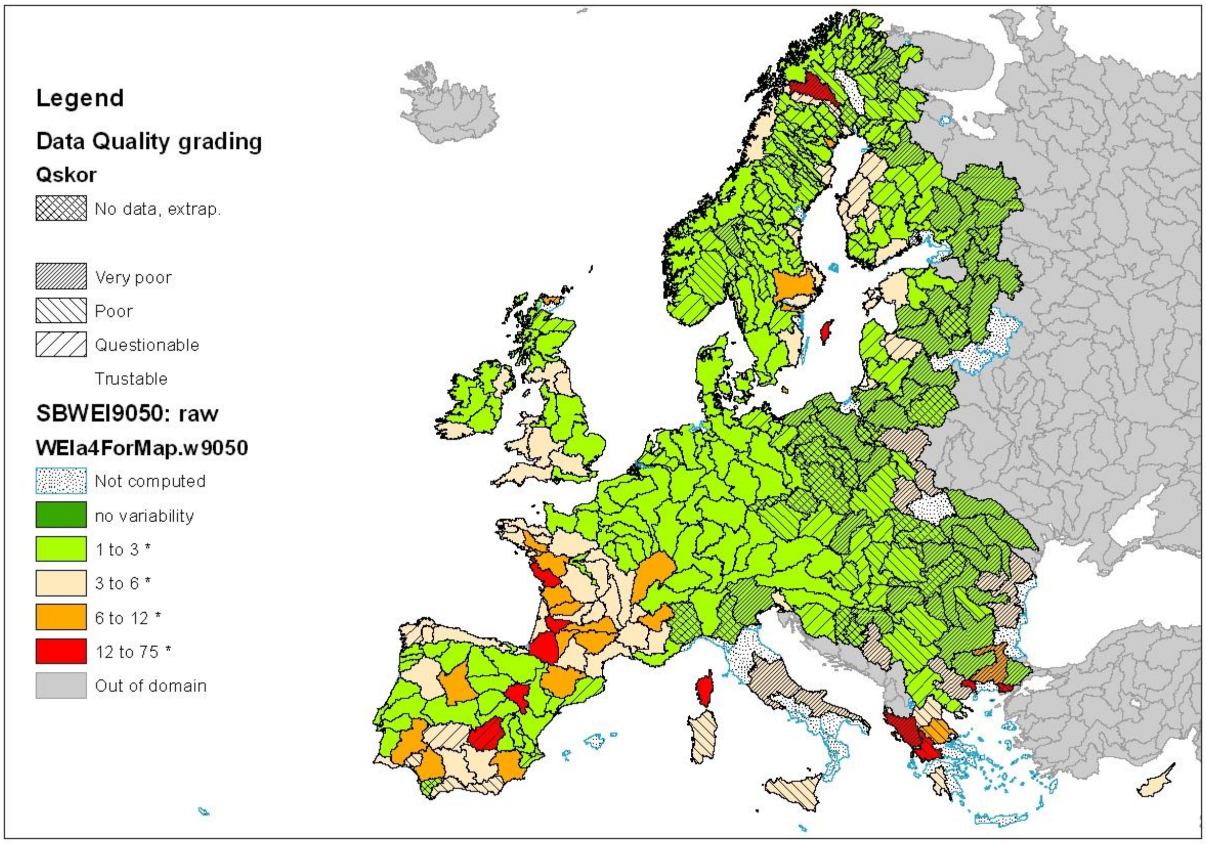
Natural variability of the water regime between modules and low water is a key element of water management. A natural variability of 10-fold to 20-fold (module/low water discharge) is considered the rule of thumb in establishing regulations for establishing the permissions to abstract water in the Atlantic. Such data are available in the water accounts data sets, however, an indicator based on nWEI was used instead, because it is simpler to compute with the data sets computed for the indicators displaying, and it displays the same information if demands are in a narrow range.

To assess variability, the ratio of the 90th percentile to the 10th percentile and the 90th50 th seemed good candidate indicators that can be mapped. The 10 th nWEI is the smallest being trespassed in 90 over 100 cases, and 90 th is the largest, being trespassed in only 10 % of the months. The second ratio was preferred because it better exhibits scarcity versus standard conditions rather than scarcity versus abundance conditions.

Such ratios can be better computed with raw WEI since the ratio of raw WEI has a very similar abstraction value (the seasonal change in demand is small compared to the seasonal change in resources).

The map shows a very high variability of ratios of WEI in the West of Europe and Mediterranean. The elements of Map 5.5 are much in line with the map: WEI ratios in the Czech Republic are in the range of ~2 when considering the same distribution threshold: annual 50 th is ~13 % one year over 2, the value of the index is 13%), and the 90 th largest WEIs are around 25 % to 30 % of consumed resources, with these independent estimates.

Map 5.5 Variability of demand as from the ratio 90 th to 50 th raw WEI



Source: WA calculations, August 2012 EEA from Pöyry data sets for DG Environment

In many areas, ratios often exceed 10, making the low water periods more prone to pose threats for the environment. This can be completed by mapping the 90 th raw WEIs that indicate which share of the initial resource is still present after exploitation.

5.3.7 Ecological share of water

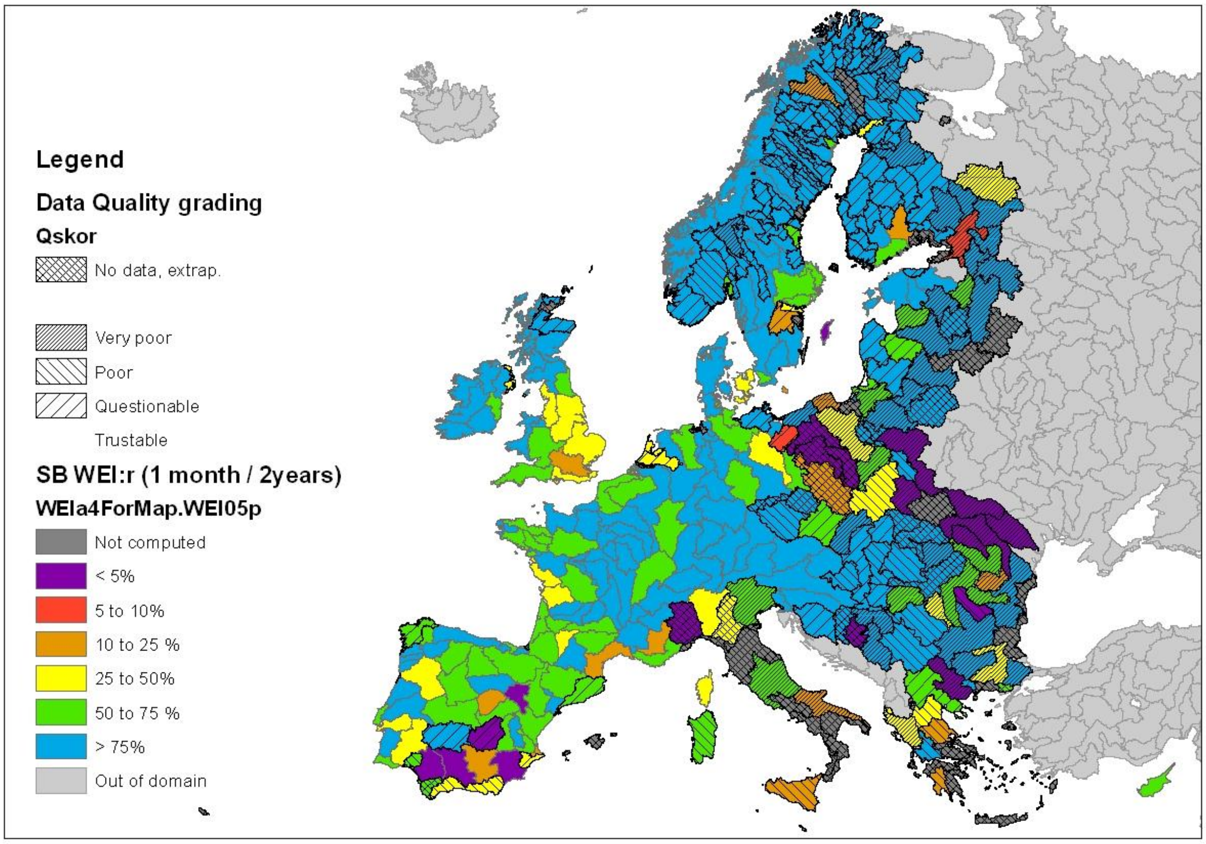
Amongst the most important issues to be discussed is indeed the ecological share of water; its policy definition and target are not discussed here. The assessment of the ecological water share is important, as it is expressed thanks to the raw WEI. The resource at the outlet incorporates the returns; hence the raw WEI expresses the quantity of natural flow used in the economy versus the quantity remaining for the downstream hydrosystems. Even the coastal systems can be impaired by lack of freshwater inputs, and hence the coastal sub-basins are considered as well.

The indicator symmetric of the WEI is the 'Ecological stress index', proposed in Subsection 3.3.5 : it shows the proportion of resource remaining after use of water (disregarding consumption, for reasons explained in Subsection 3.3.5) for the natural systems and for the downstream uses. It is not exactly the 'freedom of action' indicator proposed by Margat (1996), but is in the same spirit.

However, this indicator is rather ambiguous; it does not reveal the ecological availability of the resource, but only the economic pressures on the resource. Accurate information on ecological flows is still to be derived from water accounts background information. It should come from a statistical confrontation between discharges per segment, and waterbody status, as they are now attached to ECRINS.

The map of potential stresses, as computed from the percentage frequency (1 month every ~2 years on the average) is reported in Map 5.6.

Map 5.6 Share of untouched resource (available for ecological uses) at the 5 % frequency



Source: WA calculations, August 2012, EEA from Pöyry data sets for DG Environment.

Statistics elaborated with all the necessary precautions in relation to their representativeness are shown in Table 5.5 below. This table supplements the information reported in Map 5.6.

Table 5.5 Synthesis of the IEStr, based on averaged Quadratic means

Class	no entities	Mean Q mean	Min. (2 %)	Max. (2 %)	Min. (5 %)	Max. (5 %)	Min. (10 %)	Max. (10 %)
NC	2							
0–5 %	74	0.04	0	0.03	0	0.03	0	0.04
5–10 %	6	0.35	0.05	0.1	0.06	0.15	0.07	0.18
10–25 %	15	0.51	0.1	0.22	0.11	0.27	0.12	0.35
25–50 %	33	0.73	0.26	0.48	0.3	0.61	0.32	0.67
50–75 %	72	0.85	0.5	0.74	0.52	0.82	0.53	0.86
Over 75 %	209	0.97	0.75	1	0.77	1	0.78	1

Source: WA calculations, August 2012, EEA from Pöyry data sets for DG Environment.

Comments: Figures express the share of estimated resource being left untouched by abstractions.

Data in the table are the interannual quadratic mean of the EIStr, sorted by class of values and by categories of return period. The interpretation of the return period is not straightforward and is therefore summarised here. The smaller the percentage, the less frequent the observation is. In general, less frequent observations are connected with more stringent situations, expressed by the class values. For example, if only 10 % of the resource is free to the economy, its impact is all the more important when it happens more frequently. This is why the three return periods of 2 %, 5 % and 10 % (1 month over 50, 1 over 20 and 1 over 10) are expressed. Indeed if a lack of natural resources is permanently experienced, it is no longer a threat but rather constitutes permanent damage.

The first class, 0–5 % is possibly to some extent spurious, since it includes catchments with nil outflows; it may be an artefact of data collection. However, it can be seen that except in the most favourable class, there is a significant number of sub-basins that might experience difficult situations with frequent possible cases (one in the 10 % frequency).

This means that more attention should be given to further analysing the ecological stress of water uses and to the use of the very comprehensive database created thanks to the opportunity presented by the accounts. Moreover, the invaluable data represented by the integration of waterbody reporting in the same topological database, ECRINS, is an urgent and important task prior to the next EEA SoER.

5.4 Conclusions

Water accounts used for a water exploitation index are able to provide medium-scale overview information, which can help the further policy development, as discussed at the moment with the review of the Water Scarcity and Droughts (WS&D) policy and the 2012 Water Blueprint. However the same methodologies with more detailed, local level information can also provide detailed planning instruments for water resource management and water allocation. Therefore, the 2012 Water Blueprint explicitly considers water accounts as a tool for water resource management.

The share of consumption is significant and might be poorly sustainable in case of diminishing resources in 15 % to 25 % of catchments (including those that are at the higher limit of the median class).

The ecological situation of many sub-basins shows intense use and limited remaining flows; this is a concern given that ecological health is driven by recurrent stress situations much more than by average water availability. A difficult situation occurring on average one month per year is sufficient to prevent achievement of the targeted ecological status.

External references

EC, 2009, Communication from the Commission to the Council and the Parliament 'GDP and beyond. Measuring progress in a changing world' (COM(2009) 433 final of 20 August 2009).

EEA, 2001a, *Reporting river quality using the Water Quality Accounts methodology. Application within the Eurowaternet process*, Copenhagen, European Environment Agency.

EEA, 2001b, *Test application of quality water accounts in England and Wales*, EEA-IEPA England and Wales, Copenhagen, European Environment Agency.

EEA, 2001c, *Test application of quality water accounts in the republic of Ireland*, EEA-Irish EPA, Copenhagen, European Environment Agency.

Weber, J.-L., 1986, *Les comptes du patrimoine naturel*, Paris: INSEE.

Babillot, P., 1995, *Compte de l'eau sur le bassin de la Haute-Loire*, Institut *Français* de l'Environnement (IFEN), Orléans, 1995.

de Groot, R., Fisher, B. and Christie, M., 2010, 'Integrating the ecological and economic dimensions in biodiversity and ecosystem service valuation', in: Kumar, P. (ed.), *The Economics of Ecosystems and Biodiversity (TEEB): Ecological and Economic foundations*, TEEB, London, p. 422.

EC (DG Environment), 2007, 'Water Scarcity and Droughts. Second Interim Report' (http://ec.europa.eu/environment/water/quantity/pdf/comm_droughts/2nd_int_report.pdf) accessed 6 March 2013.

EEA, 2009, *Water resources across Europe - confronting water scarcity and drought*, EEA Report No 2/2009, European Environment Agency.

Margat, J., 1993, *Les ressources en eau. Conception, évaluation, cartographie, comptabilité*, Editions BRGM, FAO, Orléans, non public Rapport BRGM R 37267.

Lequiller, F. and Blades, D., 2006, *Understanding national accounts*, OECD, Paris.

Millennium Ecosystem Assessment, 2005, *Ecosystems and Human Well-being: Synthesis*, Island Press, Washington, D.C. (http://www.millenniumassessment.org/documents/document.356.aspx.pdf) accessed 6 March 2013.

Kaczmarek, B., 2006, *Un nouveau rôle pour les agences de l'eau? Essai pour une politique de l'eau rénovée*, Johanet, Paris.

UNSD, 2003, *Integrated Environmental and Economic Accounting. Handbook of National Accounting*, United Nations Statistics Division, United Nations, European Commission, International Monetary Fund, Organisation for Economic Co-operation and Development, World Bank, New York.

Carrasquer, B., 2012, *Integrating thermodynamics and European hydrographical information tools II. Taking into consideration the composition of rivers for a comprehensive water quality assessment through the 'exergy' accounting*, Copenhagen, Research working document.

Valero, A., Uche, J., Valero, A., Martínez, A. and Escriu, J., 2007, *Physical Hydronomics: application of the exergy analysis to the assessment of environmental costs of water bodies. The case of the Inland Basins of Catalonia*, Seminari sobre costos i comptes de l'aigua a Catalunya en relacio amb la Directiva marc de l'aigua, 18–19 June, Barcelona, Generalitat de Catalunya, p. 10.

de Marsily, G., 1995, *L'eau. Un exposé pour comprendre. Un essai pour réfléchir*, Flammarion, Paris.

UNSD, 2007, *Environmental accounting: System of Environmental-Economic Accounting for Water (Statistical Commission Background document)*, United Nations Statistics Division, UNSD Thirty-eighth session 27 February – 2 March, New York.

UNSD, 2012, *International recommendations for Water Statistics*, United Nations Statistics Division, Series M, No 91, New York.

Vogt, J., Colombo, R., Paracchini, M. L., de Jager, A. and Soille, P., 2003, *CCM River and Catchment Database*, European Commission, Ispra.

EEA, 2012, *EEA Catchments and Rivers Network System – ECRINS v1.1. Rationales, building and improving for widening uses to Water Accounts and WISE applications*, EEA Technical Report No 7/2012, European Environment Agency.

Margat, J., 1996, *Les ressources en eau. Conception, évaluation, cartographie, comptabilité*, Editions BRGM, Orléans, FAO, Vol. 28.

EEA, 2012, *Water resources in Europe in the context of vulnerability. EEA 2012 state of water assessment*, EEA Report No 11/2012, European Environment Agency.

Haylock, M. R., Hofstra, N., Klein, A. M. G., Klok, E. J., Jones, P. D. and New, M., 2008, 'A European daily high-resolution gridded data set of surface temperature and precipitation for 1950 - 2006', *Journal of Geophysical Research: Atmospheres*, 113(D20119). doi:10.1029/2008JD010201

Kurnik, B., Louwagie, G., Erhard, M., Ceglar, A. and Kajfez-Bogataj, L., (submitted for publication), 'Analysing seasonal differences between a soil water balance model and in-situ soil moisture measurements at nine locations across Europe,' *Journal of Environmental Modelling and Assessment*.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M., 1998, *Crop evapotranspiration - Guidelines for computing crop water requirements*, FAO, Rome.

FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009, *Harmonized World Soil Database (version 1.1)*, FAO, Rome, and IIASA, Laxenburg.

Saxton, K. E. and Rawls, W. J., 2006, 'Soil Water Characteristic Estimates by Texture and Organic Matter for Hydrologic Solutions', *Soil Science Society of America Journal*, 70(5), 1569–1578.

de Marsily, G., 2007, *Cours d'hydrogéologie*, Cellule TICE.

Seguin, J.-J., Mardhel, V. and Schomburgk, S., 2012, *Référentiel Hydrogéologique Français BDLISA (version beta). Présentation du référentiel. Principe de construction et mise en oeuvre. Rapport final*', BRGM/RP-61034-FR (http://www.brgm.fr/result/telechargement/telechargement.jsp?id=RSP-BRGM/RP-61034-FR) accessed 6 March 2013.

Beauzamy, B. and Zeydina, O., 2007, *Méthodes probabilistes pour la reconstruction des données manquantes*, Société de Calcul Mathématique, Paris.

Zeydina, O., 2011, *Méthodes probabilistes pour la Sûreté Nucléaire* (Doctoral thesis), Universit*é* de Bretagne-Sud (http://scmsa.eu/RMM/Olga_Zeydina_These.pdf) accessed 6 March 2013.

Wriedt, G. and Bouraoui, F., 2011, *A spatially distributed assessment of water allocation in EU27 for Year 2000*, Joint Research Centre (JRC), Publications Office of the European Union, EUR - Scientific and Technical Research Reports JRC66186. doi:10.2788/51970

Wikibooks, 2013, 'Algorithm Implementation/Strings/Levenshtein distance' **(**http://en.wikibooks.org/wiki/Algorithm_implementation/Strings/Levenshtein_distance#Visual_Basic_for_Applications_.28no_D amerau_extension.29) accessed 6 March 2013.

JRC, 2013, 'The Monitoring Agricultural ResourceS (MARS) Unit Mission', Joint Research Centre (http://mars.jrc.ec.europa.eu/) accessed 6 March 2013.

ATEAM, 2013, 'ATEAM — Advanced Terrestrial Ecosystem Analysis and Modelling' (http://www.pik-potsdam.de/ateam/) accessed 6 March 2013.

Gourbesville, P., 2009, *Le bassin versant du Var et la crue de 1994*, Groupe Interdisciplinaire de Réflexion sur les traversées sud alpines et l'aménagement du territoire Maralpin (GIR Maralpin), Hydro Europe programme (http://www.gir-maralpin.org/AmenagementTerrit/GourbesvilleHydroeuropeFleuveVar912.pdf) accessed 6 March 2013.

CIRCABC, 2012, 'Informal meeting of Water and Marine Directors of the European Union, Candidate and EFTA Countries. Copenhagen, 4-5 June 2012. Synthesis', Communication and Information Resource Centre for Administrations, Businesses and Citizens (https://circabc.europa.eu/sd/d/b5d535f4-0df2-4b7b-aa39-a71483ab8fd4/Synthesis_Water_and_Marine_Directors_Copenhagen.docx) accessed 6 March 2013.

The Club of Rome, 2013, 'The Club of Rome' (http://www.clubofrome.org/eng/about/4/) accessed 6 March 2013.

GRACE, 2013, 'GRACE — Gravity recovery and Climate Experiment' (http://www.csr.utexas.edu/grace/RL05.html) accessed 6 March 2013.

ESA, 2013, 'SMOS', European Space Agency (http://www.esa.int/Our_Activities/Observing_the_Earth/SMOS) accessed 6 March 2013.

FAO, 2013, 'Aquastat glossary', Food and Agriculture Organization of the United Nations, online 23 June 2009 (http://www.fao.org/nr/water/aquastat/data/glossary/search.html?_p=100&submitBtn=-1&keywords=Aquitard&subjectId=-1&termId=-1&submit=Search) accessed 6 March 2013.

FAO, 2013a, 'Aquastat data query', Food and Agriculture Organization of the United Nations (http://www.fao.org/nr/water/aquastat/data/query/index.html?lang=en) accessed 6 March 2013.

EEA, 2012, *EEA Catchments and Rivers Network System – ECRINS v1.1. Rationales, building and improving for widening uses to Water Accounts and WISE applications*, EEA Technical Report No 7/2012, European Environment Agency (http://www.eea.europa.eu/publications/eea-catchments-and-rivers-network) accessed 6 March 2013.

UNSD, 2013, 'Energy Statistics Database', United Nations Statistics Division (http://unstats.un.org/unsd/energy/edbase.htm) accessed 6 March 2013.

Cement Australia, 2013, 'Cement Australia — Environmental performance' (http://www.cementaustralia.com.au/wps/wcm/connect/website/cement/home/sustainable-development/environmental-performance/) accessed 6 March 2013.

Wikipedia, 2013, 'List of countries by cement production' (http://en.wikipedia.org/wiki/List_of_countries_by_cement_production) accessed 6 March 2013.

FOOTPRINT, 2013, 'FOOTPRINT :: creating tools for pesticide risk assessment and management in Europe -- The FOOTPRINT SUGAR index' (http://www.eu-footprint.org/sugar.html) accessed 7 March 2013.

() CEPI, 2013, 'CEPI — Confederation of European Paper Industries' (http://www.cepi.org/topics/statistics) accessed 6 March 2013.

de Lamberterie, B., 2011, *EIP Water Efficiency. ESTEP Contribution*, DG Env-EIP Water efficiency (http://ec.europa.eu/environment/water/innovationpartnership/pdf/15_12_2011/ESTEP%20for%20EIP%20water.pdf) accessed 6 March 2013.

Annex 1 Methodology and supporting application

1 *Scope of the methodology as supported by the application*

1.1 Targets

The objective of the methodology is to produce calibrated and tuned hydrological balances, analysed at the 'statistical unit level' under the SEEAW matrixes, presented at the monthly time and 'territory of reference (=sub-basin)' resolution, and aggregated annually for the full period.

It should be emphasised that the objective is not to apply and calibrate certain hydrological models at EU level, because this method uses rainfall as a data basis for flow instead of measurements; this could introduce a circular bias in the results.

The methodology was developed to produce reasonable water balances under the SEEAW, that is to say, assimilating and blending very different data sources possibly exhibiting large discrepancies. There are three expected outcomes of the methodology; these are as follows (in no particular order):

- produce SEEAW table and derived indicators;
- provide all relevant information related to data consistency (data are not used only for accounting);
- Provide ancillary information, that is, significant information for other processes (e.g. the WFD).

The methodology is hence addressed in this annex from the point of view of the way the supporting application applies this methodology in the production process.

1.2 The Nopolu application

Nopolu is the trademark of the application developed by Pöyry consulting firm that has been developed to carry out water accounting (under the SEEAW) as an outcome in the three main outcomes presented in the section above. This application is modular and is built as 'application modules' sharing a common structure and coded, so that the kernel module can exchange information and use same data sets across modules.

The main module for the water accounting is the Nopolu/WRA (water resource accounting) module, that allows the user to calculate water resources accounts at different restitution aggregates: administrative (NUTS1 to NUTS3), or hydrological (sub-basins or other aggregates that are set in the hydrographical reference data set), and for different years.

The specificity of the WRA module is its 'thematic data collector structure' design. Indeed, the water resources accounts are basically a huge and highly structured aggregation of data, generally not needing a sophisticated calculation process. This module, intended to be used with Nopolu, acts like a cataloguer of data sources which are too numerous to be described in this report. The pathways to reaching them through the multiple and various thematic tables of Nopolu databases are already described in Nopolu handbook (produced by the consultant and unpublished). The present description aims at highlighting the concepts and limitations of the module, in close interaction with the SEEAW and data availability.

Since several sources of information are rather uncertain or poorly populated, the making of the accounts table reflects some degree of 'scenario definition' that expresses deliberate choices. The term 'scenario' will be used in this sense hereafter.

The general understanding of the SEEAW tables is that each and every cell is populated from a certain source that exploits very different databases: industrial abstractions and snow melting are radically different objects. Populating a cell is hence the result of a specific process that is generically called a 'query'. In reality, populating a cell can range from simple data extraction (a simple SQL query) to surrogating this data under a certain scenario choice (a specific procedure).

All 'queries' are then stored in a table, and are retrieved thanks to their structured query name that explicitly refers to the cell of the matrix to be filled. Hence, the complete process can be automated (with the noticeable exception of defining previously what the query should do to populate this cell).

2 Organisation of data sets and databases

2.1 Application constraints

For both historical and practical reasons of flexibility, the Nopolu System2 application is an MS Access® database (not a project), that links to either other MS Access® databases (the application databases) or to SQL server tables in SQL server databases. In this case, it has been developed so that the heaviest computer work is carried out at server level.

2.2 Common nomenclature of tables

All data sets are structured in the same way. This systematic structuring follows Nopolu System2 logics: an entity has characteristics and time-related values. This section is hence common to almost all subsystems.

- The constant characteristics of the entity are stored in table C_xxx, where xxx is a mnemonic of the entity category (C_quan describes the gauging station, C_agri the agricultural entities, etc. entities characteristics may have connections with other entities of course, depending on the subsystem.
- The variables (time-dependent) are stored in tables V_(prefix_)xxx. The prefix is optional and depends on the complexity of the information attached to the entity.

2.3 Natural assets data sets nomenclature of tables

2.4 Water use–related nomenclatures of tables

Nopolu WA IEI water use application distinguishes three types of activity entities (xxx being the abbreviation Agri, Indus or Domes):

- **Agri** for agriculture;
- **Indus** for industries;
- **Domes** for households and connected services;

Each of this type of activity uses three structural linked tables:

1. **C_xxx** representing the entity (being a characteristic).
2. **V_act_xxx** for characterisation of activity which is time-dependent following the scenario considered (the scenario index is the DATE field, expressing activity at date DATE) and

deals with driving force of the activity. V_Act_xxx is linked to C_xxx via the field CODE that identifies the entity.

3. **V_pol_xxx** for characterisation of the resulting of the activity as water flow, pollution load, etc., organised by parameter (volume or pollutant type) which is linked to V_act_xxx via the Num field, and deals with pressure and impact.

The structure of each these tables is nearly identical for each xxx, being adjusted to the topic category.

The other tables are those related to the sewage system, identified by the field Res for the transport network system, and STEP for wastewater treatment plants (WWTP):

- **C_yyy** for network or WWTP characteristics;
- **V_yyy** for network or WWTP time-dependent emissions of flow and pollutant (scenario indicator is still the DATE field).

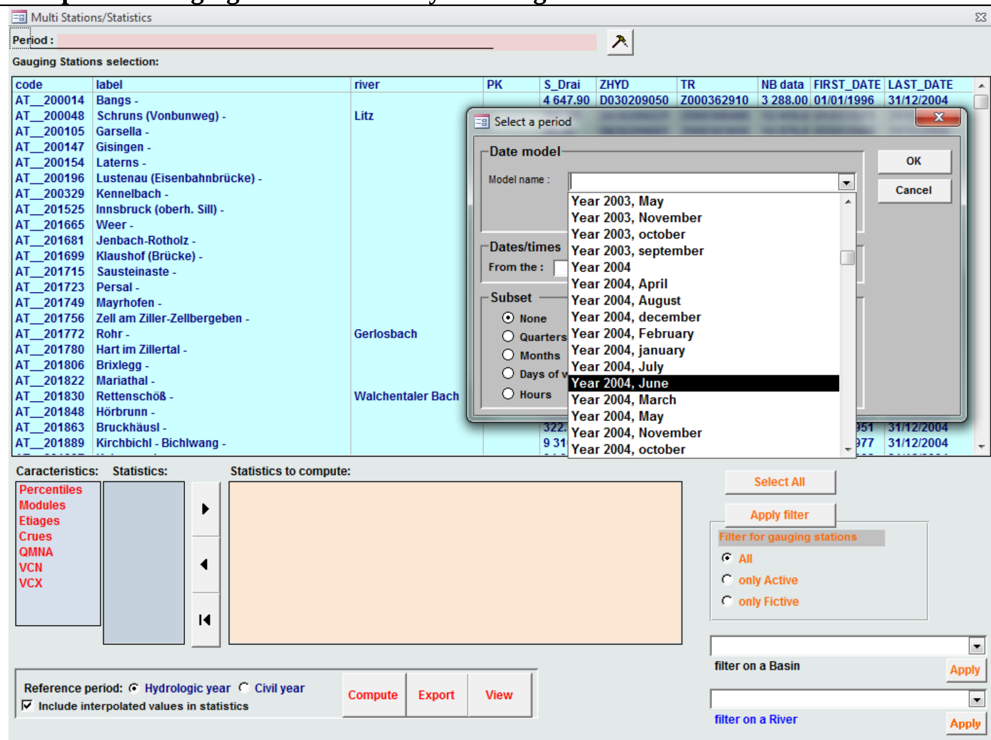
3 Linearised flow computation

3.1 Discharges and statistical discharges management at station level

The way the flows are computed is discussed in Annex 7 , section 4 in detail (algorithms). The procedure for the water accounts only is addressed here.

The handling of daily discharges is not specific to water accounting; daily discharge data are stored in SQL Waterbase (V_Quan). Daily data are aggregated per month (bridging the gaps where required, or disregarding them) and per station, thanks to generic Nopolu System2 procedures, and stored in Eu_Q.mdb (table VI_CSQ_Q_Stat, controlled by CI_Hyd_Calc that stores all the references of calculations. This organisation helps minimise the information attached to a calculation, especially if the statistics cover several years. The structure may handle several types of statistics).

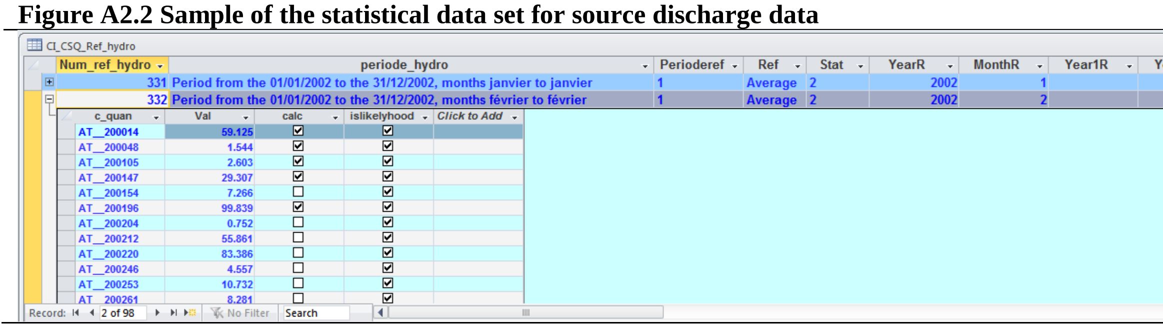
Figure A2.1 Nopolu: managing statistics on daily discharges



Source: Pöyry report, June 2012

Data storing in the target database is organised by station–year–month identification, for general purposes of flexibility. All stations and many dates can be processed in a single session. The processing module also allows for depletion curves assessment, and is hence ready for addressing GW to river inputs.

Figure A2.2 Sample of the statistical data set for source discharge data



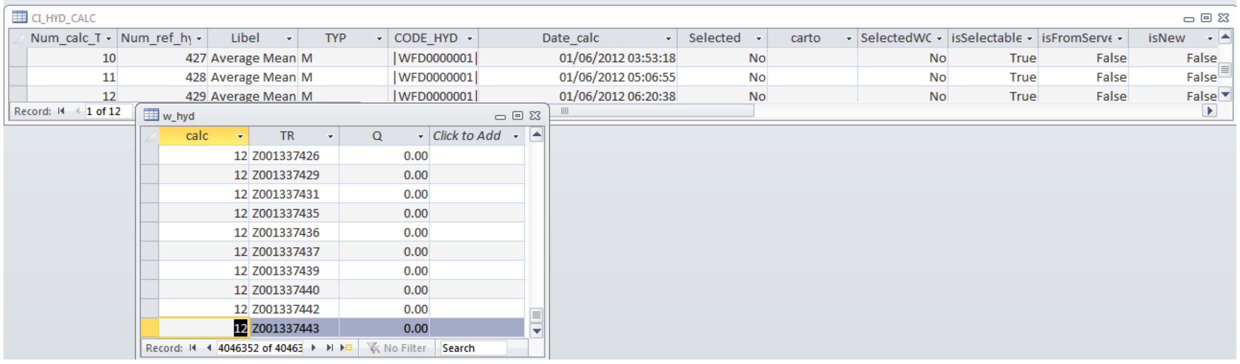
Main table is the reference of calculations and the embedded table (displayed extract) contains the values of data, that tells if (Booleans calc) data that comes from calculation of daily flows or if data are taken as such from the source data set, and the second likelihood indicates whether the data have passed the qualification tests implemented in Nopolu System2 .

This module also allows one to enter data collected as monthly averages instead of daily values, and data reconstructed by the EPH time reconstruction method (see Annex 8 for more on the reconstruction methods).

3.2 Computation of linear flows proper

Discharge is extrapolated between stations along the main drains of ECRINS to populate each and every ECRINS segment with a monthly value. The ECRINS segment represents the 'statistical unit' of the accounts for rivers. This is carried out with the Nopolu System2 hydrological algorithm: discharge is extended up to the source of river, and considers the affluent as well as reservoirs. This is why, in some sub-basins, an estimation of resources is carried out despite absence of stations, while the downstream sub-basin is populated.

Figure A2.3 Sample of discharge per river segment



CI_HYD_CALC

Num_calc_T	Num_ref_hy	Libel	TYP	CODE_HYD	Date_calc	Selected	carto	SelectedWC	isSelectable	isFromServe	isNew
10	427	Average Mean	M	|WFD0000001|	01/06/2012 03:53:18	No		No	True	False	False
11	428	Average Mean	M	|WFD0000001|	01/06/2012 05:06:55	No		No	True	False	False
12	429	Average Mean	M	|WFD0000001|	01/06/2012 06:20:38	No		No	True	False	False

Record: 1 of 12

w_hyd

calc	TR	Q	Click to Add
12	Z001337426	0.00	
12	Z001337429	0.00	
12	Z001337431	0.00	
12	Z001337435	0.00	
12	Z001337436	0.00	
12	Z001337437	0.00	
12	Z001337439	0.00	
12	Z001337440	0.00	
12	Z001337442	0.00	
12	Z001337443	0.00	

Record: 4046352 of 40463 No Filter Search

The data organisation has been set up to mitigate adverse constraints: processes are very long and should be carried out by basin (there are no interactions across basins), be easily updated (one basin is updated and not the next one), and accept some scenarios (some basins can be computed in parallel with different options). Hence data sorting and management is based on scenario results selection and data handling.

This management forms part of the global water quality accounting, since discharges per stretch are the source data for the standard river kilometre (kmcn) used for water quality accounting (UNSD, 2007). Data are handled in tables W_hyd (this is transient information) and controlled by Ci_Hyd_calc. The tables may be enormous in size and are expected to be handled at server databases for calculations (field IsFromServer). Filed CodeHyd is a memo filed (of practically unlimited length) that stores the encapsulating levels of the FECs (in this case they are districts). Referential integrity has not been implemented for reasons of performance, and is handled by the application. Since the reference systems are topologically consistent, it is feasible to map results per FEC or create videos from these data sets (as exemplified in many presentations).

3.3 Computing resource per FEC

The discharge data at stretch resolution can then be then aggregated at the sub-basin level that makes the 'domain of reference' of the accounts.

A specific database for hydraulic productivities per entities (FECs, sub-basins, basins) has been developed (EU_prodhyd.mdb). Tables in this database are named X_hyd, where X is the basin aggregation level (v_hyd: FEC level, S_Hyd sub-basins, G_hyd basin level). Their structure is common to all levels and is reported in the sample design. The structure is common to all tables as is shown in the database extract of Figure A1.4.

Figure A2.4 List of fields for hydraulic productivity per sub-basin

X_hyd

Field Name	Data Type	Description
calc	Number	Calculation number (refers to source data)
SB	Text	de entity of aggregation (e.g. sub-basin)
ProdHy	Number	Estimation of hydraulic productivity litres/second/km2 (l/s/km2)
Lame	Number	Estimation oth the water balde as mm
nbj	Number	nomber of days of the reference period
YYYY	Number	year of results
MM	Number	Month of results
isExtrapolated	Yes/No	iTRUE if extrapolated
percSurfCalc	Number	Estimation of the area covered by calculatd FECs
nbFEC	Number	nomber of FECs in teh aggregate (min is 1)
ProdHyUpst1	Number	Estimation of the upstream productivity (1st upstream) l/s/km2
ProdHyUpst2	Number	Estimation of the upstream productivity (2nd upstream) l/s/km2
ProdHyDo	Number	Estimation of the downstream productivity (1st upstream) l/s/km2

An automatic map can be produced from the Nopolu WA module, showing a synoptic view of monthly hydraulic productivities per FEC (or other entity).

4 Common processes to build the water accounts tables

4.1 Generic procedure

The main form (Figure A2.6) is built with an information tab that recalls the categories, and as many tables as are required to populate the categories stepwise. In the first tab (information), a button allows updates of the table of reference of assets (CR_assets).

Figure A2.5 sample extract CR_assets table of assets references

CR_asset

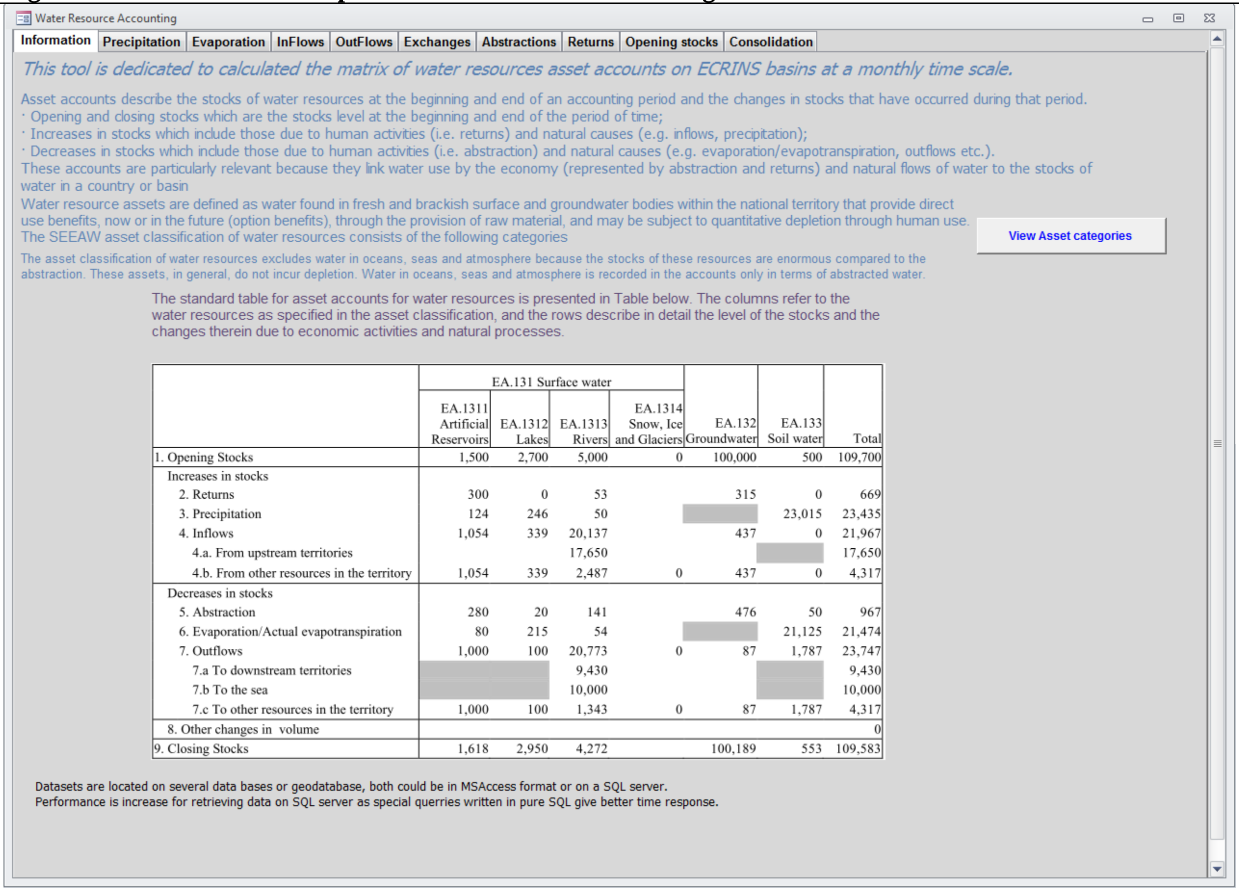
Level	Code_EA	Libel		Description	rep
1	13	Water Resources	1 3 :	Water resource assets are defined as water found in fresh and brackish surface and groundwater bodies within the national territory that provide direct use benefits, now or in the future (option benefits), through the provision of raw material, and may be subject to quantitative depletion through human use. Water resources include all inland water bodies regardless of their salinity level. Freshwater is naturally occurring water having a low concentration of salt. Brackish water has salt concentration between that of fresh and marine water. The definition of brackish and freshwater is not clear cut: the salinity levels used in the definition vary between countries. Brackish water is included in the asset boundary on the ground that this water can be (and often is) used, with or without treatment, for some industrial purposes, for example, cooling water or even for irrigation of some crops.	
2	131	Surface water	1 3 1 :	Surface water comprises all water which flows over or is stored on the ground surface (International Glossary of Hydrology, 1992). Depending on data availability and country priorities, the classification could be further disaggregated. Reservoirs can be classified according to the type of use, for example for human, agricultural, electric power generation and mixed use. Rivers can be classified on the basis of the regularity of the runoff as perennial, where water flows continuously throughout the river, or ephemeral, when water flows only in direct response to precipitation or to the flow of an intermittent spring. Namibia (Lange 1997), Moldova (Tafi and Weber 2000) and France (Margat 1986) have used such a breakdown.	
3	1311	Reservoirs	1 3 1 1	the classification could be further disaggregated to classify artificial reservoirs according to the type of use, e.g. for human, agricultural, hydroelectric power generation and mixed use	
3	1312	Lakes	1 3 1 2	boundaries between the different categories in the asset classification, such as between lakes and artificial reservoirs and rivers and lakes/reservoirs, may not always be precise.	
3	1313	Rivers	1 3 1 3	Rivers could be further classified on the basis of the regularity of the runoff as perennial rivers, where water flows continuously throughout the year, or ephemeral rivers, when water flows only as a result of precipitation or to the flow of an intermittent spring.	
3	1314	Glaciers, snow and ice	1 3 1 4	Glaciers, permanent snow fields, ice, and marine water are now part of the classification (The SEEAW asset classification expands the SEEA-2003 classification by including the categories EA.1314 Glaciers, snow and ice and EA.133 Soil water) even if water abstraction does not have an effect on the size of the stocks. In fact, it is important to understand the role of these components in the hydrological cycle and to account for them when compiling the accounts. This information is particularly relevant in the case of seasonal accounts when water stored in soil and permanent snow fields in one period is an essential resource for the following one.	
2	132	Groundwater	1 3 2 :	Groundwater comprises all water which collects in porous layers of underground formation known as aquifers. Aquifers may be unconfined, that is have a water table and an unsaturated zone or may be confined when they are between two layers of impervious or almost impervious formations. Unconfined aquifers are recharged during the water cycle by the percolation of rain or melted snow and thus hold renewable groundwater. The water in confined aquifers has accumulated over a geological time span and, because of its location, cannot be recharged at all or only over a long time span. Such water resources are non-renewable or fossil water. (Most water in lakes is also non-renewable since the replenishment rate is a small proportion of the total volume of water.)	
2	133	Soil Water	1 3 3 :	Soil water consists of water suspended in the uppermost belt of soil, or in the zone of aeration near the ground surface, that can be discharged in to the atmosphere by evapotranspiration. Soil water cannot be abstracted.	

Enr : 1 sur 8 Aucun filtre Rechercher

Source: Nopolu System2 WA application.

The elements of water accounting (first the positive ones for increasing stock, followed by negative ones) are listed line by line, and are split in water resources categories by column. It is possible to create and calculate a scenario for all of them. The last one is used for consolidation of all elements, i.e. all individual scenarios for each element. How to deal with each of these elements is described in the following sections.

Figure A2.6 Main form of Nopolu Water Resource Accounting



	EA.131 Surface water						
	EA.1311 Artificial Reservoirs	EA.1312 Lakes	EA.1313 Rivers	EA.1314 Snow, Ice and Glaciers	EA.132 Groundwater	EA.133 Soil water	Total
1. Opening Stocks	1,500	2,700	5,000	0	100,000	500	109,700
Increases in stocks							
2. Returns	300	0	53		315	0	669
3. Precipitation	124	246	50			23,015	23,435
4. Inflows	1,054	339	20,137		437	0	21,967
4.a. From upstream territories			17,650				17,650
4.b. From other resources in the territory	1,054	339	2,487	0	437	0	4,317
Decreases in stocks							
5. Abstraction	280	20	141		476	50	967
6. Evaporation/Actual evapotranspiration	80	215	54			21,125	21,474
7. Outflows	1,000	100	20,773	0	87	1,787	23,747
7.a To downstream territories			9,430				9,430
7.b To the sea			10,000				10,000
7.c To other resources in the territory	1,000	100	1,343	0	87	1,787	4,317
8. Other changes in volume							0
9. Closing Stocks	1,618	2,950	4,272		100,189	553	109,583

Source: Nopolu System2 WA application.

4.2 Data storage and preparing consolidation

The database Eu_Compt_Res hosts the pairs of data sets SCEN_AX and VR_AX, where X is a number referring to the line number in the assets matrix (Figure A2.6). For example, SCEN_A3 and VR_A3 refer to precipitations. The target data set is hence a matrix (stored as vector, to allow any number of final columns) scoring for a line of the table, the values related to the columns in the SEEAW table.

The data structure is quite similar.

- The scenario table has the scenario ID, year min., year max. (time range), unit scale (level of the statistical unit), target aggregation scale (territory of reference), a text for description and a Boolean (the selected scenario).
- The corresponding data table comprises the scenario ID (many scenarios can be stored in the same table), year of data, month of data (all accounts are considered to be at monthly resolution), ID of the spatial entity of aggregation, recall of the scale of aggregation, resource considered and volume for the time and entity. Depending on technical issues, the number of columns may vary.

This structure follows the general constraints of making accounts and consolidation that may result from an optimum combination of scenarios.

5 *Processing the natural resources components of the tables*

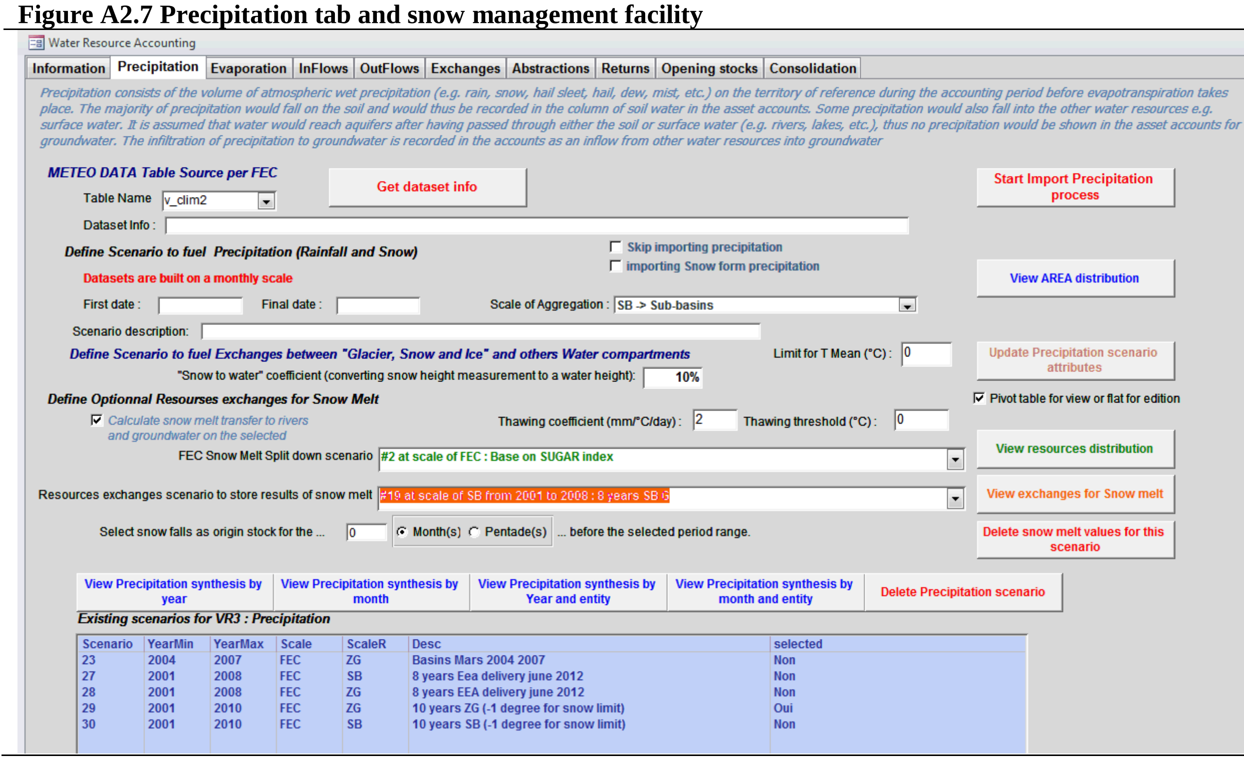
5.1 Management of precipitation data (tab ‘precipitation’)

Precipitation consists of the volume of atmospheric wet precipitation (e.g. rain, snow and hail) on the territory of reference during the accounting period before evapotranspiration takes place. Most precipitation would fall on the soil and would thus be recorded in the column of soil water in the asset accounts. Some precipitation would also fall into the other water resources, e.g. surface water. **It is assumed that water would reach aquifers after having passed through either the soil or surface water (e.g. rivers and lakes), and thus no precipitation would be shown in the asset accounts for groundwater**. The infiltration of precipitation to groundwater is recorded in the accounts as an inflow from other water resources into groundwater, and is computed separately.

The first element tab in the Main form of Nopolu Water Resource Accounting is dedicated to precipitation. The source of data is a database ([64]) containing monthly aggregated calculation of daily meteorological Ensembles data modelled by the EEA and aggregated at FEC level (see Annex 3 for details on modelling with the Ensembles data sets). The source database is EU_VClim(x).mdb, x=2 being the latest delivery.

The ‘precipitation’ tab displays scenarios: for example, one is the temperature threshold when rainfall is actually snow, as displayed in Figure A2.7, in which scenario 29 is active.

Figure A2.7 Precipitation tab and snow management facility

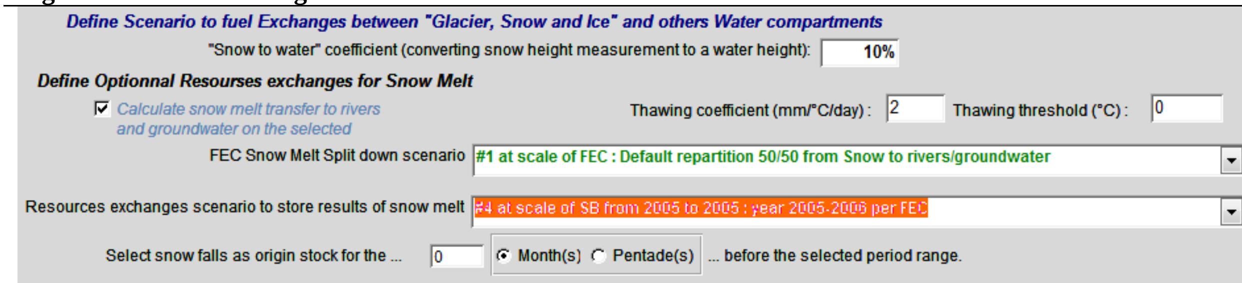


Source: Nopolu System2 WA application.

A significant part of precipitation in certain areas can be in the form of snow, and snowmelt can be calculated for approximation of transfer between the ‘Glacier, Snow and Ice’ (CODE_EA 1314) water resource category to ‘Soil Water’ (CODE_EA 133) and ‘Rivers’ (CODE_EA 1313). Computation is driven by a scenario for snowmelt broken down between those water resources categories (in green), and a resources exchanges scenario for storing the results of snowmelt transfer.

([64]) This may be a SQL server database or a classic MS Access® database. Depending on the size of the aggregated table, both can be processed; Nopolu manages the type of query process accordingly.

Figure A2.8 Snowmelting scenario selection



Source: Nopolu System2 WA application

The snowmelt processing allows snow to be considered as it is provided in the meteorological data sets, and then the snow-to-water coefficient should be provided or, as in the situation reported in this case, snow is computed from the precipitation using temperature. In this latter case, the ratio of snow (in cm) to water is disregarded. Snow melt is calculated using a Swiss-based formula that has been implemented, and uses a thawing coefficient (mm/°C/day) and a thawing threshold (°C).

The full process is carried out by the programme launched by the button 'Start import…' and creates the data sets needed for accounting (65). It populates the Scen_A3 and VR_A3 data sets. Values can be seen in a synthetic display (flat table or cross-tabulated) by clicking the appropriate buttons.

5.2 Management of evaporation data (tab 'evaporation')

Evaporation/actual evapotranspiration (66) is the amount of evaporation and actual evapotranspiration that occurs in the territory of reference during the accounting period.

Both evaporation and evapotranspiration values are computed by modelling, since they are not monitored and provided in the source Ensembles data sets; they are part of the data delivered in the V_Clim data sets along with the precipitation data. Such modelling requires information on soil capacities that has been collected separately and processed.

In this case, the data sets are SCEN_A6, completed by VR_A6 and VR_A6x. The surface run-off is now part of climatic data sets and is no longer modelled inside Nopolu System2, better represented in the relative independence of the data categories.

In source data set, evaporation is separated from actual evapotranspiration, that takes into account both the soil coverage and soil characteristics. There is a source of uncertainty in the process, since the model computes, at grid scale, the actual evapotranspiration using distinct coefficients for each land cover type and the WA procedure that secondarily breaks down the values at different scales. This risk is addressed in more detail in Annex 3 Section 5.2

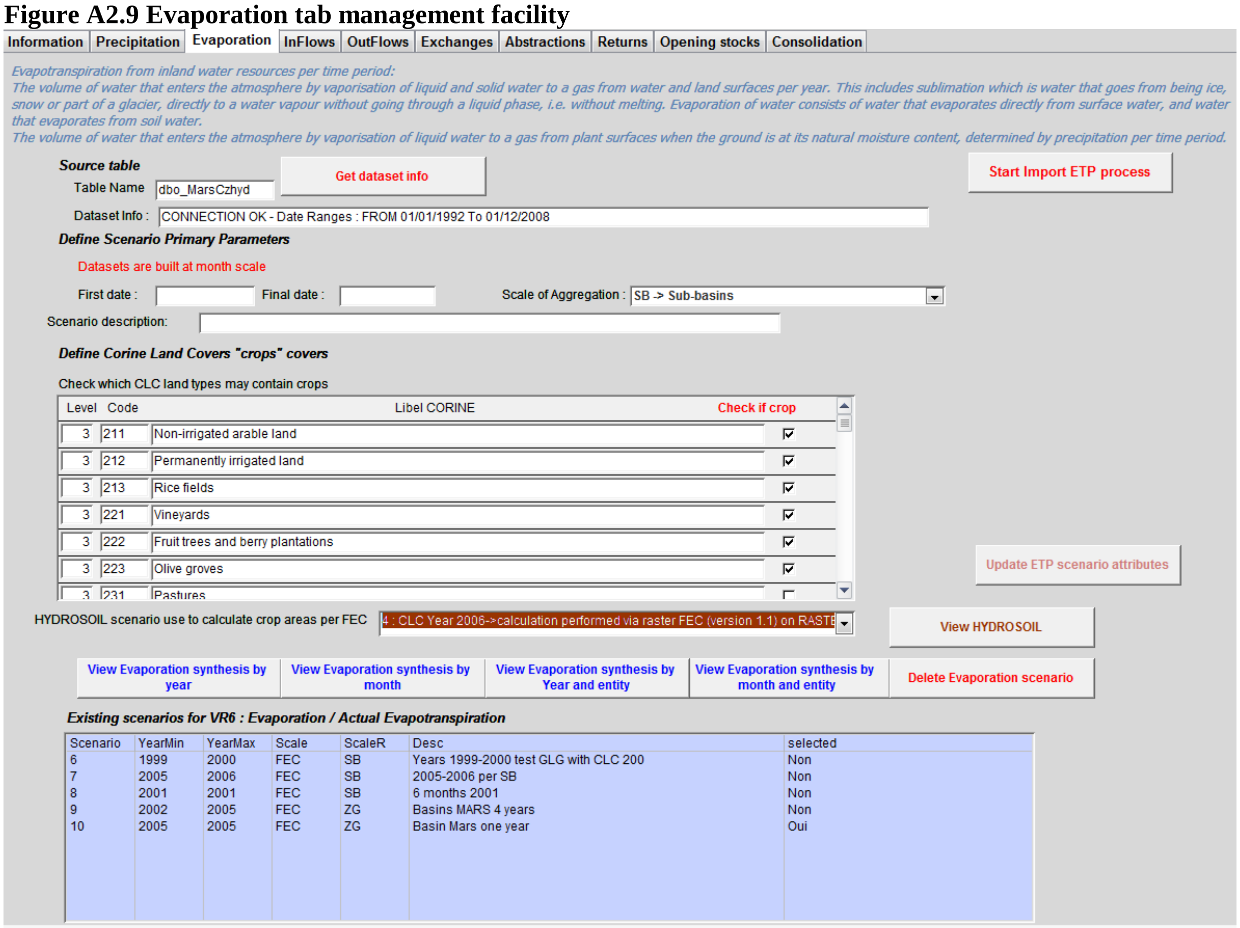
The tab allows for tuning the land coverage as it is represented in the Hydrosol table, and allocates crops accordingly.

---

(65) Despite a great deal of work on optimising the process, computations can long and hence, scenarios well balanced before testing a new one.

(66) **Evaporation** refers to the amount of water evaporated from waterbodies such as rivers, lakes, artificial reservoirs, etc. and **evapotranspiration** refers to the amount of water that is transferred from the soil to the atmosphere by evaporation and plant transpiration.

Figure A2.9 Evaporation tab management facility



Source: Nopolu System2 WA application.

5.3 Inflows management tab

Inflows represent the amount of water that flows into water resources during the accounting period. The inflows are disaggregated according to their origin: (a) inflows from other territories/countries; and (b) from other water resources within the territory.

Inflows from other territories occur with shared water resources. For example, in the case of a river that enters the territory of reference, the inflow is the total volume of water that flows into the territory at its entry point during the accounting period. If a river borders two countries without eventually entering totally into either of them, each country could claim a percentage of the flow to be attributed to their territory. If no formal convention exists, a practical solution is to attribute 50 % of the flow to each country. Inflows from other resources include transfers, both natural and man-made, between the resources within the territory.

They include, for example, flows of infiltration and seepage, as well as channels built for water diversion.

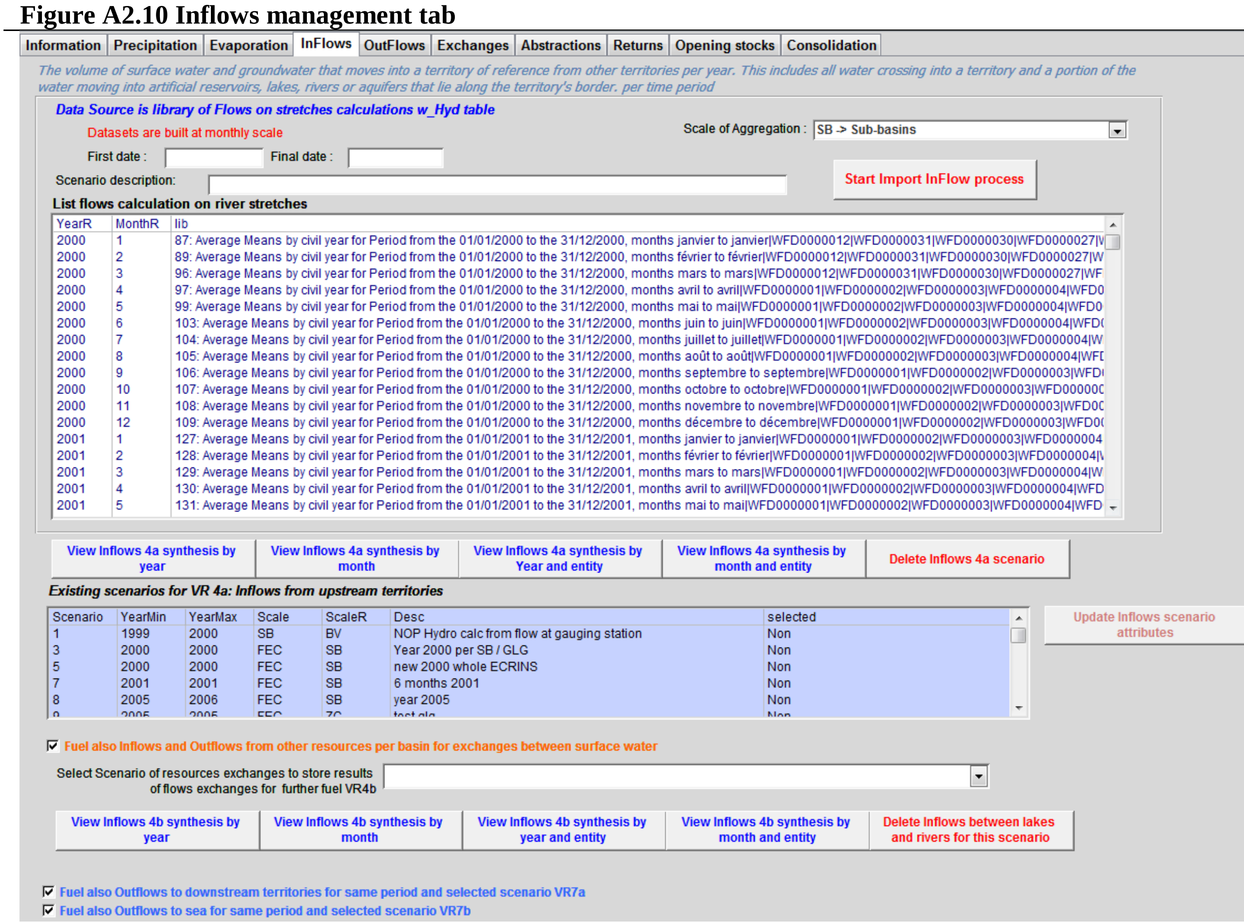
The SEEAW table is split into categories 4a (inflows from upstream) and 4b (other water resources), two distinct tables with basic structures of scenario, and data are populated as follows:

- SCEN_A4a and VR_A4a respectively for inflows from upstream territories scenario library, and values for inflows from upstream territories;
- SCEN_A4b and VR_A4b respectively for inflows from other water resources categories scenario library, and values for inflows from other water resource categories.

Inflows from upstream territories as well as inflows from other water resources **in surface water categories** are calculated from the linearised flow values per stretch obtained with the Nopolu resources module. The same tab processes both the inflows existing by definition from upstream territories, and inflows from other resources.

The results of flows per stretch are stored in table w_HYD, whose values are related to scenarios in table CI_HYD_CALC via the calc field, representing the calculation number. They are shown in the box 'List of flow calculation on river stretches' (Figure A2.10).

Figure A2.10 Inflows management tab



Source: Nopolu System2 WA application

Of course, selecting all linear discharge scenarios is achieved by choosing the starting and ending years of the time range. The selection can be itself saved as a scenario to combine with other data categories' scenarios for the final consolidation. It may seem that processing by 'scenarios' is a useless complication; in fact this reflects possible adjustments related to local issues (e.g. diversions) and this type of management is eventually more flexible for fuelling other purposes. Since computations are long, time can be saved by this process.

Inflows and outflows are better processed together (for consistency reasons, since they come from the same data source: outflow from upstream is inflow for downstream, and this is why it is so important to process within a fully topological hydrological reference system). For this reason, it is possible (and recommended) to fill in both tabs and the request for common processing by ticking the appropriate checkboxes. If this is applied, then the computation is carried out in parallel.

Inflows and outflows from other resources' (line 4b in the SEEAW tables) territories had to be prepared separately, as they represent volumes transferred between basins.

Inflows to surface water from other resources are computed with the linearised flow values. To a large extent, they are the inputs from lakes and reservoirs, a calculation method addressed in Annex 7 4.5 . In all cases, data are stored in table w_HYD, referring to a linear discharge calculation scenario stored in CI_HYD_Calc.

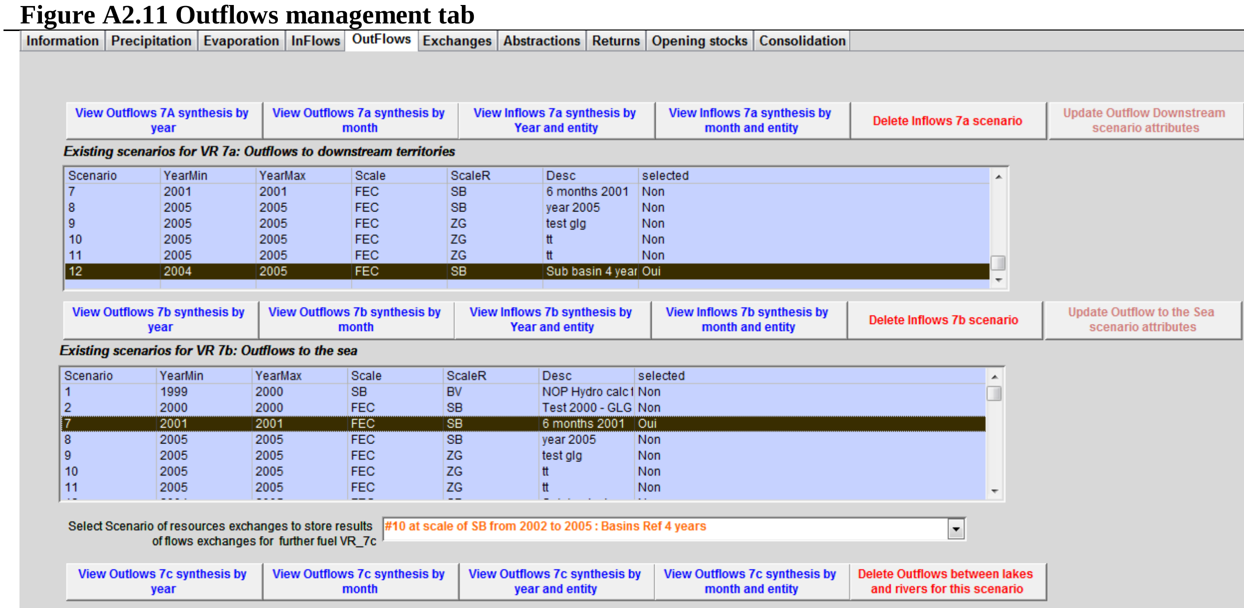
5.4 Outflows management tab

Outflows represent the amount of water that flows out of a water resources system during the accounting period. Outflows are disaggregated according to the destination of the flow: (a) to other water resources within the territory; (b) to other territories/countries; and (c) to the sea/ocean. Outflows to other water resources within the territory represent water exchanges between water resources within the territory. In particular, they include the flows of water going out of a waterbody and reaching other water resources within the territory. Outflows to other territories represent the total volume of water that flows out of the territory of reference during the accounting period. Shared rivers are a typical example of water flowing from one upstream country to a downstream country. Outflows to the sea/oceans represent the volume of water that flows into the sea/oceans.

The SEEAW table is split into categories 7a (outflows to downstream), 7b (outflows to the sea) and 7c (outflows to other water resources). These three distinct tables with basic structures of scenario and data are populated as follows:

- SCEN_A7a and VR_A7a for outflow to downstream territories scenario library and values for outflows to downstream territories, respectively;
- SCEN_A7b and VR_A7b for outflows to the sea categories scenario library and values output to the sea, respectively;
- SCEN_A7c and VR_A7c for outflows to other water resources in the territory scenario library and values for such outflows, respectively.

Figure A2.11 Outflows management tab



Source: Nopolu System2 WA application

The scenario and values for outflows to other resources are symmetrical to inflows to other resources since most of these resources address storage to reservoirs and lakes: both inflow and outflow from surface water are by default calculated simultaneously; hence there is no calculation button on this tab.

5.5 Exchanges management tab

Exchanges of water between water resources are a key component of the accounts: the accounts procedure aims at tracing the pathways of change, not only reporting the differences in stock.

They are hence stored in a separate table, subject to some scenario definition. This table, which breaks down the information in rows 4.b and 7.c of the standard table for asset accounts, provides information on the origin and destination of flows between the water resources of a territory of reference, allowing for a better understanding of the exchanges of water between resources. This table is also useful for the calculation of internal renewable water resources and for reducing the risk of double counting when assessing this indicator separately for surface water and groundwater due to the water exchanges between these resources (FAO/AQUASTAT, 2001 ([67])). The table of exchanges between water resource categories assists in identifying the contribution of groundwater to the surface flow as well as the recharge of aquifers by surface run-off as envisaged in the standard Table 6.2 (Figure 2.6, page 40 of main report).

This part of the accounts may have to incorporate some modelling, despite this contradicting the general principles, because the exchanges between resources lie beyond monitoring, on the whole,.

As the data structure is systematic, the exchange of resources is stored in a couple of tables, not numbered (not a line in the SEEAW):

- SCEN_RE and VR_resources_exchange respectively for exchanges scenario library and values for exchanges between categories.

Figure A2.12 Exchanges between resources management tab

VR_resources_exchanges

Field Name	Data Type	Description
Scenario	Number	number referenced scenario
Year	Number	Year of reference, 0 for default year
Month	Number	0 for whole year or indicate month
entity	Text	Code of the hydrologic entity (could be ZHYD or wider)
Scale	Text	Code of the entities scale
Origin	Text	Type of initial source
Destination	Text	Type of destination
rate	Number	percentage of outflow coming from origin and going to destination
volref	Number	volume of reference on which apply rate
Calcrate	Yes/No	indicate if calculation is done via rate
vol	Number	volume of outflow coming from origin and going to destination
volcalc	Number	calculated volume of outflow coming from origin and going to destination
Description	Text	Description of calculation or source of information

Source: Nopolu System2 WA application.

The scenario of exchanges marks the dates of relevance and the level of aggregation, to allow appropriate aggregation.

As the tab and the SEEAW table show, exchanges occur between source and target, making the structure of the table adapted and vector designed to allow further categories to be considered.

Figure A2.13 Structure of the table of exchanges

([67]) The FAO/AQUASTAT indicator is Internal Renewable Water Resources: 'Average annual flow of rivers and recharge of groundwater generated from endogenous precipitation'.

VR_resources_exchanges

Field Name	Data Type	Description
Scenario	Number	number referenced scenario
Year	Number	Year of reference, 0 for default year
Month	Number	0 for whole year or indicate month
entity	Text	Code of the hydrologic entity (could be ZHYD or wider)
Scale	Text	Code of the entities scale
Origin	Text	Type of initial source
Destination	Text	Type of destination
rate	Number	percentage of outflow coming from origin and going to destination
volref	Number	volume of reference on which apply rate
Calcrate	Yes/No	indicate if calculation is done via rate
vol	Number	volume of outflow coming from origin and going to destination
volcalc	Number	calculated volume of outflow coming from origin and going to destination
Description	Text	Description of calculation or source of information

Source: Nopolu System2 WA application.

The tab in Figure A2.12 proposes a set of options for fuelling and managing some of those types of exchanges, but will need to be extended (noticeably for the following types of exchanges: D1 (from surface to groundwater), D2 (from groundwater to surface), and D4 (between groundwater resources). The latter is for the time being not feasible.

The different categories of exchanges are handled separately, and the principles of calculation are presented below.

Exchanges from surface water to groundwater

It is assumed that water would reach aquifers after having passed through either the soil or surface water, and thus no precipitation to groundwater can be set in the tables. Infiltration of precipitation to groundwater is recorded in the accounts as an inflow from other water resources (soil, for example) into groundwater. Exchange from soil water to groundwater is assessed via a simple partitioning rule using two tables:

- one table (VR_FEC_Distribution, (based on SUGAR data sets)) for splitting the water flow from water soil between rivers and groundwater, and another for time-tuning monthly exchanges (table MR_Resources_exchanges).

Moreover, for each month, the surface run-off (computed in the current version at the meteorological model level on a daily basis, and approached as a volume of precipitation minus evaporation if no such data are available) for the selected scenario (precipitation & evaporation) in soil water is broken down between rivers and groundwater, based on the water resource distribution scenario coefficients, and is modulated per month following the modulation rate.

The manual contains further details; the scope of this report is only to show the main steps to be taken.

The button launches the calculation and populates the target data table accordingly. If values for exchange from soil already exist for the exchanges scenario selected, the user is asked if they want them deleted and to continue calculations.

Exchanges between surface water resources

This part deals with rivers to lakes and lakes to rivers, a subject already discussed in the inflows and outflows chapters. On the exchange tab, results obtained can be viewed by selecting D3 (Figure A1.12).

6 Managing abstractions and returns

Assessment of abstractions deals with the flows from the environment to the economy. Albeit not part of the water assets computation, there are some important points concerning the abstractions flows and the exchange of flows between the economy that can help in understanding the structure of the application. Returns flows representing flows from the economy back into the environment is the dual component of abstraction, and in fact, the Nopolu WA module has been developed so that all of these items could be assessed in a same process. The abstraction tab of the application pilots most of menu options.

For assessing these flows related to the economy, the adjusted Nopolu IEI (integrated emissions inventory, previously developed for assessing the emissions under the WFD requirements) proved suited to deal with complex relationships between the sites of water abstraction, its users, and where water is returned — which is done partly directly to the environment and partly to sewage systems. The tremendous amount of conceptual work for implementing Nopolu IEI made it possible to use this for the water accounts, since this application is built around water flows as the vector of pollution for assessment of the loads to the environment. It was then a natural idea to adjust this development to address water uses in the economy on the basis of IEI application structure.

This annex focuses only on the characteristics relating to the water uses, for the sake of understanding the mechanics behind the accounts, and does not consider elements relevant for pollution load accounting or assessment.

6.1 Economic activities classification

The SEEAW breakdown of the economic activities, classified according to ISIC Rev.4, distinguishes the following groups:

- ISIC 1-3, which includes **agriculture, forestry** and **fishing;**
- ISIC 5-33, 41-43, which includes **mining and quarrying, manufacturing** and **construction**;
- ISIC 35, which includes **electricity, gas, steam and air conditioning supply;**
- ISIC 36, which includes **water collection, treatment and supply;**
- ISIC 37, which includes **sewage system;**
- ISIC 38, 39, 45-99, which corresponds to the **service industries**.

For appropriate management of ISIC nomenclature, another important classification table is the ECO table (rev 4) implemented in table T_ECOISIC that relates to T_ECO with the fields Code - WRAcode. Since there are numerous ISIC (table T_ISIC) codes, only a sample is displayed in **Error! Reference source not found. Error! Reference source not found.**.

Table A2.1 Table T_ISIC of ISIC table (rev 4) classification

T_ISIC

Num	le	Code	Description	WRAcode	selWRA
1	1	A	Agriculture, forestry and fishing	A	☐
2	2	01	Crop and animal production, hunting and related service activities	A	☐
3	3	011	Growing of non-perennial crops	A	☐
4	4	0111	Growing of cereals (except rice), leguminous crops and oil seeds	A	☐
5	4	0112	Growing of rice	A	☐
6	4	0113	Growing of vegetables and melons, roots and tubers	A	☐
7	4	0114	Growing of sugar cane	A	☐
8	4	0115	Growing of tobacco	A	☐
9	4	0116	Growing of fibre crops	A	☐
10	4	0119	Growing of other non-perennial crops	A	☐
11	3	012	Growing of perennial crops	A	☐
12	4	0121	Growing of grapes	A	☐
13	4	0122	Growing of tropical and subtropical fruits	A	☐
14	4	0123	Growing of citrus fruits	A	☐
15	4	0124	Growing of pome fruits and stone fruits	A	☐
16	4	0125	Growing of other tree and bush fruits and nuts	A	☐
17	4	0126	Growing of oleaginous fruits	A	☐
18	4	0127	Growing of beverage crops	A	☐
19	4	0128	Growing of spices, aromatic, drug and pharmaceutical crops	A	☐
20	4	0129	Growing of other perennial crops	A	☐
21	3	013	Plant propagation	A	☐
22	4	0130	Plant propagation	A	☐
23	3	014	Animal production	A	☐
24	4	0141	Raising of cattle and buffaloes	A	☐
25	4	0142	Raising of horses and other equines	A	☐
26	4	0143	Raising of camels and camelids	A	☐

Enr : 1 sur 766 · Aucun filtre · Rechercher

Table A2.2 Table T_ECO representing the SEEAW breakdown of economic activities

T_ECO

CODE	Label	Name	selWRA	selOrigin
A	01-03	Agriculture, Forestry and Fishing	☐	☑
B	05-33/41-43	Mining and quarrying, Manufacturing and Construction	☐	☑
C	35	Electricity, gas, steam and air conditioning supply	☐	☑
D	36	Water collection, treatment and supply	☐	☑
E	37	Sewerage	☐	☑
F	38/39/45-99	Service industries	☐	☑
H	00	Households	☐	☑

Enr : 1 sur 7 · Aucun filtre · Rechercher

All industrial activities all classified upon ISIC so that the Nopolu WA application is capable of aggregation at WA level.

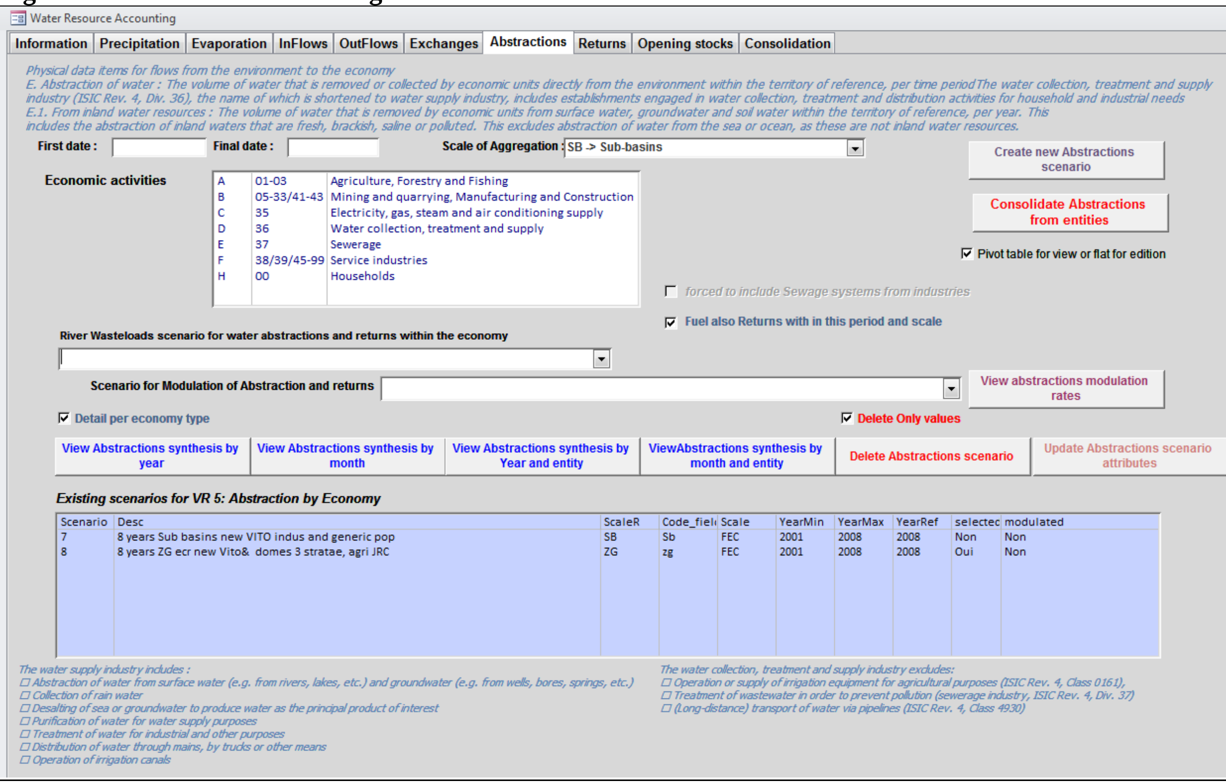
Data collection for these sectors is described in the corresponding annexes; the focus here is how the collected and organised data are inserted into the final tables.

6.2 Flows from the environment to the economy and within the economy

For incorporating abstractions and returns, some adjustments can be considered, that form a scenario which is stored in a set of pairs of tables in relation: Scen_A5/SCEN_Modul_A5 and SCEN_A2/SCEN_Modul_A2, etc.; the numbering reflects the position of the flows in the target SEEAW tables, according to this SEEAW nomenclature. Making scenarios is all the more important since most the abstractions are reconstructed to large extent, and in the best cases are obtained as annual averaged volumes. At this final level, it is possible to insert some time modulating of abstractions over the flat data collected.

The handling of information is controlled by one tab on the flow to the economy and within the economy.

Figure A2.14 Abstractions management tab



Source: Nopolu System2 WA application.

Flows from the environment to the economy involve the abstraction of water from the environment by economic units in the territory of reference for production and consumption activities. In particular, water is abstracted from the inland water resource system (which includes surface water, groundwater and soil water as defined in the asset classification), and from other sources.

Abstraction from other resources includes abstraction from the sea (for example, for direct use for cooling, or for desalination purposes) and collection of precipitation which occurs (for example, in the case of water roof harvesting or hillside dams). The supplier of these flows is the environment and the user is the economy, and more specifically, the economic agents responsible for the abstraction. It is assumed that the environment supplies all the water that is abstracted; hence there is equality between supply and use. If the same economic unit which abstracts may use it (in which case, we refer to it as **abstraction for own use**) or supplies to another agent, possibly after some treatment.

Water is abstracted either to be used by the same agent, or is supplied to other economic units (**abstraction for distribution**). The industry which abstracts, treats and supplies water as a principal activity is classified under class 36 of ISIC Rev. 4, **water collection, treatment and supply**. There may be, however, other industries which abstract and supply water as a secondary activity.

Flows within the economy are not handled as such in the water assets accounts. However, some activities, essential for the assessments, are part of this category and are therefore also processed, since flows to the economy at large could not be addressed otherwise.

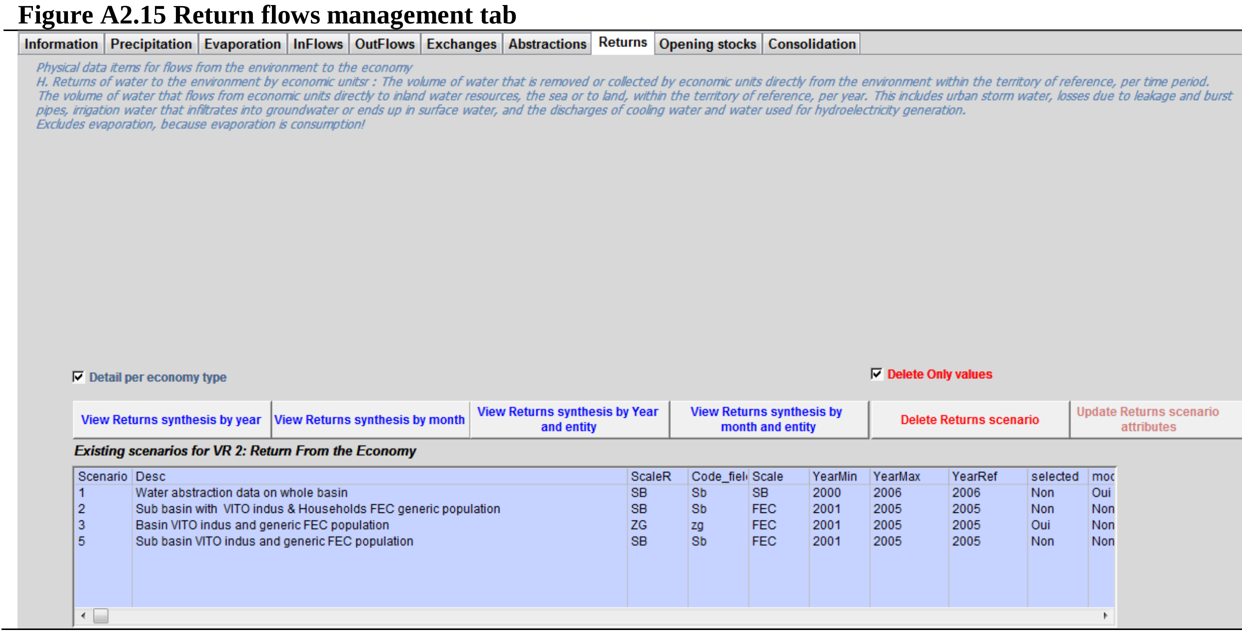
Most of the water is generally supplied by the industry ISIC 36, **water collection, treatment and supply**; however, it can also be supplied by other industries and households. This includes the cases, for example, when water is supplied by industries and households for further use or is supplied to treatment facilities before being discharged into the environment. Note that the physical supply of water by households generally represents a flow of wastewater to ISIC 37, **sewerage**.

The collection of wastewater by ISIC 37, **sewerage**, is recorded as use of wastewater by ISIC 37 and a supply of wastewater by the industry or households generating the wastewater. The corresponding monetary transaction is recorded instead in the opposite way: ISIC 37 supplies the service of wastewater collection and treatment, which is in turn used by the economic units who physically generate wastewater.

6.3 Management of return flows

Assessment of abstractions concerns flows from the economy back into the environment.

Figure A2.15 Return flows management tab



Source: Nopolu System2 WA application.

Flows from the economy back to the environment consist of discharges of water by the economy into the environment (residual flows). Thus the supplier is the economic agent responsible for the discharge (industries, households and rest of the world), and the destination of these flows is the environment. The environment is assumed to use all the water that is returned (supplied) to it. Hence, for these flows, use equals supply.

Flows from the economy to the environment are described in accounting terms in the supply table as a supply of an economic unit to the environment. Each entry represents the amount of water generated by an economic unit and discharged into the environment (in the SEEAW, discharges of water back to the environment are also referred to as **returns** or **return flows**).

Returns are classified according to the receiving media: a distinction is made between 'water resources', which include surface water, groundwater and soil water and 'other sources' such as seas or oceans.

7 Opening and closing stocks

The concept of a stock of water is related to the quantity of surface and groundwater in a territory of reference measured at a specific point in time (beginning and end of the accounting period). While for lakes, reservoirs and groundwater, the concept of a stock of water is straightforward (even though for groundwater it may be difficult to measure the total volume of water), for rivers it is not always easily defined.

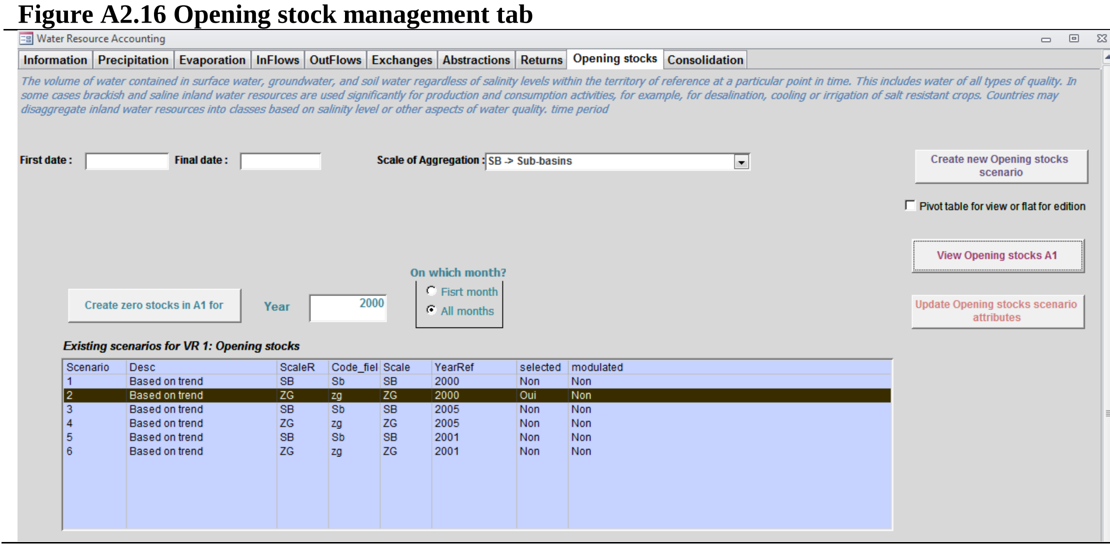
To maintain consistency with the other water resources, the stock level of a river should be computed as the volume of the active riverbed determined on the basis of the geographic profile of the riverbed and the water level. This quantity is usually very small compared to the total stocks of water resources and the annual flows of rivers, and so is provided only for facilitating the calculation of balances. However, the river profile and the water depth are important indicators for environmental and economic considerations, and hence need to be addressed accurately for other reasons. There might be cases, however, in which the stocks of river may not be meaningful for any reason.

The closing stocks are the outcome of calculations and therefore have no management tab.

7.1 Opening stock management tab

By definition, the opening stock of month M is the closing stock of month M-1. Hence, only a small number of functionalities is available on this tab. This is why the system had to offer a very flexible way to calculate opening stocks per each water resource category.

Figure A2.16 Opening stock management tab



Source: Nopolu System2 WA application.

Opening stocks are on line 1, whereas closing stocks are on line 9 of the SEEAW nomenclature. Hence, the scenarios and values for opening and closing stocks are stored in the tables:

- SCEN_A1 and VR_A1 for opening (numbering is the SEEAW nomenclature);
- VR_A9 for closing (since closing is a result, it has no scenario).

7.2 Populating the starting opening stock

The opening stock at month M is closing at month M-1, thereby posing the question 'what is the starting stock of the first of the computed months?' The response cannot be straightforward and depends on data available; it is likely the outcome of a trial and error process.

There are several options:

- populating from existing data;
- populating as a final result of a series of computations;
- setting to all values to zero (this option is supported by a button in the tab).

Another option is to set to standard values in relation with the entities, something which is not yet implemented. All options are evident in the data table VR_A1.

Figure A2.17 Opening stocks storage table VR_A1

VR_A1

Field Name	Data Type	Description
Scenario	Number	number referenced scenario
Year	Number	Year of reference
Month	Number	null for whole year or indicate month
Entity	Text	Code of the hydrologic entity (could be ZHYD or taller)
Scale	Text	Code of the entities scale
Resource	Text	Type of initial source
Vol	Number	Opening Stock (unit: million cubic meters) refered to month period
VolC	Number	Opening Stock Calculated (unit: million cubic meters) refered to month period
Calculated	Yes/No	indicate if calculation is done from Closing Stocks
Desc	Text	Description of calculation

Source: Nopolu System2 WA application.

The Boolean 'calculated' is very important to flag the origin of volumes.

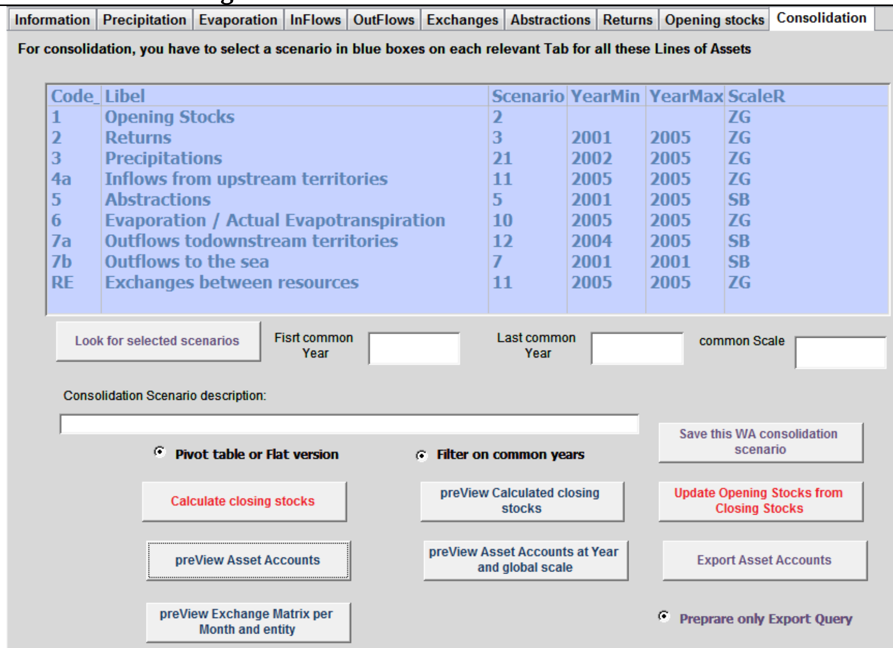
8 *Consolidation management options and tab*

Consolidation is the penultimate step of the accounting procedure, the next being exporting data so that dynamic cross-tabulation can be performed. The consolidation proper consists in putting together the appropriate scenarios and blending them so that the full series of couples opening/closing stocks can be built. By definition, consolidation can be carried out only if all aggregation levels are the same across the different categories of data (classes 1 to 8, and exchange between resources RE included). Otherwise the application warns that consolidating sub-basins with districts, for example, is nonsensical.

8.1 Consolidation management tab

The consolidation management tab is first and foremost a tool for mixing the different scenarios of other tabs, ensuring that the consolidated scenario is composed of a full set of selected scenarios that exist over a common period with the same scale of calculation.

Figure A2.18 Consolidation management tab



Source: Nopolu System2 WA application

If the scenario building has been carried out appropriately, it is possible to create a consolidation scenario with different configurations, and to view detailed results to determine where adjustments are needed, or to try other hypotheses in such a way that step by step, gaps and errors are eliminated.

For example, selecting a common scale for scenarios and common years displays only those that fit the criteria. Once the panel of source scenarios is established, closing stocks can be computed.

In the given display, the calculation will be cancelled, since the currently selected scenarios have different aggregation scales.

8.2 Computation of closing stocks

Closing stocks represent opening stocks for the next period (month). They are stored in the VR_A9 table. Once all items' corresponding scenarios have been selected and given a valid period and common scale, the calculation of closing stocks can be launched. They are calculated in one step with the button 'calculate closing stock' and can be then used to update opening stock:

- calculate the difference in volume for each entity in the period, and store results in the closing stocks table (the calculations are first given '0' as scenario numbering, until they are eventually validated with scenario 0);
- search for the first date of opening stock (in table VR_A1) selected that shares the minimum year and month with those just calculated.

Indeed, the closing stocks depend a great deal on the opening stocks, that are by definition unknown. Hence the closing stocks' computation is the outcome of a procedure of trials and improvements that cannot be fully automated (this would be an unstable procedure in many cases, and would be computing endlessly).

A preview button displays a view of the result table.

Different buttons allow visualisation of provisional results at different scales. This step requires expert judgement to assess if the closing stocks are mature enough to be used as seeds for opening stocks and to launch calculations again.

Figure A2.19 Sample view of results as provisional scenario#0, for seeding opening stocks

VR_A9_view

Scenari	Entity	Year	Month	total	1311:Reserv	1312:Lakes	1313:Rivers	1314:Glacier	132:Ground	133:Soil Water
0	WFD0000003	2005	10	-83 151.70	-2 108	-268	-1 952	0	-2 708	-76 115
0	WFD0000003	2005	11	-81 398.18	-2 244	-282	-1 241	0	-1 834	-75 797
0	WFD0000003	2005	12	-80 080.99	-2 348	-291	-825	0	-1 087	-75 530
0	WFD0000004	2005	1	-2 799.07	-289	-31	-299	23	-1 072	-1 130
0	WFD0000004	2005	2	-6 032.93	-586	-63	-522	69	-2 144	-2 787
0	WFD0000004	2005	3	-11 366.22	-899	-95	-706		-3 216	-6 449
0	WFD0000004	2005	4	-19 752.28	-1 247	-129	-896		-4 288	-13 191
0	WFD0000004	2005	5	-29 451.52	-1 607	-164	-1 088		-5 361	-21 232
0	WFD0000004	2005	6	-43 215.45	-2 009	-200	-1 275		-6 433	-33 299
0	WFD0000004	2005	7	-58 414.78	-2 363	-231	-1 453	0	-7 269	-47 099
0	WFD0000004	2005	8	-70 761.91	-2 544	-243	-1 554		-7 512	-58 909
0	WFD0000004	2005	9	-79 606.67	-2 894	-277	-1 736		-8 585	-66 114

Source: Nopolu System2 WA application.

Once the user is happy with the development of closing stocks (that in principle tend to reach a stable value along time whichever the opening stock primarily produced), the button 'Update Opening stocks from closing stocks' allows updating the opening stocks in table VR_A1. The first opening stock month updated is the second month of the common period. You can adjust this if needed with the tab related to opening stock.

9 Results exporting and managing

Water accounting tables have at least a three-dimensional structure (entity of aggregation, year and month) that is reflected in the final table and can be previewed with the appropriate buttons. These displays are targeted to expert users, and are not very end user–friendly.

Figure A2.20 Sample display of preview of the accounts (aggregation, district), values are non-final

v_Am_ac

Scale	Entity	YY	MM	Code_T	WaterAsset	total	1311 : Reserv	1312 : Lakes	1313 : Rivers	1314 : Glacie	132 : Groundv	133 : Soil Wat
ZG	WFD0000001	2005	1	1	1 : Opening Stocks	0	0.0	0.0	0.0	0.0	0.0	0.0
ZG	WFD0000001	2005	1	2	2 : Returns	92	1.5	0.8	87.9			1.5
ZG	WFD0000001	2005	1	3	3 : Precipitations	7 565	16.3	6.6	10.8	171.1		7 360.0
ZG	WFD0000001	2005	1	4b	4b : Inflows from resources in the territor	4 693			2 929.5		1 763.4	
ZG	WFD0000001	2005	1	5	5 : Abstractions	514	92.8	10.6	69.4		341.3	
ZG	WFD0000001	2005	1	6	6 : Evaporation / Actual Evapotranspiratio	2 158	5.4	2.1	3.4			2 147.3
ZG	WFD0000001	2005	1	7b	7b : Outflows to the sea	3 779			3 779.3			
ZG	WFD0000001	2005	1	7c	7c : Outflows to other resources in the te	4 693						4 692.9
ZG	WFD0000001	2005	2	1	1 : Opening Stocks	0	0.0	0.0	0.0	0.0	0.0	0.0
ZG	WFD0000001	2005	2	2	2 : Returns	87	1.4	0.7	83.1			1.4
ZG	WFD0000001	2005	2	3	3 : Precipitations	4 524	8.0	3.7	6.2	299.4		4 206.4
ZG	WFD0000001	2005	2	4b	4b : Inflows from resources in the territor	682			425.5		256.1	
ZG	WFD0000001	2005	2	5	5 : Abstractions	507	91.1	10.3	69.0		336.5	
ZG	WFD0000001	2005	2	6	6 : Evaporation / Actual Evapotranspiratio	3 467	8.8	3.5	5.7			3 449.3

Source: Nopolu System2 WA application.

The table is displayed by month. and can be summed by year (not displayed).

A more important feature is the exporting as Excel tables. A dialog form allows selection of the name and target of the workbook which is populated in the same format as the access table in Figure A2.20.

Copying or linking to preprogrammed Excel workbooks allows a dynamic cross-tab facility to display or map the results.

The same procedure applies to exchanges between water resources that can be viewed and exported to Excel workbooks.

Annex 2 Rivers and catchment reference system: special adjustments

1 The reference system in the process

Water accounts methodology, as described in the general overview (Figure 2.2, page 28) involves computing water exchanges between the elementary statistical units, within and between the different territories of references.

These computations demand a fully connected reference system delineating the statistical units, the territories of reference and the connection between the 'inland water resource system' and the 'economy'.

2 The reference system and its components

2.1 Short description

ECRINS ([68]), developed by the EEA, is a fully connected system of watersheds called FECs with rivers, stretches, lakes and reservoirs, monitoring stations (quality and quantity) and dams. The system is derived from the JRC CCM2.1. Compared to the CCM, ECRINS offers a smaller number of elementary catchments with 181 071 FECs instead of the more than 2 000 000 elementary catchments within the CCM v1, and a bit less in the second version .

ECRINS v1.0 is the reference system to which all hydrological features are related. Relations are provided by the EEA but significant supplementary work was needed for implementing ECRINS into the Nopolu WA calculation module, and for checking the relationships between the different GIS objects. Many lagging errors that were identified have been corrected.

The latest ECRINS v1.0 allows the aggregation of FECs at several levels, namely sub-basins (Strahler-based), natural sub-basins, basin, river basin districts in line with the WFD and their hydrological surrogates, the 'functional river basin districts', NUTS0 and NUTS2.

In addition, contrasting with the previous version of ECRINS (v0), ECRINS v1 now has stable river stretches identification, and a correspondence with Member States' river bodies for WFD reporting is available.

The map in the main text (Map 3.1, page 45) shows both the sub-basins level and the extent of the 27 countries of EU representing the study area.

2.2 Entities used as statistical units and territories of reference

The ECRINS entities used as statistical units (understood as being populated with a single representative value for a certain time) are as follows.

- The FEC as spatial entity: the FEC can be subdivided if required, depending on its land coverage, to meet SEEW requirements.
- The river segment from main drains as a linear entity; any river segment is fully contained in a FEC. River segments from secondary drains (those drains not connecting FECs together) are not considered in the accounts.

([68]) See the EEA report for a full description of ECRINS http://eea.eionet.europa.eu/Public/irc/eionet-circle/ecrins/home

- The lake as a free waterbody. Any lake is within a certain (or several) FECs. Lakes can be natural or artificial (resulting from damming). Wherever possible, lakes are documented with their volume, because volume is key information in relation to water accounts, since volume allows storing across seasons.

The territories of references are the natural sub-basins (nicknamed SBs). All SBs are made of FECs that must be hydrologically consistent and belong to the same FRBD. FRBDs have been built into ECRINS as the closest set of FECs, mimicking the true RBD and having hydrological consistency.

When making the SBs, the simple rules taken into consideration were:

1. if the SB belongs to continental masses, hence it cannot have several seas as outlet recipients;
2. if the SB is part of an island, this rule can be breached if the island is significantly smaller than the target area for the SB;
3. the SBs should lie in a reasonably narrow range of sizes and be the catchment of parts of river catchment or affluent catchment.

The rules are both constraining and fuzzy enough to provide some flexibility. The making of the SBs is still partly automated, because no algorithm could be developed that is 100 % free of error. The manual construction resulted in some errors which were corrected only at the final delivery stage. In some cases, applying the second rule was not straightforward, since the implementing of the definition of 'island' presents some errors in ECRINS v1.

The version of SBs used in making the accounts comprised 636 SBs, of which 412 are involved in the computation area. Map of the SBs involved is part of exploitation data (Map 3.1) and is not represented here.

2.3 ECRINS v1.0: statistics and key features

The source CCM comprised raw elementary catchments and drainage segments as an outcome of the processing of the ERTS 1989 mission. This data set has been taken on by the EEA for its processing, due to the area covered, the topological relationships and the absence of licensing.

The EEA development consisted in fully reprocessing the source data sets, adding complements, checking and correcting errors, and making it usable for modelling, reporting and hosting purposes. The features of ECRINS are as follows.

- A set of catchments. CCM comprises 1 409 644 CCM elementary catchments apportioned by 'basins', sets of catchments having the same outlet and Strahler hierarchy.

  ECRINS is made of a layer of 181 071 FECs created by clustering the CCM elementary catchments within a narrowed size range.

  The FECs are organised by:

  - marine shore (based on the Marine Strategy); the islands are fully revised and reallocated;
  - clusters of RBDs, functional districts;
  - Strahler sub-basins and hydrological sub-basins;
  - country and region.
- A set of river elements. CCM has a layer of 1 348 163 'river segments' and nodes, connected to elementary catchments.

  During work on ECRINS , ~160 000 spurious segments were removed from the database (segments in the range of 100 m and inside 1 ha of elementary catchments attached to larger rivers; only nodes were kept) and correct several topological errors (not all corrected). The supplementary information inserted is:

  - main drains, that connect FECs together (other are drains inside a FEC), hence allowing for analysis of 'main rivers';
  - routes, that define 'dummy rivers', from spring to outlet or confluence (with distance to sea);
  - rivers, understood as sets of drains with the same name (~22 000 rivers created).

- A newly created set of lakes based on CLC (validated against the ERM and Article 13 deliveries), locally completed by the water layer computed with CCM. This data set counts 70 847 lakes, connected to river segments (inlets and outlets if relevant), connected to dams known by the EEA (currently about 3 000 — the figure is constantly changing) and, for the largest, completed by hydrographical information taken from external sources: Article 13 and mostly Wikipedia (depth, volume of natural lakes), completing volume information provided by the dams database Eldred2.
- The simplified groundwater systems, made from BGR transboundary aquifers in Europe is yet complete, but shall be added to ECRINS V1.5 under preparation (it is not part of the submission).

3 Topological problems and corrections

3.1 Issues raised

ECRINS has many errors that are geometrical and therefore topological. This was expected since it covers a wide area and has been built from modelled catchments. Several improvement processes were undertaken, included during the making phase. The consultant responsible for making the water accounts for DG Environment nevertheless encountered several issues, which are summarised here.

3.2 Main problems identified

The FEC coverage does not fully match the different areas to address. The main errors are related to the building process of FECs in islands. This issue is well documented in the report describing ECRINS making.

Another error that could not be corrected in ECRINS v11 is the catchment of endorheic systems that does not comprise the lake itself; this is scheduled to be fixed in the next release.

The main problems relate to river topology and snapping the gauging stations (see Annex 7 , page 161) and delineation of sub-basins.

4 Delineation of sub-basins

4.1 Example of errors related to FECs

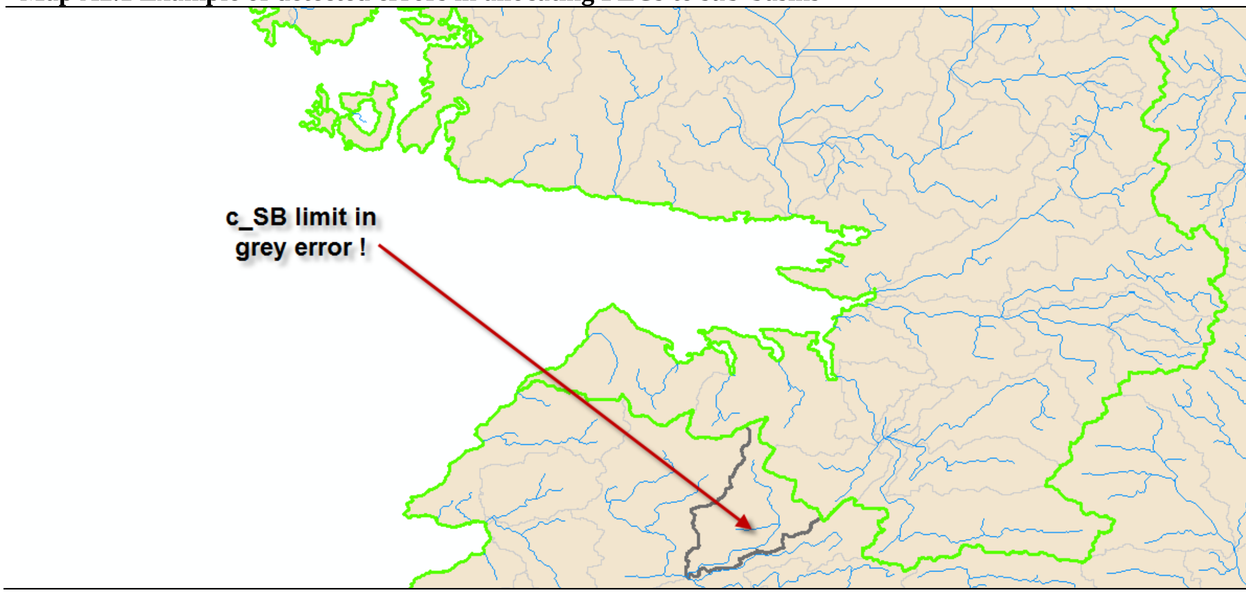
Two types of errors were observed: pieces of land not having FECs, and FECs not allocated to countries. In the latter case (for example, the Isle of Man, the full island happened being not allocated and the related country was set to the 'Joker' ID. The number of such errors was limited and hence did not call for fundamental corrections, being a small cause of uncertainty compared to other error sources.

4.2 Errors related to sub-basins

During the first phases of the development, many inconsistencies were found where FECs were erroneously allocated to the wrong sub-basin, creating hydrological inconsistencies. In the given example, one FEC was incorrectly attached to the neighbouring sub-basin. This anomaly has been corrected (along with several others). The source of anomalies is the non-automated procedure for creating sub-basins.

After having collected all errors and remarks by Member States, a new set of sub-basins has been created and used to aggregate data. This was possible because the 'statistical units', constituted by FECs, are a stable layer and all calculations are carried out at this level.

Map A2.1 Example of detected errors in allocating FECs to sub-basins



Source: Pöyry, final report 1, July 2012.

4.3 Mitigating topological errors

During the process of computation, some discrepancies between the controlled area provided by gauging stations and the catchment area in relation to the placement of this station in ECRINS were detected. There were many sources of error: those related to incorrect information of stations of wrong snapping are discussed in Annex 7 . Some errors are related to incorrect topology making the rivers branch incorrectly.

5 *Delineation of 'territories of reference'*

The territory of reference is the heart of the aggregation of the water accounts. The natural sub-basin has been selected to this end. The experience of processing with sub-basins raised specific issues that could not be dealt with and might require specific adjustment in further processing.

In line with the FEC definitions, a sub-basin can be defined as either coastal (having no upstream catchment and possibly several outlets) or non-coastal (having an upstream sub-basin and a single outlet). This definition, inherited from the topological principles, posed practical problems in areas where the rivers are contained in long and very narrow parallel catchments emptying into the sea: this is typically the case for Swedish catchments. Along the coast, catchments that are coastal, then non-coastal, and then coastal again are found, preventing the dissolving of these sub-basins into larger ones, with many outlets, but reciprocally preventing from discriminating their upstream part.

The consequence of this specific geographical situation is that many of these rivers are not monitored, and that the scoring computed per sub-basin deeply contradicts the scoring at the country level.

There is no simple solution to such configurations; the only possibility is to revisit the process of:

- data reconstruction (this would be relatively easy, thanks to the dense monitoring network and homogeneity of water regimes);
- data aggregation, applying procedures developed (not applied in this exercise) to aggregated water accounts at non-hydrological levels, e.g. NUTS.

6 Improvements to schedule and implement

The improvements required to make water balances under the SEEAW and their associated production of indicators more accurate and hence more useful, are the following:

1. eliminate all topological errors in relation with **river connectivity, this will result when shifting from ECRINS v1.0 to ECRINS V1.2 or v2.0;**
2. **eliminate the substantial geometrical errors that prevented accurate snapping of gauging stations, and implement contextual QA for gauging stations;**
3. **revisit the delineation of sub-basins so that aggregation basins could be a set of the next non-coastal basins (hence having the possibility to have upstream sub-basins), to mitigate the specific cases of narrow parallel sub-basins that are poorly documented (and that do not need to be documented).**

Annex 3 Meteorological data

1 Place in the process

1.1 Place of climate data

Climate data are essential to the water balances production. Meteorological events are the primary source of water and should be broken down in their flows and allocated per location (as sketched in Figure 2.2, page 28); the corresponding values populate the cells of Table 6.1 of the assets.

Climatic information is important data for the assets.

1. Soil water is one of the key compartments of water and is understood as primary recipient for effective rainfall (see Section 1.2 of Annex 3). This compartment is identified as such in the SEEAW. The SEEAW notes that the assessment of soil water, despite being extremely important, is rather difficult to achieve, and accepts that this item could be left empty (the EEA took the opposite view, for reasons explained below (see Annex 4 page 145)).
2. Soil water is the water hub for surface water and groundwater, through soil infiltration, and both pathways are part of the SEEAW tables.
3. Soil water is the main source of rainfed agriculture, and hence this is an important natural asset with economic worth. Similarly, natural water sources for other vegetation patterns are an important side-product of the accounts (e.g. to fuel carbon accounts);
4. Precipitation and evaporation over the reservoirs, lakes, rivers and glaciers.

However, climatic data are likely to be uncertain (data come from models or at least patchy monitoring that is spatially extended) and are submitted for validation versus the river outflows, which in turn contributes to quality assurance of climatic information.

1.2 Operational definitions

The meteorological sources of water focus on 'effective rainfall' that has three different definitions:

1. the amount of rain that reaches the soil, after leaf interception has captured part of the rain (meteorologists' definition);
2. the amount of rain that is available for plants' evapotranspiration (agronomists' definition);
3. the amount of rain that contributes to groundwater replenishment or surface run-off (hydrologists' definition)

In this annex, the first definition is used for describing the meteorological model applied to compute values.

1.3 Historical sources of climatic data

Several sources of information were mobilised during the development of water accounts at the EEA level. During the first WA implementation, carried out at the EEA's initiative, data were generated from the only available source. The ATEAM grid provided densely reallocated data (10' × 10', i.e. about 8 km x 16 km resolution, yielding several data per FEC) with limited parameters, but at a monthly time period, making the actual evapotranspiration more than uncertain. The ATEAM source had to be abandoned despite its high spatial granularity, because no computation beyond the year 2000 was scheduled and because of the inappropriate time resolution.

In a second improvement, the MARS data sets provided by the JRC were substituted. The MARS data have daily time resolution, and in theory comprise all required parameters (even though snow is often not populated), but they are based on a 50 km × 50 km grid. Disaggregation to FECs and aggregation on time (a 5-day period and monthly) was SQL base–processed and directly connected to the Nopolu System2 water account application. Analysis of the water accounts pilot studies results strongly suggest that:

- monthly aggregation of potential evapotranspiration is not capable of providing accurate balance results, and that actual evaporation pre-computed in the MARS datasets, could not be controlled since soil components taken into account by the authors of MARS were unknown;
- higher spatial resolutions and more accurate coverage than that provided by MARS were needed (see below in this appendix);
- actual evapotranspiration (ETP) should be computed from daily data and take into account the soil capacity and wilting point characteristics, as suggested by scientific evidence and supported by the SEEAW;
- data availability could not be secured on the long time for data ownership issues;

1.4 Currently used data sources ([69])

For achieving a more precise series of water balances calculations, a full process of building a monthly climatic database and distributing it at the FEC level has been developed by the EEA. The computations were carried out so that both precipitation (as liquid and as snow) and evaporation from different surfaces (potentially applied to liquid areas and actual evapotranspiration from vegetation) could be obtained to populated SEEAW cells.

All data are derived from exploiting the soil–water balance model outcomes (Kurnik et al.). The study region includes the European continent (1.3) area (between 35° to 70° N, -25° to 30° E), excluding previous Soviet Union countries but including Turkey (=35° to 40° N, 30° to 45° E). The study domain covers the EEA member and collaborating countries (but due to the lack of sufficient input data, Iceland is not included).

The data sources used in the provision of data incorporated in the water accounts is manifold: there are source precipitation and temperature data (primary source), complementary meteorological data for the climatic components of modelling (secondary source, e.g. wind speed) and finally, vegetation and soil characteristics (tertiary source).

The primary source is the E-OBS (version 5) gridded meteorological data that have been used in this study. The collection of data was primarily carried out by the Royal Netherlands Meteorological Institute (KNMI), which also hosts the European Climate Assessment and Data set (ECA&D), based on daily observations that were compiled for precipitation, and minimum, maximum and mean surface temperature covering the time period 1950 to 2011 (version 5). The ECA&D set of observing stations served as the starting point for the ENSEMBLES project. E-OBS version 5 uses around 2 500 stations across Europe; however, stations are unevenly distributed, with the highest density in west and central Europe and the lowest density in the east part of the analysed domain. Station data have been interpolated into 25-kilometre grid using universal kriging.

The secondary source comprises gridded daily wind speed, global radiation and relative humidity that have been obtained from the MARS - STAT database (JRC) which contains meteorological

([69]) Substantial part of text taken from report and publications provided by Blaz Kurnik, who carried out this modelling as EEA staff.

interpolated data from 1975, covering the EU Member States, the central eastern European countries, and the Mediterranean countries including Turkey. Meteorological stations are located mainly in the agriculture zones, due to the main objective of the database construction: agricultural modelling. Originally, in the MARS DB, data refer to more than 6 000 stations distributed in 48 countries, but of these, less than one third present an adequate level of reliability and regularity in providing data. Also, in this database, data have been interpolated into 25-kilimetre grids using universal kriging.

The final group of sources relate to crop coefficients, vegetation index and soil characteristics. Crop coefficients are estimated from the grouping of the 44 original classes of the 2006 CLC into 10 different groups for the purpose of this modelling study (Table A3.1). This latest release of CLC was used as a basis for calculation of crop coefficients.

The crop coefficients are used to convert reference evapotranspiration into potential evapotranspiration (ETP_{pot}=kc* ETP_{ref}). The values in the reported Table A3.1 were compiled by FAO from hundreds of observations of different crops (Allen et al., 1998); the table is taken from Kurnik et al.

Table A3.1 Crop coefficients in non-water stressed conditions and for well-managed land cover in semi-humid regions

Land cover aggregate	Kc Ini beginning of growing season	Kc Mid middle of growing season	Kc End end of growing season	Kc yearly yearly aggregate
1. Artificial surfaces and low transpiring vegetation	0.20	0.20	0.20	0.20
2. Pastures and grasslands	0.35	0.90	0.80	0.68
3. Arable land and permanent crops	0.46	1.11	0.71	0.76
4. Broad-leaved forest	0.60	0.75	0.20	0.52
5. Coniferous forest	1.00	1.00	1.00	1.00
6. Mixed forest	0.80	0.88	0.60	0.76
7. Agroforestry	0.63	0.99	0.65	0.76
8. Transitional woodland and scrub	0.24	0.26	0.18	0.23
9. Wetlands	0.77	1.18	0.74	0.90
10. Inland water	1.05	1.05	1.05	1.05

Source: Allen et al., 1998, cited and recompiled by Kurnik et al.

The last data set in this group relates to soil properties. Computation of evapotranspiration results from the storage in soil (requiring the soil capacity to store water) and wilting point (expressing capacity of soil to deliver water to vegetation in water stress conditions). The hydrological soil properties (namely point of saturation, field capacity, and wilting point) have been retrieved primarily from the European Soil Database (JRC-EC, 2010), combined with the FAO soil map of the world, to produce a homogenised map of soil texture and organic matters, prepared earlier by the EEA ETC/LUSI, and finally completed for the modelling by including other sources of information that are not detailed here (see Kurnik et al. for details).

The final computations from the source soil data sets yielded soil water content at selected tensions, notably 1 500 hPa (field capacity), 33 hPa (wilting point) and 0.1 hPa (soil saturation point) by linear regression with soil characteristics and mineral contents. The procedure and sources of knowledge are in the publication where the modelled data are checked against the observed soil moisture at experimental sites (Kurnik et al.).

Summary details about the model are given in Section 3 of this annex.

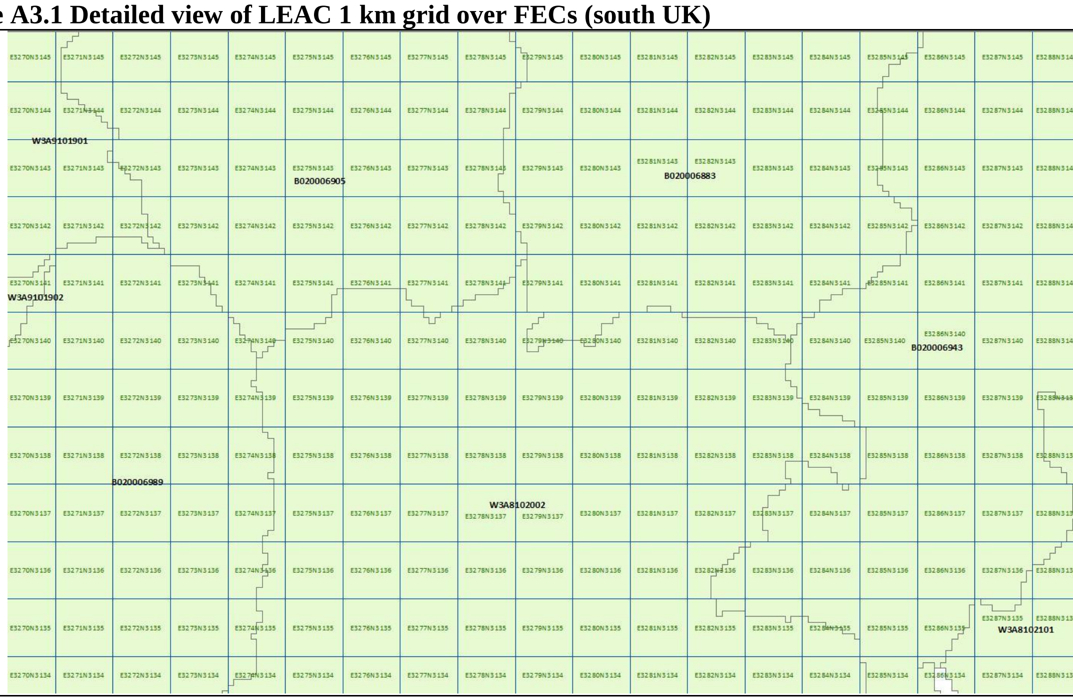
2 Structuring data sets and populating information

2.1 Information expected from the water accounts module

To populate the SEEAW cells, the water accounts module requires, per statistical unit (=FEC) and within a FEC, per category of land cover type (crops, waterbodies, rivers, etc.), a flow of water that is taken from the source data produced as output of the simulation model described in Section 3 of this annex.

The complete data pathway is quite complex, and responds to the best fit between modelling constraints, and SEEAW specifications, themselves adjusted to the structuring of the WA application of Nopolu System2 .

Figure A3.1 Detailed view of LEAC 1 km grid over FECs (south UK)



Source: Pöyry report, July 2012.

The data are hence produced according to:

a. At the EEA level: the surface water balance model computes, per grid element (25 × 25 km), and based on daily data, monthly aggregates of rainfall (raw), potential evapotranspiration (ETP) actual evapotranspiration for the weighted soil coverage of the grid (ETR), surface water balance (SWB) and run-off, and difference between rain and ETR.

b. At the EEA level: the source grid data is exploded to the LEAC grid (1 × 1 km) and stored in the EEA_quan SQL-S database for further aggregation. At this stage, there is no interpolation whatsoever. Disaggregation of all spatial data is carried out at the LEAC ([70]) grid level (1 km^2 — see Figure A3.1), so that any further reallocation should be made from this standard level, that can be managed as a central component of the system data sets and that could be immediately available for the ecosystem accounts.

([70]) The Land and Ecosystem ACcount (LEAC) is the standard Inspire-based kilometric grid to which all continental data sets refer.

c. At the EEA level: given a target set of polygons (FECs of ECRINS v1.0 in this case), all data from the LEAC grid are aggregated at the FEC level, thanks to the appropriate procedure that extrapolates at the FEC area missing data. However, to be extrapolated at the FEC level, no LEAC cell in which the 'no data' joker had been erroneously replaced by '0' can be accepted. This problem was discovered recently in the water accounting process, since many missing data were substituted by '0' at an unidentified place in the process.

d. At water accounting process level: meteorological data aggregated per FEC is disaggregated inside the FEC per category of land cover (land proper and waterbodies, according to SEEAW classification).

e. At water accounting process level: once disaggregated, rainfall is possibly recomputed as snow, if the temperature is below the scenario threshold (below 0°C during the month is the standard value). Rainfall is considered to be snow that is kept and cumulated as rainfall until the temperature triggers thawing (no sublimation computed during the period).

f. Rainfall and disaggregated ET are placed in the appropriate boxes; the run-off (also computed by the EEA) is apportioned between surface water and groundwater using the SUGAR partitioning coefficient (see Annex 4

Considering the duration of computations and the potential worth of such data sets for other purposes, a routine process has been implemented that carries out aggregation of any timely gridded variables to FECs coverage using a transfer Matrix in a SQL Server database.

However, the current process, which is quite time-consuming in execution, is not fully satisfactory since it involves land cover being used to compute the ETR, applying the crop coefficients (item a above). However, in item d proportion of CLC within each LEAC cell is identical to the source data distribution (the size is 25 × 25 km). This means that some uncertainty is added, since in item e, the reverse process is carried out.

At an improvement stage, the primary disaggregation should be envisaged with disaggregation based on CLC (or any other land cover categorisation) to keep untouched source gridding when transferring to the central LEAC data set.

2.2 Implementation in the water accounts calculation module

The climate table has FEC × months in the chronicle dimension, assuming a set of ~200 000 FECs, and 20 years, the number of lines is 48 million lines, making the table a SQL-S table and not a MS Access® one. In the current accounts, the size was less (not full ECRINS coverage, less years) and reached only 11 023 800 records; this notwithstanding, the database has a size of 1.7 GB, which is close to the practical limit of an operational MS Access® database ([71]).

The data set structure (V_Clim2 table) has been designed to fasten queries. Flexibility given by parameter coding (pile table instead of matrix table) has been abandoned in order to deal with capacity and processing speed issues. This is possible since the meteorological parameters are stable enough to cope with a column structure. The compromise chosen and applied to data sets exported from EEA_quan, and subsequently used for the accounts is as follows.

- Each record comprises all the relevant variables. A record is for a place (cell/FEC) and a type of aggregate (in the event less than monthly values should be taken in the future).
- According to the number of different aggregates that have been defined, there are as many records as aggregates defined. For example, monthly average (sum) is an aggregate;

([71]) The physical limit is 2 GB, but in practice, a smaller size is safer, unless the database is in read only status.

monthly maxima are another, monthly minima a third, etc. In the current export for the accounts, only the sum month (SM) aggregate is present.

As suggested by some Mediterranean countries, the following water accounting should comprise more years and the working meteorological table should be processed as a SQL-S data set rather than as a single MS Access. This does not change the WA application.

2.3 Specific procedures and complements in the water balance calculations

Snowmelt is an important temporary storage that is potentially of significant importance when considering monthly data (whereas it was considered negligible in the standard SEAAW). It can be calculated as an approximation of transfer between 'glacier, snow and ice' (CODE_EA 1314) water resource category to 'soil water' (CODE_EA 133) and 'rivers'(CODE_EA 1313).

However, snowfalls are not part of the meteorological data; they were partly supported by MARS data sets, but were oddly populated and hence not reliable.

The way snow was considered has changed over time.

- In past calculations, when snow was provided as snow layer thickness, snow-to-water equivalence had to be set; the possibility to make this operational again is retained in the current version, since the snow thickness information might become available in the years to come and could hence be introduced again.
- Snow is now computed as a type of rain, and is expressed as mm, that can thaw or not thaw ([72]).

Snow melt is calculated using a Swiss-based formula that uses a thawing coefficient (mm/°C/day) and a thawing threshold (°C) set by default to 0°C. The mobilisation of snow is part of the scenario and is not modelled: water from snow is hence directly apportioned between surface water of soil water and does not adjust the evaporation after melting. This might be a source of uncertainty, but cannot be taken into consideration in the current meteorological modelling as this would go far beyond the frame of the water accounts.

For the first month of the considered period, snow stocks can be mobilised for snowmelt, by decision of the operator of water balance scenario. Despite many attempts to optimise the speed, the process is lengthy. Besides, the speed of calculation process depends on the version of MS Access used.

The precipitation as snow and the gap between precipitation as snow and snow melting is not yet documented, and was modelled by the Nopolu System2 application.

3 Water balance model used to provide data

3.1 Summary of the surface water balance model for the water accounts

Soil moisture is the cornerstone of the modelling, and it is carried out using an algorithm for calculating the water balance at the surface and in the sub-surface horizon. The surface water balance model for the water accounts (swbEWA) model is based on the Food and Agriculture Organization (FAO) recommendations for estimating the available water content in soils.

([72]) This difficulty is mitigated by use of coefficients. Snow coverage per month (in snow layers (cm)), is transformed into its water equivalent afterwards (if snow is provided as water, a 1 factor shall be used). Concerning ice coverage (for the time being ice is poorly managed, because of missing data), ice is managed like snow, except that the water content is different (snow is by default assumed to be 1 mm water per cm of snow, and ice should be 0.9 mm water per mm of ice).

The procedure is typically a water balance process that calculates soil moisture by adding and subtracting losses and gains of the various components of the soil water budget, expressed in terms of water column (mm). Hence, the soil water balance (SWB) can be represented by:

Equation A3.1 Soil water balance fundamental equation

$$D\frac{\Delta SWB}{\Delta t} = RR(t) - ETa(t) - SRO(t) - DP(t)$$

In this equation, D [mm] is the depth of the modelled soil profile (root zone), RR is the amount of precipitation at the surface, ETA is the actual evapotranspiration, SRO is the surface runoff, and DP is the deep percolation. All values are in mm/day.

The quantity SWB [m^3/m^3] is the change of the water volume over an area having depth D between two consecutive steps expressed by time t.

The important meaning of this equation is the dynamic expression of water balance that is refilled by rain and emptied by all other factors over a time period and with radically different dynamics. All monthly rainfall can be concentrated at the beginning or end of the month, changing the monthly balance radically. This is why the computations are carried out at a daily level.

Assuming that rain that arrives on the soil surface is immediately incorporated in the process ([73]), the SWB in the root zone is driven by the pair of fundamental soil hydrological properties, wilting point (WP) and field capacity (FC), corresponding to the aforementioned tensions of 1 500 hPa ([74]) and 33 hPa, respectively. The pressure values indicated are those suction pressures that must be exerted to extract water from soil. In the case, soil is saturated in water, and a very small suction (0.1 hPa) is enough. When the soil is close to field capacity, roots need limited suction to extract water, and a more when water contents drops. The value of 1 500 hPa is a lower limit for wilting point that may reach up to 15 000 hPa in very low granularity. The maximum amount of storable water in the root zone is limited to FC and any amount exceeding this capacity is evacuated in the same time step (on the same day, for example).

In addition, ETa will gradually reduce the amount of soil water in the root zone, and in the absence of precipitation, the soil water balance will gradually reach WP. After this point, no water in the root zone will be available for evapotranspiration and plant uptake. Hence, actual evapotranspiration reaches a ceiling because the the water available for evaporation, and not due to potential demand.

The evapotranspiration estimate is not a straightforward calculation. The process involves first computing potential (ETp), then turning it into actual (Eta) evapotranspiration.

Equation A3.2 Relationship between potential and actual evapotranspiration

$$ETa = k_s \times k_c \times Etp$$

Where k_s and k_c are respectively stress factors and crop coefficients. Crop coefficients were discussed in the source data sets section. Stress factors refer to the fact that vegetation extracts water less easily from dry soils than from wet soils. Thus the model imposes a non-linear

([73]) This is wrong if hard rain falls on desiccated and sloppy soil; in this case, a large share of rainfall runs off, and until the soil is wetted again, does not follow the theoretical water balance equation.

([74]) hPa = hectoPascal, the standard pressure unit. One pascal (P) is the pressure exerted by one Newton over one squared metre. The standard atmospheric pressure at sea level is 1 013 hPa.

relationship between demand from plants and water provision; it attenuates the actual evapotranspiration when water availability tends to the wilting point.

The ETp (mm/day) is the potential evapotranspiration, which is calculated using the FAO Penman–Monteith equation (Allen et al., 1998).

3.2 Possible uncertainties

All values are the result of complex modelling described below. There are important, albeit not calculable, sources of uncertainty. The primary source is the climatic data, provided as observations at meteorological stations. The accuracy of the result directly depends on the density of stations. The density is very different from country to country and is for sure insufficient in Spain and in France. This is addressed in the scoring section.

The second source of uncertainty was elucidated during the first comparisons with river discharges: the computation of ETp provided systematic underestimation of around 10 %, resulting in underestimation of effective rainfall by 25 %. This bias has been corrected, but the source rainfall is likely underestimated, especially in Spain. The final results are discussed with the consolidation of accounts, in the main report (Chapter 5 , page 76 and thereafter).

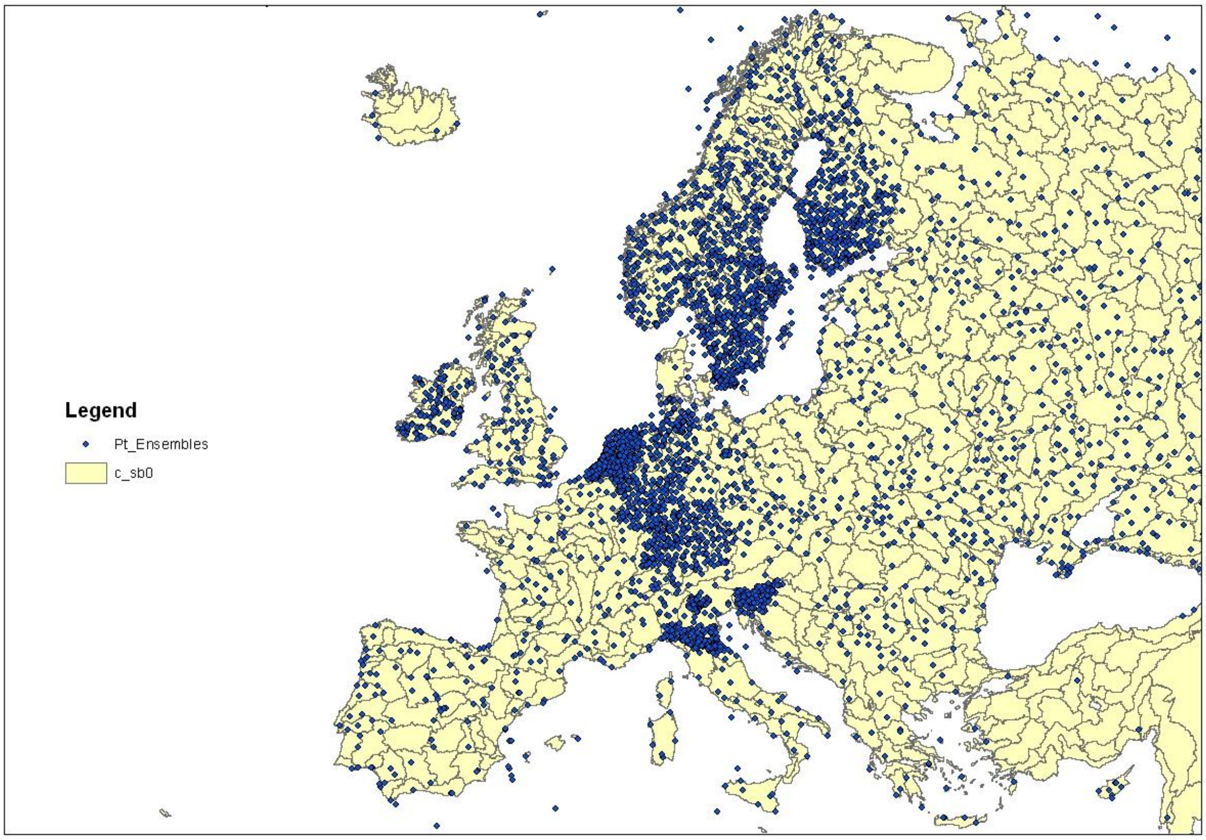
The issue of the complex integration chains and the likely added uncertainties related to the aggregation/disaggregation of the evaporation could be mitigated at the expense of a possible significant increase in computation time and unresolved programming issues.

4 Scoring

4.1 Scoring method and rationales

The full set of stations comprises 4 396 points that are very oddly distributed (**Error! Reference source not found.**). The method for interpolating data at stations is finely described in the reference publication (EC, 2007).

Map A3.1 Source network of rainfall stations incorporated in the source grid



Source: Last set of processed Ensembles stations.

The authors specify that the ideal density of stations would be **'at least one station per 25 × 25 km',** i.e. 625 km^2, and this has been used for reference in the scoring approach. The scoring method has been significantly improved since the presentation of the first scores to the Member States. In this last presentation, only the presence or absence of data in a FEC were making up the score. This method, being simplistic and incorrect, has been replaced by the one described next.

The map shows that some catchments appear to be overlaid by many stations and many others by none.

Analysing the map of climatic stations used for the calculations (the number of stations has been somehow improved since the 2008 reference publication), reveals that **Error! Reference source not found.Error! Reference source not found.** is based on a specific delivery of stations coordinate that is slightly different from the map in the reference document by Haycock et al.

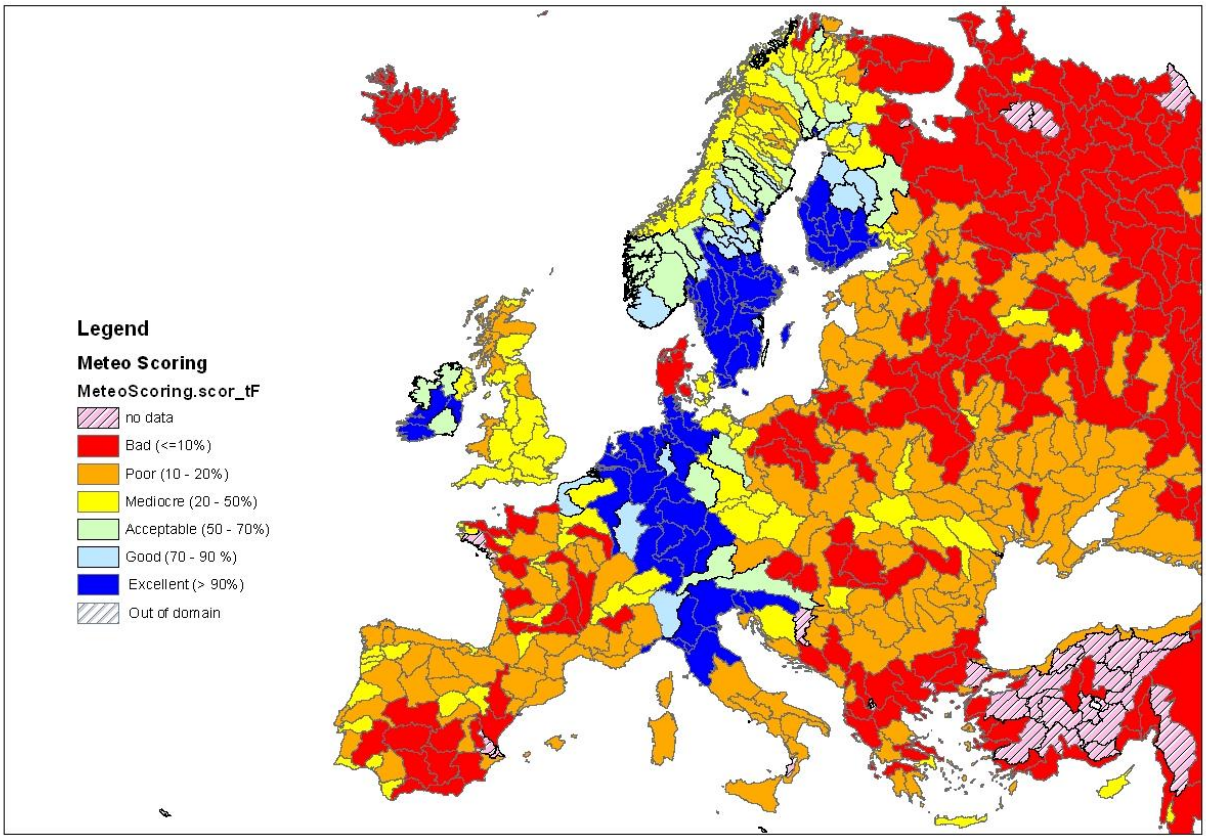
The points were turned into circles of radius 28 km (circle area= 4 × 625 km^2) to consider as documented any sub-basin sharing a common area within the station's radius, to capture any part of a catchment at a distance that could be in the target coverage area of 625 km^2. Since the stations are buffered, in case of high density of stations, the sum of intersecting areas may overpass the total area of the target catchment. This is why the sum is capped: in the densest monitored areas the sum of capture overpasses the total catchment area. This makes it possible to capture sub-basins close to station placement, regardless of whether the station is at the edge or inside or even at short distance outside of any border. Conversely, the captured area should then be divided by 4 to ensure not the coverage has not been overestimated. This is reflected in the scoring formula.

The overlay sub-basin – station has been computed and a primary indicator computed as:

Equation A3.3 Computation of the equivalent number of stations for climatic grid (scoring)

$$St_{eq} = 1 + \frac{\min(\{\sum intersected\ areas/4\}, \{BV\ area\})}{625}$$

Map A3.2 Scoring meteorological source data used by ENSEMBLES for climatic daily data provision



Source: EEA computations from the Ensembles stations data set.

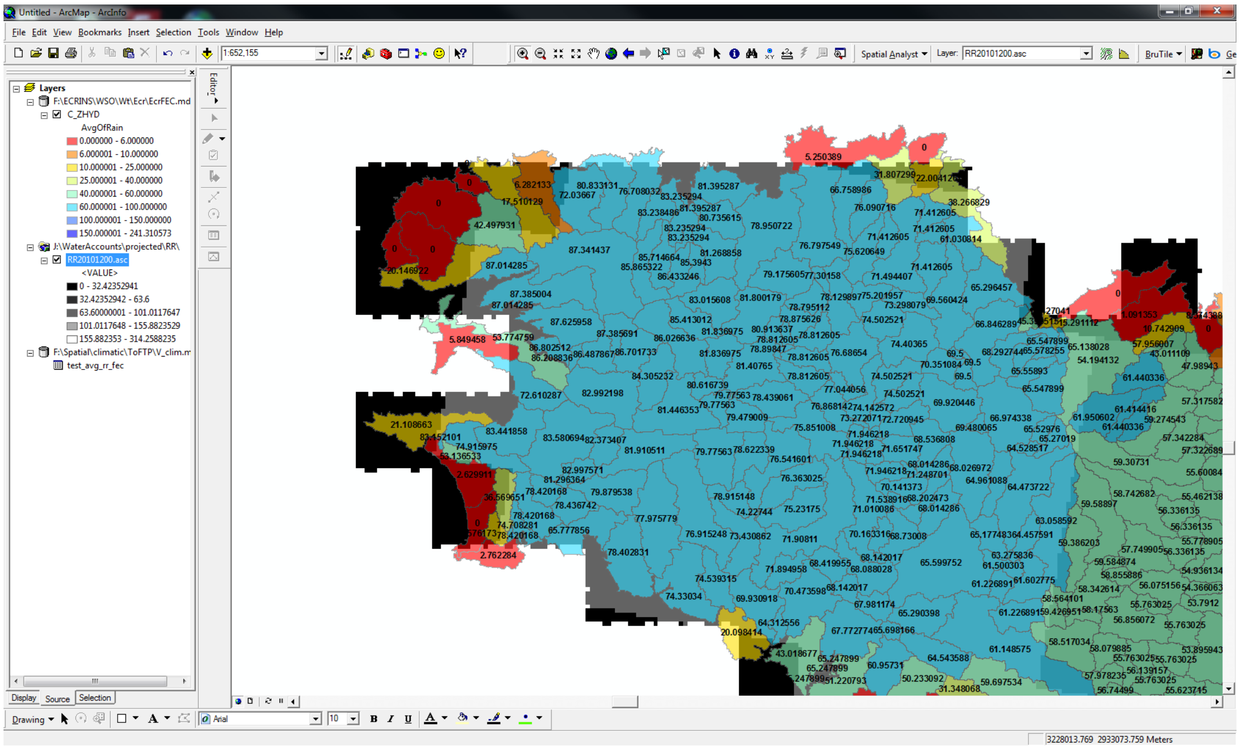
This formula provides the equivalent number of stations, with reference to a target number (625 km^2, that is recommended by the authors (Haylock, 2008); one station is considered as default per catchment, since all the catchments are populated by the gridding. Conversely, the target number is computed considering a minimum number of 1 station (1+(BV area/625 km^2)) for reasons of homogeneity. It becomes obvious that in those areas where rainfall accuracy was questioned by Member States (e.g. Spain), the response is to be found on the map: with such poor data provision, the result cannot be extremely accurate.

Moreover, some catchments have an even worse coverage, because the gridding does not fully overlay the coastal catchment, and because errors in data (jokers replaced by zeroes) result in erroneous interpolating in the last stage of processing. This creates a second element of uncertainty, mentioned in the data assessment report provided to Member States. These issues are reported per country in Table A3.2.

Table A3.2 Coverage of coastal FECs per country

CTY	No of FECs	T: Total area of FECs	R: Area with rain data	E: Area with ETP real data	Score as ratio of the smaller of E or R /T
AL	23	1 358.09	1 175.43	1 331.62	0.866
BE	6	1 576.06	1 576.06	1 576.06	1.000
BG	23	1 461.11	1 045.36	1 452.89	0.715
CY	40	4 246.53	3 288.49	4 223.85	0.774
DE	93	10 377	9 402.56	10 210.2	0.906
DK	209	23 571.67	20 746.4	23 008.56	0.880
EE	97	7 127.58	6 110.02	6 945.18	0.857
ES	333	21 944.06	14 060.72	15 739.63	0.641
FI	440	18 380.55	16 800.1	17 833.46	0.914
FR	278	21 872.46	18 779.76	21 320.94	0.859
GR	511	28 896.93	17 208.6	26 300.56	0.596
HR	138	6 097.42	3 943.08	5 086.08	0.647
IE	174	15 215.15	14 137.56	15 077.86	0.929
IS	316	34 408.57	0	0	0.000
IT	470	28 039.88	24 567.14	27 274.32	0.876
LT	5	568.86	266.77	568.86	0.469
LV	32	2 679.32	25 45.57	2 679.3	0.950
MK	9	165.06	165.06	165.06	1.000
NL	92	17 984.39	17 416.45	17 915.81	0.968
NO	886	70 633.18	64 497.53	68 231.59	0.913
PL	45	3 473.21	3 277.95	3 466.39	0.944
PT	70	5 445.06	4 546.14	4 834.72	0.835
RO	16	1 649.12	1 571.85	1 611.52	0.953
RU	570	62 877.81	1 371.71	1 401.53	0.022
SE	556	28 642.48	25 738.28	27 612.44	0.899
SI	2	95.42	95.42	95.42	1.000
TR	402	42 516.7	28 917.54	31 713.76	0.680
UA	137	18 675.2	215.98	215.98	0.012
UK	580	49 089.41	42 621.69	48 273.34	0.868

Map A3.3 Sample map of erroneous calculations of rainfall into coastal FECs



Source: Final report by Pöyry, July 2012.

Comments: all areas in yellow to red indicate FECs where the rain value is jeopardised by coastal incorrect coverage. Figures are the average rainfall in the period.

4.2 Reported insufficiencies

The incorrect scoring of some sub-basins can generate substantial errors in climatic inputs. The Spanish representatives mentioned in a report delivered at the Member States' consultation meeting:

'En cualquier caso, los resultados obtenidos para España de precipitación y escorrentía superficial total son muy inferiores a las estimaciones habituales. El volumen de precipitación anual media estimada en España es de 348000 hm³/año, mientras que en el documento se indica que son 250126 hm³/año. La escorrentía superficial media anual se estima en unos 82000 hm³/año, mientras que en el documento se evalúa en 28795 hm³/año. Ambas cifras, como se ve, muy alejadas de la realidad. En cambio la ETR media anual en España se sitúa en unos 234000 hm³/año, mientras que en el documento se proporciona un valor de 218637 hm³/año, que se encuentra menos alejado del valor real.

La diferencia tan importante en precipitación parece indicar que la información de partida de precipitación no es adecuada. En cambio, el que partiendo de una precipitación tan desviada de la realidad se obtenga una ETR encajada y una escorrentía mucho más alejada de la realidad que la precipitación, parece apuntar a que el procedimiento de modelación utilizado no reproduce adecuadamente el comportamiento hidrológico en el caso de España. Otra posible justificación es que el período de modelación (como se ha dicho anteriormente, se indica que son 10 años pero no se detalla cuales) no sea hidrológicamente representativo, correspondiendo a un período muy seco, aunque es difícil atribuirle a ese aspecto una influencia tan elevada ».

This translates as follows: 'In such a case, the precipitation and superficial run-off results obtained for Spain are much below the usual estimates. The median annual volume of precipitation estimated for Spain is 348 000 hm^3 whereas the [water accounts] report only 250 126 hm^3. The median superficial run-off is estimated at 82 000 hm^3, whereas the [water accounts] report 28 795 hm^3. Both figures, as can be seen, are much below reality. By contrast, the median ETR in Spain is situated around 234 000 hm^3, whereas the [water accounts] reports a value of 218 637 hm^3, which is just slightly below the real value.

Such a difference in precipitation strongly suggests that the information of precipitation distribution is not adequate. Contrasting with such underestimated precipitation, one finds a value of ETR more realistic and a run-off totally disconnected from reality.'

The rest of the comments suggest that either the modelling process is not correct or that the time lag of modelling should be enlarged beyond 10 years.

The factual observations are entirely correct; we observe that, after a first modelling where the ETR was much underestimated, the correction generated a rather reasonable estimate (the difference with the Spain data is ~6.5 %). Rainfall and run-off that are modelled less (only extrapolated over areas for the first and deduced rain-ETR for the second) are both very inaccurate, because of the rainfall underestimation (rain is 31 % underestimated!). As a result, the run-off deficit is amplified, precisely because the ETR modelling is accurate.

The reason for this is the lack of stations' coverage for Spain: more data would solve the issue and their obtaining it is an urgent priority.

5 Improvements to schedule and implement (climatic)

5.1 Upgrading the crop coefficients

Vegetation activity information is based on the normalised difference vegetation index (NDVI), which is widely used as a proxy of the vegetation status and health. NDVI calculations have been systematically carried out by ETC/LUSI and are currently being quality checked for suitability.

NVDI is derived from the red and the near infrared bands recorded by satellite sensor following the following equation:

Equation A3.4 NDVI equation

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)}$$

In which NIR and RED are the amount of near infrared and red light reflected by the object (here: vegetation) and this indicator is understood as potentially capable of monitoring land cover changes and land degradation. This index ranges between -1 to +1 and relies on the following: i) the absorption in the red part of the light spectrum due to the chlorophyll contained in the leaves; and ii) the scattering of the NIR by the mesophyll's cells of the leaves. The more green and turgescent the leaves are, the closer to 1 is the NDVI. Values below zero present non-vegetation land cover.

A major issue in using such indicators regularly is the need for correcting and filtering data for calibration, view geometry, volcanic aerosols, clouds, water masses (and field inundations) and other effects not related to vegetation change.

In order to check the relevance of NDVI for such categories of uses, the EEA has carried out a project with SCM (under a framework contract) to analyse the data quality, using forested areas (in which vegetation is considered stable between years, excepting accidents) as a benchmark. In parallel, attempts to replace the lumped and quarterly crop coefficients by timely and monitored values derived from NDVI have been tested and seem very promising.

5.2 Procedural improvements

The analysis of issues related to climatic data call for three classes of improvements that relate to different domains.

1. Regarding source data, the current density of source data is low compared to requirements; this low density is reflected in the very different scoring and explains the very different matches between effective rainfall per sub-basin and river discharges observed at the outlet. This and the other uses of climatic data sets call for a deep improvement of data provision, undertaken jointly with the JRC.

2. Regarding climatic data processing, the following is noted.
    a. The first phase of creating gridded data should aim at reducing grid size so that the grid area is on average significantly smaller than sub-basin size. This process is currently being carried out by external meteorological specialists and should be included in the EEA/JRC requirements.
    b. The second phase is the computation of the relevant parameters at the gridded level, including the land cover and soil characteristics. This task, that has been carried out internally by the EEA, is not secured for the future. It should be revised to incorporate NVDI information (detailed and data-based, including seasonality effects instead of the quarterly and inclusive parameter of seasonality); this should become a standard process in data production and be regularly updated as well as computed according to the different instances of CLC.
3. Regarding climatic data post-processing for the accounts, the following is noted.
    a. The current post-processing converts the gridded source data into the LEAC grid, disregarding the land cover distribution. The programme should be adapted to make this happen.
    b. Aggregation per FEC could consider providing the land cover differentiation directly instead of lumped characteristics. Both actions require in-depth revision of the procedure to store the same land-cover data sets as those used in previous phases (data processing).

Annex 4 Land use and soil-related data

1 Place in the process

Land use information has a limited direct place in the process. By contrast, land cover and soil information are ancillary variables of paramount importance in many processes leading to the provision of the appropriate data.

Direct inputs are the apportioning of FECs' coverage and dedicating the share of FECs to certain activities, namely agriculture.

2 Land cover processing

2.1 Data structuring

The land cover of each FEC is documented and stored in table Hydrosol. The standard process for building the Hydrosol table is based on the intersection of CLC with both the FECs and NUTS5 feature sets in order to create an administrative layer for hydrological land cover computation.

Due to the fact that no NUTS5-level references were needed, the structure was retained and the administrative layer just populated with the FEC ID. The structure of the Hydrosol table relates to the library of crossing (table Hydrosolib) as shown in the sample display of the table.

Figure A4.1 Sample display of Hydrosol table as related to its log table

HYDROSOLlib

numHYDRO:	libel	desc
4	CLC Year 2006	calculation performed via vector CLC inter FEC (version 1.1) on RASTER CLC (version 2006)

CM	ZHYD	I_HC	SURFACE	LINEAIRE	SUR_HC	SUR_COM	F_ZI	F_IZ	111	112	121
A020000004	A020000004		16.7	-1	16.7	-1	100.00%	-1		1.01997	
A020000005	A020000005		10.27	-1	10.27	-1	100.00%	-1			
A020000007	A020000007		3.22	-1	3.22	-1	100.00%	-1			
A020000008	A020000008		6.58	-1	6.58	-1	100.00%	-1		1.26501	0.01656
A020000009	A020000009		8.15	-1	8.15	-1	100.00%	-1			
A020000010	A020000010		25.41	-1	25.41	-1	100.00%	-1			
A020000011	A020000011		76.94	-1	76.94	-1	100.00%	-1			
A020000012	A020000012		61.14	-1	61.14	-1	100.00%	-1		0.11247	

The table contains as many fields as there are CLC categories. Only categories 111, 112 and 121 are displayed in the sample display.

2.2 Documenting agricultural uses

Agricultural uses must be documented, since the rainfed volumes and the irrigation areas are taken from the agricultural areas per FECs. These agricultural areas correspond to the 2.x.x codes, for example, plus others in the mixed groups.

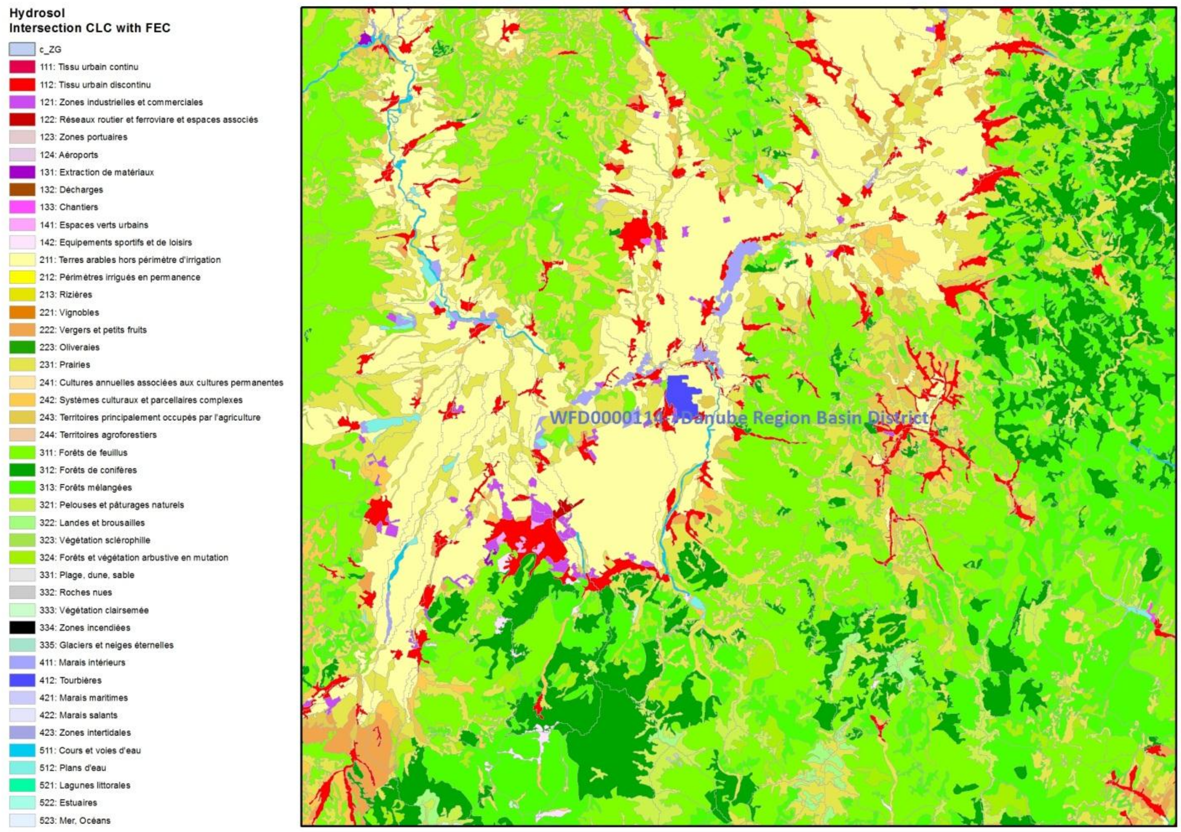
Precise land cover references for each CLC type per FEC have been built from reference to the 2006 CLC.

The figure expresses that within a single FEC there can be several different types of land use; however, aggregation through the Hydrosol computation is the single affordable way to take into account the apportionment of the different soil coverage per elementary catchment.

The main issue (whose impacts on uncertainty cannot be assed for the time being) is the difference in approach for computing the evapotranspiration and it further re-allocating during the water accounts process.

The authors are conscious that this appears an illogical method, but it is the only one that is feasible, considering the duration of meteorological modelling (expressed in weeks). A sounder way would just prevent any calculation from being carried out.

Figure A4.2 Map of Hydrosol source on the Danube River Basin District (over Romania)



3 Making urban delineations

This issue is discussed in Annex 9 2.1

4 Soil data sets computed

Soil, as mentioned in the annex dealing with climatic data (Annex 3), is essential for modelling characteristics of actual evaporation. Two parameters are required: soil capacity and wilting points, both expressed in mm of water.

4.1 Data sources used

Two data sources were explored and processed before final production. First, the JRC soil data centre was solicited and test modelling was carried out by the ETC/SIA. The test concluded that soil information, as recorded in databases, could be re-modelled to derive such characteristics from standard modelling equations.

However, the JRC data centre covers only the EU, and was therefore not fully suited for climatic modelling (that requires comprehensive coverage), so the Harmonised World Soil database (HWSD) was used instead. This coverage takes stock of the JRC data set in Europe (the JRC is one of the authors), but is seamless across borders.

This database (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2009) is free for use outside commercial purposes; the EEA made assessments and computations and downloaded and recomputed the source database at the kilometric grid (LEAC) resolution.

4.2 Production and transformation process

Hydrological soil properties (namely soil saturation point, field capacity, and wilting point) were first calculated and equations tested on the European Soil map (ESDB version 2.0 (JRC, 2013a), by the ETC/SIA. The full calculations, alongside the HWSD, were carried out inside the EEA.

In both cases, relevant information was needed for this modelling (soil texture, percentages of sand (S), silt and clay (C) and soil organic matter (OM)). The equations developed by different authors, primarily Saxton (2006), were used for this purpose. Calculation details and specifics of testing the results against observation points are presented in an EEA publication (Kurnik et al.).

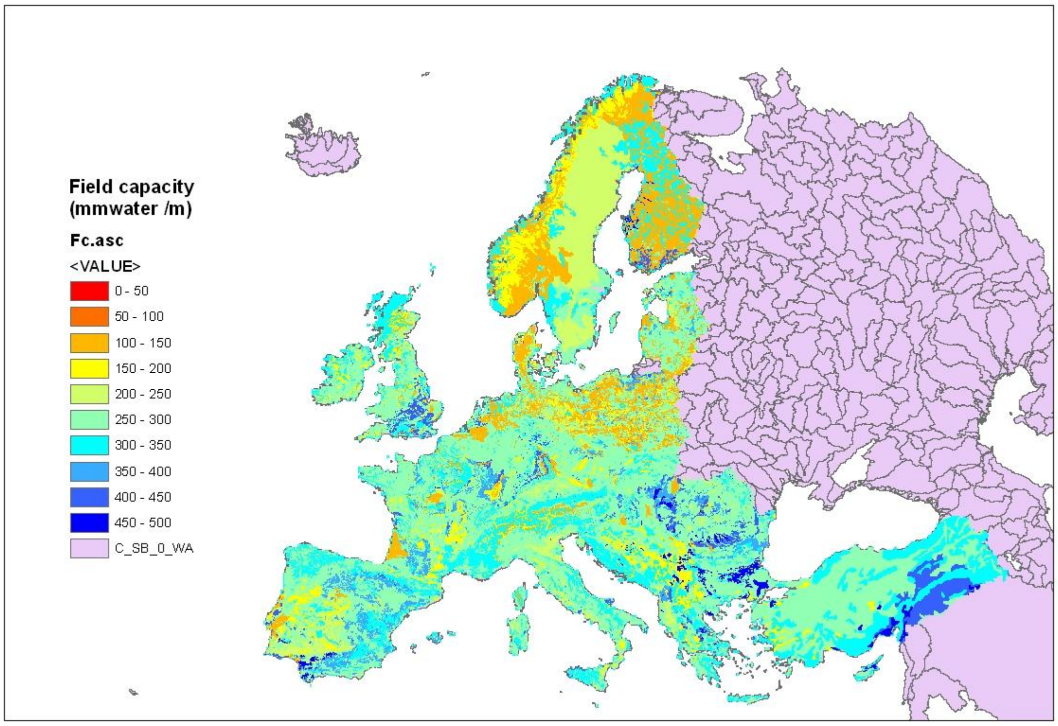
The system of equations aims at estimating the wilting point (soil humidity below which plants cannot withdraw water), corresponding to suction pressure of 1 500 hPa, field capacity (suction pressure of 33 hPa) and soil saturation point (0.1 hPa). The equations (not detailed here) are sets of linear combinations of C, S and OM of the general form:

$$T_i=a_i\times S+b_i\times C+c_i\times OM+d_i\times(S\times OM)+e_i\times(C\times OM)+f_i\times(S\times C)+g_{i,}$$

where the indices i are the different pressures, each pressure having its own set of constants a_i, b_i, etc., the resulting values being linear of parabolic combination of Ti, for a given index. The whole predictive equations are hence the result of correlations that were defined from samples taken and prepared in the United States. The authors note that the equations were developed outside the salinity context, which may make salty soils deviating from the predicted results.

Apart from this, the predictive capacity of the system (limited to three variables) is reasonably good, albeit decreasing from wilting point to saturation, the capacity at 33 hPa of intermediate quality being a bit uncertain (~2/3 of variability explained) and non-linear. Coupled with uncertainty in rainfall described in the climatic Annex 3 , this may result in uncertainty in the water accounts balances that is intrinsic to the process.

Map A4.1 Soil capacity, as recomputed by the EEA from the HWSD

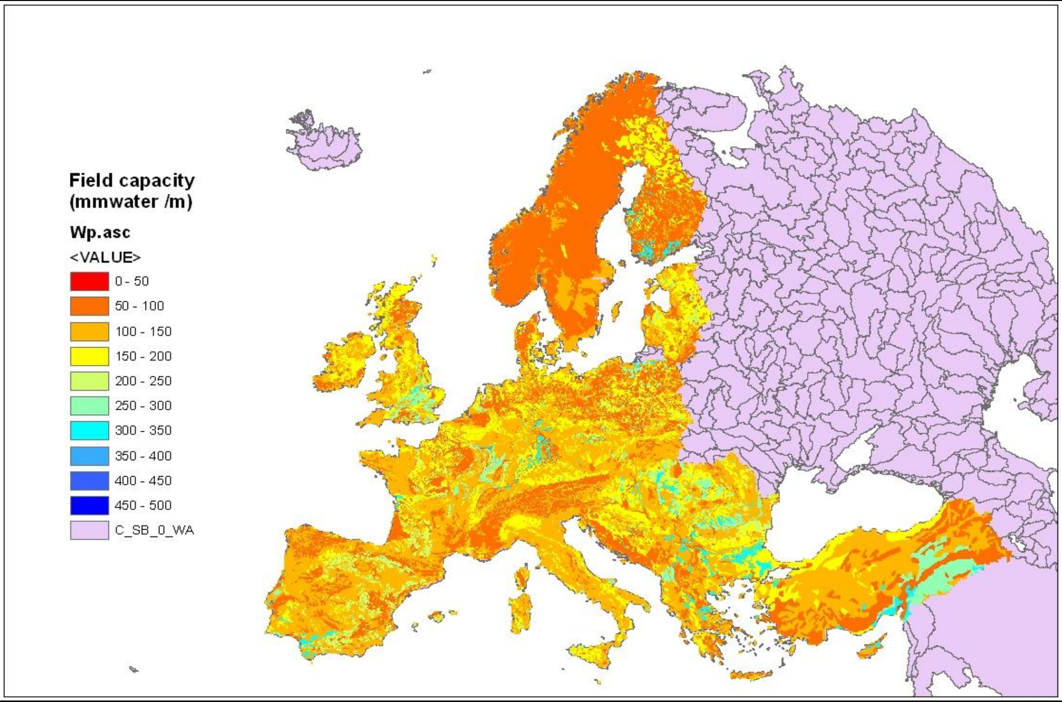


Source: Soil values modelled by EEA.

The spatial distribution of soil capacities and wilting points are displayed in the following maps, to show the importance of soil characteristics, since they differ widely from place to place.

Despite uncertainties, it is important to note that very sharp differences in field capacity may exist within a close radius, hence justifying the efforts devoted to developing locally defined climatic models.

Map A4.2 Wilting point, as recomputed by the EEA from the HWSD



Source: Soil values modelled by EEA.

5 *Suggestions for improvements*

As mentioned, the estimates of soil hydrological characteristics rely largely on EU data for soil structure, but on an American regression model for modelling. With the development of EU soil data centre, it would be advisable to check ascertain the relevance of these modelling equations in EU situations.

In their assessment, Kurnik et al. had only nine pedological references with which to test the time-series of soil moisture and dry periods. Soil hydrological characteristics are obviously part of the differences between predictions and observations, but they are certainly not the single factor of errors and uncertainties .

Annex 5 Reservoir and lake data

1 Place in the process

Lakes and reservoirs (a reservoir is lake resulting from damming) are the major surface water storage facilities. In the accounting process, lakes are free-surface waterbodies that:

- receive rainfall and evaporate to the atmosphere,
- store water from and deliver water to other compartments of the natural water cycle and to the economy.

They are built-in parts of the water accounts table and play a significant role in the regulated water storage and availability. In many areas, the river discharge during summer is just water stored in rainier periods of the year.

2 Feature-dependent information and structural relationships

2.1 Lakes and dams as components of ECRINS

Lakes are part of ECRINS, and are polygon features related to FECs and to river segments. Dams are also part of ECRINS, and are point features in relation to a lake and placed on a river segment. The detailed structure of lakes and dams inside ECRINS is reported in the ECRINS report (Vogt et al., 2003).

The current release comprises about 72 000 lakes. This is given as rounded figure because the accurate number is misleading. The lakes layer is the outcome of a complex compromise between acceptation of lakes as individual entities, data provision and following licensing rules. ECRINS provides the two major tables of relevance for the accounts:

1. apportionment of lake polygons per FEC (to address rainfall and evaporation);
2. inlet (may be nil to several) and outlet (must be unique, may be nil) river segments.

Both main and secondary drains are considered in allocating river segments to lakes. It may be that a lake lacks an inlet or outlet for many reasons, most frequently because the source of the lake's polygon has a finer resolution than the rivers have, hence the rivers attached to the lake may be not represented because of this difference in resolution. However, it may also happen that a lake has no outlet if it is endorheic or a coastal lagoon.

Inlets and outlets are major features to address the presence of lake on a system and for computing the changes in reserve by difference in river discharge.

2.2 Lake volumes and dams

The source of lakes' polygons provide only the area of the mirror, and not any information related to volume, which is essential information when considering a lake as a potential water supply. Several sources of information were mobilised to assess lake volume. These sources are discussed from the volume perspective but are also used to complete names, areas, etc.

1. If a lake is created (or modified) by a dam recorded in the Eldred2 database, which is the source for dams in ECRINS, and provided the dam can be attached to ECRINS([75]), then

([75]) That is to say that the dams position has been set and that this position could be attached to a lake and a river segment. Procedures are described in the ECRINS report and are not that straightforward. Dam coordinate are not in the usual catalogues of dams.

the volume is information that is present in most cases; if volume is not proposed, then the maximum depth is estimated as being 90 % of the height of the dam. Values are added to ECRINS.

2. If a lake has been reported under Article 13 and its volume attached, this information is added to ECRINS.
3. If a lake is recorded on Wikipedia and is harvested, volume, average depth and maximum depths are added to ECRINS;
4. If lake has only area and maximum depth, an estimated volume is computed.

Computing proxy average volume is done by assimilating the lake to a spherical cap of area equal to the lake's area, and arrow equal to the maximum depths. If lake is dammed, volume is halved.

In the general case, lake area is equivalent to a circle

$$A = \pi r^2$$

And spherical cap volume is: $V = \frac{\pi}{2} Z \left(\frac{z^2}{3} + r^2\right)$, where r is $\sqrt{\frac{A}{\pi}}$, following lake's equivalent area.

Hence the calculation can be done in a single run of the next formula: $V = \frac{\pi}{2} Z \left(\frac{z^2}{3} + \frac{A}{\pi}\right)$;

Z represents the estimate of the maximum depth of the lake.

3 Improvements to schedule

The current development of integrating lakes and dams into ECRINS is satisfactory, but nevertheless could use some improvements.

3.1 Correcting ECRINS defects as lake geometry

Lakes' geometry has some imperfections in the current stage of ECRINS; these have to do with the many sources of geometry that have resulted in overlapping polygons in some cases. The geometry will be 'cleaned' when making version 2.

Another flaw stems from the CLC source of polygons that in some circumstances do not correctly separate rivers and narrow lakes resulting from damming, and possibly merge such lakes together.

The main issue of collecting lakes' volumes remains unresolved: the use of Wikipedia cannot provide representative support for this item, and it is time-consuming to operate.

3.2 Streamlining the update of dams

The main current data source for dams is the International Commission on Large Dams (ICOLD), with whom the EEA started to cooperate in the period from 2004 to 2007. This source has been the central component of the Eldred2 database, which fully reorganised the ICOLD source (that is a catalogue) into a relational database. The EEA integrated the 2003 catalogue of large dams, but no formalisation of updates was set in place. This was due to three factors: first, lack of dam positioning in ICOLD (mitigated by the development of DAMPOS ([76]), that took much longer to populate than envisaged, as there were cooperation issues with the national ICOLD contact point), second, delays in development and installing of ECRINS (the natural recipient for dams), and last,

([76]) Dam positioning, a web facility based on Google Earth open-source tools, allows users to collect dam coordinates by clicking a dam on its seen position. DAMPOS is harvested by the Eldred2 administrator and update rights are given to experts on request.

because ICOLD was reorganised and its revised priorities did not include this project (dams are not being built in Europe as much).

Revision of the cooperation with ICOLD should envisaged and could be developed under the SEIS / Copernicus auspices. Streamlining information on reserve variation

Some countries (Spain and Portugal) provide online services of reservoir variations in reserve. This information is contrastingly becoming more hidden in other countries. In fact, where the reservoir (generally a dam) is operated for energy generation, some degree of commercial secrecy is applied, and prevents the provision of such data, even on a retrospective scale. However, regulation reservoirs (for low-flow enhancement for example) are reported in 'hydrological bulletins', albeit not systematically recorded.

The ETC/ICM has included the reservoir change in reserve in the data to be collected; this objective was not fulfilled to the standards required by the accounts, and should be restarted with more specific data collection, by negotiating with countries on which reservoirs to collect.

Annex 6 Groundwater data

1 Place in the process

Groundwater is one of the assets' recipients (SEEAW Table 6.1, Figure 2.5, page 39) and may receive water from returns and soil, and deliver water mainly to rivers (other recipients are not forbidden — the only impossible pathway is precipitation to groundwater. In the real world, this does exist but it is extremely rare).

Groundwater is by definition contained in subsurface recipients, the aquifers (definitions are provided in the main text, Subsection 3.2.3 , page 46). This annex summarises developments and explains how the passage of surface water to groundwater had to be extremely simplified due to the inadequate reference systems, and which mitigating processes were implemented.

2 Groundwater reference systems developed for the accounts

2.1 Minimum specifications

Hydrogeology is a complex science, and groundwater systems are the result of the geological substrate deeply modified by water circulation in bedrock. From the water accounting process point of view, the important issues are the possibilities to identify, per elementary area (term kept fuzzy deliberately) a potential volume for water storage (to compute stocks, that are capped by a maximum volume and floored by a minimum volume).

In this case, the range of volumes addresses not the volume of rock but the effective porosity × volume of rock. Hydrogeologists distinguish at the minimum level of knowledge the total porosity (volume of vacuums within the rock) and effective porosity, which is the capacity in which water can move. Aquifer capacity primarily depends on the porosity and faults in the rock. The capacity of water to move depends on the structure of the rock, the porosity, and the specific area of pores, that drive the balance between capillarity and gravity movements (de Marsily, 2007).

This basic information is the minimum data to be incorporated in a reference system, along with delineation and rock depth (thickness). In the real world, there are many aquifer systems on a single vertical cut, and there may be unconfined on top of confined ones, and possibly fossil aquifers (confined with no more refilling). The first target of including groundwater in the accounts was to consider a delineation of the first layer of aquifers, with the basic characteristics mentioned above.

2.2 Source for reference systems for the accounts

There is no European coverage of aquifers for the time being. The Unesco/BGR ([77]) map of European aquifers is far from ready (de Marsily, 2007), since only 5 sheets of the 30 scheduled were ready in 2007, and they had limited access. The BGR had coordinated the compound map of 'transboundary aquifers in Europe' at the 1:1.5M resolution. This map has been deeply processed and improved, albeit still not reaching a degree of achievement compatible with processing for the accounts.

The only possible developments with the available resources and the general constraints intrinsic to all EEA developments (i.e. public sources so that the results can be disseminated) made it necessary to build the process as follows:

([77]) Bundesanstalt für Geowissenschaften und Rohstoffe; the Federal Institute for Geosciences and Natural Resources is situated in Hannover, Germany. Its French counterpart is Bureau de Recherches Géologiques et Minières (BRGM).

1. use a source of aquifers from the Unesco/BGR map, as modified by the BGR (International hydrogeological map of Europe: IHME);
1. populate the aquifer information with porosity data derived from rock type information from the processing of the International Geological Map of Europe and Adjacent Areas 1:5.000.000 (IGME5000 and OneGeology Europe 1:1 million;
2. derive rock thickness as much as possible from the information obtained from groundwater abstraction points (and any other suitable source).

First, the BGR transformed the Unesco/BGR bitmap into a vector map. This transformation was complex: the source bitmap is more or less the outcome of scanning paper maps that contained a great deal of unnecessary information (names of cities, country boundaries, etc.). The filtration method had been jointly developed by the EEA, and was completed and applied by the BGR.

At this stage, the ETC/SIA core team and the partner CMA-REDIAM took the layers and homogenised them so that the geometry would correspond with the ECRINS frame.

Available data from OneGeology-Europe 1:1 million were reviewed: coordinate reference systems, data models, the degree of completion of fields, lithology and age classes, geometric consistency, and 22 data sets from different countries were analysed.

One of the main issues found was the large share of the 'undifferentiated' value in the main information filed 'Petro1' in IGME5000. This value is a common value in geology maps for those layers with an undefined type of rock. However, comparing this database with other sources, a number of zones occupied by polygons defined as 'undifferentiated' have a known lithology associated. The 'undifferentiated' value occupies massive zones in central Europe and Baltic countries, while minor zones are represented with this value in many other countries. Countries with extensive 'undifferentiated' zones are Denmark, the Netherlands, Germany, Poland, Lithuania, Latvia, the former Yugoslav Republic of Macedonia and Turkey, where a large share of groundwater is exploited as a resource.

In addition, IGME5000 doesn't include information about lithology for quaternary deposits. Quaternary deposits' lithology is derived first from the fact that quaternary aquifers are identified as such in IHME, and second, in those cases the OneGeology map is used, or an average estimation when OneGeology is not feasible.

A second issue is the poor quality of the shapefiles in the 1:1M OneGeology for Europe map: the dilemma lies between a geological map with good features but lacking a great deal of information, and another where the geological documentation is better, but it is highly unexploitable because of the geometrical errors (splinters, non-adjacent polygons, etc.).

The huge cleaning process carried out made it impossible to meet the time deadlines for it to integrated into the water accounts.

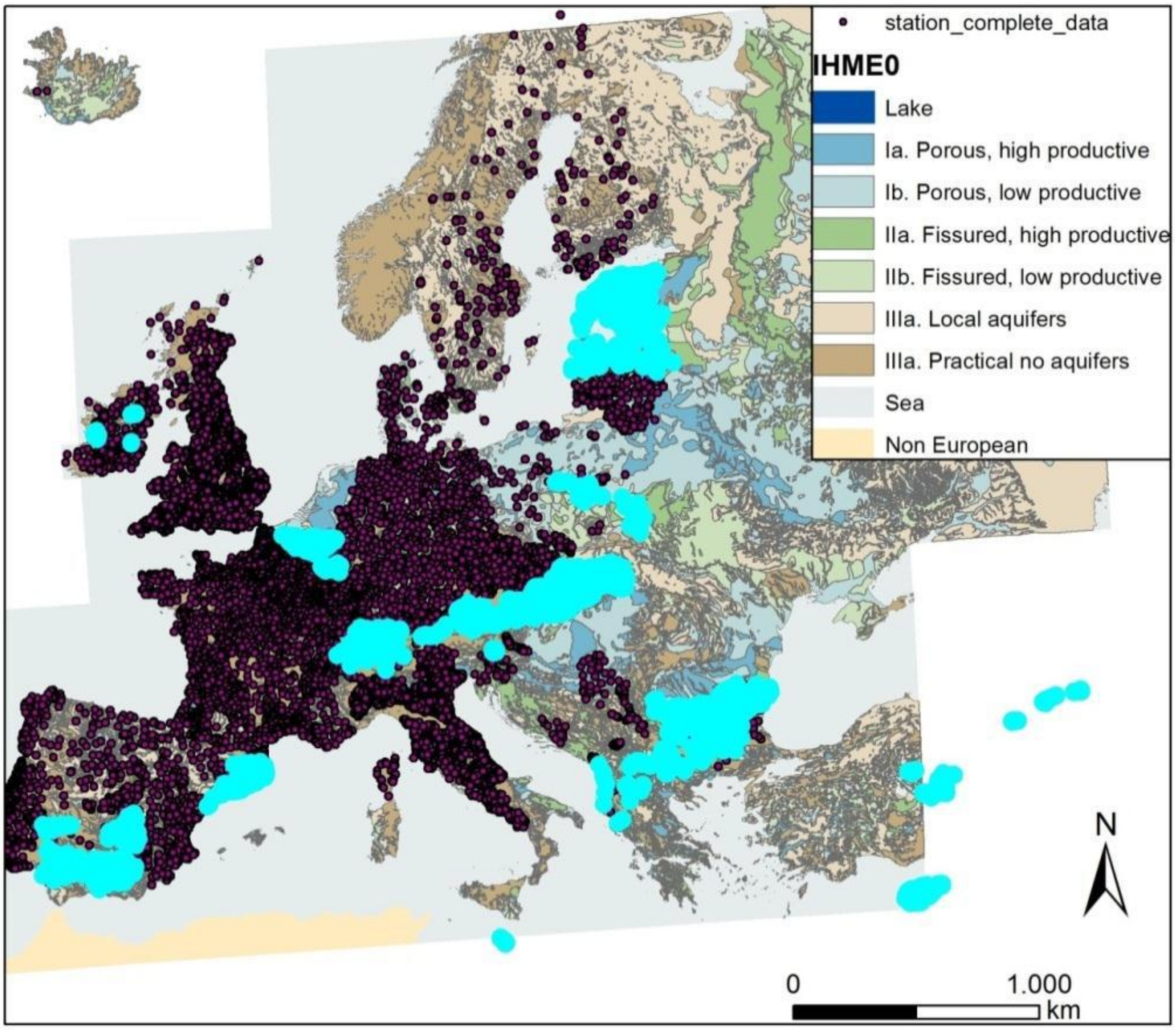
The last piece of information required is bedrock thickness; this has been sought in the reporting information related to wells. However, only a non-representative distribution of such information has been reported under priority data flows.

The thickness information gaps in the groundwater bodies layer are as follows.

- amongst the 7 226 groundwater bodies reported, only 424 have measures of thickness (minimum, mean and/or maximum); in other words, 3 604 station of 20 607 (17 %) reported by member and collaborating countries have thickness data (see the blue points in Map A6.1);

- thickness data reported show a measure per main aquifer or lithology (one data per groundwater body); we can find quite a lot of points with the same thickness data, making the usable number of points even smaller.

Map A6.1 Wells with exploitable information



Source: ETC/SIA report: Underground and groundwater integrated data set, 14 September 2012.

Comment: the underground map is the IHME, after vectorisation.

In addition to this, the groundwater thickness was also sought in the deliveries from Article 13 of the WFD. The completeness of the thickness attribute is again poor (fully informed only for Sweden, the Netherlands and Lithuania, and only partially for Austria and Bulgaria), which jeopardises the potential estimation of groundwater bodies' capacity.

Table A6.1 Summary information in WFD Article 13 reporting, in relation to groundwater bodies

Member State	Figures in no of WB with the mentioned information present.					
	Number GWB	Depth Range Code	Average Thickness	Capacity	Average Depth	**% with thickness**
Total no documented	11 607	3 285	3 270	3 202	3 200	**28 %**
% of the total	100,00	28,30	28,17	27,59	27,57	
AT	136	54	69	0	58	**51 %**
BE	42	42	0	0	0	**0 %**
BG	177	53	70	0	34	**40 %**
CZ	173	0	14	0	0	**8 %**
DE	989	4	0	0	0	**0 %**
FI	3 804	0	0	0	0	**0 %**
FR	574	71	53	53	53	**9 %**
GR	236	0	0	105	0	**0 %**
IE	756	0	0	0	0	**0 %**
LT	20	20	20	0	20	**100 %**
NE	23	0	23	23	11	**100 %**
PL	7	0	0	0	0	**0 %**
RO	142	0	0	0	0	**0 %**
SK	101	0	0	0	0	**0 %**
ES	683	20	0	0	3	**0 %**
SE	3 021	3 021	3 021	3 021	3 021	**100 %**
UK	723	0	0	0	0	**0 %**

Source: EEA computations on reporting from Member States/

3 Developing an alternate source of aquifers delineations

French authorities (BRGM and Onema ([78])) released in July 2012 a revolutionary data set, the *base de données des limites de systèmes aquifères* or database of aquifer systems delineation (BDLISA) (Seguin et al., 2012) that is a reference aquifers layer for France, built with three degrees of simplification: national, regional and local. The BDLISA is the new French hydrogeological reference system made to respond to the requirements set up by the WFD. It covers metropolitan France plus the four overseas departments. It exhibits very important and up-to-date characteristics:

1. it is almost a 3D referential, since it takes into account deep units (albeit thicknesses are not yet included);
2. it delineates units (aquifers and semi-pervious and impervious units) at three different scales: local, regional and national;

([78]) BRGM: Bureau de recherches géologiques et minières (which German homologous is BGR). Onema: Office national de l'eau et des milieu aquatiques, French office for water and aquatic environment.

3. being controlled by a management model, the combination of local units is possible, whilst retaining topological consistency in the grouping of units;
4. last but not least, it is free for use and downloadable from the Eau-France website. Combined with the Adès database of hydrogeological events (level, quality, composition, etc.) it makes for a comprehensive underground information system.

With respect to the water accounts that operate from a 1:250 K resolution (the ECRINS resolution), BDLISA that has three levels of aggregation (local at 1:50 K, regional at 1:250 K and national at 1:1M) and proposes an important alternative to current developments (limited to France) and a reference benchmark to check and validate the developments mentioned in the section above.

4 Current processing of soil to groundwater

4.1 Data structure in Nopolu System2

Water can reach aquifers only after having passed through either the soil or surface water, and thus no precipitation directly reaches groundwater. In fact, the infiltration of precipitation to groundwater is recorded in the accounts as an inflow from other water resources into groundwater. Exchange from soil water to groundwater is assessed via a simple rule, using two tables.

1. Table VR_FEC_Distribution (that may have many scenarios, related to which resource is partitioned) should be read with reference to SCEN-FEC_distribution, considering the SUGAR-derived data sets is used for splitting the water flow from water soil between rivers and groundwater.
2. Table MR_resources_exchanges defines time modulations of monthly exchanges.

Both tables have a large scope. They don't address groundwater exclusively but handle snowmelt similarly, for example. The first table (SUGAR scenario) has been constructed from the source SUGAR raster, by simple overlay then populated the FEC with the weighted average of the intersecting (based on areas).

Each month, the run-off at the FEC level (from meteorological data) is apportioned between surface water and groundwater by application of the SUGAR coefficient (x % to surface, 100-x % to groundwater).

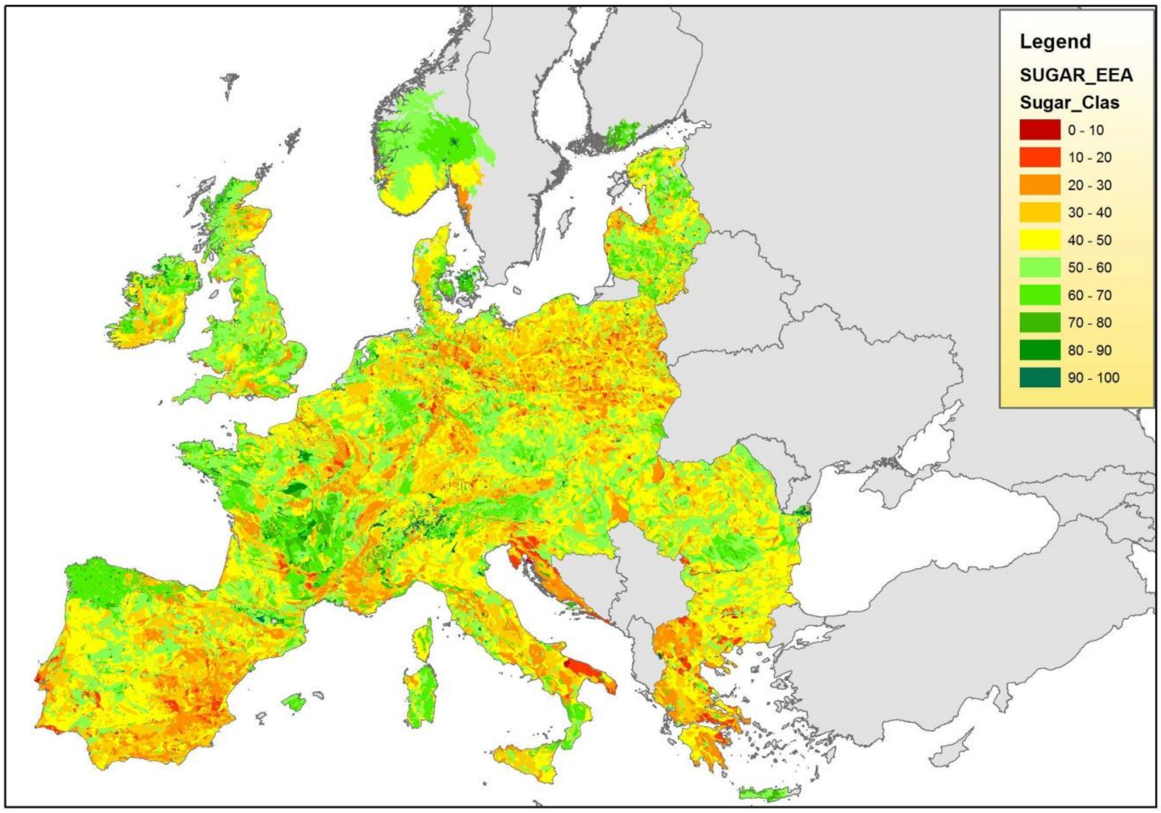
The second table allows applying (for an entity belonging to the group indicated, the year and the month, the origin and target) a factor that modulates the previous table. This possibility has not been used in this case; however, this second table has a very general scope and can apply, for example, from lake to river, from groundwater to rivers, etc., depending on the coding.

4.2 Data sources for allocating soil water to groundwater

Following much research and deliberation it appeared that the SUGAR index presented the most suited solution. The BGR, as ETC/ICM partner and under the EEA's request, checked whether the SUGAR index might be a suitable approach. The comparison carried out by the ETC/ICM provided no negative information that might cause the SUGAR data set to be rejected. Positive proof cannot be given, however.

Unfortunately, then availability of the index does not fully cover Europe ([79]), as shown by **Error! Reference source not found.**.

Map A6.2 Sugar indices as contribution to groundwater



Source: SUGAR public values map prepared by author (Igor Dubus)

SUGAR is an innovative index which indicates whether water falling on a particular zone mostly contributes to groundwater recharge (i.e. infiltrating areas) or to discharge to surface water that relies only on measured data. Being the outcome of a European research project ([80]) coordinated by BRGM, France means that the attached information has been compiled for the whole of Europe and is distributed in the form of national data sets on its website (FOOTPRINT, 2013).

The authors of SUGAR recommend caution when using the index, as it is somewhat 'semi-quantitative'. This is not an issue for the time being, since many other sources of uncertainty exist, and because it is the single wide coverage data set that has some European legitimacy and a homogeneous protocol of production.

([79]) Cyprus, Finland, Malta, Sweden, Turkey, Iceland and Norway do not grant full access to the data sets. Bosnia, Serbia-Montenegro and Albania are not part of the project.

([80]) At the origin, the target is to model pesticide effects on the environment: Functional TOOls for Pesticide RIsk assessmenNt and managemenT (FOOTPRINT) is a research project funded by the European Commission as part of its Sixth Framework Programme for Research and Technological Development (FP6).

The project aims at developing computer tools to evaluate and reduce the risk of pesticides impacting on water resources in the EU (both surface water and groundwater). The SUGAR index is a necessary intermediate developed for this purpose.

In summary, the groundwater issue is extremely complex, and the WA project managers involved are dealing with two opposed constraints:

1. regarding the scientifically sound balance in pilot areas, even a detailed process model's relevance (such as MODFLOW) could be denied by specialists;
2. a complex model would breach accounting principles and be beyond any resource achievable, so there is no other way round this than to integrate, even with huge uncertainties, groundwater as water storage and surrogate resource (compared to surface water).

5 Current processing of groundwater to surface water

Since there are no groundwater reference systems, the groundwater-to-river exchange cannot be based on the usual rationale of depletion curve analysis. However, in order to avoid some drift in exchanges balance, a procedure has been implemented as a set of queries to achieve equilibrium per entity modelled over full period. The rationale is to analyse the result at the end of the period for each water compartment, and then apportion the value for groundwater (being positive) between other negative surface water compartments. Values are spread monthly over the full period and inserted in the final tables. The details of the queries would be out of the scope of this report and form part of the Nopolu manual. Moreover, this procedure should be considered provisional since the normal Nopolu procedure is estimating groundwater to river from observation.

The set of provisional queries compute the global volume of exceeding water from groundwater to be transferred to compartments in surface water stated in water deficit. Then, a final query fills the table VR_resources_exchanges (this table is mentioned in Annex 1) in order to spread the monthly groundwater-exceeding volumes over the modelled period.

6 Suggestions for improvement

They are four categories of improvements related to reference system, stock follow-up, and surface-to-groundwater inputs and groundwater-to-surface outflows

6.1 Groundwater reference system

The lack of a reference system is a major gap in the accounts, as well as in the implementation of the WFD.

The developments carried out by the ETC/SIA in using the UNRSCO/BGR map where no substitute is available and populating it with the important information are almost archived, but were not delivered in enough time for the consultant to incorporate them.

This map, where it cannot be substituted by local synthetic map (as for example the BD LISA which is a more effective substitute to Unesco/BGR) should be compared to groundwater bodies information to identify, where possible, the category of aquifer at stake. This still constitutes a significant difficulty, since the latest deliveries of groundwater bodies (November 2012) contain many errors (ETC/SIA analysis).

The work is not expected to produce usable data before the end of first quarter 2013, when integration with ECRINS will have to be carried out.

6.2 Stock follow-up

The stock can be estimated only if the reference is populated enough and if level information on wells is documented with regularity (even though monthly measures are not required).

Better information, especially the depth of wells, is required to assess a reasonable value of the exploitable stock. Indeed in the case of groundwater, the resource is but the one that can be mobilised and hence is largely defined by the depths of drillings and not by the aquifer's properties only.

6.3 Surface to groundwater transfer

The SUGAR method has no practical concurrent for the time being; it is limited to areas where the indices has been produced and to those countries accepting dissemination.

In first step, the countries where data are not available should be contacted, within Eionet, to make these data available, possibly with control on further dissemination.

The integration of groundwater reference system to ECRINS will require improvement of the Nopolu calculation procedure.

Without a detailed modelling process between groundwater and surface water, we had to implement a simple solution for assessing exchanges between resources. In order to access the order of magnitude of transfer from groundwater to surface water, Nopolu was completed with a procedure based on the idea that over long periods, a kind of equilibrium should be reached for surface water. The procedure is that for each surface water compartment, the application calculates how much water could have been lost over several years, and calculates the ratio of each compartment over the total of surface water deficit. For each basin, the groundwater volume gain over the period is then apportioned with these ratios, in order to access how much water can be transferred for each compartment from groundwater.

Annex 7 Surface run-off data collection

1 Place of river discharge in the process

1.1 Detailed uses of river discharge data

River discharge data (noted as Q) are paramount in the water accounting process; they make many essential contributions:

- the river at the outlet is the only fully controlled information that gives the remaining resource after abstraction and uses;
- river discharge in the catchment is the lumped result of surface run-off and exchanges with groundwater, and hence is the touchstone of the accuracy of the rainfall/evapotranspiration balance (this is more complex in its details but is mentioned here as a principle),
- river discharge, thanks to the analysis of the depletion curves (not yet carried out in this implementation) is the most rational way to assess the groundwater-to-surface water inputs (and hence fills cell EA132 -> EA1313 in Table 6.2 (Figure 2.6, page 40));
- river discharge computed upstream and downstream of lakes and reservoirs is the most systematic way for estimating variation of volume (in the general absence of dedicated data).

Since most recordings of river discharge are actual observations and not reconstructed natural discharge, it can be stated that:

Q observed = (surface input + groundwater inputs+ rainfall on rivers+ inputs from reservoirs + returns)

-(abstractions + groundwater recharge + storage in reservoirs + evaporation from rivers).

The degree of accuracy of river discharges per segment and the total value at the outlets depend however on the space and time resolution of the values inserted into the accounts calculation.

This manifold role of Q data makes the river discharge to be considered as irreplaceable data in the water accounting process: where the data are lower than required in time and space resolution there is some risk of producing somewhat inaccurate accounts, and having limited capability of identifying the sources of uncertainty; this is why so much emphasis is placed on scoring river discharge data.

1.2 Categories and formatting of river discharge data

Since the water accounts are compiled at the monthly time resolution, monthly averages are the appropriate aggregate. The values of monthly river discharge are, by definition of the statistical unit, required for each and every river segment of the main drains in the sub-basins.

This requirement demands extending data collected at the gauging stations all along these segments ([81]). The making of depletion curves demands daily data since the changes in trend can be analysed only at the daily time pace. Considering that hydrographical services (that provide river

([81]) This term used to call all hydrological monitoring stations, being gauged or not, for simplification purposes; not all discharge stations are gauging stations that mean that water discharge is gauged regularly to derive a rating curve. Some discharge monitoring need not gauging, by construction. Gauging is required if the station is natural (not a notched weir for example), making the Q = f(h) relationship not straightforward.

discharge) usually don't deliver aggregates over periods in which even a single day is missing, all monthly aggregates have to be computed (with reconstructed gaps if required) from daily data.

2 *Placing and controlling the gauging stations on drains*

2.1 Checking delivered data sets

River run-off data are collected under two possible variables: water elevation ('H' : parameter) and discharge ('Q': parameter 1421 for daily, 11421 for monthly averages). Water elevation was not processed but it could be supplementary information in case gauging is abandoned at a site, and Q values are no longer computed. Some national Member States' web-based applications provide both pieces of information online.

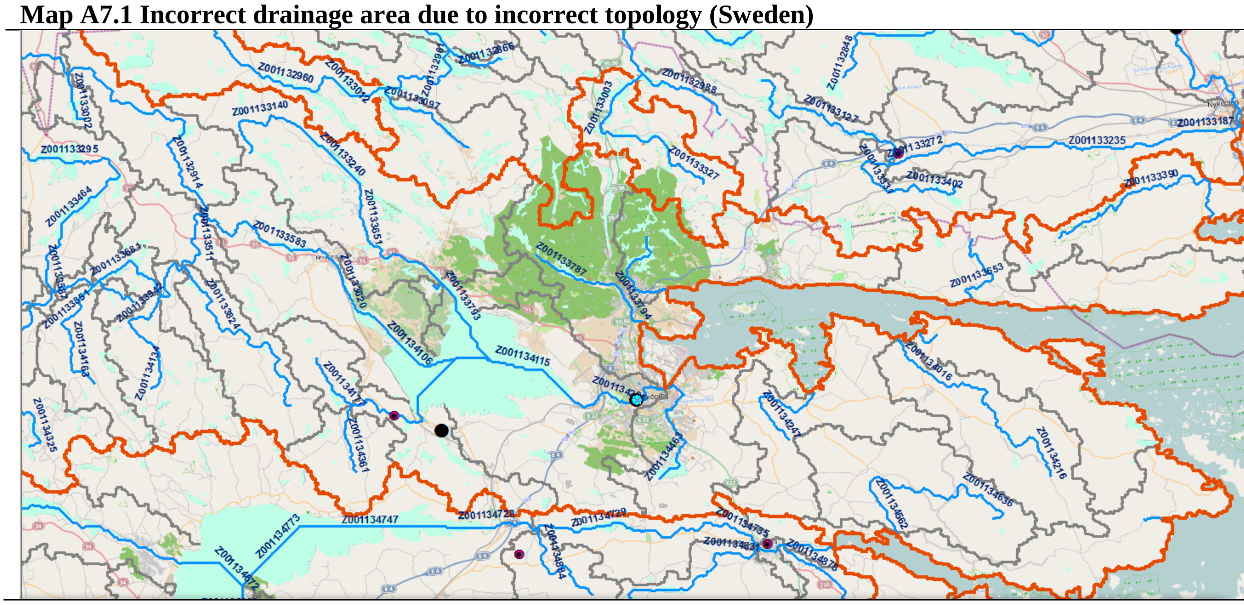
River discharge is fast-changing information that is continuously recorded. Records provide near real-time data, i.e. the instantaneous values are compiled over 10 minutes, 1 hour, etc.

The most widely report aggregate is the day, meaning that the instantaneous discharges are integrated over the 0:00:00 to 23:59:59 period. This daily integrated value is not a discharge but a volume which is expressed as its averaged discharge in m^3/s over the 86 400 seconds, the conventional duration of a standard day.

2.2 Gap-bridging and corrections

Most of the corrections concerning the drained area of gauging stations data set were made in Spain, where it seems that an incorrect corresponding field was imported for this country. We can assume this because there were several gauging stations where the drained area field was set to the same value, in fact corresponding to the area of the whole basin.

Map A7.1 Incorrect drainage area due to incorrect topology (Sweden)



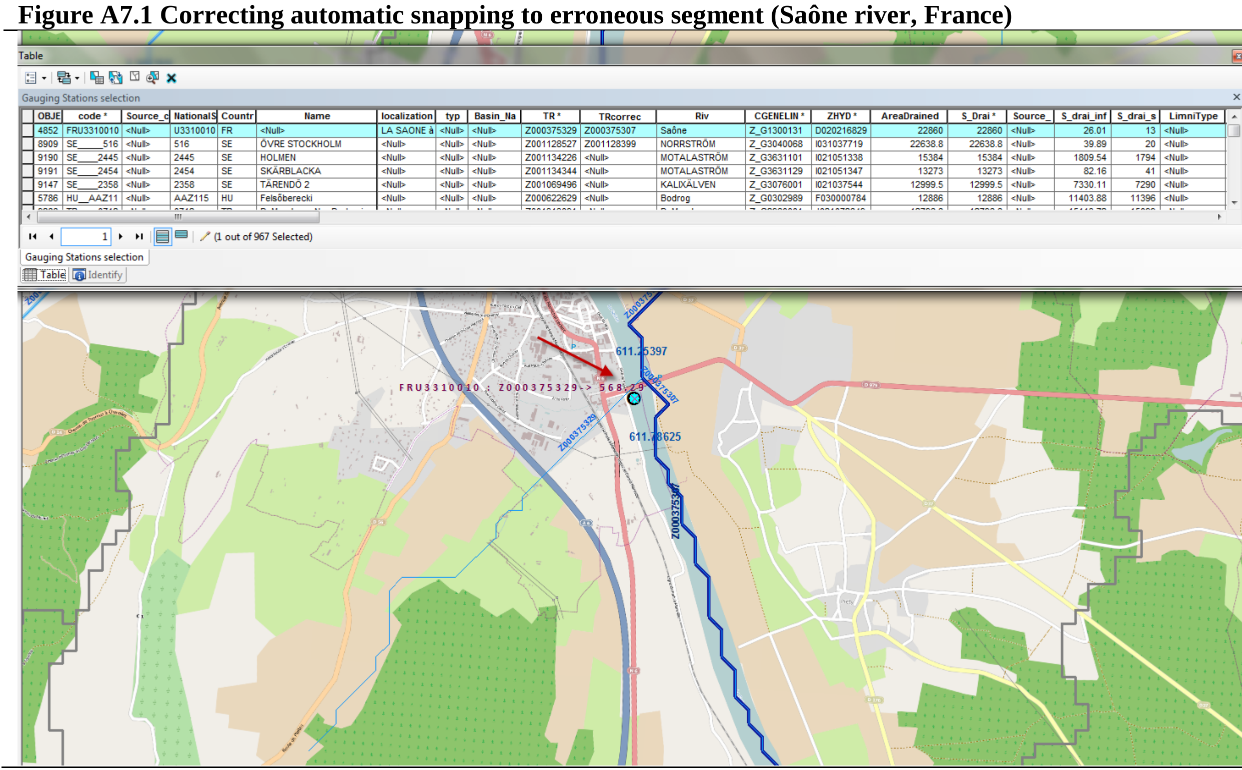
Source: Pöyry report.

The figure illustrates one of the corrections made for many locally significant placement errors or topological errors in ECRINS that render the catchment area incorrect, and hence result in erroneous productivity. The stations initially placed at the black spot had to be moved to the bright blue spot at the other outlet of the lake. Errors in gauging stations placement are critical both for the processing of data and for their use as information for reconstructing missing data.

Correcting inappropriate placement due to topological errors is also illustrated in the next figure (Figure A7.1); it also shows the need for the planned revision of ECRINS on the one hand, and the importance of full documentation of gauging stations with attribute data on the other hand.

The ECRINS main drain of the Saône river (France) is very close to the actual river; the station falls closer to a possible spurious affluent (segment Z000375329). Geoprocessing automatically provides a wrong link to gauging station FRU310010. The correction was assisted by the river name and the catchment area of the station. The Saône segment allowed correcting snapping to the ECRINS main river stretch Z000375307. Corrections are stored in the database under the file TRCorrec.

Figure A7.1 Correcting automatic snapping to erroneous segment (Saône river, France)



Source: Pöyry report to DG Environment.

3 *Data collection and organisation of data sets*

3.1 Data collection proper

The production of river discharge values is generally carried out by local and regional services. Making the relevant and significant gauging data demands an immediate response and availability from highly specialised experts so that relevant events are monitored. Local services always produce their estimates, while finalisation may require time and corrections. The fact that each gauging record contributes to the rating of the Q=f(h) makes Q values susceptible to having to be updated with time. This makes it necessary to regularly reload historical data. For example, in the event that the rating curve becomes incorrect (if the river profile has changed, for example, because of a flood), observed values have to be calibrated again.

Depending on the country's organisation and policy, final and assumed validated Q data are provided from the local producers, and centralised into a national data bank.

Data collection has been carried out in multiple ways.

1. The EEA had directly collated data from several countries during the previous steps of the water accounts development; the ETC/ICM (NTUA) has been requested to complete and update data sets.
2. A specific procedure has been set up, by informing Eionet that a huge data collection was envisaged for the sake of water accounts, especially requesting the following: i) comprehensive maps of gauging networks; and ii) access to daily data (the past 10 years at minimum) for selected stations (i.e. all those on main drains and all those on drains of the upstream and coastal FECs). The ETC/ICM has been requested to carry out data collating under this procedure. The success of the process is limited, since some countries have not even provided maps of gauging networks.
3. Continuation of specific direct contacts established by the EEA (e.g. France and Switzerland);
4. Extensive use of open websites ([82]) from which Q data can be downloaded (by group of stations/group of years, to one station/one year); this has been mainly carried out by Pöyry consultants.

The possibility to cooperate with the GRDC in order to share daily flows over European countries where the above scheme had not produced sufficient responses from Member States has been tried and unfortunately failed, the narrowest interpretation of WMO Members' restriction rules being applied by the GRDC. As a result not even the monthly statistics could be obtained.

To nevertheless take the maximum stock of existing data, places where access to data was denied (or just not provided) were completed as far as possible by doing the following.

1. Using and joining GRDC reference networks (stations) to the set of stations collected by the EEA.
2. Collecting and re-processing the modelled discharge values produced by the JRC, from GRDC data (to which JRC has access) and reconstructing data from this source. JRC modelling, however, provides values for cells and not for rivers. This point is discussed in Annex 8 , page 178.

All collected data sets at the required time resolution have been uploaded to the SQL database thanks to the application developed by the EEA.

3.2 Databases

In the water accounts, the final data are stored in the V_Quan ([83]) table, part of database 'Waterbase ([84])' on the EEA Bison ([85]) SQL server, because the size of the data set.

The V_Quan structure is as follows. For accounts computation purposes, a specific table for monthly average flow has been set up to make computations faster, and since some of the reconstructed data were directly estimated at the monthly step:

- station ID (ECRINS), controlled by C_Quan with references;
- date and time; convention for daily and monthly data;

([82]) Mainly http://snirh.pt; http://hercules.cedex.es/general/default.htm; http://vattenweb.smhi.se/; and http://www.hydro.eaufrance.fr/ (see online).
([83]) The table-naming syntax is 'C_' + suffix for characteristics, and 'V_'+ same suffix for attached values.
([84]) This is **not** the Waterbase as disseminated on the EEA data service.
([85]) At the time the data were processed and the report written; the name of the database inside the EEA may change with time and IT upgrades.

- parameter code, to indicate if data are Q (daily average), monthly average or h (water level). The parameter table manages the units, in this case m^3/s, which is not depending on the time aggregate (daily or monthly flows are systematically expressed as discharge rate and not as volumes);
- value. The unit is referenced in the parameter code;
- flags (Rem, Sup and Interp indicating respectively value status, if larger or smaller that the value reported (in case river actual flow was out of monitoring range) or if data had been modelled or extrapolated).

3.3 Scoring data deliveries and availability

River discharge is the essential information to capture the value of the water balance as well as the touchstone of the other components. Hence, scoring is an essential tool to analyse the quality of the proposed results. Scoring is based on a double categorisation:

1. Is the density of information adequate?
2. Is the quality and number of data per reported point satisfactory?

The first category is addressed by analysing the density of stations reported by all countries, considering only those that are on main drains (with limited exceptions since the stations not on main drains are not processed in general) and propose a target as density of station per 1 000 km of main drain within the scoring domain. Two analyses have been carried out, with the same principle: at country level (for simple discussion with countries and negotiating further deliveries) and at sub-basin level (for analysing the validity of computed balances). There are some differences between the approaches: if a country has a huge density of stations in a limited part of its territory, the country score can be excellent and the composite score from the sub-basins can be rather less. This is the case for Sweden, for example, where some sub-basins are very poorly monitored (probably because there are not any water issues to justify the expense of monitoring!).

To address this issue comparatively, the following calculations were carried out.

1. Compute the length of main drains (as ECRINS) per country/sub-basin ('entity'); main drains are the dummy rivers on which computations are carried out.
2. Analyse the number of station known per entity and sorted out per category of drain on which the station is snapped (disregarding their accurate positioning and possible errors — some stations may be assigned to the wrong drains category) and analyse as density per 1 000 km of main drain.
3. Propose a range of target stations. To this end, the number of stations between percentiles 75 % and 50 % across all entities has been computed. These values (as density) are used to estimate a reasonable range of number of stations per country, after rounding off. Percentiles have been computed over the whole ECRINS area and cover non-EEA countries as well as countries having not reported any stations, thus lowering the target number of stations.
4. If the number of stations known is larger than the target, no further station is required; otherwise, a supplementary number of stations should be collected.
5. Only those entities whose number of stations provided is less than the lowest threshold are considered as insufficiently documented under the criterion 'number of gauging stations'

Computed targets at the 50 % percentiles are 3.5 stations/1 000 km at country level and 4.5 stations/1 000 km at sub-basin level.

Many entities do match and largely overpass the minimum target. However, several entities have huge gaps, compared to the target: this is not necessarily abnormal (it may depend on the hydrological structure of rivers in the country). In most cases, this is just because country has not reported any information.

To address data delivery, three cases have been considered. Monthly data is the aggregate that is required for the accounts; this aggregate is preferably computed from daily averages. Three categories of data have been prepared and then computed.

1. Source data (daily data flows computed as monthly averages) (best scoring: 1).
2. Time-reconstructed monthly averages, when local gaps were identified. This reconstruction is carried out with recorded values at other stations correlated (see Annex 8 , page 178) with the stations' data to reconstruct (intermediate scoring : 0.75).
3. Space-reconstructed data. These data have been reconstructed with a probabilistic approach (see Annex 8 , page 178), to avoid any discharge modelling, using existing stations and modelled data provided by the JRC. This later method was applied where no data had been provided, and only targeted to documented gauging stations. No virtual station was created (low scoring: 0.25).

Scoring has been carried out based on how many data were expected (target no of stations * no of expected data per station) versus currently obtained no of data, sorted per quality of data.

- The target number of data is 10 years *12 months * target number of stations (low/high).
- This target is compared to the current number of data; the number of data is scored according to its quality. The scores per category are capped to 1 (if more data that the general target are present, it does not count for more). This method tends to lower the demand, since the target number of stations is lowered by the countries having provided minimum information on their gauging network.
- Scoring is computed in two runs, because of the very large number of reconstructed data in some cases, despite the provision of source data meeting the targets. Indeed, if the source data are already enough, considering the reconstructed data (having a lesser score) in the same run jeopardises the entity where both conditions of many stations and many data are met. The first run is hence a computed score resulting from raw and time-reconstructed (normal gap bridging), which then completes the range 1- first score with the space-reconstructed data. If no data are provided, the whole score is fully driven by space-reconstructed data.

It is important to note that only existing gauging stations are taken into account. As mentioned above, these stations can be delivered or found in other sources, as GRDC. No dummy stations were created in the reconstruction process. This explains the very low score of many entities.

The map of river discharge scoring per sub-basin is reported in Subsection 4.2.5 , page 69.

The summary of data (as month-station) eventually collected, extrapolated and compared to targets (per country) is reported in

Table A7.1.

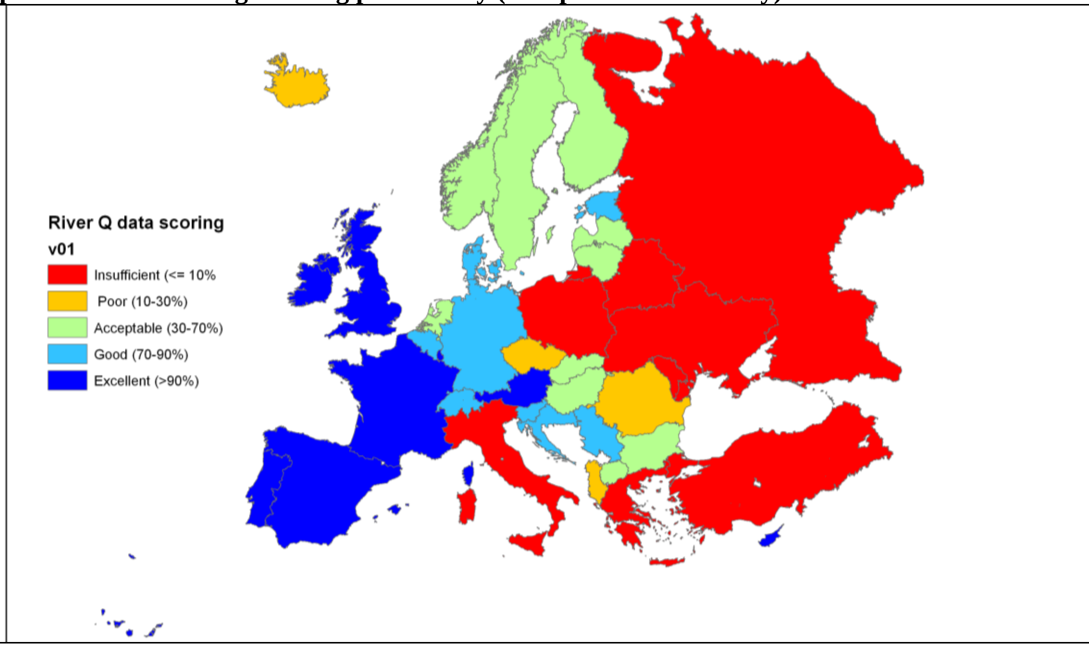
Table A7.1 Summary of month-station delivered, reconstructed per country

Country	Score (50 %)	Score (75 %)	Total month-station	Source no of months	Reconstruction (months) Time	Space	Target low in M-S	target high in M-S	Missing low in M-S	Missing high in M-S
AL	0.17	0.07	2096	0	0	2 096	3 120	8 040	1 024	5 944
AT	0.96	0.96	247 929	210 927	36 740	262	8 280	21 480	0	0
BE	0.80	0.79	15 228	6 696	6 960	1 572	2 160	5 400	0	0
BG	0.31	0.12	6 308	4 212	0	2 096	9 600	24 720	3 292	18 412
BY	0.08	0.03	4 978	0	0	4 978	14 760	38 280	9 782	33 302
CH	0.74	0.74	71 662	756	70 644	262	3 840	9 840	0	0
CY	1.00	0.84	1 512	1 512	0	0	720	1 800	0	288
CZ	0.16	0.06	1 392	0	1 392	0	6 600	17 040	5 208	15 648
DE	0.72	0.65	125 556	19 224	101 616	4 716	27 000	70 080	0	0
DK	0.88	0.88	51 433	30 084	19 777	1 572	1 920	5 040	0	0
EE	0.79	0.67	16 486	3 348	12 876	262	3 000	7 920	0	0
ES	0.98	0.98	510 402	497 841	509	12 052	42 360	109 920	0	0
FI	0.59	0.27	37 816	6 480	27 144	4 192	23 040	59 640	0	21 824
FR	0.98	0.95	845 331	787 259	49 426	8 646	44 520	115 320	0	0
GR	0.08	0.03	3 084	1 512	0	1 572	12 000	31 080	8 916	27 996
HR	0.88	0.46	13 308	6 696	6 612	0	4 920	12 600	0	0
HU	0.57	0.22	9 284	3 888	4 872	524	6 240	16 200	0	6 916
IE	0.98	0.97	222 332	215 352	1 740	5 240	4 920	12 840	0	0
IS	0.18	0.07	4 008	864	0	3 144	4 560	11 760	552	7 752
IT	0.06	0.02	7 244	432	0	6 812	27 480	71 280	20 236	64 036
LI	0.75	0.75	348	0	348	0	120	120	0	0
LT	0.54	0.21	8 964	3 564	3 828	1 572	5 040	12 960	0	3 996
LU	1.00	0.60	432	432	0	0	240	720	0	288
LV	0.58	0.22	5 758	972	4524	262	4 920	12 840	0	7 082
MD	0.02	0.01	262	0	0	262	3 000	7 800	2 738	7 538
MK	0.69	0.30	3 108	324	2 784	0	2 400	6 360	0	3 252
NL	0.66	0.60	11 230	1 335	7 275	2 620	1 080	2 760	0	0
NO	0.49	0.25	70 208	7 452	32 364	30 392	23 520	60 840	0	0
PL	0.09	0.04	8 384	0	0	8 384	22 080	57 360	13 696	48 976
PT	0.95	0.94	113 112	105 952	348	6 812	7 200	18 600	0	0
RO	0.11	0.04	8 384	0	0	8 384	19 440	50 400	11 056	42 016
RS	0.83	0.60	16 996	11 772	4 176	1 048	5 760	14 880	0	0
RU	0.05	0.02	45 850	0	0	45 850	253 920	657 840	208 070	611 990
SE	0.65	0.25	55 894	26 568	27 492	1 834	3 5160	91 200	0	35 306
SI	0.72	0.69	13 522	1 080	12 180	262	1920	5 040	0	0
SK	0.55	0.21	4 560	1 080	3 480	0	4 080	10 680	0	6 120
TR	0.01	0.00	2 932	1 188	696	1 048	55 440	143 760	52 508	140 828
UA	0.02	0.01	3 406	0	0	3 406	40 800	105 600	37 394	102 194
UK	0.95	0.95	119 577	112 726	5 279	1 572	17 280	44 880	0	0

Source: EEA calculations from Pöyry-delivered results.

Mapping shows geographical differences.

Map A7.2 River discharge scoring per country (to express data delivery)



Source: EEA computation, based on median (least demanding) data delivery.

Differences in scoring between country level (all stations lumped together) and scoring by sub-basin (reporting territories of reference, only those stations inside the basin) are very instructive on two issues:

1. poor country scores requires strong action next to the country's NRC to obtain better access to the station;
2. a large discrepancy between country score and sun-basin scores questions the relevance of sub-basin delineation and the way data is being reconstructed to carry out accounting.

Major cases of discrepancy are those where the country is apportioned in many sub-basins, for geographical reasons that oblige making such apportionment and if many of these are not monitored. This issue has been raised but no operational solution has been found yet.

Methods for bridging gaps are shown in Annex 8 , page 178 and thereafter.

4 *Computing linearised data*

4.1 Purpose

The purpose of river discharge linearisation is to produce reference information. It is not hydrological modelling. The imposed constraint to linearising is the use of only those data that are directly accessible: reference discharge at gauging stations (daily data are not required, they are just needed to compute reference discharge) and drained areas at these stations.

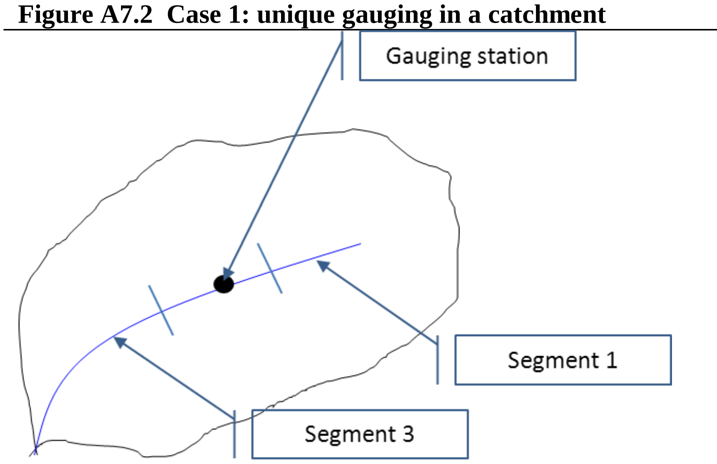
Many possible cases are envisaged, depending on the respective placement of gauging stations and drains within the calculation catchment. The target is to populate each river segment (there can be many within an elementary catchment), either for water asset accounts or as a source of the kmcn (normalised kilometre, for quality accounting). These cases are discussed in the following figures (A 7.2 to A 7.6).

4.2 Analysis of cases

The gauging station is on the drain

The discharge downstream of the elementary catchment (coded ZH understood as FEC in the application) is computed from discharge at the gauging stations, assuming that the specific discharge is constant (where production of water is proportional to the drained area) within the elementary catchment.

Figure A7.2 Case 1: unique gauging in a catchment



It is assumed that available data are:

- Drained area at the gauging station,
- Area of the elementary catchment ZH,
- If applicable the cumulated area of all ZH upstream the current one (SurfC)

Source: Pöyry documentation.

The discharge downstream the ZH is computed as being from the recorded discharge at the gauging station. The simple hypothesis of constant specific discharge is applied: discharge at the outlet is proportional (for one moment) to the productivity of the catchment, extrapolated to the outlet area.

In the real world, discharge varies continuously along the drain. Since the target of the calculation is merely to obtain a close proxy of the discharge at the end of each segment, the continuous discharge development function is replaced by a 'ladder-type' function, assuming a constant discharge within a segment. When possible, the estimate of this discharge is the most likely value in the middle of the segment that is understood as mimicking the average value of the discharge along the segment computed from linear approximation.

The method consists in first computing discharge Q_{S1} at the exit segment, based on drained areas, and using the known discharge at the gauging station.

Equation A7.1 Discharge end of elementary catchment (unique gauging station)

$$Q_{S1} = Q_{S2} * \frac{SURF\_tr1}{s\_drai}$$

In the next steps, the hydrological productivity is computed; it is noted as P.

Equation A7.2 Hydrological productivity at the gauging stations

$$P = \frac{Q_{S2}}{s\_drai}$$

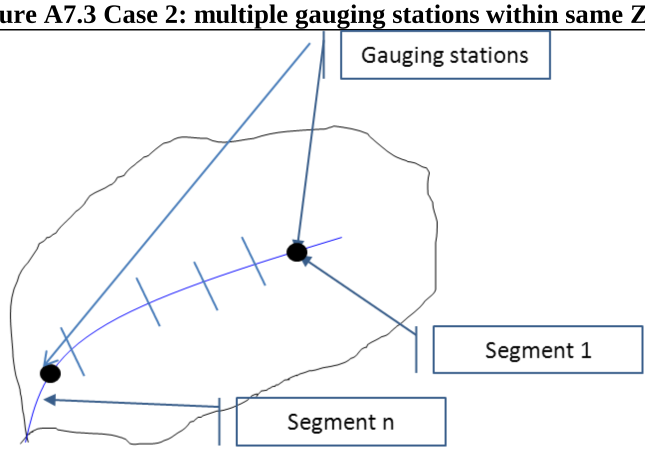
Assuming that ending segments Tr1 and Tr3 have known or recomputed discharges, all segments between can be attributed proxy values, computed from the length of segments. The length is used as a proxy of the drained area (within a FEC, more than the elementary areas composing a FEC are unknown). Segment length is a feature of ECRINS, and hence the distance can be computed. Discharge is reported for the middle, and this is the reason why the half sum of lengths is in Equation A7.3.

Equation A7.3 Discharge per segment (unique gauging in a catchment)

$$d = \frac{lg\_tr1 + lg\_trn}{2} + \sum_{i=2}^{i=n-1} lg\_tri$$

In the second case, discharge linearisation shall be carried out between segments tr1 and trn (where gauging stations are).

Figure A7.3 Case 2: multiple gauging stations within same ZH



The discharge in any segment between Segment_1 and Segment _n is expressed as a function of distance in relation to Tr1, since the sub-elementary areas are unknown inside a ZH (same as above).

Source: Pöyry documentation.

Equation A7.4 Discharge per segment (if more than one station within a ZH)

$$Q\_trx = Q_{S1} + \frac{Q_{S2} - Q_{S1}}{d} * lg\, x$$

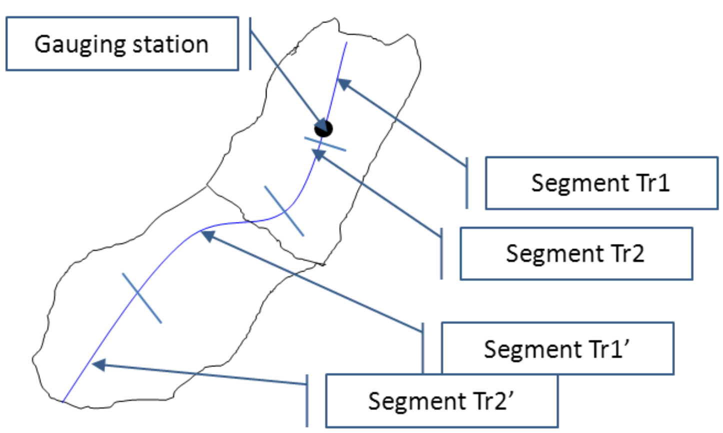
Source: Pöyry documentation.

Equation A7.4 makes explicit the rule illustrated by Figure A7.3 above.

Gauging station drives several ZH

Of course, there is not a station (or even several) in each and every elementary catchment. Discharge values have hence to be propagated to upstream or downstream segments.

Figure A7.4 Case 3: propagating discharge to the downstream segments inside downstream ZH



The discharge of segment Tr1 is computed as in Case 1, and Trn is expressed as a function of distance in relation to Tr1, since the sub-elementary areas are unknown inside a ZH (same as above).

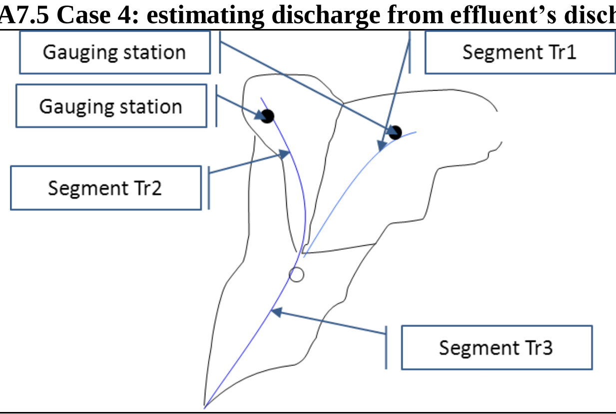
Source: Pöyry documentation

The example in Figure A7.4 illustrates all cases where the linearisation process could be applied. As a FEC in ECRINS may contain several stretches, the Nopolu System2 calculation has been upgraded so that interpolation can be carried out at the stretch level inside a FEC, since the intermediates catchment areas are available.

Several gauging stations and confluence

In the case of confluence, discharge is known or extrapolated for the two segments making confluence. Having used methods presented above, discharges for segments Tr1 and Tr2 are ready. Discharge downstream is hence the sum of QS1 and QS2; QS3=QS1 + QS2. In reality, the situation is more complex, and the number of intermediate segments is quite large. All computations are a blend of the different cases that are handled by the application.

Figure A7.5 Case 4: estimating discharge from effluent's discharge



If one of the two branches has no gauging stations, productivity is first computed as from Equation A7.2, whose results are then applied to the appropriate segments, considering the respective drainage areas on each of the upstream branches.

Source: Pöyry documentation.

4.3 Special cases

For reasons of accuracy, no calculations are carried out where the main drains are insufficiently documented. This could be improved if the data reconstruction procedure allows such data provisioning (see Annex 8 , Section 7).

Some segments may host two gauging stations. Only one can be used in a stretch (albeit several can be used inside a single FEC, provided on different stretches). For example, on a large FEC on the Trieux basin (coastal in northern Brittany), there are 6 stretches whose drained areas increase from 285 upstream stretch (reference Q 14.62 m^3/s) to 446 downstream stretch (reference Q 24.26 m^3/s). The information on the drained area at stretch level has allowed extrapolating discharges proportionally to the area instead of using the proportion of stretch length that can be largely disconnected from the drained area inside a FEC.

4.4 Limitations

In the current version of both ECRINS and Nopolu System2, defluences are not considered. It is, however, possible to take them into account, at the expense of specific definition of the application, imposing the implementation of fictitious gauging stations.

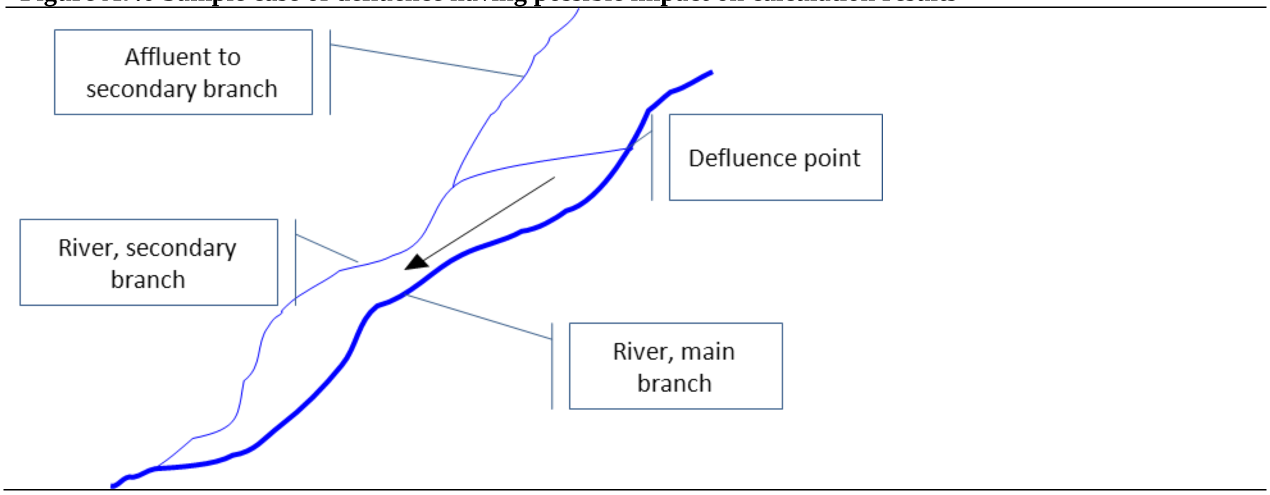
The issue of defluences is technically important in the following cases, and should be addressed later in the scheduled ECRINS v2, that will impose in turn some adjustments to the current water accounting implementation. These cases are described below.

- Big deltas, in which several branches empty into distinct areas. The key (yet unsolved) issue is to determine the true drainage area.

- Long islands in the river system. When the island is long, each river branch may receive affluent that modifies the water balance. This is an issue when the length of the island is significantly bigger than the length inside the elementary catchment.
- Diversions making canals that themselves receive affluent of diverted large shares of water.

Apart the delta issues that create errors in productivity (especially in gauging stations that are present on the different arms of the delta), the main problem relates to quality assessment, since one of the branches may receive significant pollution loads that are diluted in a part of the total discharge.

Figure A7.6 Sample case of defluence having possible impact on calculation results



The issue is all the more relevant for water quality accounting, and it has been analysed in this framework. Many suggestions and improvements were proposed, but these have not been implemented yet due to resources constraints.

4.5 Inputs from other resources

This specific part of calculation helps fuel line 4b of the SEEAW, 'inputs from other resources', to a large extent. The introduction into tables is dealt with in Annex 1 5.3

5 *Summary of obtained results*

5.1 Hydrometric stations and complementary stations snapped

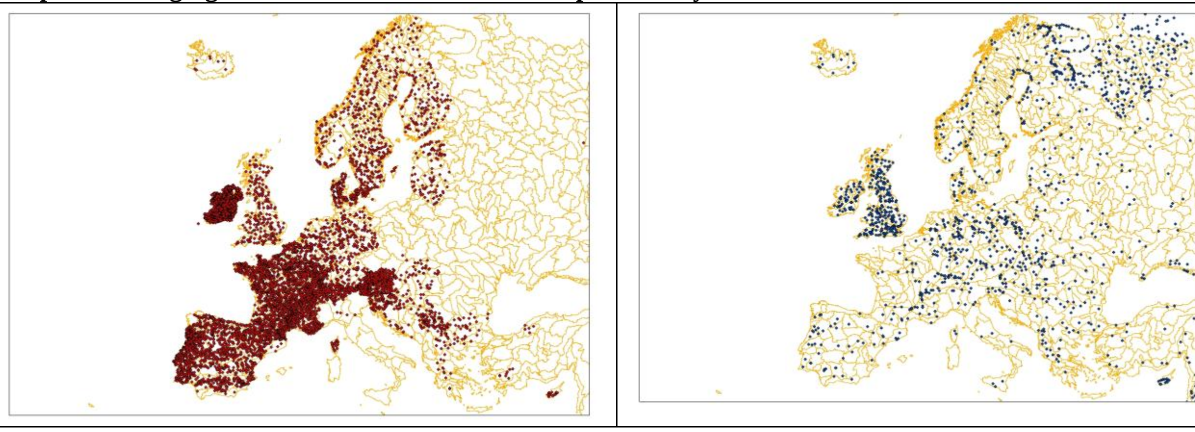
Hydrometric stations (nicknamed 'gauging stations') come from several sources. The first source, which is a blend of country deliveries at different stages of the process, constitute the baseline data set included in Waterbase and documented in table C_Quan. Since some countries have reported all stations, and others provided only a small sub-selection, the map of these stations has very varied density.

Two complementary sources of gauging stations were explored: the GRDC that manages 7 530 stations worldwide (1 406 in Europe) and the European Water Archive (EWA Euro-Friend, (formerly known as Friend-Ahmy)). Both were carried out by Pöyry in 2010, under an EEA contract.

The European Water Archive is a central feature of European FRIEND as part of the IHP of Unesco and the HWRP of WMO. Since its inception in 1985, this archive has been one of the most comprehensive hydrological archives in Europe, containing daily flow data and station metadata for more than 3 700 river gauging stations in around 30 countries. It unfortunately exclusively serves the scientific research of the members of the European FRIEND projects. Since 2004, the GRDC maintains EWA in support of the European FRIEND programme. EWA has 3 834 stations in Europe (including the European part of Russia, but excluding the southernmost ex-USSR Republics and Turkey).

These data sets are of importance for locating stations, and eventually for reconstructing data. However none of these organisations agreed to provide any information, daily or statistical, to the EEA. Since the EEA is not part of the FRIEND projects, it is not eligible for data provision under any circumstances.

Map A7.3 Gauging stations' source: EEA and complementary



Comment: left (reddish dots): source C_Quan base; right (blue dots): GRDC stations.

Source: Final Pöyry report for DG Environment, July 2012.

Similarly, the GRDC could not provide any data to the EEA, under strict interpretation of resolution 215 of the WMO ([86]). This restrictive position made it necessary to develop, from the gridded data modelled by the JRC, a rather complex process of data reconstruction that gave rather questionable results. This lack of data seriously jeopardised the quality of the water balances in poorly documented areas.

However, it is important to consider that GRDC data set could not replace the data collection for the sake of water accounts: it could only provide complementary data where the Eionet process failed. The density of GRDC stations serves global purposes of assessing final run-off and is hence rather low, much lower than required for carrying out water accounts at sub-basin level.

But the information related to gauging stations was usable, and it is presented in the maps.

These two maps are extremely instructive and indicate the following.

([86]) By personal communication dated 25 January 2012, the Director of GRDC responded to the EEA request in the following terms: 'As you know, the GRDC is operating under the auspices of the WMO and as such bound by WMO resolution 25 (Cg XIII-1999). Data provided to the GRDC by WMO members do have certain restrictions and as long as we have not been mandated by the individual WMO member states to supply "their" data to the EEA for the water accounts, we are not in a position to do this. The aggregation of daily data into a monthly value at station level is not considered by the GRDC (in recent consultation with the WMO) as a derived data product, so these monthly values are subject to the same restrictions as the original daily data. However, results from a gridded product, such as developed by the JRC can be used, as the gridded data have been derived via a specific method from the station based daily data.'

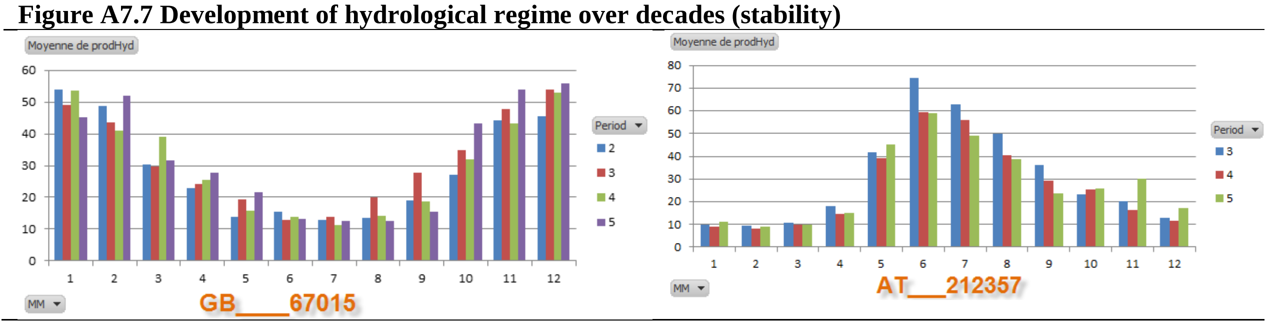
- The Eionet data collection perfectly fits (more than needed) western and northern Europe and the Swiss–Austria central Europe, plus some local patches; however, much of eastern and southern Europe, plus Turkey, is poorly documented.
- GRDC stations are more widespread, but are not regularly placed in the regions well documented by Eionet (except for Great Britain and Ireland) and some catchments in Poland. There are many commons stations with Eionet, but also many differences, that pose problems in snapping stations to rivers.
- If the current GRDC stations selection was the single source of discharge data, would its representativeness for Eionet-documented regions would be very poor and exacerbate the current situation? This is discussed and analysed in the comparative scoring, presented in Map A 7.4, section 6 of this Annex.

5.2 Data collected at stations and data checking

The data collected varied greatly in length of chronicles and number of stations from country to country. In many countries, data have been collected for much longer periods than that required for the water accounts (10 years). A statistical insight into the number of stations, data and range of years documented is provided in the main body of report (Table 4.1, page 67).

In some countries, data collection started as early as 1900. This allowed analysis of station productivity distributions along time, to check substantial changes. Several queries and tools (like Excel crosstab) were developed, in order to assess evolution of productivity per station over century (a period covering 25 years, beginning in 1900 and numbered from 1 to 5, from older to recent) and seasonally (monthly productivity) for sample gauging station as shown in Figures A7.7 and A7.8.

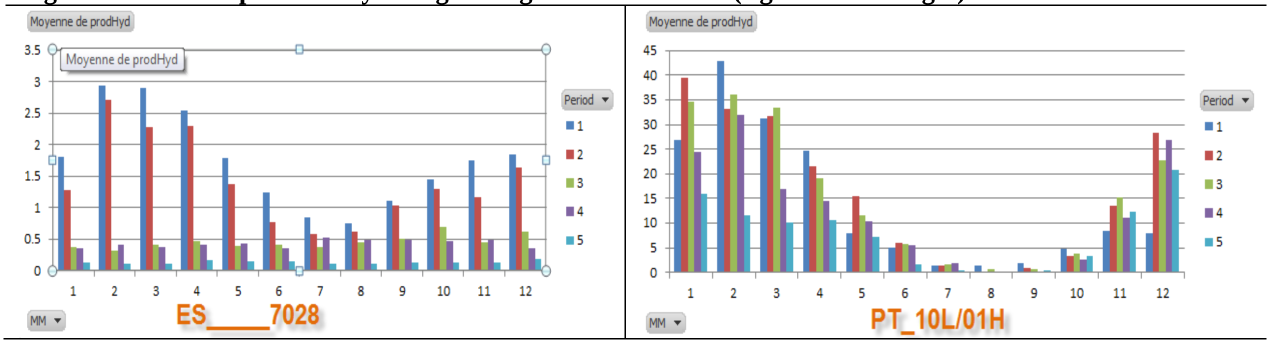
Figure A7.7 Development of hydrological regime over decades (stability)



Source: Final report Pöyry for DG Environment, July 2012.

In the two stations presented, one in Great Britain, other in Austria, the flow regime is radically different although the monthly productivity is rather constant with time. By contrast, Spanish and Portuguese stations, presented in Figure A7.8, show significant changes in productivity. In the Spanish station ES_7028, catchment productivity collapses in the last period, suggesting a dam construction and massive water transfer from the dam (the station hence presenting only base flow after withdrawals). Portuguese stations PT_10L/01H also suggest comparable changes over time.

Figure A7.8 Development of hydrological regime over decades (significant changes)



Source: Final report Pöyry for DG Environment, July 2012.

The examples presented support the remarks made by Spain concerning the water account results that, at least in Spain, a 10-year period is not enough to carry out significant water balances.

6 Suggestions for data collection improvement

6.1 Pros and cons of the current data collection scheme

The main inconvenience of current data collection scheme is managing a large spectrum of data sources, each with more or less different formats. This has resulted in some data errors, for example as relate to different units of reporting (litres/second versus m^3/second).

As there are many data sources, there are also many steps to update data sets. A special procedure has been developed to upload data into the table V_Quan, but the process is tedious and time-consuming.

The advantage is the range of freedom in deciding which data to upload, especially when free web access is granted.

6.2 Pros and cons of the replacement by GRDC data source

The GRDC manages two big data sets, as mentioned above: GRDC proper (under Resolution 25 of WMO) and AWA (under a joint mandate of WMO and Unesco). Current discussions between EU institutions (DG Environment, the JRC and the EEA) aim at broadening the scope of Resolution 25 and allowing EU institutions to freely obtain data for non-commercial purposes, and participate in GRDC governance.

The main advantage of this (envisaged since 2009, during the pilot implementation of the accounts at the EEA level) is that all data from the different providers would be pre-processed by a specialised institution which would carry out preliminary data quality checks (e.g. ensuring that no outliers are present). Moreover the GRDC has installed procedures for re-estimating monthly averages in case of missing data.

The main risk is in the very loose density of the current GRDC set of gauging stations (see Map A7.3 above) and their rather bad matching with rivers. Regarding the loose network, a scoring has been computed per sub-basin, assuming the very favourable hypothesis that all current GRDC stations are actually on main drains and 100 % populated with daily data (a very optimistic hypothesis).

Comparison of the global scoring related to water accounts collection on the one hand and current GRDC stations yields the following values.

Table A7.2 Comparison of global scoring of WA and GRDC data collection

Source	Score	Length covered (km)	No of sub-basins
WA data collection	0.4617	908 106	328
GRDC database	0.1831	803 788	275
Full domain	NA	1 024 409	412

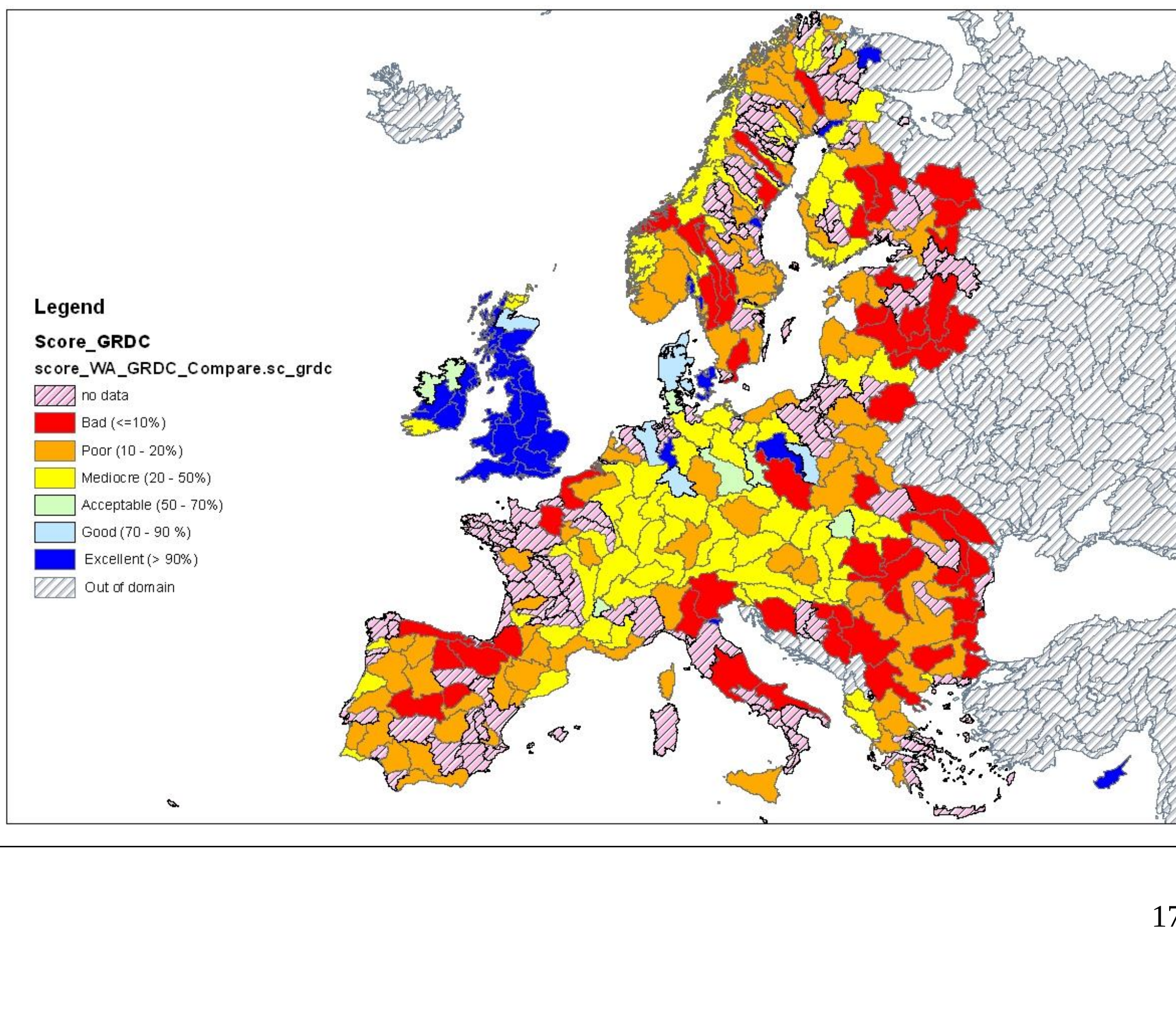
Source: EEA computations

This low scoring compared to current WA data collection indicates that the current stations' selection scheme applied by GRDC is not suited to water accounting. This means that many precautions should be taken in replacing the current (imperfect and rather not effective data collection process) by GRDC, since their current data collection scheme is targeted to 'global run-off assessments', that does not require as many stations as needed for WA purposes and focuses mainly of 'end of basin' selection methods.

On the other hand, there is an obvious complementarity between both data collection forms: in some countries, both are good (United Kingdom, Ireland). In others, the WA (France, Spain) or the GRDC (Eastern Europe) are better.

The current political discussions related to the role of the GRDC (WMO / EEA and country – GRDC discussions) are expected to provide more flexibility in the obtaining of data. However, since the number of data (mainly high density of stations) required by the accounts is much larger/ than the current collection procedure by the GRDC, its scheme should be deeply revised (at the European scale) before the GRDC could become the single source of river discharge for the EEA and the production of water accounts.

Map A7.4 Compared scoring between current data collection and GRDC hypothetical data collection 3



Annex 8 Bridging surface run-off data gaps

1 Rationales

1.1 Typology of errors

Collected data sets present several categories of errors.

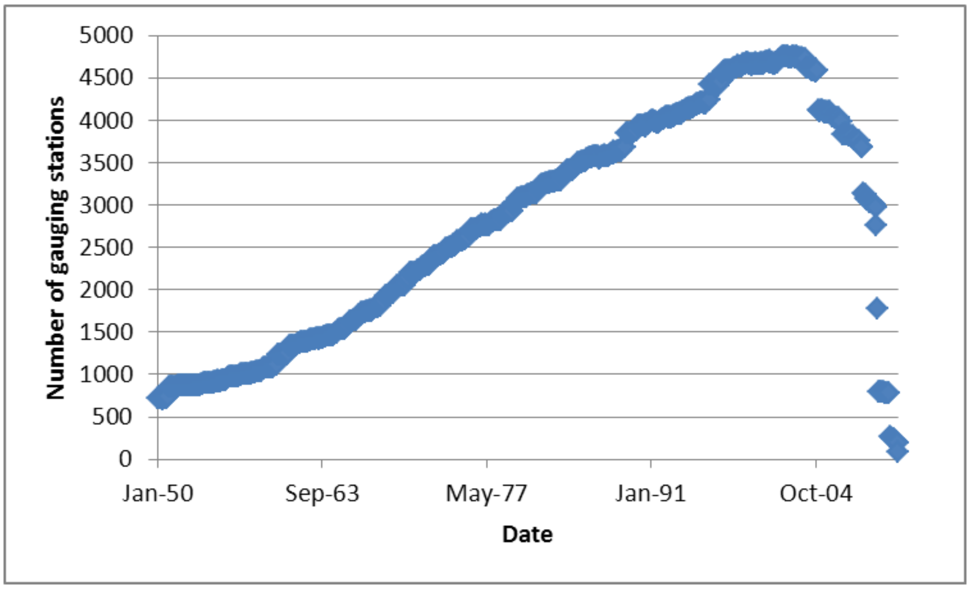
1. Data errors proper: the value is wrong. The erroneous data are mainly identified through outlier identification. The local acceptation of 'outlier' ([87]) is any data or set of data whose value is very unlikely to represent the true value. The presence of outliers may result from many reasons, most being data transmission of interpretation errors. For instance, some data were delivered in litre/s and entered as m^3/s because the entering procedure could not read the information (processed by outlier detection).
2. Time gaps: there are missing data in a certain station between date 1 and date 2. For example, the sensor had an error, the paper record (in case of H is registered by paper, historical data are dealt with) was damaged, or the pen ran out of ink, etc. (processed by time reconstruction).
3. There are no data at all for the stations. (processed by EPH reconstruction).
4. Data are correct but are analysed as erroneous because the information attached to the gauging station is wrong and creates abnormal productivities. This case is not addressed in this annex.

1.2 Size of the work remaining

Water discharge data that are stored in the V_Quan data set range between 1900 and 2011 deal with 5 711 stations of C_Quan data set. Only 61 % stations of C_Quan data set have data in V_Quan data set. The number of gauging stations having at least one value in the year per date is given in Figure A 8.1

[87] This extensive definition of 'outlier' does not follow the statistical definition of an outlier as any value which x time the standard deviation away from the average. Such a definition is not meaningful in this case.

Figure A8.1 Number of gauging stations from C_quan having data per date



This figure delivers the important message that there is a substantial number of stations with data extending backwards in time since ~1980, and that the recent years are quite poorly documented, with the acme being around ~2005, so the past 5 years are rather poorly documented.

Moreover, the number of stations documented in the areas with a small number of stations raises concerns for a need for large spatial reconstruction. These two characteristics, coupled to the need to carry out the discharge analysis at the sub-basin level, explain the very patchy scoring reported on Map 4.1, page 70.

2 Detecting outliers

If erroneous data are not checked and corrected or removed, they will be used as reference data in the data reconstruction step, and then the final data set will not be reliable. A systematic outlier's detection has been carried out by SCM, at the opportunity of the DG Environment-driven work.

This quite fuzzy acceptation of 'outlier' in this report is based on a set of criteria that were systematically applied. Under this assumption , when parts of data within a set of data for a certain station are considered as outliers, then the full set of data for the station is rejected as a whole because of possible systematic outlier data (until otherwise documented). The categories of candidate outliers that can be identified with reasonable likelihood are:

1. repeated data, when exactly the same value is found for more than X consecutive dates; this situation makes raises suspicions of processing/transmission errors;
2. data that have singular or unlikely values, and place them out of a likely trend (variation); this is the closest to the classical definition of outlier;
3. stations/data whose discharge values are likely too high, based on productivity criteria (helps in detecting primarily unit errors);
4. stations whose drainage area is invalid, jeopardising the computation of specific discharge (=productivity).

Criteria for excessive apparent productivity ([88]) are defined from the maximum productivity (MP) (as largest of monthly productivity) and should meet the following criteria:

- Drained area < =1 000 km^2 and MP> 1 000 l/s/km^2;
- Drained area > 1 000 and <=10 00 km^2 and MP> 600 l/s/km^2;
- Drained > 10 000 and <=50 000 km^2 and MP> 400 l/s/km^2;
- Drained area > 50 000 km^2 and MP>250 l/s/km^2.

These criteria do not claim to have hydrological relevance: they are warnings for likely errors in units. Only the first criterion was verified for 30 stations; in some cases the error is related to an incorrect drainage area because the station is on a defluence which is not documented with the proper drainage (this issue remains unresolved; it requires currently unfeasible structural improvements). In some other cases, the excess productivity may relate to the exceptional event on very small catchments (e.g. 1 400 l/s/km^2 for 41 km^2). Some stations were hence removed because of excessive apparent productivity.

Criteria for drainage are defined from a quick statistical analysis on the possible errors in drainage areas, that suggested the following criteria that were applied by checking the range of the station. Any station snapped to a FEC has a lower ECRINS area (all FECs upstream cumulated) and higher area (upstream cumulated plus the current FEC). This analysis primarily checks stations misplacement or incorrect documentation by comparing the declared drained area (in C_quan) to the found ECRINS drainage, resulting from snapping to the river system. The calculation uses the simple Equation A8.1.

Equation A8.1 Likelihood of drainage area

Lowest ECRINS area –X% <= Drained Area<= Upper ECRINS Area+ X%

- Where:
- X= 20% if Drained Area <=1 000 km^2
- X= 10% if Drained Area >1 000 km^2 and <=10 000 km^2
- X= 5% if Drained Area >1 000 km^2 and <=50 000 km^2
- X= 1.5% otherwise

A set of 2 230 C_Quan stations have been eventually set aside (Table A8.1 gives some examples of stations that do not verify previous exclusion criteria).

Table A8.1 Sample stations that do not verify the area criteria

Station code	Country code	Name	River name	Drained are a(as in C_quan)	Downstream area (FEC of attachment)	Upstream area (FEC of attachment)
HR____3211	HR	Županja stepenica	SAVA	340	66 373	66 363.95
FRK4180010	FR	La Loire à Gien		35 914.5	15	0.09
PT_07M/02A	PT	VALEIRA (EDP)	DOURO	85 641.4		
SI____2010	SI	HE Dravograd		1 082	12 630	12 629.74
ES____8047	ES	HUERTO DE GOIG	JUCAR	17 986	21 555	21 529.58

Source: SCM for Pöyry.

([88])'Excessive apparent productivity' because the computed data do not preclude any hydrological phenomena. Calculation just mentions that productivity=discharge/area (hence expressed in l/s/km2) is out of reasonable limits, putting suspicion of error on the value of discharge wrong unit of reporting?, missing decimal point?) or on the value of the area (error?, ill-snapped gauging station?) or both.

In most cases, the station area mismatches with ECRINS. For example, the two first in the list are indeed misplaced: the first is placed on a large river receiving the small affluent where the station is and the second first exhibits the opposite error, the station actually on the large river is snapped to a small affluent. Checking the correct value is important since in many cases the drainage area reported for the station is the basin area and not the area monitored by the station.

These examples are given to emphasise that quality assurance of data must be contextual. The sources of errors can be many: erroneous areas documented for the station, wrong units for the data or incorrect coordinates, all eventually generating and incorrect geometry of ECRINS. In all circumstances, the contextual analysis detects a large share of errors and helps find the appropriate action.

The rejected stations could not be processed, since they do not meet inclusion criteria; however once the corrections are applied, their data could be reused in further calculations.

3 Completing time-series

3.1 Rationales

River discharge series in V_Quan database are incomplete and present time gaps. Time gaps can be reconstructed assuming that other time series exist for the missing range of data and that common time spans are correlated; then it is possible to complete missing data.

Reconstruction is carried out using a probabilistic method of robust data reconstructions based on the conditional expectations that are computed for each case. This approach in two steps eliminates the questioning related to 'which statistical law applies to this or that station', since the condition law is just built from existing data for the pair (or sets of stations) involved. The details of the method and their theoretical justifications are discussed in the reference literature (Beauzamy and Zeydina, 2007).

The reconstruction process comprises the following steps

1. Data from stations are compared by pairs in order to identify the dates when the both stations have at least one value.
2. Computation of the correlation coefficient between stations in order to check a minimum likelihood.
3. Establishing the conditional probability law between the flow rates of the target station and those of the reference station. This law expresses the flow of the target stations as a function of flow rates of the reference stations. The expectations of this law allow filling gaps of the target stations.

3.2 First step: correlation between two data series

The method conventionally used to determine the correlation between two data series (X= x_1, x_2, … x_n and Y=y_1, y_2, … y_n), based on a linear relationship (which helps defining the likelihood of station X being possibility surrogated by Station Y) is computed with:

$$\boldsymbol{X = x_1 \ldots x_i \ldots x_n Y = y_1 \ldots y_i \ldots y_n}$$

Equation A8.2 Linear correlation coefficient applied to series with different numbers

$$\rho(X,Y) = \frac{\sum_i x_i y_i}{\sqrt{\sum_i x_i^2}\ \sqrt{\sum_i x_i^2}} \rho_X(Y) = \frac{\sum_{i\in D(X)\cap D(Y)} x_i y_i}{\sqrt{\sum_{i\in D(X)} x_i^2}\ \sqrt{\sum_{i\in D(X)\cap D(Y)} y_i^2}}$$

This formula applies to samples with different sizes (which is the case precisely because of missing data). The station best correlated is used to reconstruct the data for the target station; if there are still missing data, the second best correlated stations is used, and so on. A correlation coefficient threshold could be defined, which includes in the reconstitution only the best correlated stations.

This coefficient, especially built for the purpose of reconstructing missing data (demonstrated in (Beauzamy and Zeydina (2007)) has no value for computing missing values and is used only for sorting out the respective contributions of the candidates for source data for replacement. It is indeed rather unlikely that the relationship between Station X and all candidates Station Y1, …, Yn would be linear.

3.3 Second step: conditional probability law computation

To rebuild data of station X with data from station Y, the conditional probability law for X, knowing Y is used. The procedure requires dividing the range of data into smaller ranges, hence defining 'cutting' values that are borders of the intervals. The Minimal Number of Distinct Values (NMVD) (known as *Nombre Minimal de Valeurs Distinctes* in French) method is used. This method consists in dividing the value range in intervals each having the same number of separate data. It may be one or more separate data intervals. This number is improved by simulating reconstructions of known data and comparing the corresponding quality indices.

The successive sub-steps of the method are:

- calculating the optimal NMVD;
- constructing intervals for cutting;
- calculating the conditional expectation of knowing that belongs to a given interval;
- reconstructing missing data from conditional expectations.

3.4 Exemplifying application of the method

A simple example composed of 5 stations and 20 dates is used for illustrating the process. This is composed of a selected set of data to make the process as instructive as possible. This data set presents missing values, marked by 'NaN' (Not a Number).

Table A8.2 Sample data set used to illustrate reconstruction methodology

DATE	Station 1	Station 2	Station 3	Station 4	Station 5
D1	2	3,3	2	2,5	2,1
D2	3	NaN	3,1	3,325	3,09
D3	4	NaN	NaN	1	0,3
D4	2,3	3,75	2	2,5	2,1
D5	4,2	6,6	4	4	NaN
D6	6,2	9,6	6	5,5	5,7
D7	1,56	2,64	2	2,5	2,1
D8	6,3	9,75	3	3,25	3
D9	4,3	6,75	4	4	3,9
D10	NaN	NaN	3	3,25	3
D11	NaN	6,3	3,9	3,925	3,81
D12	6,2	9,6	6,1	NaN	5,79
D13	2,6	4,2	3	NaN	3
D14	5,4	8,4	6	NaN	5,7
D15	2	3,3	2,3	2,725	2,37

Table A8.3 Correlation coefficients $\rho_X(Y)$ for the sample data set

Station	S. 1	S. 2	S. 3	S. 4	S. 5
Station 1		0.9891	0.8304	0.4761	0.6786
Station 2	0.9999		0.8286	0.6704	0.8272
Station 3	0.824	0.8184		0.7166	0.997
Station 4	0.5484	0.6135	0.7818		0.9746
Station 5	0.6761	0.6891	0.8415	0.754	

Source: Reformatted from SCM for Pöyry, July 2012.

Comment: Station in lines correlates with station in column$\rho_X(Y)$

If accepting for example a threshold for inclusion of 0.6, all except two of the stations to station coefficients can be accepted, making it possible to select the stations taken for reconstruction.

Table A8.4 Ranking of correlation coefficients $\rho_X(Y)$ for the sample data set

Station as X	Best correlated stations			
Station 1	2	3	5	(4)
Station 2	1	3	5	4
Station 3	5	1	2	4
Station 4	5	3	2	(1)
Station 5	3	4	2	1

This method allows maximising the number of reconstructed data, at the expense of possible larger uncertainty compared to reconstructing a gap with data coming from several stations at the same time.

3.5 Reconstruction of missing values

Stations missing data (marked by 'NaN' in Table A8.2) reconstruction starts, for example, by fuelling Station 1 with expectations from Station 2. In this case, the optimum is a NMVD equal to 1: this means that there is a distinct value for each interval. The interval bounds are as follows.

Table A8.5 Conditional expectations of Station 1 vs. Station 2 (sample data)

Conditional expectation of Station 1 ('X')	Intervals of records at Station 2 ('Y')
1.56	[2.64 ; 3.30[
2.00	[3.30 ; 3.75[
2.30	[3.75 ; 4.20[
2.60	**[4.20 ; 6.60[**
4.20	[6.60 ; 6.75[
4.30	[6.75 ; 8.40[
5.40	[8.40 ; 9.60[
6.20	[9.60 ; 9.75[
6.30	[9.75 ; +∞[

NB: the notation [v1; v2[means larger or equal to v1 and strictly smaller to v2

In Table A8.6, 1.56 is the expectation of X given that the value of Y is in the range [2.64, 3.30[. To reconstruct the date D10 and D11 of X, considering that Y value is 6.3 for D11, that ranges in interval [4.20, 6.60[, makes the value 2.60 being set to X for this date D11.

By contrast, for date D10, there is no data for station Y = 2, hence data for X cannot be reconstructed: the missing data for Station 1 must be estimated from the next best correlated stations, Station 3. In this case, the optimum is a NMVD equal to 1 as well. This means that there is a distinct value for each interval.

Table A8.7 Conditional expectations of Station 1 vs. Station 3 (sample data)

Conditional expectation of Station 1 ('X')	Intervals of records at Station 2 ('Y')
1.9533	[2.00 ; 2.30[
2.0000	[2.30 ; 3.00[
4.4500	**[3.00 ; 3.10[**
3.0000	[3.10 ; 3.90[
2.6000	[3.90 ; 4.00[
4.2500	[4.00 ; 6.00[
5.8000	[6.00 ; 6.10[
6.2000	[6.10 ; +∞[

For the date D10, Y value is 3. This value is in the range [3.00, 3.10 [.Therefore the value 4.45 is set into X for the date D10. Thus, step by step, the missing data on all stations are being reconstructed; sample results are given in Table A8.8.

3.6 Production of completed data

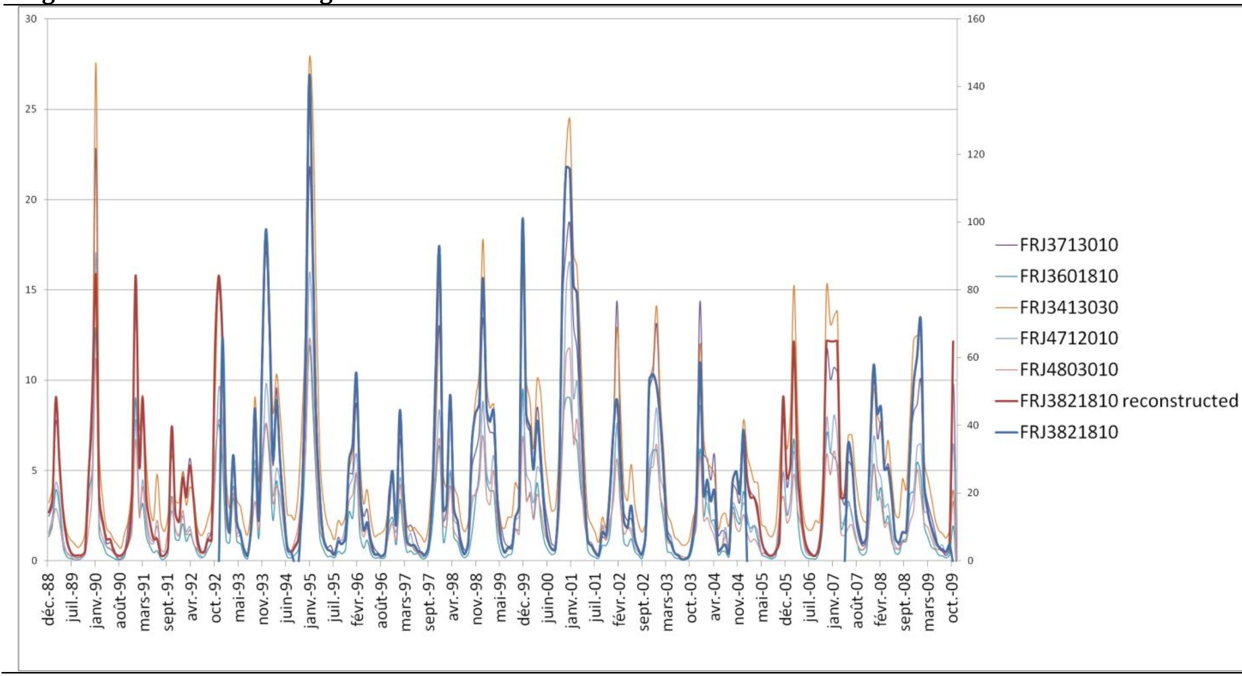
Take, for example, the station FRJ3821810. For the last 21 years, 149 months were missing, which corresponds to 59 % of the values.

The station best correlating with Station FRJ3821810 is station FRJ3713010 (Correlation index: 0.975), which is the one used for the reconstruction.

Table A8.8 Results of reconstruction of station FRJ3821810 from station FRJ3713010

Gauging station	Number of real measures	Number of data after reconstitution	Number of data still missing	Number of reconstructed data	Percentage of reconstructed data
FRJ3213010	108	251	1	143	99,3

Figure A8.2 Reconstructing FRJ3821810 is station FRJ3713010



Source: SCM for Pöyry, July 2012

4 Reconstructing spatial gaps

4.1 Reasons for spatially reconstructed data

The previous step of the work was to bridge time gaps when the gaps in the station where data should be reconstructed have reference values that can connect, for other periods, to reference stations documented for the time gaps at the target station.

When no data at all are present, the challenge of reconstructing data is much more difficult and cannot guarantee high-quality results. However, since water accounting deals with resources, and not with risks of flooding, for example (in other words, since the data to reconstruct is highly smoothed and aims at obtaining monthly averages), the challenge was taken on, provided that reconstruction of totally missing data would not eat into the total resources of the project.

The complete lack of data results from two different classes of reasons.

1. The provider does not accept providing data; this issue is matter to discussion and may possibly be solved once acceptation is granted. During the negotiations and until solved, missing data requires best possible reconstruction. In this case, priority should be given to reconstructing at existing gauging stations that are expected to be provided in a near future).
2. The area is not monitored, for any reason. In this case (which is typically the issue with some Swedish sub-catchments), reconstruction should be carried out in all circumstances, and at least should address the outlets in the system. In this case, dummy stations should be created, since there is reason for having historical references.

The best matching process that was selected is the probabilistic model called Experimental Probabilistic Hypersurface (EPH); this is conceptually radically different from the previous one. The idea of the EPH is to propagate the information, depending on the (generalised) distance between the available measurement points to any unknown one. The 'speed' of this propagation depends on one parameter which is tuned by the rule of maximal entropy.

The challenge supported by this method was all the more complex that in certain areas, the missing stations are surrounded by rather well-documented (even though partially time-bridged) stations. In most areas where spatial reconstruction is required, only estimates from gridded modelling could be obtained (from JRC models, these results are not restricted in use, by contrast with the source data that was used ([89])) and a poorly documented stations network is available.

The result of the reconstruction (propagation) is given in a form of a probability law where the value of the mathematical expectation is kept as a reconstructed value.

4.2 Rationales of the using of the EPH method

The EPH starts from the hypothesis that unknown information at point X is, with a certain probability, depending on information at points A_1, A_2, … A_n situated at respective distances d_1, d_2, d_n of X . If point X is very close to any of the A_i points, its value is close to the combined value recorded at all the A_i, hence depending to some extend (the closer weights the more) on the information captured on the other points.

In the EPH, 'distance' should be understood as a generalised distance in a n-dimensional space, the dimensions being the parameters taken into account (in this case, it can be longitude (W – E axis), altitude, longitude (expressing N –S axis, one dimension of climate), catchment area, etc.).

The key issue to fully understand is that EPH is a rather 'blind' method that cannot and does not aim at replacing complex rain-discharge models: it there to provide the cheapest possible surrogate information in the minimum time for several dozen non-documented stations.

The target is hence twofold:

1. provide surrogate information where, for any reason, data collection failed at providing values;
2. demonstrate the consequences of restriction on data accessibility on developing assessment for public policies.

Three major factors are expected to jeopardise the quality of data reconstruction.

- The bad quality of the reference data and its possible misplacement (many reference data are themselves outcomes of gridded modelling).
- Incorrect assessment of 'distance' between stations. The importance of this second point could be estimated incorrectly. This is addressed in the next paragraph where a simple demonstration of the method is given.
- Inappropriate evaluation of the value to reconstruct, which may result from insufficient documentation on the factors related to distance.

4.3 Selecting the value to surrogate

Based on the fundamentals, the most likely parameter to reconstruct is the specific discharge: at point X, it can be expected that specific discharge on this river is rather well explained by specific discharges at nearest points.

([89]) The reason being that the JRC is considered by some data providers as a 'research institute' (and hence eligible to accessing these data), while the EEA is not.

This rather logical consideration (that is contrasting with the parameter to reconstruct with the time gap reconstruction, where discharge is reconstructed) is breached by the lack of the indispensable data to carry this out: the catchment area. In many cases the station catchment is missing and hence the decision was taken by the consultant to attempt to directly reconstruct the discharge, unfortunately, providing quite erratic results.

Similarly, the parameters taken into account in the distance require more assessment and testing and possibly stricter selection before reconstruction. The experience of the systematic reconstruction of missing catchments becomes a stepwise adjustment for reconstructing poorly equipped catchment surrounded by well-monitored ones. This scope of use of the method is expected becoming the main niche for the implementation on this method in the years to come. Exemplifying the methodology

In this example, it is assumed that discharge can be directly computed (it is an example) only to demonstrate the 'mechanics' of the method.

N is the number of stations and *K* is the number of parameters describing each of the stations. Two stations are fully documented (*N*=2), and the three parameters are (latitude, longitude and altitude, presented in simplified values to make hand calculation possible). Data for each station are A_1={1, 2, 5} and A_2{3, 4, 6}, with respective values of 3.4 and 10.8 (units assumed m3/s, but this has no importance in the example). The values are noted $\xi_k^{(n)}$ for the k^{th} parameter of the n^{th} station.

The question is the reconstructing value for station X ={2, 3, 4}, supposedly having the same range of values, if they are observed.

To reconstruct the value, V(X) is the result of the calculation of the probability law, extending between two bounds V_{min} and V_{max}, that should be established from expert knowledge (these bounds are used to discretise the interval). In this case, the range V_{min} and V_{max} is taken [0.2 – 15].

In a general case, where K parameters are considered and N the number of stations, each source station is a point in a K dimensional space and the value taken at a certain moment is noted as $V(A_n)$, ..n=1,K, N by convention

Between the two bounds (V_{min} and $V_{max)}$, the interval is discretised, based for example on a chosen interval τ = 0.2 (this value is arbitrary and is related to the expected uncertainty on the reconstructed result and on the possible values of it). The role of the discretisation is very similar to the discretisation taken for time reconstruction.

In this case, the interval V_{min} and V_{max} is divided into ν elements, ν = (V_{min} and $V_{max)}$/ τ. The probability law for V(X) will therefore be given by *v* points in this interval that are noted by F_j.

Having the baseline elements, the distances are first normalised as $(x_k-x_{min})/(x_{max}-x_{min})$. After this normalisation, the values of parameters become:

$$A_1 = \{1, 2, 5\} \text{->} \{0.1, 0.2, 0.5\}$$

$$A_{2 =} \{3, 4, 6\} \text{->} \{0.3, 0.4, 0.6\}$$

$$X = \{2, 3, 4\} \text{->} \{0.2, 0.3, 0.4\}$$

The normalisation is justified by the fact that some ranges may be quite different: some parameters are between 0 and 0.1, whereas others are between 1 000 and 10 000, for instance. Performing this normalisation gives equal preliminary importance to all parameters. This option may be incorrect in some cases, but it poses the question of the justification for using different weighting.

The theory behind the EPH (Zeydina, 2011) requires computing a coefficient of propagation l which expresses the 'speed ' at which information propagates in the considered space. The expression for the parameter λ is the following:

$$\lambda = \frac{Log(\nu + 1)}{d_{max}}$$

Where ν is a width of the subdivision and d_{max} is the largest distance between the most distant station and the one of the space corners ([90]). This means that at this most distant point the probability law should be uniform at that place.

In the given example, the stations are points in the three-dimensional space. The distance between such points is the usual Euclidean one (this would apply to any n-dimensional space).

So, the distance between X and A_1 is equal to:

$$=d_1 = d(X, A_1) = \sqrt{\left(x_1 - \xi_1^{(1)}\right)^2 + \left(x_2 - \xi_2^{(1)}\right)^2 + \left(x_2 - \xi_3^{(1)}\right)^2} = 0.17$$

Application to point A_2 gives distance= 0.24

Now the step consists in building the discretised probability law in X, by computing the probability of each of the discrete values from the contribution of the N stations to the value in X, the general form is:

$$p_j(X) = p_j(x_1, x_2, x_3) = \gamma_1(X) \times p_{1,j}(X) + \cdots + \gamma_N(X) \times p_{N,j}(X)$$

In which $p_{n,j}(X)$ represents the contribution of the n^{th} station, to the unknown station X, for the j^{th} interval.

The $p_{n,j}(X)$ is given by the formula:

$$p_{n,j}(X) = \frac{c_n \tau}{\sigma_n \sqrt{2\pi}} \, .exp\left(-\frac{(V_j - V(A_n))^2}{2\sigma_n^2}\right)$$

In which $\sigma_n = \frac{\tau e^{\lambda d_n}}{\sqrt{2\pi e}}$

In these formulas, that represent the Gaussian distribution of probability around each source data set, for one value, the variables are:

- d_n (n= 1 … N) are the distances between X and A_n
- c_n is a normalisation coefficients so that the sum of $p_{n,j}$=1 for all n
- τ is the discretisation interval of the range of values, hence depending on the selection of range
- the coefficients $\gamma_n(X)$ have their sum equal to 1, each being of the form: $\gamma_n = \frac{d_n^{-K}}{\sum_1^N d_i^{-K}}$, K being the total number of parameters involved
- the propagation coefficient λ that depends on N and ν.

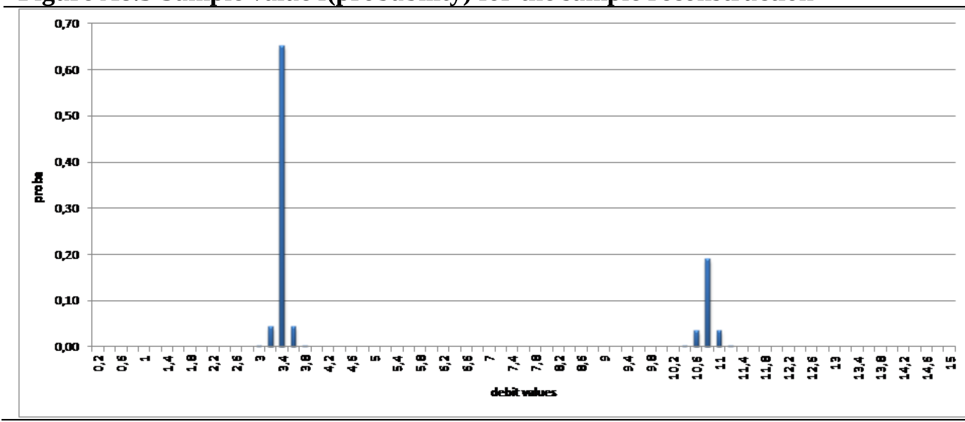
([90]) The 'space corner' is the point which is farthest from all measurement points. For instance, in dimension 3, after normalisation we have a cube, with some measurement points inside, and the biggest distance is to one of the corners.

The different values are known, and the calculation can be carried out for each searched value, with data related to all sources: if there are many sources, and the intervals narrow with large ranges, the calculations can be very long.

For example, the first coefficient $\gamma 1 = (0.17^{-3})/(0.17^{-3} + 0.24^{-3}) = 0.737792$, and the second is 0.262208, with the provided sample distances.

Once computed, the distribution of each value versus probability can be displayed (this is for instructional purposes) and the final values estimated from the sum of products of the probability attached to each of the values by each of the values.

Figure A8.3 Sample value f(probability) for the sample reconstruction



Source: SCM for Pöyry report.

The final result, $V(X) = \sum_{i=1}^{\tau} p_i . v_i$ gives in this case 5.33, because of the intermediate distances between A_1, A_2 and X.

5 *Critical analysis of the computed results during the project*

During the development of the project, hundreds of data reconstructions were carried out, using catchment areas (where available) as elements of the distance. It was the first time such an ambitious reconstruction attempt was carried out, and many problems arose. The lessons from the production efforts are very important, even though the outcomes are not optimum.

5.1 Choice of parameters

The chosen parameters must be independent from each other and have an important influence on outflow observations. Four parameters are used: drained area; two position parameters (latitude and longitude; and altitude, making a four-dimensional analysis space.

The quality of the source database is far from perfect, and it has evolved along the way. There are many erroneous values in relation with 'altitude' and 'drained area'. The parameter 'drained area' has many missing data. It means that the stations with unspecified 'drained area' cannot be taken into account. Due to this fact, not all missing stations can be reconstructed using the four parameters. In order to correct this situation, reconstruction was carried out in two steps: first, we consider only those stations having all four parameters and we reconstruct the flow values for stations with the four parameters as well (reconstructed values are marked in red). The same procedure is repeated for the remaining missing stations, but considering only the three parameters: latitude, longitude and altitude.

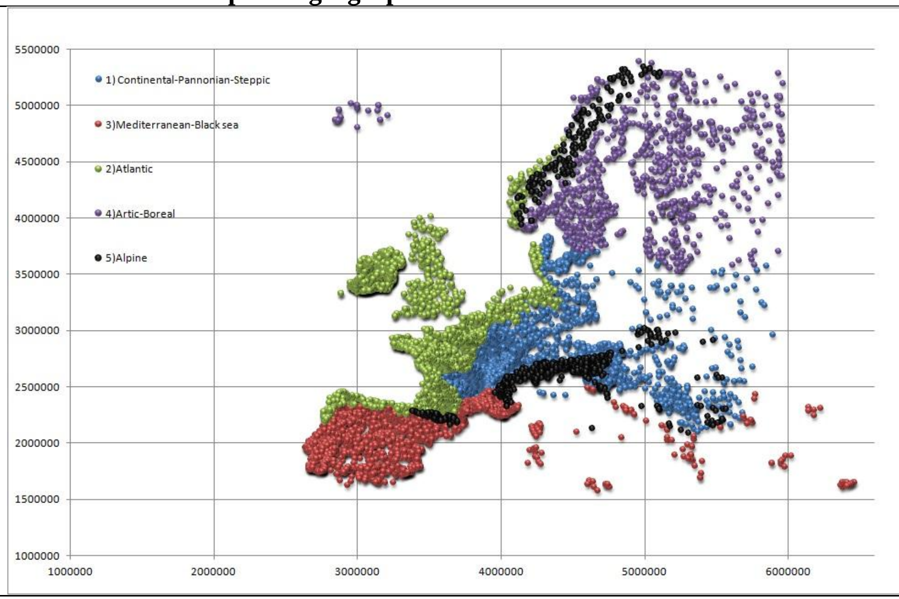
This seemingly logical procedure led in fact to an incorrect grasp of the issue and drove attempts to reconstruct discharge instead of specific discharge, which is not fully within reach of the method.

The filtering and results checking could not at the end prevent erratic data, and jeopardised the utility of the reconstruction.

5.2 Selecting the area scoping

The biogeographical region was used as selector for stations: only those stations located in a biogeographical region were used to reconstruct a station in the same region.

Map A8.1 Distribution of stations per biogeographical area



Source: SCM in Pöyry report

Further experience on forest versus water productivity suggested that this selection was not sufficient, and fostered complements in ECRINS to compute the share of the catchment upstream station apportioned by biogeographical region. For example, a station on the Rhone at Arles (Mediterranean) is in fact driven by Alpine conditions, and should not be considered as Mediterranean. This remark seems to stating the obvious, but the calculation of station dominance could not be carried out in time. Moreover, in certain situations, such filtering could be counter-productive (e.g. in Sweden where catchments are systematically shared between Alpine upstream and Boreal downstream).

5.3 Processing the distance and range referential

As mentioned in Section 0 above, the normalised distance is the key criterion to define the probability field. Because of the reasons invoked in relation with the reconstruction choices, the ranges of values taken were quite large, and yielded two opposing results: many source stations selected (potentially positive) and too large ranges (hence making distance poorly selective.

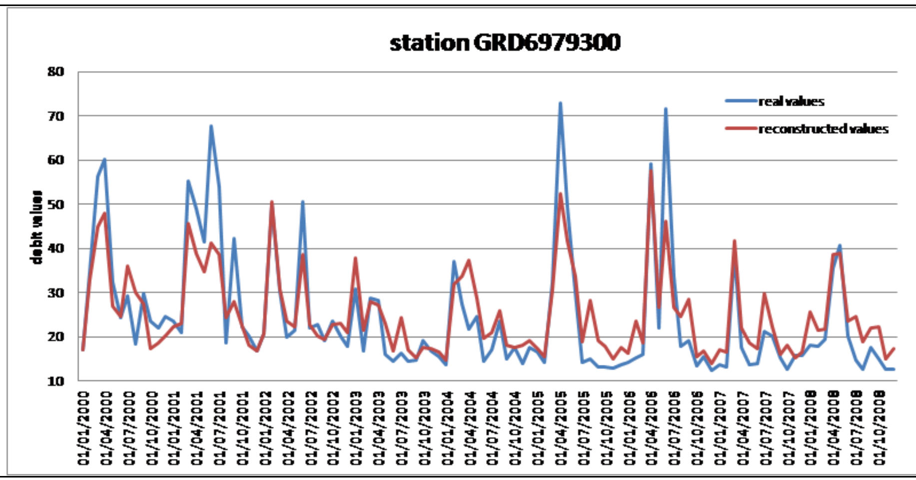
Using the catchment area as an important factor for distance was expected to be sufficiently discriminating and to avoid assigning similar water discharges to very distant (as range of area) stations. As tested in the development of the method, the methodology is quite efficient if a target is located near those sources, having similar parameters and complete information: in that case, the probability law is rather concentrated, rendering the information certain; otherwise, the probability law tends to uniform law, and the reconstruction is considered of bad quality (uncertain), and in the worst case, totally unpredictable. This allows assessing the baseline uncertainty (interpreted as likelihood for suitability).

The discretisation of values to reconstruct was taken rather large, by analysing the monthly distribution ranges.

5.4 Validating results

In order to check the quality of the reconstructed information, reconstructing already existing data and then comparing them with the really observed values was carried out, and gave mitigated results. First, station GRD6979300 which belongs to the Continental zone was reconstructed and values plotted against reconstructed discharges over 108 months (parameters: latitude: 5670092; longitude: 3253118; altitude: 145; and drained area: 2600).

Figure A8.4 Results of reconstruction of station GRD6979300

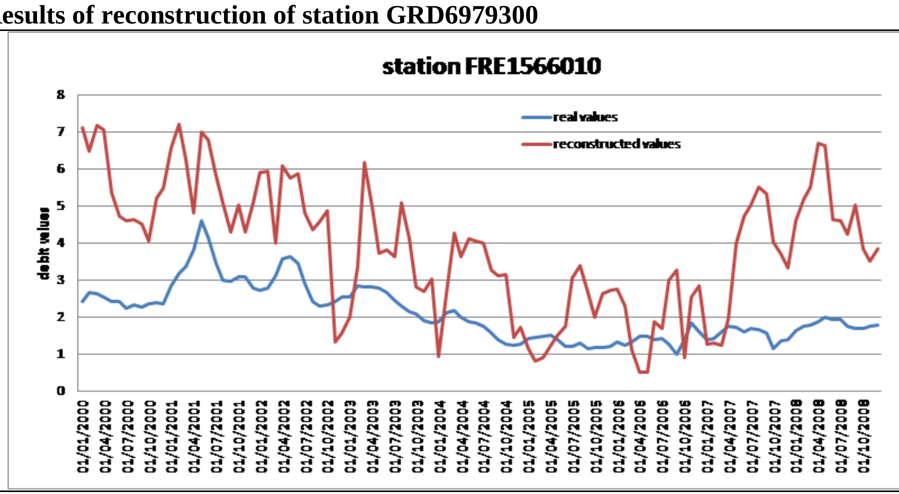


Source: SCM in Pöyry report .

The average of the reconstructed values equals **25.21** m^3/s. The two confidence intervals (percentiles 5 % and 95 %) of the probability law of each reconstructed month. The averages of these percentiles are equal to **0.005** m^3/s and **46.10** m^3/s accordingly. In order to compare the observed and reconstructed values, the absolute and the relative difference for each month were computed for each month, yielding respectively **4.81** m^3/s and to **2.6 %.**

By contrast, and to identify the reasons why some reconstructions don't operate, another station with poor accordance with observed values is presented.

Figure A8.5 Results of reconstruction of station GRD6979300

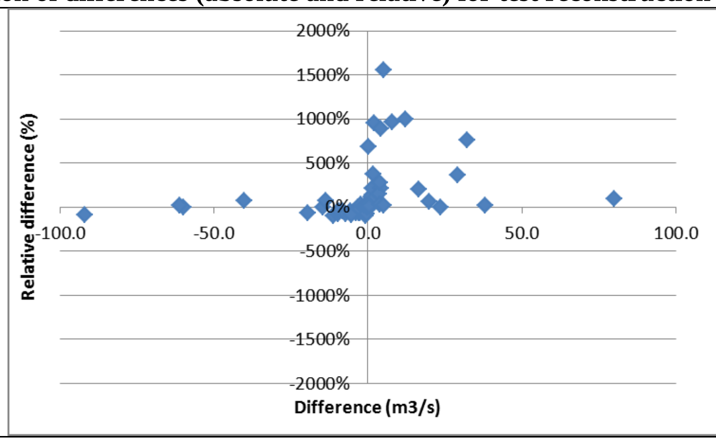


Source: SCM in Pöyry report.

The absolute difference equals 1.8 m^3/s; the relative one reaches 86 %, making the reconstruction ineffective. Similar validation tests were performed for each climatic zone. The test stations were chosen randomly and their quantity depends on the size of analysed region.

Overall quality analysis was carried out (after filtering the obviously erroneous stations) and using random reconstruction of stations. The result is that in most cases, the reconstruction is acceptable and in a significant proportion, is totally erratic and unusable. The key issue is that reasons for such unpredictable outcomes are unclear, and they make the full set of results questionable.

Figure A8.6 Distribution of differences (absolute and relative) for test reconstruction



Source: SCM in Pöyry report.

Ultimately, over a total number of 7 442 stations:

- 4 852 stations (65 %) are documented (already known or reconstructed by the previous method);
- 2 590 stations (35 %) were reconstructed using the EPH, of which:
  - Continental: 1 759 stations (1 484 (84 %) documented and 275 (16 %) reconstructed);
  - Atlantic: 2 646 stations (1 415 (53 %) documented and 1 231 (47 %) reconstructed);
  - Mediterranean: 1,763 stations (899 (51 %) documented and 864 (49 %) reconstructed);
  - Arctic: 626 stations (516 (82 %) documented and 110 (28 %) reconstructed);
  - Alpine: 648 stations (538 (83 %) documented and 110 (17 %) reconstructed).

6 Improving the relevance of the method

6.1 Rationales

The reconstruction provides results that are good in the majority of cases, but extremely bad under unpredictable conditions in a minority of cases. This minority, however, is still too much to consider the process of reconstruction as part of the discharge production process.

The large share of acceptably reconstructed data pleads to the recognition that the EPH is a good candidate for reconstructing data since it meets the accounting requirements of reconstructing information inside a category (no risk of circularity).

However, to be a reference method it must be improved and its effectiveness substantially upgraded, so that calculations would be substantially shorter, by revisiting the conditions under which the EPH is being applied. After having improved the baseline data sets, it became possible to set up and test a different approach based on two categories of conditions:

1. only and exclusively monthly specific discharges are reconstructed (instead of discharges);
2. explicit educated filtering of conditions for including stations as seeds for reconstruction and reducing drastically their number to propose only those stations that are likely to provide effective information.

The question of adapting the normalised distance has been raised with SCM; after considering the pros and cons, it was preferred to not touch the normalised distance but to make it more effective by appropriate selection of source stations, so that the targets are reconstructed only from a limited albeit relevant set of information. In the previous application the normalised distance was in fact given a twin role of weighting the most probable values (the more distant the lesser the weight) and implicit filtering thanks to longitude, latitude, altitude and catchment area values.

The improved procedure explicitly separates these functions

6.2 Filtering and selecting parameters

The target of filtering is selecting only those source stations that are likely to provide the best possible estimate of the target station to reconstruct and second to drop stations which contribution is likely marginal or which selection cumulates the inconvenient of adding noise and not signal on the one hand and making computations longer on the other hand.

The parameters belong to different categories: hydrological relevance, climatic proximity and geographical proximity

Hydrological relevance is driven by catchment area and belonging to same to the similar watershed systems. Climatic proximity is driven by the biogeographical region (the use of GRDC hydrographical provinces has been envisaged by they are too small and not yet part of ECRINS, and hence abandoned). Last, the geographical proximity is driven by direct distance computed from the coordinates.

The role of catchment area is twin. It is first a selection parameter, by considering only those stations in acceptable range of catchments as seeds and a role as component of distance since the area of the watershed is itself a factor of productivity (summarising, under same conditions, the smaller the catchment, the larger the productivity). As a rule of thumb, that could be improved later on, the catchment areas of the seed stations should fall within 25 % to 4 times the area of the target station. For example, a target station draining 10000 km^2 could be reconstructed by any station meeting the other criteria, provided draining between 2500 to 40000 km^2 is eligible.

Source and target stations should belong to the similar watershed systems. This definition is deliberately fuzzy since the EPH aims at reconstructing discharges where limited data is present. As working definition, watersheds are understood similar if they empty into the same shore (by default into the same sea), as coded in ECRINS.

The climatic proximity is defined by two characteristics; biogeographical region and altitude that is a strong climatic factor. These two factors should be taken with some flexibility. For example, a station in bioregion B, downstream bioregion A may use stations from region A if ‘not too far’.

6.3 Filtering application

The fuzzy logics factors are turned into programmed application by a set of computations and scoring that avoids abrupt excluding. The application is set as an ancillary component of WERC (the application that makes ECRINS) because it calls for essential components of ECRINS that are fully available in this application. It has been developed to be capable of testing thanks to parameters that operate in three steps:

1. preselecting candidate source stations;

2. scoring the candidates and eliminating those below half of the threshold for inclusion;
3. populating the remaining candidates with their informative characteristics (presence of monthly discharges in the period chosen for reconstruction) and finalising the score and elimination.

Taking the presence of data in last step is targeted to possibly ask for supplementary data if needed to reconstruct essential stations and secondarily to fasten the computations.

Starting from the list of target stations (having eliminated those with no information) , and processing one by one with, a short list of preselected source stations is prepared based on the altitude range, area range and distance. All factors and thresholds can be adjusted.

- Source station is candidate is its altitude (given or estimated from FEC mean altitude otherwise) falls between target altitude × f_{A0} and target altitude f_{A1}. Test is carried out with f_{A0} =0.5 and f_{A1} =2.0. To avoid scaling issues, the minimum range is {0 – 100m} and otherwise the difference target – candidate is limited to 500m.

  For example, if the target is situated at an altitude of 250 m, all candidates within 125 m to 500 m are preselected; in contrast, if the target is at 1 500 m, all 750 to 2 000 (and not 3 000 m) are preselected. Reciprocally, if target is at 20 m all between 0 m and 100m are preselected (and not only those in the range of 10 m to 40 m).

- Source station is candidate if its catchment lays between 0.25 × to 4× target catchment area.
- Source station is not one of the target stations.

Once pre-selection is finished for a target station, information on sources is completed by sea shore and biogeoregions that cover first, second and third the area (this has been computed separately with the cumulative function of Werc) and with the distance between this source and the target. Then scoring starts based on the following rules, for each pair of source - target:

- If source and targets are in the same biogeoregion as dominant type, target distance is the basic distance (is parameter, set to 250 km in the test). Otherwise, if first dominant of one is second dominant of the other, target distance for different bioregions is taken instead s target distance (is set to 100 km in the test).
- If matching is only for the third dominant, the target distance is the target distance is the target for different shortened by a penalty factor (set to .7). If not any similar bioregions, then target distance is again shortened by squaring the penalty factor (making it possible to take as source very close stations with no relationship ([91]).
- Sharing same sea is favourable factor of proximity: if stations share the same shore, the target distance is multiplied by 1/penalty (penalty always < 1), if both stations drain to the same sea (a sea contains many shores), this is neutral and otherwise, the penalty factor applies to reduce target distance.

Temporary score is then computed as target / distance, having set a minimum distance of 100m in case of the use of clone stations which distance is nil. The larger the score, the better the possible fit. At this stage, all couples target source which score is larger than half ([92]) the exclusion threshold

([91]) In mountains, an upstream station can be '100 % alpine' and the downstream 'alpine 20 %; boreal 80 %', this case hence accepts more distance than if there is not any common biogeoregion.

([92]) Since records excluded are lost until next calculation, it seemed wiser to keep them and use the results to better tune the scoring and the parameters for inclusion.

are kept and added to the final table, before computing the presence of discharge values. Discharge values are essential components that are analysed in the absolute (all years present) and for the targeted period of reconstruction. They are hence introduced in the last step of calculations because well scored sources without values (in the period or for any period) result from either discontinuing measurements or insufficient data provision.

For the reconstruction at one moment in time however, the quantity of data adjusts the score by multiplying it by number of months/number of expected month. A fully populated station is hence neutral since the ratio is 1.

7 Test computation with improved information

7.1 Data set and scoring results

The set of 76 stations used for testing have been scored and processed in the way indicated above. The possibilities are quite limited:

- Only 53 stations have a non-null score (smallest 0.28);
- Amongst the 53, 48 stations have usable sources (4 have 0 source ready, despite possible scores), this may result from no data available in the period for example

Table A8.9 Distribution of the number of source stations that can be used for reconstructing (test)

	>=50	>=25	>=10	>=5	<5 sources	
Number matching test	16	21	29	39	9	
Number in the class	16	5	8	10	9	48

Source: EEA WERC development, list provided by SCM, data Pöyry.

7.2 Obtained results and conclusions

The test was carried out in December 2012 and suggests a significant improvement of the reconstruction, albeit not yet fully satisfactory. In between the data required for the four target stations mentioned above was documented. The results are analysed into four categories:

1. Unusable: Difference observed – reconstructed more than 100 % on the annual average (9 stations)
2. Questionable: Difference between 50 % and 100 % (11 stations)
3. Acceptable: Difference 30 % to 50 % (11 stations)
4. Good: Difference less than 30 % (22 stations)

When considering the results, and especially those in the categories ‘acceptable’ and ‘questionable’, it comes that taking too many seeds can be counterproductive.

Interesting example is station DEe0000069. This station (159 300 km^2, interannual productivity 14.82 $l/s/km^2$) is reconstructed from 13 stations and the result is only 8.22 $l/s/km^2$. This seems clearly resulting from the selection of stations: the closest (as area and distance has productivity 16.42, but all other (and no closer ones were found!) have much smaller productivities of 7 and below 7. They are however both farther (over 400 km) and in a different biogeoregion. It would probably be more effective to put higher penalty on distance and accept a wider range of drainage area; this would deserve being reanalysed.

This situation is rather systematic in these categories. Similarly, even in the best category, there is no observable relationship between the number of seeds and the quality of the result; this militates for reducing the number of seeds by raising the inclusion threshold.

Up to now (for data reasons), the hydrological category of the placement of stations has not been considered: this would probably avoid the errors mentioned above.

The conclusions analyse in different categories.

- Major differences (first group), are suspected important anomalies in data, with very low productivities (less than 0.1 $l/s/km^2$ at the annual level are not trustable in temperate areas); in all cases, the observed values are much less than the reconstructed ones, whichever the number of seed stations is.
- There is no apparent relationship between the number of seed stations (that ranges between 783 and 2!) and the quality of results. As a conclusion, the filtering could be still more strict and limit to c.a. 20 best suited stations for a target and possibly filter stricter according to the scoring; this would still more hurry the calculations; the change in method has divided by close to 10 the computer time needed (from one week to less than a day, and possibly reaching a couple of hours with further filtering).
- The time saved by improving the calculations should be devoted to stricter analysis of the attributes of the target stations, its local conditions (more or less altered) and on the likelihood of the seed stations, what has not been done here.
- There is no formal relationship between drainage area and quality of reconstruction; however, amongst 4 large stations reconstructed, one fall into each of the categories; indeed no reconstruction method should be applied to large drainage areas that are not likely to have good 'seed' stations.

Documentation of station is extremely important. In the case of the largest catchments there is a possible difficulty, related to the current ECRINS data model that defluences are not considered. The largest catchments may present defluences and the station being on a branch. In such case, the area upstream is not representing the productivity at the station and the conceptual issue is not well analysed. It seems possible that such cases would not be addressable by probabilistic reconstruction methods.

Annex 9 Domestic and urban uses data collecting

1 Place in the process

Generic features in relation with all water uses are presented in the main report in section 4.3 71. Domestic and urban uses are major water users which complete the following parts of water cycle:

- Water abstraction from surface and groundwater (exceptionally desalination),
- Transferring water for the abstraction place (or places) and distribution, possibly resell water to other users;
- Collecting and hopefully processing wastewater and return them to final recipient, close or distant.

2 Mobilised data sources

2.1 Delineation of cities

The following sources were used:

- The urban audit (Eurostat) that covers 31 countries in Europe (under its extended data set)
- UMZ (urban Morphological Zones) derived from CLC by ETC/LUSI. There are 124,623 urban morphological zones (UMZ) which area ranges between 0.03 to 202 km^2 hosting populating from several hundred inhabitants to several millions.

The UMZs are defined as ‘contiguous urban areas laying less than 200 metres apart’; this definition, and the selection of the CLC classes that connect the urban fabric, make UMZs highlight urban sprawl. This results sometimes in large UMZs, with several nucleuses, with large number of branches and interconnected, which doesn’t fit the purpose of water accounts, as the cardinality with UWWTPs grows.

Because of that, some large UMZs were split according to their morphology and population. As this had to be done manually, the methodology followed was:

- First, select automatically the UMZ candidates to be split
- Second, visit each of them, decide whether they should be split, and manually split them through connecting areas, with the support of satellite imagery.

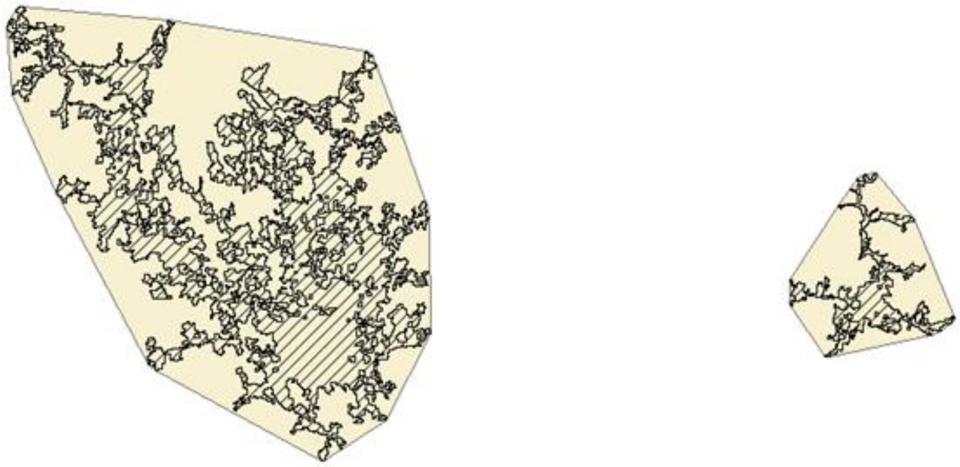
The UMZs to be split are first selected with homogeneous, automatic criteria, which require the previous calculation of the following elements for each UMZ:

- Population: derived from LandScan 2010, attributed to the UMZs using the standard ArcGIS® Spatial Analyst extension Zonal Statistics tool
- Compactness: based on existing compactness indices in the literature, several compactness indices were tried, but all of them showed a very strong correlation with the surface of the city. Based on any of them, big cities are always less compact than small cities, regardless of their shape. Because of this, a novel, very simple compactness index was defined as Compactness index = UMZ area / UMZ convex-hull area, the hull is the area encapsulating the UMZ (see Figure A9.1)
- Core areas: core areas are defined from CLC classes density in 1 square Km (CORILIS), calculated with a smoothing of 5 km; when residential and urban dense land cover classes (CLC classes 111 and 112) density is bigger than 50 hectares per square Kilometre, then we

have a core area; core areas are individualized using ArcGIS Spatial Analyst Region Group tool, and then by intersecting and summarising the resulting table the number of core areas for one UMZ is calculated.

These concepts are illustrated by the next figures.

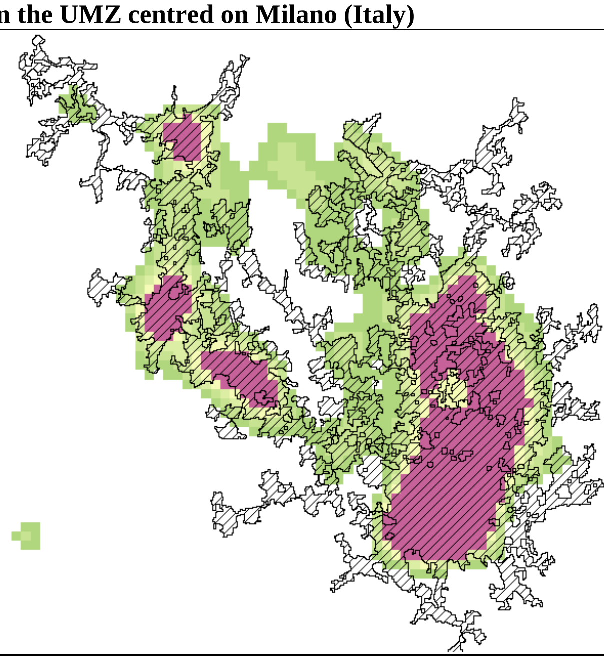
Figure A9.1 Illustration of the compactness figure name in annex



Source: EEA. Left is Milano area and right is Brescia

The next Figure A9.2 illustrates the core area (reddish) concept applied to the UMZ in the left part of Figure A9.1. The colours indicate a decreasing density of the urban CLC classes from reddish to green

Figure A9.2 Core areas in the UMZ centred on Milano (Italy)



Source: EEA processing.

Then, the following criteria are applied for identifying the list of UMZs to be considered as candidates for splitting

- *Population larger than 1,000,000 inhabitants (derived from LandScan 2010) and Compactness index below mean value – standard deviation = 0.476493*

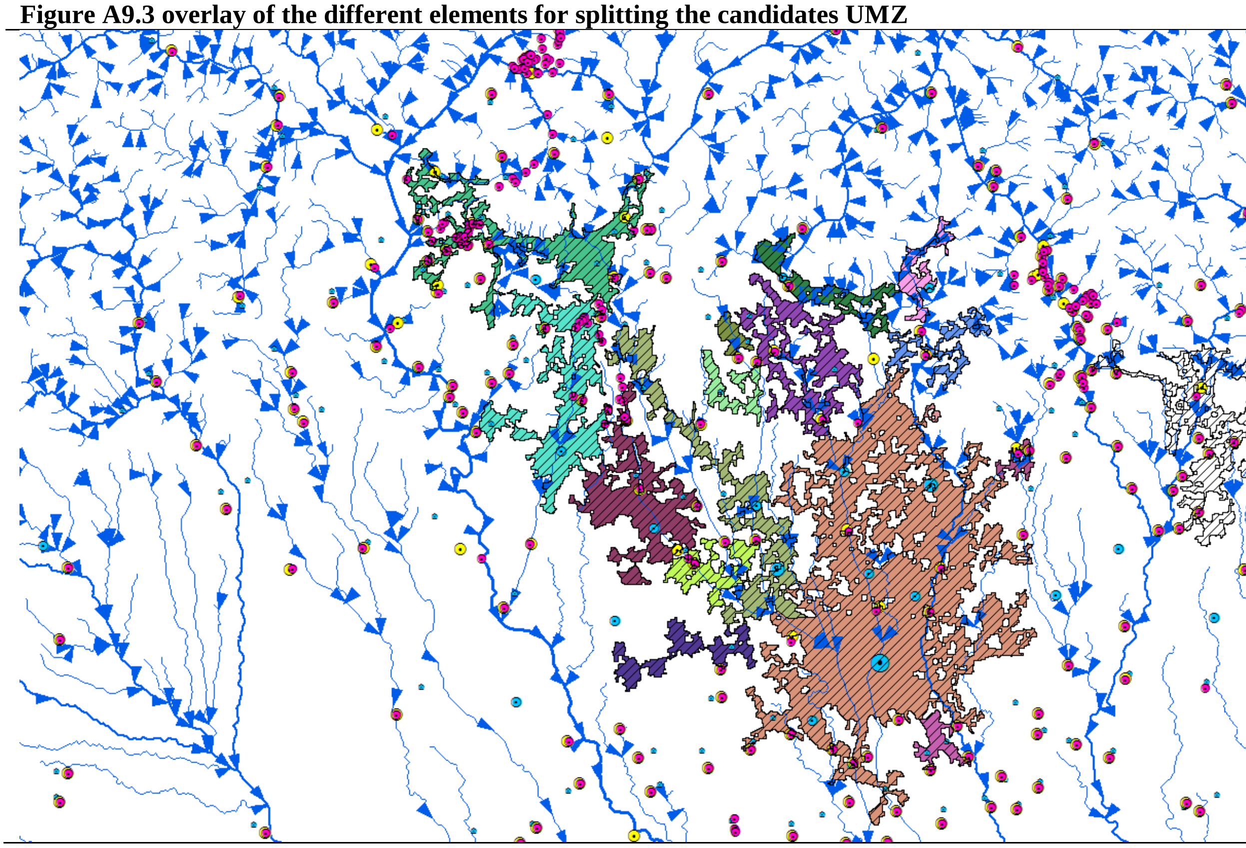
or

- *Having more than one core area.*

By applying this set of criteria, 135 UMZs were selected. These UMZ were then manually edited. The visual inspection of each and every candidate resulted in apportioning the set into those having more than a one UWWTP, with special regard of the number of river systems involved on the one hand and those having more than one core area or having more than one UWWTD 'agglomeration' (this notion is defined in this annex) on the other hand.

The candidates can then be apportioned, without splitting residential areas (using connecting elements like roads, rivers, green urban areas, industrial areas, or infrastructures like ports or airports). The support information used for this includes ECRINS' rivers with their flow direction, CLC, core areas as defined above, UWWTPs, and Bing Maps aerial and topographic maps, as illustrated in the next

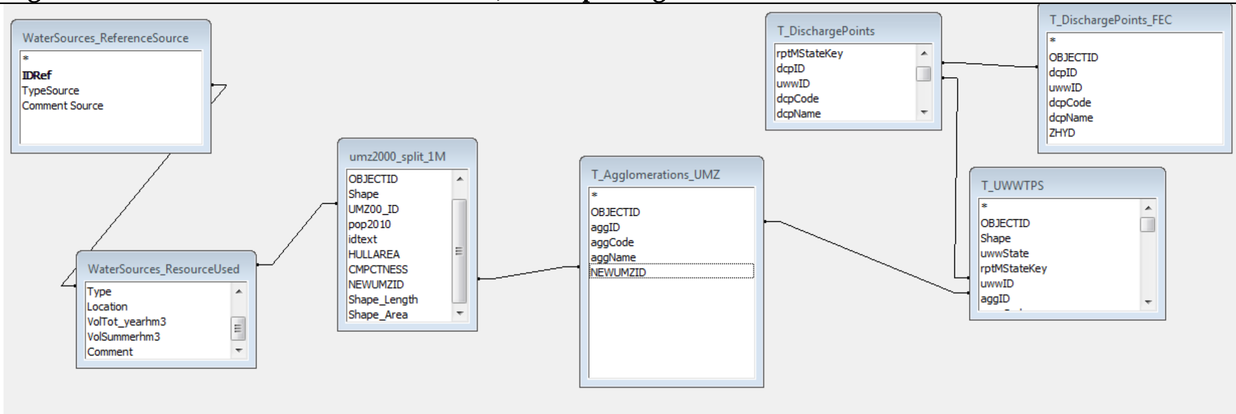
Figure A9.3 overlay of the different elements for splitting the candidates UMZ



Source: EEA

After doing the split, a final data model is assembled and populated. This data model explicitly defines the relationship between this new 'sub UMZs' and the UWWTD agglomerations (and subsequently with UWWTPs and discharge points, which have also been linked to ECRINS river segments). It also keeps the relationship of each new 'sub UMZ' with the original UMZ for the complete consistency of the data model.

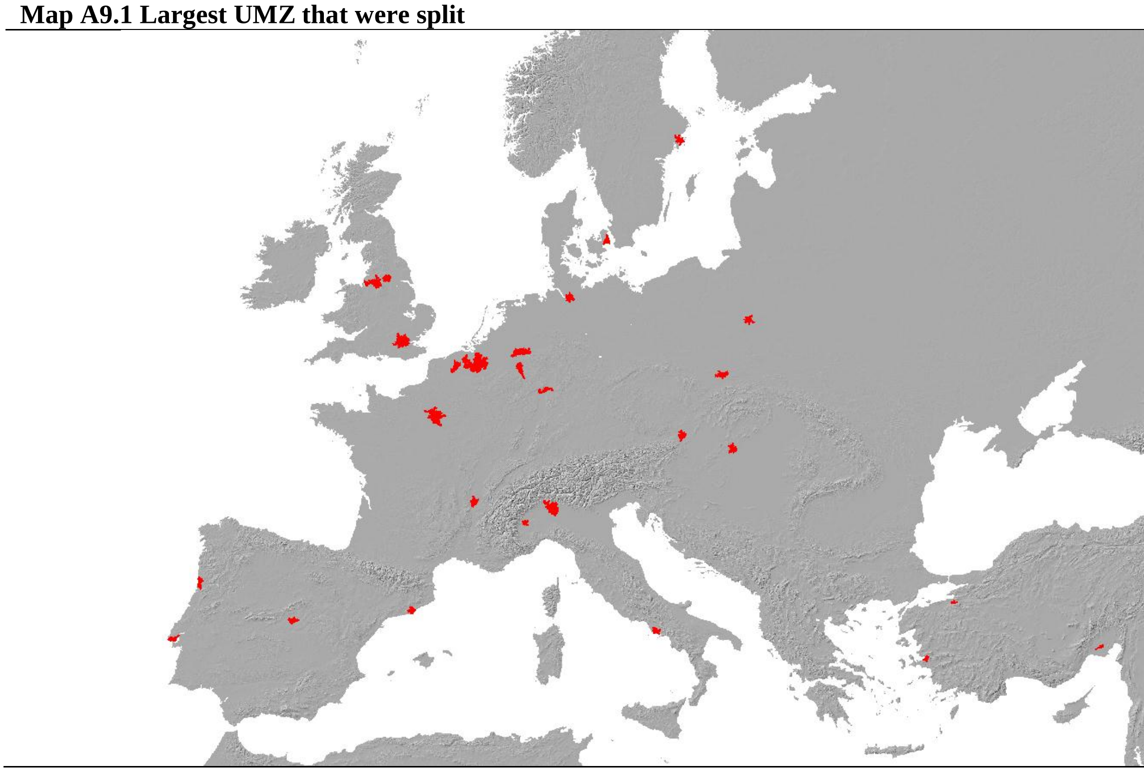
Figure A9.4 Data model for the final UMZs, after splitting



Source: EEA

As a result, many significant areas were split and apportioned into more functional entities, the largest being visible on the next

Map A9.1 Largest UMZ that were split



Source: EEA

2.2 Population numbers

LandScan data, re-sampled to meet licensing was processed by the EEA and used systematically.

2.3 Water supply

The following sources were used:

- The urban audit (Eurostat)
- In France, public data from 4 water agencies
- Searches on the Web (city sites, Wikipedia)

2.4 Sewage collection and purification

Reporting under the Urban Waste Water treatment Plants directive has been used as single source of information. This source however misses information from several countries, from complete absence of reporting to incomplete documentation.

Table A9.1 Population and UWWTP capacity declared

Country	EU	Pop counted (1)	UWWTP declared capacity (PE) (2)	no UWWTP	Apparent % served as from declarations ((2)/(1)
BG	yes	7435887		254	ND
RO	yes	22111653	14596382	2631	66 %
AT	yes	8436086	20608099	641	244 %
BE	yes	10723790	10653921	655	99 %
CY	yes	1088733		13	ND
CZ	yes	10440720			ND
DK	yes	5905778	12233454	422	207 %
FI	yes	6123180	7348800	207	120 %
FR	yes	64864013	4716550	3025	7 %
DE	yes	84832180	1.46E+08	4114	172 %
GR	yes	10514382			ND
HU	yes	10259321	9635638	515	94 %
IE	yes	4830761	4916376	466	102 %
IT	yes	57741597	87101835	4181	151 %
LU	yes	484863	990350	42	204 %
NL	yes	17447188	24034633	414	138 %
PL	yes	39834539	38819392	673	97 %
PT	yes	10291642	13694568	450	133 %
SK	yes	5511640	7865237	276	143 %
SI	yes	1935018	1824981	137	94 %
SE	yes	10169317		347	ND
EE	yes	1410403	1821628	48	129 %
LV	yes	2416966	4187694	135	173 %
LT	yes	3904714	4064222	97	104 %
MT	yes	376678	120000	7	32 %
IS	No	309055			ND
NO	No	5240621			ND
TR	No	73501378			ND
CH	No	8217965			ND
LI	No	32075			ND

Source: population from re-allocated population as used in the accounts, UWWTP capacity and number of UWWTPs, as summed from reporting, % served is just for information and does not reflect country's reality.

The questions raised in this table are discussed in a next section.

3 Description of data sets related to domestic water uses

All data sets dealing with water uses are structured in the same way. This systematic structuring follows Nopolu System2 logics and is informed in Annex 1 , § 2.1 . All tables are in database EU_Uses.mdb.

3.1 Data sets involved in the urban and domestic water uses

Domestic and urban activities demand many different tables to be properly described, related and the data populated so that the appropriate cells in the SEEAW I/O tables are populated. The conceptual model of cities is that the city (statistical unit) is a constant object (more precisely a spatial superstructure (slowly changing) related to spatial infrastructure (not changing at considered scales)) having two activities: hosting persons and making services. Regarding water, these activity use, consume and dispose of wastewater which supplying is an industrial activity (as ISIC nomenclature) but which quantity is driven by the cities own activities volumes that are estimated thanks to the number of inhabitants.

The volume of water is hence, for a city or a group of cities, etc. the result of a simple calculation:

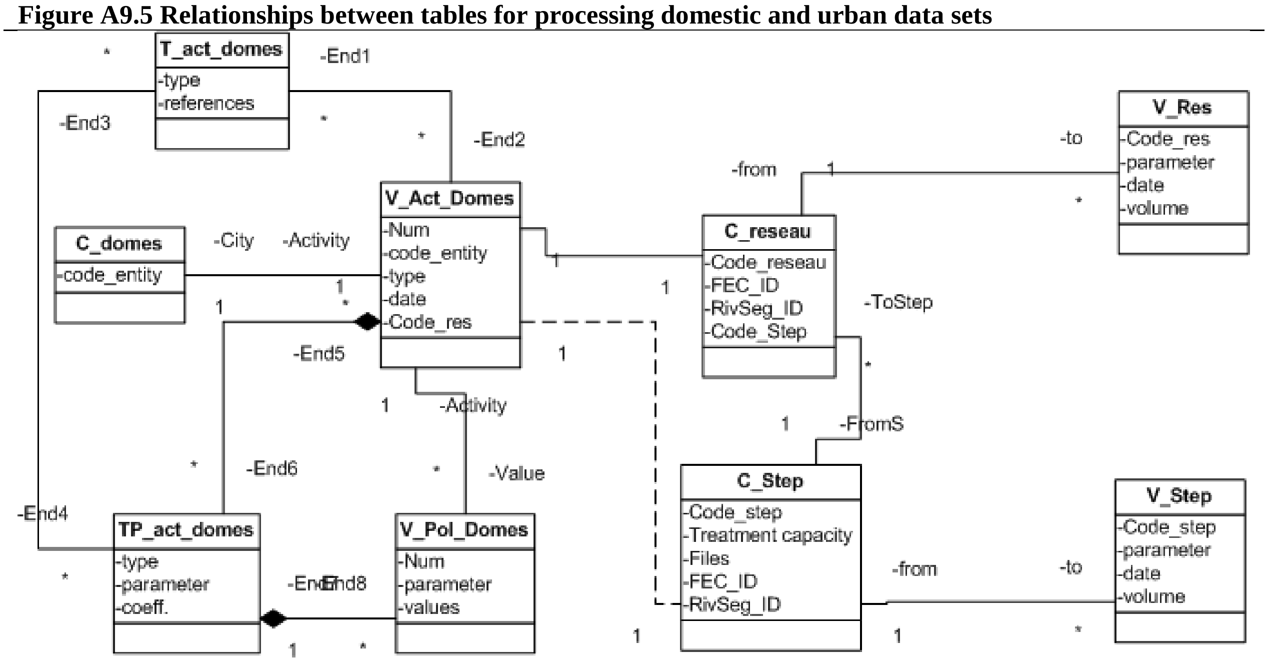
Equation A.9.1 generic calculation of values from activity

$$vol_{a,c,t} = coef_{a,c} \times volact_{c,t},$$

where a, c and t represent respectively the activity type, the city and time (date).

All organisations of data and data processing aim at populating the volumes of activity for the well-designated cities and estimate the technical coefficients. These can be the outcome of specific data for a certain city or coming from different statistics. At the end, the data storage is the same.

Figure A9.5 Relationships between tables for processing domestic and urban data sets



Source: Simplified by EEA from Nopolu System2 databases and Pöyry report

Notes: It is important to note that a sewer network is related to an activity and not to a city. This allows processing differently (if data is present) different networks within a same city, serving different shares of activity. Similarly, a wastewater treatment plant is connected to as many networks as necessary. Dotted lines indicate possibility to process urban activity related with no network to WWTP (can be the case of village processed by pulp-mill lagoon for instance).

This construction is the simplification closest to the reality. The exceptional case where several networks serve several WWTP is possible, at the expense of creating clone networks. Despite this simple structure, some data issues have needed assumptions that are recalled in next sections before being summarised in the recommendations.

3.2 Domestic entities

Table C_domes regroup each urban entity and characterise the activity of population (households) and connected services.

The Households refer to the anthropogenic activities of population and concern the use of water coming mostly from public water unit supply, but also from own unitary household abstraction.

The connected services refer to the consumption of water for all the services connected to public water unit supply.

Figure A9.6 sample extract of table C_Domes

C_domes

ID	AREA	PERIME	Code	Ins	Libel	Country	Surf	Perc_IN	Perc_Zl	Pop	Pop_urb	Pop_rur	Densi	C_	Stratae	ZHYD
189781	0	0	A020005098		S+	DE		0.0%	0.0%		0	0			S	A020005098
189782	0	0	A020005102		S+	DE		0.0%	0.0%		0	13521			S	A020005102
189783	0	0	A020005329		S+	DE		0.0%	0.0%		4411	9635			S	A020005329
189784	0	0	A030000948		S+	DE		0.0%	0.0%		15515	34636			S	A030000948
189785	0	0	A030001201		S+	DE		0.0%	0.0%		36930	63849			S	A030001201
75	0	0	AT001C		Wien	AT		0.0%	0.0%		1674909				L	F020006078
233	0	0	AT002C		Graz	AT		0.0%	0.0%		250738				L	F030003258
257	0	0	AT003C		Linz	AT		0.0%	0.0%		188593				L	F030001413
314	0	0	AT004C		Salzburg	AT		0.0%	0.0%		147169				L	F020008745
392	0	0	AT005C		Innsbruck	AT		0.0%	0.0%		117150				L	F030003195
189786	0	0	B020000105		S+	UK		0.0%	0.0%		0	0			S	B020000105
189787	0	0	B020000112		S+	UK		0.0%	0.0%		0	0			S	B020000112
189788	0	0	B020000280		S+	UK		0.0%	0.0%		0	49			S	B020000280
189789	0	0	B020001068		S+	UK		0.0%	0.0%		0	0			S	B020001068

Comment: two different types of domestic entities can be seen: large cities, having an ID specific to city and a name, and small entities which Id (field code) is the FEC (field ZHYD) where they are.

3.3 Domestic volumes

Each urban entity in C_domes is split up into two types of water use, the 'Households activity' and the 'Connected services' activity. Both activity volumes are populated in the V_act_domes table, field type. Urban water uses are a certain type of industrial activity that is described by a type and by a volume of activity that is simply the number of inhabitants

By contrast, the technical coefficients are entered into table TP_Act_Domes, with the source of information and must be referred into T_Act_domes that ensures absence of duplicates of type. The syntax of type field is any, and is hence a mnemonic. Water supply and sewerage are industrial activities and are hence not dealt with in the _domes tables. In turn V_act_domes relates to the V_pol_domes, that contains the values for the activity.

Figure A9.7 sample extract of table C_Domes with its depending tables

C_domes

ID	AREA	PERIME	Code	Ins	Libel	Country	Surf	Perc_IN	Perc_Zl	Pop	Pop_urb	Pop_rur	Densi	C_	Stratae	ZHYD
189781	0	0	A020005098		S+	DE		0.0%	0.0%		0	0			S	A020005098
189782	0	0	A020005102		S+	DE		0.0%	0.0%		0	13521			S	A020005102
189783	0	0	A020005329		S+	DE		0.0%	0.0%		4411	9635			S	A020005329
189784	0	0	A030000948		S+	DE		0.0%	0.0%		15515	34636			S	A030000948
189785	0	0	A030001201		S+	DE		0.0%	0.0%		36930	63849			S	A030001201
75	0	0	AT001C		Wien	AT		0.0%	0.0%		1674909				L	F020006078

Num	Type	Date	Libel	ZHYD	INDUS	WOrigin	WDestinatic	TR	PK	Ci	Cc	Rac_in	Code_res	Rac	Vol_activ	Water_ir
2311603	AT_dom	01/01/2001	Wien	F020006078		D	E	dom	0.00			0.00%	641	100.00%	1674909	2250
2312497	AT_pub	01/01/2001	Wien	F020006078		D	E	pub	0.00			0.00%	641	100.00%	1674909	7343
(New)									0.00			0.00%		0.00%	0	

Code_para	F_Prod	F_Prod_m	Perc_Elim	F_Elim_A	F_Elim_S	F_Brut_A	F_Brut_S	F_Brut	F_Brut_mes	F_Net_A	F_Net_S	F_Net	F_Net_mes	Conc_equi	Impact	Ecart_Obj	Cl
1,421	73,430.6		0.0%	0.0	0.0	73,430.6	0.0	73,430.6		73,430.6	0.0	73,430.6					
0	0.0		0.0%					0.0				0.0		0.0000	0.00	0.00	

Comments: C_domes record for city of Graz (AT) points to v_act_domes (green extract) and for its public activity, computed from Austrian coefficient, produces a final volume in the purple extract.

It can be seen in Figure A9.7 that values relate to type At_pub, indicating that a global coefficient for Austria and public services are used for computation. The coefficients are in reference table TP_Act_domes, presented in Chapter 5

3.4 *Sewage networks and wastewater treatment plants*

In the water accounts, the current structure has been used to a minimum: a sewage networks connects a source of sewage to a WWTP. To a large extend, the network is fictitious since it is in this first application a way to keep an application procedure in the absence of information, in order to allow the calculation of leakages that can be computed only at the network level by Nopolu System2.

In the WA, the process is quite simplified: WWTP capacity is not used (and is not populated in many countries in the reported data sets under the UWWTP directive).

The names of the UWWTP are those from the directive report and are not that explicit. The important relationship is with the source data set The C_step table has a twin ID system: internal for consistency and processing and documentary that is the mirroring of the source data. However, source data refers to the Id inside the UWWTP and possibly does not link to country's internal coding. In future work, it might be convenient to add another field to this end, to foster data validation and exchanges.

4 Analysis per stratum

Data processing from the different sources has not been identical in the different strata: the degree of processing has been adjusted to the targeted accuracy per stratum.

4.1 Large cities

The stratum 'L' contains the largest European cities (544 cities in the 27 European countries members). The data required for the water account calculation on each city is:

- Population
- The urban wastewater treatment plants associated with each city
- The sewage networks
- FEC
- Abstraction volumes, type of resources and localization of the resources into the FECs.

Population

The population data for the items in stratum 'L' comes from the urban audit. This population is located in the core of the cities delimited by the administrative border.

The urban audit data file available (Eurostat_Table_tgs00079FlagDesc.xls) on the Eurostat website contains only 371 cities. This information was completed with data taken in the shapefile (URAU_CITY_RG_01M_2007.shp), which contains the delimitation of the core of the urban audit cities. This file contains 679 cities and only records the population for the year 2001.

Urban wastewater treatment plants

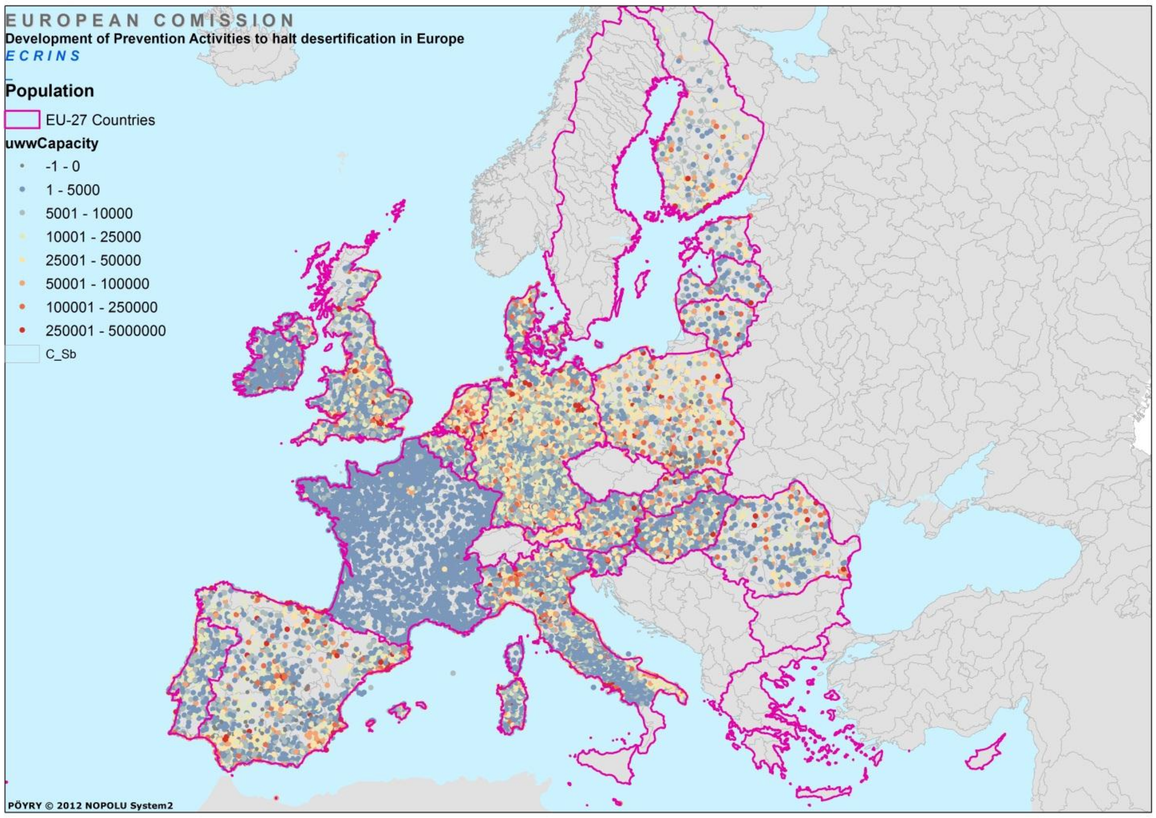
The Urban Waste Water Treatment Directive addresses the collection, treatment and discharge of urban wastewater along with those from certain industrial sectors. A report of the facilities has been provided by the Member States as part of the UWWTD implementation (24 countries of the European Union, not including the Czech Republic, Bulgaria and Greece). The non-EU countries have provided no data as well.

The map below shows the location of the UWWTPS of the EU.

The UWWTPS were reported on the agglomerations that generate a pollution load larger than 2000 P.E ([93]). This database contains a table with the 'agglomerations' and a table with the UWWTPS. Each UWWTPS is linked with an agglomeration. The problem to solve is creating a relationship between an agglomeration and the part or set of cities they serve.

Table A9.1 above is very informative and witnesses the reasons exposed in further paragraphs on the way to attach a UWWTP to a city: the UWWTP reporting is aggregated per 'agglomeration' that have no reported relationship with the administrative entities (or part of entities) that are drained by it.

Map A9.2 Location of the reported UWWTP in Europe



Source: Pöyry interim report, April 2012 from UWWTP reported data. Many French stations appear in grey because most have no capacity attached.

The 'agglomeration' is a reporting object, spatially located (as point) that is not detailed in the reporting. The following matching method has hence been applied, to establish the link between each city of the stratum L and the UWWTPS.

- A correspondence table has been created to link the agglomerations and the UMZ 2000. And a correspondence table has also been created between the cities of the strata L and the UMZ. First we used these tables to create another correspondence table between the cities, the agglomeration and the UWWTPS. This table was the basis for the following work.

([93]) PE : « person equivalent » quantity of pollution equal (from biodegradable carbon load) to the one generated by a standard person's activity. An agglomeration normally hosts a number of persons <= its PE figure.

- The correspondences of every large European city has been checked, with special validation over stratum L cities in France (where agglomeration – cities relationships could be established with certainty) in order to verify the matching method. Supplementary checks were carried out thanks to Web harvested information to correct the list of wastewater treatment plants linked to the core of each city. This work is exhaustive and precise for the main cities of Europe (more than 900 000 inhabitants).
- At this stage, it was noticed that the name of the agglomeration linked to many UWWTPS often had the same name as the city. This hypothesis was used to sort out the UWWTPS which agglomeration has the same name as the city. This hypothesis allows a more effective verification process for all the smaller cities.
- It happened that the names of the agglomerations were different to the names of the cities. In that case, the localization of the city, the UWWTPS and the agglomeration was checked by spatial overlaying with ArcGIS. The UWWTPS that appeared being the most likely was selected; this assignment is sometimes arbitrary.
- As counter checking, the number of cities linked to one or many UWWTPS was calculated and shown that many cities of the stratum L were missing. So the agglomerations that could be related to the missing cities using their name were checked. When the name of a city didn't have a correspondence with the agglomeration names, the UWWTPS was assigned using the mapping on ArcGIS. This work was also sometimes arbitrary.
- Finally, all the cities of the stratum L have been linked with one or multiple UWWTPS, except 34 cities whose countries haven't participated in the UWWTD report. 2 large cities are missing: Prague (CZ) and Athens (GR).
- This correspondence work is more accurate for large cities and for France. For all the smaller cities, some errors may be found even if all the cities were checked individually.
- Eventually, the UWWTPS were assigned a FEC thanks to cartographic intersection on ArcGIS.

The UWWTPS empty processed waters at a point which coordinates are given and that was related to the appropriate ECRINS river segment.

Sewage network

A sewage network connects domestic entities represented by the population and the UWWTPS. To simplify the calculations, and because no data related to the sewage networks is being collected, it was assumed that each UWWTPS is connected to a single sewage network. This assumption is factually incorrect but does not pose any kind of problem regarding the water assets accounts. The sewage network entity is assigned the same FEC as the UWWTPS.

Placement of each domestic entity into the FEC

The large urban entities possibly extend over several FECS. Each domestic entity was assigned to the largest FEC identified by cartographic intersection between the delimitation of the city (table: URAU_CITY_RG_01M_2007.shp) and the delimitation of the FECs.

Water abstraction of the stratum L

The following data sources were mobilised.

Data from French water agencies

We used data from 4 basin agencies: Rhône-Méditerranée-Corse, Adour-Garonne, Loire-Bretagne, Artois-Picardie that freely give access to abstraction points and volumes. The data from Seine-Normandie and Rhin-Meuse basins are missing ([94]).

The link between each city of the stratum L and the water abstraction points is not included in the data sets that address only the abstraction point. The name of the abstraction points and the name of the owners were used to find the correspondences with the cities ([95]), the Web was mobilised as well to verify this work.

Water abstraction points and the FEC were overlaid to localize them accurately (in the Loire Bretagne basin, point have no coordinates, they were placed to the most likely place using the map of French elementary catchments).

Internet data

A database has been created specifically by the EEA ([96]) to compile data found on Internet for the largest European cities. This database was first seeded from the urban audit and completed with the water abstraction volume, the type of resource, the FEC. This database is quite exhaustive for the larger cities of the stratum L (94 cities). However, data found are very uneven and had to be consolidated in a post-processing phase.

This data consolidation was carried out by testing the technical coefficients (e.g. vol/hab/day) utilised from each resource. This made it possible to calculate the volume abstracted by the population of the city and the consumption ratio for public services (of which drinking water production) of the country.

Hypothesis for the poorly documented cities of stratum L

Smaller cities of the stratum L (426 cities) do not have information about the volume, the type of resource and the localization of the resource, they were completed using statistics of water resources found by country on Eurostat website. The types of resources are Surface water, Groundwater and Sea water.

Surface water must be apportioned between reservoirs, and rivers. The allocating of a source has been made (in the absence of dedicated information) with the next hypothesis.

Surface water source is associated with a lake (or reservoir) if a large lake (lake from table C_lak, whose area is beyond 100 hectares) is located within 10 km radius around a city (type 'lake' if it is natural, otherwise type is 'reservoir').

Surface water source is associated with a river if there is no lake within a radius of 10 km around the city.

([94]) The four first water agencies mentioned consider that these data are public, based on the water law, and so they provide these data on publicly accessible websites. Seine-Normandie has different outlook, and considers that these data are covered by statistical secrecy rules, so it refuses to provide them, except to statistical services. Rhin-Meuse considers that these data should be disseminated to external institutions only through centralised and official data providers and we were denied access for this reason.

([95]) This is not a trivial issue, since in data sets lacking unique identifiers, the syntax for names is unstable. This issue was also experienced in the EPRTR, where names reported differed slightly from names from other sources (and IDs don't match, even though there is certainty of addressing the same object). Algorithms based on the Levenshtein distance (Wikibooks, 2013) were developed and implemented and used in many parts of this project, for example in checking the identity of lakes from Wikipedia and lakes in ECRINS.

([96]) And populated by a guest scientist, Gaëtan Marchesini, in summer 2011.

4.2 Specific work and methodology applied for urban stratum M

The stratum « M » contains the intermediate sized European cities. The domestic entities of the stratum M come from the UMZ (CLC 2006). Since United Kingdom and Greece 2006 layer was missing when the work was carried out, UMZ 2000 has been used for these two countries.

The data required for the water account calculation on each city of this stratum is the same as for the stratum 'L', the single difference being that source of volumes is modelled and not individually collected. The items in this stratum are capped by those already taken in stratum 'L' and the lowest items are those limited by the UWWTP directive, 2000 PE. UMZs having a population smaller fall hence in the next stratum 'S'. Hence, after allocation population there were 12,104 UMZ in the stratum M.

Population

The population of the UMZ was calculated with the LandScan 2010. Population in each UMZ was summed by the zonal statistic tool on ArcGIS. The LandScan grid cells (raster data set) have a resolution of 30 seconds, making the area depending on latitude. The polygons representing the UMZ are projected and hence made it necessary to compute first an averaged population density per grid cell (Land scan gives a number of persons per grid cell), associate to each UMZ the average of population density per grid cell and compute the population of the UMZ by multiplying the mean density of the UMZ with the UMZ area.

This method does not take into account the population of the grid cell partly outside the UMZ. Consequently, a part of the population can be ignored in the calculation. However, this is not an important issue because the population that has not been taken into account inside the UMZ will be taken into account inside the rural population of the FEC (explained at the next paragraph about the method of the stratum S). The possible errors are all the more important that the UMZ is small; however the smaller are eliminated by application of the threshold criterion mentioned above. Smaller UMZs and scattered population are in stratum 'S', making the total number of inhabitants kept.

The largest FEC (larger area intersected) associated to the UMZ is the reference FEC for the UMZ that is the statistical unit of urban activities in stratum M.

Urban Waste Water Treatment Plants (UWWTPS)

The closest UMZ has been associated with UWWTPS, this method implies that an UMZ can be associated with many UWWTPS, which is often the case in reality.

Water abstractions for public facilities encompass drinking water for households and other connected services. The principles applied for stratum L were applied, based on statistics of water resources found by country on Eurostat website to model each UMZ.

4.3 Specific work & methodology applied for urban strata 'S'

The stratum « S » contains the rural cities whose population is aggregated at the FEC level which is hence a pseudo-city which are is FEC area. Hence, the domestic entities of the stratum S have the same ID as the FEC and have only Rural population.

Population

The total population of the FEC was calculated with the LandScan data from 2010, with the same method described in the stratum M methodology paragraph, applied at FEC minus non-scattered population area. A special correction had to be developed to correct the population of the FECs located on the coastal area, to take into consideration the mismatches between ECRINS and LandScan grid and the risks of missing population grid cells that fall outside the coastal line.

Rural population is eventually computed by subtracting the urban population (strata L and M) from the total population (LandScan 2010) on each FEC. Somewhere, negative values appeared, that were all set to zero. . When the rural population of a FEC was negative, we attributed the value 0. The negative difference has been equally redistributed into all the other FEC from each sub-basin (SB level of the table C_ZHYD), to mitigate the double counting between urban and rural population.

Wastewater disposal

Wastewater from rural population (being by definition related to no UWWTP since the threshold for rural are precisely all the urban areas below the threshold for reporting UWWTP) is the FEC.

We also utilised the rural population of the FEC and the abstraction ratio for public services by country to calculate the volumes abstracted (see **Error! Reference source not found.**).

5 Statistical information related to domestic and urban results.

The overall results and intermediate steps are presented in the next tables. These tables are presented as mainly sourced from EU statistics; in the reality many sources have been mobilised (Eionet, Web, etc.) as indicated above. For simplicity reasons, the full procedure is exemplified with synthetic tables.

5.1 Summarised volumes at stake

The volumes used are composed of volumes abstracted and distributed by public systems and self-provisioning. The volumes abstracted are composed of volume to be distributed and volumes needed for preparing the volumes to be distributed (making tap water requires water).

A part of abstracted volumes is possibly lost (leakage) and the rest apportioned between households (domestic uses) and other uses (urban uses and not identified industrial uses in cities). Once used (and to some extend consumed, because evaporated or incorporated into products), wastewater is disposed in sewerage and eventually returned, in a lesser volume ([97]).

Most of these data are not available and must be reconstructed in the most transparent and accurate (at least reproducible) way, hopefully taking stock of reference and framing information.

5.2 Population connected to public water system

In the absence of otherwise information, it was considered that 100 % of the urban population is connected to a public drinking water supply system, so in that case, the water origin is 'water collection treatment and supply' that corresponds to ISIC code 3600.

The rural population is partly connected to a public drinking water supply system. The other part of the rural population is supplied by unitary household abstraction (individual drillings) whose water comes from the groundwater resource. To the end, the connecting ratio of the population to the public water supply network for each European countries found on Eurostat website reported in Table A9.2.

([97]) This is a principle scheme: in the real world some sewers act as drainage and return more water than was entered…

Table A9.2 Percentage of population served by public water supply network by country

Country	% of population linked with public water supply	Year	Sources and hypothesis
Austria	95 %	2008	Eurostat: env_wat_pop
Belgium	100 %	2009	Eurostat: env_wat_pop
Bulgaria	99 %	2009	Eurostat: env_wat_pop
Cyprus	100 %	2009	Eurostat: env_wat_pop
Czech Republic	92 %	2007	Eurostat: env_wat_pop
Germany	99 %	2007	Eurostat: env_wat_pop
Denmark	97 %	2002	Eurostat: env_wat_pop
Estonia	80 %	2009	Eurostat: env_wat_pop
Spain	94 %	2008	Estimation with Portugal data
Finland	85 %	2007	Estimation with Sweden data
France	99 %	2001	Eurostat: env_wat_pop
Greece	94 %	2007	Eurostat: env_wat_pop
Hungary	95 %	2009	Eurostat: env_wat_pop
Ireland	85 %	2007	Eurostat: env_wat_pop
Italy	94 %	2008	estimation with Portugal data
Lithuania	76 %	2009	Eurostat: env_wat_pop
Luxembourg	100 %	2009	Eurostat: env_wat_pop
Latvia	76 %	2009	Estimation with Lithuania data
Malta	100 %	2009	Eurostat: env_wat_pop
Netherlands	100 %	2008	Eurostat: env_wat_pop
Poland	87 %	2009	Eurostat: env_wat_pop
Portugal	94 %	2008	Eurostat: env_wat_pop
Romania	55 %	2009	Eurostat: env_wat_pop
Sweden	85 %	2007	Eurostat: env_wat_pop
Slovenia	91 %	2002	Eurostat: env_wat_pop
Slovakia	86 %	2009	Eurostat: env_wat_pop
United Kingdom	97 %	2002	Estimation with Denmark data

Source: reworked from Excel tables backing Pöyry interim report 1, April 2012.

5.3 Computation of technical coefficients

To model volumes per cities (the case of directly documented large cities is set aside), technical coefficients must be found and tested. They are based on exploiting abstraction data in relation with the different water resources. The table below shows the data found.

Table A9.3 Water resources abstracted by public sector

Country	Surface water (hm^3/y)	Year	Source of information	GW hm^3/y	Year	Source of information	Sea water (hm^3/Y)	Year	Source of information	Total abstraction (hm^3/Y)
Austria	1	2008	Eurostat	607	2008	Eurostat	0	ND	Eurostat	608
Belgium	249	2009	Eurostat	482	2009	Eurostat	0	ND	Eurostat	731
Bulgaria	517	2009	Eurostat	461	2009	Eurostat	0	ND	Eurostat	978
Cyprus	8	2009	Eurostat	11	2009	Eurostat	49	2009	Eurostat	68
Czech Republic	357	2009	Eurostat	315	2009	Eurostat	0	ND	Eurostat	672
Germany	1 547	2007	Eurostat	3 581	2007	Eurostat	0	ND	Eurostat	5 128
Denmark	0	2009	Eurostat	385	2009	Eurostat	0	ND	Eurostat	385
Estonia	29	2009	Eurostat	34	2009	Eurostat	0	ND	Eurostat	63
Spain	4 208	2008	Eurostat	1 557	2008	Eurostat	240	2004	Eurostat	6 005
Finland	165	2005	Eurostat	239	2005	Eurostat	0	ND	Eurostat	404
France	2 161	2007	Eurostat	3 614	2007	Eurostat	0	ND	Eurostat	5 775
Greece	648	2007	Eurostat	198	2007	Eurostat	0	ND	Eurostat	846
Hungary	264	2009	Eurostat	369	2009	Eurostat	0	ND	Eurostat	633

Country	Surface water (hm^3/y)	Year	Source of information	GW hm^3/y	Year	Source of information	Sea water (hm^3/Y)	Year	Source of information	Total abstraction (hm^3/Y)
Ireland	489	2007	Eurostat	120	2007	Eurostat	0	ND	Eurostat	609
Italy	1 366	2008	Eurostat	7 729	2008	Eurostat	0	ND	Eurostat	9 095
Lithuania	0	2009	Eurostat	130	2009	Eurostat	0	ND	Eurostat	130
Luxembourg	20	2009	Eurostat	23	2009	Eurostat	0	ND	Eurostat	43
Latvia	0	0	estimation with Estonia	0	0	estimation with Estonia	0	0	estimation with Estonia	0
Malta	0	0	Oieau ('fiches pays')	13	2009	Eurostat	17	2009	Eurostat	30
Netherlands	490	2008	Eurostat	762	2008	Eurostat	0	ND	Eurostat	1 252
Poland	649	2009	Eurostat	1 418	2009	Eurostat	0	ND	Eurostat	2 067
Portugal	590	2008	Eurostat	317	2008	Eurostat	1	2008	Eurostat	908
Romania	1 032	2009	Eurostat	473	2009	Eurostat	0	ND	Eurostat	1 505
Sweden	689	2007	Eurostat	202	2007	Eurostat	0	ND	Eurostat	891
Slovenia	4	2009	Eurostat	161	2009	Eurostat	0	ND	Eurostat	165
Slovakia	51	2009	Eurostat	267	2009	Eurostat	0	ND	Eurostat	318
United Kingdom	5 545	2005	Eurostat	1874	2005	Eurostat	0	ND	Eurostat	7 419
Together (hm^3/Y)	**21 079**	**(45 %)**		**25342**	**(54 %)**		**307**	**(1 %)**		**46 728**

Source: Data from Pöyry interim report 1, report June 2012.

The percentages surface water to total water range from 0 % to 80 % and from 16 % to 100 % for groundwater to total water. Sea water supplies are anecdotic at the EU wide level, but reach 72 % in Cyprus and 57 % in Malta. These data represent the volumes abstracted, to be compared with volumes delivered by the public sector. In a second time, the volumes of water supplied by public sector are taken mainly on the Eurostat website. The difference between both volumes has been considered as water losses. The table below shows the data found:

Table A9.4 Public water supplied and comparison with abstraction

Country	Total abstraction (Mm^3/year)	Total public water supply (Mm^3/year)	Year	% of water lost
Austria	608	***0***	NS	0 %
Belgium	731	700	2009	4 %
Bulgaria	978	385	2009	61 %
Cyprus	68	86	2008	-26 %
Czech Republic	672	520	2008	23 %
Germany	5 128	4 544	2007	11 %
Denmark	385	386	2009	0 %
Estonia	63	47	2009	25 %
Spain	6 005	3 827	2008	36 %
Finland	404	***0***	NS	0 %
France	5 775	5 685	2001	2 %
Greece	846	626	2007	26 %
Hungary	633	482	2009	24 %
Ireland	609	609	2007	0 %
Italy	9 095	5 533	2008	39 %
Lithuania	130	101	2009	22 %
Luxembourg	43	***0***	NS	0 %
Latvia	***0***	248	2007	0 %

Country	Total abstraction (Mm^3/year)	Total public water supply (Mm^3/year)	Year	% of water lost
Malta	30	29	2009	3 %
Netherlands	1 252	1 135	2008	9 %
Poland	2 067	1 544	2009	25 %
Portugal	908	671	2008	26 %
Romania	1 505	942	2009	37 %
Sweden	891	737	2007	17 %
Slovenia	165	118	2009	28 %
Slovakia	318	322	2007	-1 %
United Kingdom	7 419	6 109	2007	18 %
Together	**46 976**	**36 441**		**22 %**

Source: Data from Pöyry interim report 1, report June 2012, reformatted.

When no data are provided, a '0' in bold italic is placed instead of missing value. However, in this case, the value is replaced by its surrogate in the total line 'together'. For example, total abstraction '0' is estimated at first glance as at least equal to total supply. Reciprocal is true as well (supply counted at least equal to abstraction).

In few countries, negative water losses are observed, suggesting data inconsistency. In these cases, public water supply volumes are the reference data to calculate the abstracted volumes. Hence, the abstraction volumes of Slovakia, Denmark, and Cyprus are recalculated using the percentage of water losses of neighbouring countries, or countries understood as comparable. The final values used are presented in the table hereunder for the 27 EU countries.

Table A9.5 Final ratios used for abstracted water and supplied water by inhabitant and by day

Country	Total public water supply ratio by inhabitant (m^3/year/person) (raw)	Final total public water supply ratio by inhabitant (m^3/day/person)	Sources and hypothesis	Total abstraction ratio by inhabitant (m^3/year/person) (raw)	Final total abstraction ratio by inhabitant (m^3/day/person)	Sources and hypothesis
Austria	NS	0.18	Estimation with Germany losses (11 %)	0.20	0.20	Eurostat
Belgium	0.18	0.18	Eurostat	0.19	0.19	Eurostat
Bulgaria	0.14	0.14	Eurostat	0.35	0.35	Eurostat
Cyprus	0.30	0.30	Eurostat	0.24	0.40	Estimation with Greece losses (26 %)
Czech Republic	0.14	0.14	Eurostat	0.17	0.17	Eurostat
Germany	0.15	0.15	Eurostat	0.17	0.17	Eurostat
Denmark	0.20	0.20	Eurostat	0.20	0.22	Estimation with Netherlands losses (9 %)
Estonia	0.09	0.09	Eurostat	0.13	0.13	Eurostat
Spain	0.23	0.23	Eurostat	0.36	0.36	Eurostat
Finland	NS	0.17	Estimation with Sweden losses (17 %)	0.21	0.21	Eurostat

Country	Total public water supply ratio by inhabitant (m^3/year/person) (raw)	Final total public water supply ratio by inhabitant (m^3/day/person)	Sources and hypothesis	Total abstraction ratio by inhabitant (m^3/year/person) (raw)	Final total abstraction ratio by inhabitant (m^3/day/person)	Sources and hypothesis
France	0.25	0.19	Calculation with percentage of water consumed by domestic activity only and consumption ratio by inhabitants for domestic activity only (0.165m^3/day)	0.25	0.25	Eurostat
Greece	0.15	0.15	Eurostat	0.21	0.21	Eurostat
Hungary	0.13	0.13	Eurostat	0.17	0.17	Eurostat
Ireland	0.43	0.43	Eurostat	0.43	0.43	Eurostat
Italy	0.25	0.25	Eurostat	0.42	0.42	Eurostat
Lithuania	0.08	0.08	Eurostat	0.11	0.11	Eurostat
Luxembourg	NS	0.23	Estimation with Belgium losses (4 %)	0.24	0.24	Eurostat
Latvia	0.30	0.30	Eurostat	0.00	0.38	Estimation with Lithuania losses (22 %)
Malta	0.19	0.19	Eurostat	0.20	0.20	Eurostat
Netherlands	0.19	0.19	Eurostat	0.21	0.21	Eurostat
Poland	0.11	0.11	Eurostat	0.15	0.15	Eurostat
Portugal	0.17	0.17	Eurostat	0.23	0.23	Eurostat
Romania	0.12	0.12	Eurostat	0.19	0.19	Eurostat
Sweden	0.22	0.22	Eurostat	0.26	0.26	Eurostat
Slovenia	0.16	0.16	Eurostat	0.22	0.22	Eurostat
Slovakia	0.16	0.16	Eurostat	0.16	0.21	Estimation with Czech republic losses (23 %)
United Kingdom	0.28	0.28	Eurostat	0.34	0.34	Eurostat

Source: Data from Pöyry interim report 1, report June 2012, reformatted.

Notes: Each piece of information on ratios is reported in two columns, the first (marked '(raw') is the source data processed, and the second, completed by a column of sources, is the final information used in computing volumes, after estimating the most likely surrogate source when needed.

The next step is assessing the percentage of total public water supplied to households; this data was used to deduce the ratio of supplied water to connected services.

Table A9.6 Ratios used for households vs. connected services for supply water

Country	Households consumption /total public water supplied (%)	Sources and hypothesis	Households consumption ratio (m^3/day/person)	Connected services consumption ratio (m^3/day/person)
Austria	0.75	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.13	0.04
Belgium	0.68	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.12	0.06
Bulgaria	0.64	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.09	0.05

Country	Households consumption /total public water supplied (%)	Sources and hypothesis	Households consumption ratio (m³/day/person)	Connected services consumption ratio (m³/day/person)
Cyprus	0.67	estimation with Spain data	0.20	0.10
Czech Republic	0.63	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.09	0.05
Germany	0.79	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.12	0.03
Denmark	0.61	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.12	0.08
Estonia	0.78	Estimation with Poland data	0.07	0.02
Spain	0.67	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.15	0.08
Finland	1.00	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.17	0.00
France	0.87	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.17	0.02
Greece	0.67	Estimation with Spain data	0.10	0.05
Hungary	0.69	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.09	0.04
Ireland	0.61	Estimation with Denmark data	0.26	0.17
Italy	0.67	Estimation with Spain data	0.17	0.08
Lithuania	0.78	Estimation with Poland data	0.06	0.02
Luxembour g	0.64	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.15	0.09
Latvia	0.78	Estimation with Poland data	0.23	0.06
Malta	0.67	Estimation with Spain data	0.13	0.06
Netherlands	0.59	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.11	0.08
Poland	0.78	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.09	0.02
Portugal	0.67	Estimation with Spain data	0.12	0.06
Romania	0.42	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.05	0.07
Sweden	0.73	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.16	0.06
Slovenia	0.81	http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-NQ-03-013/FR/KS-NQ-03-013-FR.PDF	0.13	0.03
Slovakia	0.63	Estimation with Czech Republic data	0.10	0.06
United Kingdom	0.61	Estimation with Denmark data	0.17	0.11

Source: Data from Pöyry interim report 1, report June 2012, reformatted.

Comment, when ration has been re-estimated from another country source (assimilating foreign country's practice to country X), the line is marked in red, to pinpoint the absence of local source.

From the previous table and total volumes in Table A9.3, it comes that the ratios in the Table A9.6 are estimated from foreign sources in 42 % of cases (as weighted by total abstracted volumes). This means that much supplementary information should be sought for to improve the accuracy of the accounts.

The data from Table A9.6 are then entered into reference table TP_Act_domes which use id sketched in Figure A9.5, page 202 in this annex. For the time being this table is built around a country based coefficient; would more data become available more detailed (e.g. NUTs level coefficients) could be used as well.

5.4 Activity volumes related to entities

The volume of activity (population) of each urban entity was counter calibrated with the capacities of the wastewater treatment plant connected, when such information is provided. Indeed, each domestic activity is understood as linked to a sewage network. To calculate the level of population connected to each sewage network, the percentage of wastewater treatment plant capacity connected to each urban entity was used.

During the data processing, many scenarios were possible:

1. for an urban entity for which all the capacities of the wastewater treatment plants were populated, each treatment plant was allocated a share of population based on its relative capacity;
2. for an urban entity for which the capacities of the wastewater treatment plants were not given, the population was apportioned to the number of treatment plants.
3. for an urban entity in which some capacities were given and others were not, two scenarios were possible.
    - Population is larger than the sum of the capacity of the given wastewater treatment plants. In this case, an estimation of the capacity of the remaining wastewater treatment plants using the population difference was computed using the percentages with that new capacity breakdown. Then these percentages were used to apportion the population of the domestic entities connected to each sewage network ([98]).
    - population is less than the sum of the capacity of the given UWWTPS. Since the breakdown of population amongst WWTP is unknown, the extra capacity of wastewater treatment plants was assumed useless for the calculations.

The urban population is assumed being 100 % connected to a drinking water treatment plant, so in that case, the water source is 'water collection treatment and supply' that corresponds to the ISIC code 3600.

By contrast, rural population is understood being only partly connected to a drinking water treatment plant. The other part of the rural population is supplied by unitary household abstraction (individual drillings) whose water comes from the groundwater resource. So, the connecting ratio of the population to the public water supply network for each European countries found on Eurostat website was used in this case.

Table A9.7 Ratios of population connected to public supply network

Country	% of population served by public water supply	Year	Sources and hypothesis
Austria	0.95	2008	Eurostat: env_wat_pop
Belgium	1	2009	Eurostat: env_wat_pop
Bulgaria	0.99	2009	Eurostat: env_wat_pop
Cyprus	1	2009	Eurostat: env_wat_pop
Czech Republic	0.92	2007	Eurostat: env_wat_pop
Germany	0.99	2007	Eurostat: env_wat_pop
Denmark	0.97	2002	Eurostat: env_wat_pop
Estonia	0.8	2009	Eurostat: env_wat_pop

([98]) This rather complex procedure was set up to mitigate the case where several WWTP have a cumulated capacity less than the connected population and where several other WWTP serve as well the same persons but don't have their capacities populated. To allocate persons over all WWTP (with and without capacity) first the known persons breakdown has been used (WWTP with data) and the non-served persons were allocated to the non-documented WWTP.

Country	% of population served by public water supply	Year	Sources and hypothesis
Spain	0.94	2008	estimation with Portugal data
Finland	0.85	2007	estimation with Sweden data
France	0.99	2001	Eurostat: env_wat_pop
Greece	0.94	2007	Eurostat: env_wat_pop
Hungary	0.95	2009	Eurostat: env_wat_pop
Ireland	0.85	2007	Eurostat: env_wat_pop
Italy	0.94	2008	estimation with Portugal data
Lithuania	0.76	2009	Eurostat: env_wat_pop
Luxembourg	1	2009	Eurostat: env_wat_pop
Latvia	0.76	2009	estimation with Lithuania data
Malta	1	2009	Eurostat: env_wat_pop
Netherlands	1	2008	Eurostat: env_wat_pop
Poland	0.87	2009	Eurostat: env_wat_pop
Portugal	0.94	2008	Eurostat: env_wat_pop
Romania	0.55	2009	Eurostat: env_wat_pop
Sweden	0.85	2007	Eurostat: env_wat_pop
Slovenia	0.91	2002	Eurostat: env_wat_pop
Slovakia	0.86	2009	Eurostat: env_wat_pop
United Kingdom	0.97	2002	estimation with Denmark data

Source: Data from Pöyry interim report 1, report June 2012, reformatted.

5.5 Specific volumes and corrections

For the time being, leakages from the sewage network to the environment are not included in the consolidation process due to lack of data but the structure in table V_reseau (field F_S_MN) has been set in anticipation and computation could take stock of this information since these tables are used in the calculation to transport water and pollution to the connected.

The abstraction volumes produced for public water supply are filled in the table V_act_indus. The type of industry correspond to the ISIC code 3600 'Water collection, treatment and supply'. As industrial process, water supply uses water as intermediate fluid: the water treatment process produces sludge that has to be treated. Specific abstraction has been incorporated for that specific treatment. The water comes from the same origin as the water abstracted for the public supply, but the water destination is the type 'rivers'. The water abstracted for the sludge treatment represents 1.5 % of the water abstracted for the public supply. Two examples of water treatment plants in France have been used to determine this percentage (two water treatment plants in the local council called SIDECM, France); the percentage could hence be improved in further calculations.

6 Suggestions for data collection improvement

6.1 Strengths and weaknesses of the current data collection scheme

Water uses by domestic and urban is undoubtedly the most achieved water uses set of information of the whole process. Being the most achieved does not mean fully satisfactory.

Main strengths are that population benefits from a rather accurate spatialised reference of activities (population density). It is however a pity to have to back on a US data sets because EU data sets (more accurate and more precise) do not cover the EEA area.

The second positive point is that the stratified approach allowed allocating the data collection effort the most effectively: largest entities are probably best documented.

The weaknesses are many and categorise as follows.

- Spatial delineation of cities still improvable; this could be well achieved in fully integration the urban Atlas, the Urban audit and the UMZ approaches to end-up with common understanding of the largest cities and hence providing a seamless reference of cities for strata L and M.
- At the opportunity of integrating, consider in the data provisioning the water sources and location of supplies. There is there an important conceptual modelling, well started but not yet implemented in the current Eionet questionnaires, for example.
- To this end, redefine (with SES) the structure data collection for all spatial superstructure (see definition of time-series, spatial superstructure and spatial infrastructure data sets in Subsection 1.3.2 page 13 and thereafter).
- Deeply revise the data collection scheme for water uses (see below).
- Deeply improve the data collection scheme for UWWTP (see below).

However, there are important improvements that are categorised in the next sections.

6.2 Revising the data collection scheme for water uses

Current water uses collection scheme for urban and domestic is based on aggregated statistics at certain level (country for Eurostat, can be RBD or hopefully sub-units for Eionet). In the second cases, data is not populated.

The main assumption is that Eionet is not a good technical file for such data: good sources are cities networks and professionals. Discussions with European Federation of National Associations of Water Services (EUREAU) representatives suggest that while professionals have the information (more detailed than requested), they do not own this information. In contrast, Eionet owns information that it does not have.

The development could be carried out in two steps:

1. Propose to Eionet a list of cities in stratum L and in stratum M based on the following criteria:
    a. That are processed as L all cities in L or M by size if the detailed data sources are already present in the country ([99]);
    b. all cities over xxx inhabitants (e.g. over 250 000 inhabitants, value to synchronise with the Urban projects) belong to L and should be documented individually;
    c. all cities not L and hence M (their size is larger or equal to inclusion in the UWWTP directive) are identified (spatially defined as superstructure and hence related to the spatial infrastructure) and their technical information modelled. Volume of activity is inhabitants, and technical coefficients preferably derived from regional information(see below);
2. Under the aegis of Eionet, cooperate with EUREAU to populate required information. The information is again apportioned in three classes:
    a. Stratum L document population, and sources (with monthly volumes, direct or reconstructed thanks to specialist's advice),

([99]) This could be the case, for example, in France where ONEMA is developing, populating and maintaining an open database on water supplies and uses. It is not fully operational yet.

b. Stratum M, use statistics from professionals (jointly with Eurostat) to create these data (possibly seasonalised)

c. Stratum S: to be computed by difference.

6.3 Revising the data collection scheme for UWWTP

The UWWTP directive reporting presents three categories of gaps.

1. There is no clear relationship between the 'agglomeration' that captures sewage and domestic entities, such as 'cities'. Clarification on this point is essential to simplify the use of data and avoid difficult, time-consuming and poorly accurate guessing.
2. Some countries report only the main elements and not the installed capacity; some countries don't report at all. These gaps should be bridged.
3. Some countries do not have to report (e.g. Switzerland and Norway) and have this information. A special ad hoc canal of information should be created within Eionet for such cases. This is a systematic gap in Eionet when considering the data that are supposedly reported through directive canals.

There is no update mechanism for these superstructures that are quickly changing. This is all the more important that some countries have required that the Water accounting should develop more over time to capture mid-term changes. In many countries, the installed capacity has grown by a four- to five-fold factor since 1980.

Annex 10 Manufacturing and cooling water data collection

1 *Scope of data production*

1.1 Common rationales

Close to 60 % of total water abstractions in Europe are used for energy production and manufacturing industries ([100]). This figure seems to contradict the lay-person perception that agriculture is the main user. However, cooling using once-through technology involves massive volumes that are immediately returned, often at temperatures above the ambient receiving water. Energy production is the activity with the largest abstraction: 44 % of the total in Europe, mostly using these volumes for cooling.

It was first envisaged to separate water uses in relation to cooling (understood as cooling energy production plants) from the manufacturing industry. During the processing of data, it was discovered that such a separation would rather complicate the reporting and would add some confusion: many manufacturing activities use quantities of water for cooling as well.

There are two principle uses of water in manufacturing and energy production — the cooling of industrial installations, including large combustion plants and nuclear power plants, and the processing of materials to derive products.

The processing of available data to derive water volumes has been carried out largely based upon the data sources. All non-nuclear data are primarily sourced from the European Pollutant Release and Transfer Register (E-PRTR) database (completed by branch-specific data sources). By contrast, for the Nuclear Energy sector, which is not included in the E-PRTR database, the Platts database was used instead. As a result, the analysis is presented by industrial sectors.

1.2 Meeting the stratification principles

The stratification principles should apply to all industrial issues. In practice, stratification requires a reference statistical population that can be broken down according to certain criteria and processed accordingly. Primary criterions are *mutatis mutandis*, quantity of parameter and location — based on criteria that apply well to urban and domestic uses.

For the time being, there is no such population in the industrial sector. Available information is shared across three blocks: 1) statistical aggregates providing total volumes by sector for example (i.e. Eurostat statistics); 2) list of sites with no suitable information for stratification, since they are not populated and truncated (i.e. E-PRTR); and 3) sectoral information from certain professional associations.

The only source that shares some specifications that could help in stratifying the plants is the Platts database ([101]); however, the low quality of positioning largely prevents such use. Information used from this database meets licensing requirements.

1.3 Common operational definitions

([100]) Annex widely inspired from the consultant's report: all data and most technical comments are from P. Campling. Conclusions and suggestion as well as framing information are from EEA drafter.
([101]) Purchased from http://www.platts.com/, comprises information on the energy production plants: nuclear, combustion and hydropower. Many locations are highly erroneous since they are based on geographical names that are ambiguous due to truncation and/or misspelling.

The data to be reported are volume abstracted and resource used (surface, sea and groundwater), volumes used (as evaporated or incorporated) and volumes returned. If points of abstraction and return are distant, their position should be mentioned. **Water use** is a measure of the amount of water that is withdrawn from an adjacent waterbody (lakes, streams, rivers, estuaries, etc.). **Water consumption** refers to water that is lost, typically through evaporation. Due to the evaporation losses, closed-loop cooling technology consumes roughly twice the amount of water consumed by a once-through cooling system. **Water return** refers to water that is returned to the river system. This water does not need treating, because it is only used for cooling; however, in some cases the temperature of the returning water is higher than that of the river water, which can have an adverse effect on the ecology.

In most cases, the required information is not available and must be reconstructed. There are two major issues in addressing water for cooling since, depending on the technology, once-through or cooling tower (also called recirculating systems) processes are used. In the first case, large volumes are abstracted (and hence require sufficient resources), heated and returned close to the abstraction point. By contrast, in cooling tower technology, small volumes are abstracted; these are largely evaporated and concentrated refuse is rejected.

When no other information is provided, and if the water quality required for a process is 'drinking water standard', then groundwater is set as the resource abstracted.

2 Common data sources to cooling and manufacturing

2.1 European Pollutant Release and Transfer Register (E-PRTR)

The E-PRTR is the new Europe-wide register that provides environmental data from industrial facilities in European Union Member States and in Iceland, Liechtenstein, Norway, Serbia and Switzerland (http://prtr.ec.europa.eu/). It contains data reported annually by some 28 000 industrial facilities covering 65 economic activities across Europe. For each facility, information is provided concerning the amounts of pollutant releases to air, water and land, as well as off-site transfers of waste and of pollutants in wastewater.

Additional information includes number of employees, NACE code and geographic position of the facility (Latitude, Longitude). Unfortunately, for the calculation of water accounts the information on wastewater is scarce (it is not obligatory), and there is no information on water use and production levels (not asked for in the Protocol on Pollutant Release and Transfer Registers (2003)).

The E-PRTR database was the primary source of information for the following reasons:

- Water use for cooling is an important component of water use by some industrial sectors, with some industrial complexes with high energy demands having their own LCPs, with a rated thermal input equal to or greater than 50 MW. The consultant ([102]) has developed an approach that estimates the amount of water needed to cool installations on the basis of the quantity of CO_2 emissions (reported in the E-PRTR database).
- The geographic position of the facility with the amount of emissions common to an industrial sector can be used to weight and spatialise water use estimates calculated at the national or even European level.

However, there are uncertainties with this approach. Firstly, the reported geographic position of the facility may be the position of the company's administrative headquarters — and not where the

([102]) Vito, under the Pöyry lead contract with the Environment DG.

industrial activities are taking place. Secondly, the facility's activities may be varied and diverse — therefore including more than one NACE activity. Last but not least, there is not a direct relationship between CO_2 emissions and water use, especially if the cooling file is uncertain.

2.2 Best Available Techniques (BAT) Reference documents (BREFs)

BREFs are the main reference documents used by competent authorities in Member States when issuing operating permits for the installations that represent a significant pollution potential in Europe. These BREFS are produced and hosted by the European IPPC Bureau, which is an action of the Sustainable Production and Consumption Unit of the Institute for Prospective Technological Studies (IPTS), Seville, Spain (http://eippcb.jrc.es/index.html).

Individual BREFs are used for the following reasons:

- The BREFs provide valuable information on water use coefficients from different industrial facilities that are representative for the particular sector — so that water use estimates can be calculated if production levels are known.
- In some cases, they also provide information on whether the water needs to be of drinking water quality (e.g. for the food processing sector), meaning that groundwater is the preferred source, or whether water is only needed for cooling, meaning that large surface water sources are needed (e.g. for LCPs).
- If no statistical information on production levels can be obtained from statistical sources, then the BREFs also provide an overview of European production levels for the sector; however, this information is often out-dated (pre-2000).

2.3 Statistical data on industrial production levels

Statistical data on industrial production levels are hard to come by because of confidentiality issues. Eurostat provides mainly production information in terms of the value and not the quantity. Therefore, a web search has been carried out to look for data from industry representatives, FAO and UNECE.

2.4 The Platts database

The EEA has purchased the Platts database and the European Topic Centre Spatial Integrated Assessments has spent time and resources working with this database to ensure that the information is properly geocoded and linked with ECRINS. For each facility in the Platts database the total energy capacity is reported in MWs.

This database has been used as well to help in locating LCPs to cross-check or complement the EPER database.

2.5 Specific associations

Supplementary data has been obtained and processed from the following professional associations, either through publications from professional association websites or personal contact with representatives:

- Paper and pulp industry: CEPI (Confederation of European Paper Industries, Avenue Louise, 250 Box 80, B-1050 Brussels, Belgium, Tel.: +32 26274911, Fax: +32 26468137) http://www.cepi.org/about-us/how-cepi-works
- Oil refineries: CONCAWE (Conservation of clean air and water in Europe, Boulevard du Souverain 165, B-1160 Brussels, Belgium, Tel.: +32 25669160, Fax: +32 25669181) http://www.concawe.be/Content/Default.asp?

- Cement industry: CEMBUREAU (The European Cement Association, Rue d'Arlon 55, BE-040 Brussels, Belgium, Tel.: +32 22341011) http://www.cembureau.be/
- Electric power industry: Eurelectric (Union of the Electricity Industry, Boulevard de l'Impératrice, 66, B-1000 Brussels, Belgium, Tel.: +32 25151000) http://www.eurelectric.org/
- Steel industry: EUROFER (The European Steel Association, Av. de Cortenbergh, 172, B-1000 Brussels, Belgium, Tel.: +32 27387920) http://www.eurofer.org/index.php/eng/Home and WorldSteel (World Steel Association, Rue Colonel Bourg 120, B-1140 Brussels, Belgium, Tel.: +32 27028900) http://www.worldsteel.org/

3 Methodology

The approaches deal with three groups of industrial water uses, considered from data source constraints: LCPs (mainly for energy production), nuclear power plants (electricity production) and manufacturing industry (mainly production of goods, cooling being within the process). The following sectors represent 90 % of the water used in the manufacturing industry:

1. Basic metals (NACE 27) (Production and Cooling)
2. Chemicals and chemical products (NACE 24) (Production and Cooling)
3. Pulp, paper and paper products (NACE 21) (Production only)
4. Food products and beverages (NACE 15) (Production and Cooling)
5. Non-metal mineral products (NACE 26) (Production and Cooling)
6. Textiles (NACE 17) (Production only)
7. Fabricated metal products (NACE 28) (Production and Cooling)
8. Coke, refined petroleum products and nuclear fuel (NACE 23) (Production and Cooling)
9. Mining and quarrying (NACE 14) (Production only).

The work carried out by the consultant focuses on these sectors as they represent the majority of water use accounts for the manufacturing industry.

The OECD/Eurostat Joint Questionnaire (JQ) on Inland Waters is the only source of European-wide information on water use by industry. Table 2 of the JQ is the Annual water abstraction by source and sector and seeks to establish the volumes of water abstracted by countries from the different sources for different sectors of water use. In this table a distinction is made between the abstraction of water by different sources (fresh surface water and fresh groundwater; non-freshwater resources, desalinated water and re-used water), and the total volume of water supplied is distinguished by means of different supply categories (public supply, self-supply, other supply). However, the data from the JQ have major problems of 'data gaps and inconsistencies (e.g. sub-category water use larger than main category' (Wriedt and Bouraoui, 2011). The JRC uses interpolation techniques to fill these gaps and disaggregate information to NUTS3 areas by weighting coefficients such as gross domestic product (GDP) and population density. This approach is suited to gridded modelling and not applicable to water accounting rules, since the 'territory of reference' must be a hydrologically consistent sub-basin.

Another approach has been implemented, which can be regarded as a complementary more bottom up, SEEAW-suited approach to derive 'replaceable' data — where surrogate information is used when actual information is not obtainable (due to confidentiality issues), or the necessary policy framework is in place to oblige industrial facilities to report on water uses.

The starting point of the SEEAW-suited approach is the E-PRTR database, which provides data concerning emissions and waste from 28 000 facilities across Europe (IPPC installations). However, this provides no useful information on volume of either activities or water use, although geo-referenced information allows for a more precise connection to water stocks and flows than the JRC approach. The challenge is therefore to relate emissions information to activity and/or water use. A further problem is that much of the information is voluntary and the percentage rate of reporting is low.

The approach focuses on the main water-consuming industries (listed above), and for each industrial sector surrogate information is selected to either estimate water use or volume of activities. This has been done by reviewing the BAT information from BREFS published by the JRC-IPTS and using water use coefficients related to currently published volume activities. The concerns with the latter approach are that ranges can be wide, there are broad assumptions on technology used, and there are major problems in finding activity data on production levels.

It is fully acknowledged that this approach may be questioned. It is for the time being the only one feasible and capable of providing values that can be entered into the SEEAW processing. The strength is that final data tables are populated with acceptable statistical units, volumes of which are likely to be updated. The weakness is that volumes are likely to be somehow incorrect. As a result, this method is expected to foster further cooperation with the ad hoc entities and help in defining a better spatial information system of industrial activities in relation to the environment

4 Processing the cooling for energy production and as ancillary demand

4.1 Large combustion plants (LCPs)

There are two important pieces of information that are not known, concerning water use for cooling at IPPC facilities:

1. Is the fuel used in the facility gas or coal?
2. Is the cooling system a once-through system or a recirculating system?

The assumptions are carried out in two steps: first assume that the facility is gas-fired — and that the cooling system is either once-through or a recirculating system. In a next iteration differentiate between coal- and gas-fired industrial systems on the basis of SO_x emissions.

Since a high percentage of the water used for cooling in IPPC facilities is returned to its source, distinctions are made between use, consumption and return. The following tables provide the assumptions used for estimating water use for cooling (Table A10.1 and Table A10.2).

Table A10.1 Efficiency assumptions for industrial cooling at IPPC facilities

Type of fuel to power plants	Electricity Efficiency	Steam generator losses	Losses to coolant water
Gas-powered LCPs	55 %	10 %	40 %

Table A10.2 Approach to estimating power based on CO_2 emissions

Gas-fired LCP	55 %	Efficiency
CO_2 factor	56	kg CO_2/GJ from gas power
	1	MWh electricity
	367	kg CO_2
	0,35	MWh to be cooled
	1,047	kg CO_2 per MWh to be cooled

During the processing of data, supplementary clues were obtained from Flanders plants. On the basis of information received, the factor used to calculate the litres of water per MWh has been revised from 72 608 [l/MWhel] ([103]) to 43 000 [l/MWhel], final data being used for data processing.

Table A10.3 Formulas to estimate water use, water consumption and water return for both gas-fired once-through technology for industrial cooling and gas-fired wet cooling tower technology for industrial cooling

Step in process	Technology	Formula
Water use	Gas (once-through technology) litres of water per year	= 43 000 [l/MWh_{el}] (a) x CO_2 [kg] / 367 (c) [kg/MWh_{el}] *
Water use	Gas (wet cooling tower technology) litres of water per year	= 3 930 [l/MWh_{el}] (a) x CO_2 [kg] / 367 (c) [kg/MWh_{el}] *
Water consumption	Gas (wet cooling tower technology) litres of water per year	= 2 751 [l/MWh_{el}] (b) x CO_2 [kg] / 367(c) [kg/MWh_{el}] *
Water return	Gas (once-through technology) litres of water per year	= Water use
Water return	Gas (wet cooling tower technology) litres of water per year	= Water use - Water consumption

Comments:

(a): factor for water use obtained from annual water use for a typical LCP, cooling factors and efficiencies for once-through and wet cooling tower technologies (BREF Industrial Cooling systems);

(b): factor for water consumption obtained from annual water use for a typical LCP, cooling factors and efficiencies for once-through and wet cooling tower technologies (BREF Industrial Cooling systems);

(c) 367 obtained from 56 kg CO_2/GJ gas burned / 55% * 3.6 (converting from GJ to MWh).

There is no information available on the type of cooling technology used at the facility level in the source data. Up until now the wet cooling tower technology was assumed to represent minimum water use and the once-through technology to represent maximum water use. In addition, whether the water used is fresh or brackish ([104]) is unknown. This is guessed using a buffer analysis technique in GIS to assert whether a facility is within 5 km of the coast — if the facility is within 5 km of the coast line (as defined by the land boundary of each Member State's Economic Exclusion Zones), then the assumption is that a) the water source for use and return is brackish surface water, and b) the cooling technology is once-through. If the facility is not within this 5 km buffer zone, then the assumption is that a) the water source for use and return is fresh surface water, and b) the cooling technology is wet tower.

4.2 Nuclear power plants

Nuclear power plants require that specific factor assumptions are adopted (see Table A10.4 and Table A10.5).

Table A10.4 Efficiency assumptions for industrial cooling at nuclear energy facilities

Type of fuel to power plants	Electricity Efficiency	Steam generator losses	Losses to coolant water
Nuclear energy plants	35 %	0 %	65 %

Comment: the energy yield is less than in Table A10.1

([103]) Megawatts-hour of produced electricity.
([104]) 'Brackish' is a fuzzy term. In English, it covers salinity ranges from 0.5 to 30 g/l; however, its French translation 'saumâtre' should apply to waters in the range 1 to 10 g/l. In this report it means 'salt water, any concentration, below or equal to maximum sea salinity'. It does not encompass brines (French 'saumures'), the salinity of which is beyond 50 g/l.

Table A10.5 Formulas to estimate Water Use, Water Consumption and Water Return for nuclear energy plants

Step in process	Technology	Formula
Water use	Nuclear (once-through technology) litres of water per year	= 168 814 [l/MWh$_{el}$] (a) * Power Sum [MW$_{el}$] * 7 000 [hours]
Water use	Nuclear (wet cooling tower technology) litres of water per year	= 9 137 [l/MWh$_{el}$] (a) * Power Sum [MW$_{el}$] * 7 000 [hours]
Water consumption	Nuclear (once-through technology) litres of water per year	= 3 198 [l/MWh$_{el}$] (b) * Power Sum [MW$_{el}$] * 7 000 [hours]
Water consumption	Nuclear (wet cooling tower technology) litres of water per year	= 6 396 [l/MWh$_{el}$] (b) * Power Sum [MW$_{el}$] * 7 000 [hours]
Water return	Nuclear (once-through technology) litres of water per year	= Water use - Water consumption
Water return	Nuclear (wet cooling tower technology) litres of water per year	= Water use - Water consumption

Comments:

(a) factors for water use obtained from annual water use for a typical nuclear energy plant, cooling factors and efficiencies for once-through and wet cooling tower technologies (BREF Industrial Cooling systems and http://www.world-nuclear.org/);

(b) factors for water consumption obtained from annual water consumption for a typical nuclear energy plant, cooling factors and efficiencies for once-through and wet cooling tower technologies (BREF Industrial Cooling systems and http://www.world-nuclear.org/).

The same buffer analysis technique in GIS, as above, was applied to assert whether a facility is cooled using once-through or wet tower cooling technologies, and whether the water source for use and return is surface brackish or surface freshwater.

4.3 Distribution of large combustion plants (Electricity, gas, steam and hot water supply and other sectors) across Europe

The majority of LCPs (769 or 74 %) in the E-PRTR database fall under NACE Code 40 - Electricity, gas, steam and hot water supply (Table A10.6). However, there are 20 other NACE sectors that also have facilities that include LCPs (Table A10.6); the most important of these are: NACE Code 24 - Manufacture of chemicals and chemical products (70 or 7 %); NACE Code 11 - Extraction of crude petroleum and natural gas (61 or 6 %); NACE Code 21 - Manufacture of pulp, paper and paper products (40 or 4 %); NACE Code 23 - Manufacture of coke, refined petroleum products and nuclear fuel (25 or 2 %); and, NACE Code 15 - Manufacture of food products and beverages (15 or 1 %).

Table A10.6 Industrial sectors with LCPs, according to the E-PRTR database 2004

NACE Sector Code	NACE Sector Name	Number of industrial complexes with LCPs	Percentage of total number of LCPs
10	Mining of coal and lignite; extraction of peat	5	0.5 %
11	Extraction of crude petroleum and natural gas	61	5.9 %
14	Other mining and quarrying	3	0.3 %
15	Manufacture of food products and beverages	15	1.4 %
20	Manufacture of wood and of products of wood and cork, except furniture;	9	0.9 %
21	Manufacture of pulp, paper and paper products	40	3.8 %
22	Publishing, printing and reproduction of recorded media	1	0.1 %
23	Manufacture of coke, refined petroleum products and nuclear fuel	25	2.4 %
24	Manufacture of chemicals and chemical products	70	6.7 %
25	Manufacture of rubber and plastic products	1	0.1 %
26	Manufacture of other non-metallic mineral products	2	0.2 %

NACE Sector Code	NACE Sector Name	Number of industrial complexes with LCPs	Percentage of total number of LCPs
27	Manufacture of basic metals	18	1.7 %
28	Manufacture of fabricated metal products, except machinery and equipment	1	0.1 %
34	Manufacture of motor vehicles, trailers and semi-trailers	1	0.1 %
40	Electricity, gas, steam and hot water supply	769	73.9 %
60	Land transport; transport via pipelines	10	1.0 %
63	Supporting and auxiliary transport activities; activities of travel agencies	1	0.1 %
70	Real estate activities	1	0.1 %
74	Other business activities	4	0.4 %
85	Health and social work	1	0.1 %
90	Sewage and refuse disposal, sanitation and similar activities	3	0.3 %
	Total LCPs	1 041	

Comment: greyed lines recall the major activities mentioned in text.

All these categories use to some extent water for cooling and are addressed in the next section.

4.4 Distribution of manufacturing sectors with cooling facilities, but without LCPs

Some manufacturing sites do not have their own LCPs, but they still include industrial cooling facilities — the same approach for estimating the water needs for cooling LCPs is used.

Table A10.7 Industrial sectors with industrial cooling facilities, but without LCPs

Annex I Activity Name	NACE Sector Code	NACE Sector Name	Number of facilities	%
Coal gasification and liquefaction plants	11	Extraction of crude petroleum and natural gas	4	0.41 %
Slaughterhouses, milk, animal and vegetable raw materials	15	Manufacture of food products and beverages	25	2.59 %
Surface treatment or products using organic solvents	17	Manufacture of textiles	2	0.21 %
Pulp, paper or board production	21	Manufacture of pulp, paper and paper products	64	6.63 %
Coke ovens	23	Manufacture of coke, refined petroleum products and nuclear fuel	13	1.35 %
Mineral oil and gas refineries	23	Manufacture of coke, refined petroleum products and nuclear fuel	81	8.39 %
Basic inorganic chemicals or fertilisers	24	Manufacture of chemicals and chemical products	55	5.69 %
Basic organic chemicals	24	Manufacture of chemicals and chemical products	74	7.66 %
Biocides and explosives	24	Manufacture of chemicals and chemical products	1	0.10 %
Pharmaceutical products	24	Manufacture of chemicals and chemical products	3	0.31 %
Cement, lime, glass, mineral substances or ceramic products	26	Manufacture of other non-metallic mineral products	364	37.68 %
Metal industry	27	Manufacture of basic metals	134	13.87 %
Production of carbon or graphite	27	Manufacture of basic metals	2	0.21 %
Disposal of non-hazardous waste and landfills	90	Sewage and refuse disposal, sanitation and similar activities	22	2.28 %
Disposal or recycling of animal carcasses and animal waste	90	Sewage and refuse disposal, sanitation and similar activities	1	0.10 %
Disposal/recovery of hazardous or municipal waste	90	Sewage and refuse disposal, sanitation and similar activities	121	12.53 %
		Total IPPC Facilities	966	

Comment: greyed lines recall the major activities mentioned in text.

The most important sectors with cooling facilities, but without LCPs, are NACE 26 - Manufacture of other non-metallic mineral products (364 or 38 %); NACE Code 90 - Sewage and refuse disposal, sanitation and similar activities (121 or 13 %); NACE Code 27 - Manufacture of basic metals (136 or 14 %); NACE Code 24 - Manufacture of chemicals and chemical products (133 or 14 %); and, NACE Code 23 - Manufacture of coke, refined petroleum products and nuclear fuel (94 or 10 %) (Table A10.7).

5 Summary estimates of water use for cooling

5.1 Results found from estimations

The level of water use, water consumption and water return depends on the type of cooling technology. The once-through technology uses much more water than the wet cooling tower technology, but nearly all of the water is returned directly to where the water was abstracted, albeit at a significantly higher temperature.

From the E-PRTR data set it is not possible to know which cooling technology is being used — **this would be a useful additional question to be asked of the facility**. It is most likely that facilities located close to the sea and large rivers — with ample supplies of fresh or brackish water — will use once-through technologies. Buffer zone analysis has been applied to guess whether the water source for use and return is brackish or fresh surface water, and from this information assume that a brackish source means once-through cooling technology and a fresh source means wet tower cooling technology.

Table A10.8 provides an overview of water use, water consumption and water return volumes for all LCPs in the EU-27. The analysis indicates that 31 % of all LCP facilities are located close to the coast (i.e. within 5 km), meaning that more than 80 % of water taken in by these LCP facilities is brackish rather than freshwater.

Table A10.8 Global overview of Water Use (intake), Water Consumption (loss) and Water Return for cooling all the Large Combustion Plants in the EU-27, for both brackish and fresh surface water sources

Item	value	Unit
Number of facilities using brackish river water	321	count
Number of facilities using fresh river water	720	count
Water Use (intake) - Brackish	46 009	hm^3/y
Water Use (intake) - Fresh	10 775	hm^3/y
Water Consumption (loss) - Brackish	0	hm^3/y
Water Consumption (loss) - Fresh	7 542	hm^3/y
Water Return - Brackish	46 009	hm^3/y
Water Return - Fresh	3 232	hm^3/y

Table A10.9 indicates that for the energy producing sector (NACE 40) 29 % of all LCP facilities are located close to the coast (i.e. within 5 km), and 80 % of water used for cooling is brackish rather than freshwater, whereas for the oil refining industry (NACE 23) 38 % of all LCP facilities are located close to the coast (i.e. within 5 km), and 86 % of water used for cooling is brackish rather than freshwater (Table A10.10).

Table A10.9 Global overview of Water Use (intake), Water Consumption (loss) and Water Return for cooling Large Combustion Plants used to produce electricity (NACE 40) in the EU-27, for both brackish and fresh surface water sources

Item	Value	Unit
Number of facilities using brackish river water	221	count
Number of facilities using fresh river water	548	count
Water Use (intake) - Brackish	39 144	hm^3/y
Water Use (intake) - Fresh	9 525	hm^3/y

Water Consumption (loss) - Brackish	0	hm^3/y
Water Consumption (loss) - Fresh	6 668	hm^3/y
Water Return - Brackish	39 144	hm^3/y
Water Return - Fresh	2 858	hm^3/y

Table A10.10 Global overview of Water Use (intake), Water Consumption (loss) and Water Return for cooling Large Combustion Plants used at oil refinery installations (NACE 23) in the EU-27, for both brackish and fresh surface water sources

Item	Value	Unit
Number of facilities using brackish river water	9	count
Number of facilities using fresh river water	15	count
Water Use (intake) - Brackish	1 687	hm^3/y
Water Use (intake) - Fresh	273	hm^3/y
Water Consumption (loss) - Brackish	0	hm^3/y
Water Consumption (loss) - Fresh	191	hm^3/y
Water Return - Brackish	1 687	hm^3/y
Water Return - Fresh	82	hm^3/y

The average volume consumed per facility is roughly the same in both tables, making important the assessment of oil refineries since their local impact can be large, albeit their total volume is small.

The results of the water accounts analysis of the nuclear energy producing sector show that 62 % of the facilities are close to the coast and are therefore likely to use once-through cooling technology with brackish water for cooling purposes. A total estimate of 68 920 hm^3 of brackish water is used annually for cooling with 63 720 hm^3 (or 92 %) returning back to brackish water. For facilities not close to the coast the estimate is 5 704 hm^3 taken in for cooling, of which 3 993 hm^3 (or 70 %) is lost through evaporation and 1 711 hm^3 is returned to fresh surface water.

Table A10.11 Global overview of Water Use (intake), Water Consumption (loss) and Water Return for cooling nuclear energy production facilities in the EU-27, for both brackish and fresh surface water sources

Item	Value	Unit
Total number of facilities – Once-through, brackish	41	count
Total number of facilities - Wet tower, freshwater	26	count
Water Use (intake) - Brackish	68 920	hm^3/y
Water Use (intake) - Fresh	5 704	hm^3/y
Water Consumption (loss) - Brackish	5 190	hm^3/y
Water Consumption (loss) - Fresh	3 993	hm^3/y
Water Return - Brackish	63 730	hm^3/y
Water Return - Fresh	1 711	hm^3/y

5.2 Critical analysis

The outcomes of the process were presented to Member States and relevant stakeholders during a meeting organised by DG Environment, 7 September 2012. This consultation had been prepared by two documents drafted by the EEA and presenting, respectively, natural data sets issues (June 2012) and economic uses of water (September 2012).

A few educated comments were received, including those prepared by the Eurelectric association with relevant proposals, the main substance of which is summarised below.

These comments state on their perception of 'Inappropriate focus on water use in the cooling of Nuclear Power Plants Nuclear-fuelled power plants, with around 65 plants in Europe, make up only a fraction of the 10932 other non-nuclear thermal power plants where cooling systems used water. …, there would appear to be no reason to highlight separately the use of water in nuclear power plant. Thus, it is proposed to group all thermal generation means (e.g. nuclear, coal, gas, biomass …) in one table or to provide several tables for each type of energy …'

These comments suggest that Eurelectric has a detailed database of electricity production plants that might help (just considering the figures) in defining a stratified approach to populating the water uses at adequate locations.

As indicated in sections above, the single reason for such separation is the complete difference in data sources used to indirectly assess both the way cooling is carried out (typically: sea water possibly designated as 'brackish' or 'sea', to separate from freshwater involved in the hydrological cycle, hence not participating in the water balance, surface water with open or tower coolers). Even if the cases are few, this separation is kept for consistency with future calculations.

The second class of remarks is indeed important and was anticipated. The weakness of data sources is a major jeopardy for the accounts and the major target of the consultation. Eurelectric mentions, as a conclusion, that the range of uncertainty could be 0.5 to 8 factor, the factor itself being rather uncertain.

The use of CO_2 emissions as a surrogate variable for estimating water uses is questioned; however, this is the single proxy that could be found. This confirms the assessment in this report that current data flows are quite unrealistic for populating the SEEAW-demanded information.

As a summary, the analysis of remarks confirm that the main issues of data gaps in relation to electricity production and that apply to other categories of industrial activities are:

1. Absence of reference population, with a double criterion of activity size (that is in relation to potential abstraction and consumption) on the one hand, and spatial location (that is in relation to density of pressures ([105])) on the other hand.
2. Absence of appropriate information on both the volumes and sources of supply (and in this case activity values are not required, albeit they might be useful to calibrate technical coefficients), either activity volumes closely related to water volumes (and better than CO_2 emissions, for example) and preferably source of supply.

In its analysis, Eurelectric has not mentioned the turbined volumes; they represent information of secondary importance except where turbining diverts water from the basin of origin of water. By contrast, if the 'water wearing' indicator is to be set in place, this information becomes important. It can be reconstructed from simple information (total production and head loss).

It would be very effective if the sound and rational criticisms by Eurelectric would offspring tight cooperation and support in collecting accurate and comprehensive volumes of water abstracted/consumed and returned (including major turbining), considering as well which proxies can be accepted by the industry.

6 Estimates of water use for processing in the paper and pulp industry

6.1 First step: country aggregates estimation

In the absence of sufficient information in the E-PRTR, a side approach based on professional organisation statistics was carried out and then disaggregated at the relevant levels. This procedure

([105]) The author is conscious that the terms 'density of pressures' may seem pleonastic since a pressure is a density. What is meant is that in some areas many small pressures can together constitute a strong final pressure because they are close to each other and hence functionally act as a single lumped site.

is indeed tedious and indirect, raising possibilities of inaccurate data that will require improvement for systematic water accounts production.

The Members of the Confederation of European Paper Industries (CEPI, from the public website) account for about 95 % of the total number of paper mills in the EU, and an even higher share of EU pulp and paper production (CEPI, 2009). In Europe, about 18 % of all mills in the pulp and paper industry are integrated mills producing both virgin pulp and paper, although different degrees of integration occur (CEPI, 2009). Water is used for processing, cooling and boiler feed water. Much of the water used is recycled. According to CEPI (2010), 90 % of water is from surface water, 9.2 % of water is from groundwater and 0.8 % of water is from municipal water supply. From the Reference Document on BAT in the Pulp and Paper Industry (IPPC, 2001), water used to produce 1 tonne of pulp per year is on average 32.5 m^3 and water used to produce 1 tonne of paper and paper board per year is on average 10.3 m^3.

According to the Pulp and Paper Industry BREF, if a facility quotes a figure of more than 50 m^3/t/annum for pulp production this indicates that this water use combines water use for cooling and for processing, meaning that we cannot use this information to distinguish between coefficients for cooling and for processing.

Table A10.12 Water consumption coefficients for industrial processing of Pulp, Paper and Paperboard

NACE Sector	NACE CODE	Water Use m^3/t/year (min.)	Water Use m^3/t/year (max.)	Water Use m^3/t/year (average)
Manufacture of pulp	21.11	15	50	32.5
Manufacture of paper and paperboard	21.12			10.3

The Key Statistics 2010 publication from the European Pulp and Paper Industry (CEPI) indicates that between 2000 and 2010 the production levels for pulp production is quite stable, averaging around 40 000 Ktonnes per year. However, paper and paper board production declined from more than 92 000 Ktonnes per year in 2000 to 36 000 Ktonnes per year in 2009. This means that water use rates per year are also likely to vary for this particular sector.

Table A10.13 Key production figures for Pulp production and Paper and Board Production in Europe

NACE Sector	NACE CODE	2000 Ktonnes	2005 Ktonnes	2009 Ktonnes	2010 Ktonnes	Average Ktonnes
Manufacture of pulp	21.11	40 207	41 865	36 120	39 169	39 340
Manufacture of paper and paperboard	21.12	92 603	41 865	36 120	39 169	52 439

Source: Key Statistics 2010, European Pulp and Paper Industry, CEPI (CEPI, 2013).

Final processing combines the information on average water use coefficients with reported production levels at the national level to estimate annual water use for the manufacture of pulp and the manufacture of paper and paperboard (Table A10.13). The water used to manufacture pulp is in total 1 168 hm^3 and the water used to manufacture paper and paper board is 539 hm^3.

Table A10.14 Key production figures for Pulp production in Member States and estimated water use for processing

Manufacture of pulp	Production %	Production KT	Average water use (hm^3/y)
Sweden	30.3	11 920	387.4
Finland	26.8	10 543	342.7
Germany	7.1	2 793	90.8
Portugal	6.6	2 596	84.4
France	4.9	1 928	62.6
Norway	4.8	1 888	6.1
Spain	4.8	1 888	6.1

Manufacture of pulp	Production %	Production KT	Average water use (hm^3/y)
Austria	4.3	1 692	55
Rest of CEPI	10.4	4 091	133
Total	100	39 340	1 168.1

As can be seen, this estimate, from the existing data, does not meet the SEEAW requirement: only country-level aggregates can be produced, which will be disaggregated at plant location with a post-processing step (described in a next section).

Table A10.15 Key production figures for Paper and Paperboard production in Member States and estimated water use for processing

Manufacture of paper and cardboard	Production %	Production KT	Average water use (hm^3/y)
Germany	23.9	12 533	129.1
Finland	12.2	6 398	65.9
Sweden	11.8	6 188	63.7
Italy	9.3	4 877	50.2
France	9.1	4 772	49.2
Spain	6.4	3 356	34.6
Austria	5.2	2 727	28.1
United Kingdom	4.5	2 360	24.3
Rest of CEPI	17.5	9 177	94.5
Total	99.9	52 387	539.6

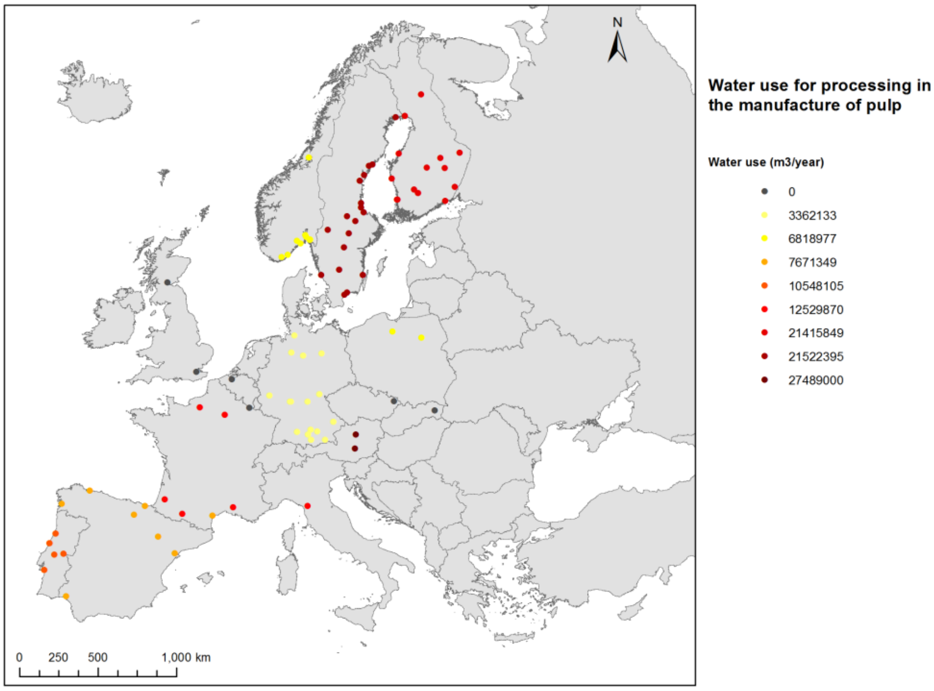
6.2 Second step: spatial disaggregation

It was initially expected that we would be able to spatialise national water use to facilities on the basis of using CO_2 emissions from the E-PRTR databases to weight volumes. The attempts to spatialise national aggregates back on the proposals exposed in the Inception Report. The purpose was to use the emissions reported by the E-PRTR Facilities to rank facilities according to their emissions and then weight them and nationally calculate for water use according to the level of emissions.

Unfortunately, the reporting performance in the E-PRTR for CO_2 emissions is only 29 %, with a range from 6 % to 100 % — this creates a problem for spatialisation.

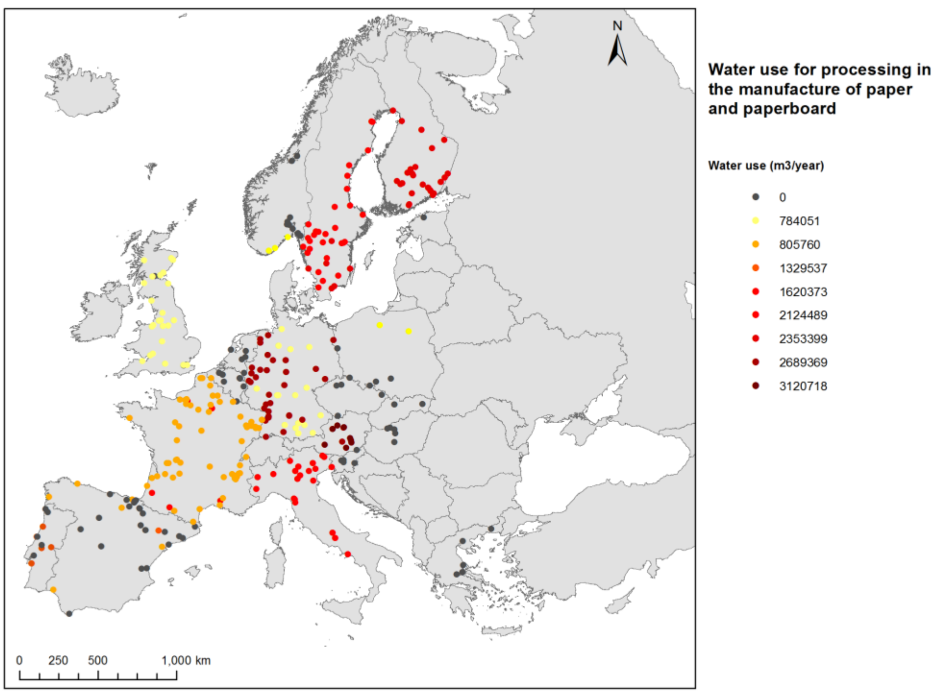
Due to the weak information contained in the E-PRTR, the national volumes could only be equally apportioned between all the Pulp and Paper facilities (respectively, Map A10.1 and Map A10.2); this is unrealistic and is not likely to give accurate information for any of the facilities — but it is the only way to flag up where demand for water from the Pulp, Paper and Paperboard sector is occurring at the local level. Until better information is received this is the only way to proceed with spatialising water use for the manufacturing sectors. This procedure unfortunately had to be repeated for other industrial branches.

Map A10.1 Spatial location of pulp facilities



Source: Vito in Pöyry final report, June 2012.

Map A10.2 Spatial location of paper and cardboard facilities



Source: Vito in Pöyry final report, June 2012.

6.3 Suggestions for improvement

Despite the rigorous approach undertaken, the result reflects that inconsistency of data cannot be overcome, meaning that satisfactory results are not presently possible:

1. total volumes are based on estimates of country-level production with standard reconstructed technical coefficients;
2. spatialisation cannot be accurately carried out because of the poor and inconsistent reporting in E-PRTR (even using very indirect proxy), whereas the objective information on water use is not requested in a pollution registry;
3. the very old information collected amongst countries makes it possible that some significant plants are missing.

A closer cooperation between CEPI and Eionet is required, and further discussions at country levels are needed, because it is clear that data does exist to produce the water use statistics, published by CEPI, at national level. CEPI is not party to the underlying data that is used to develop these annual statistics as these are sent in by each national association.

7 Estimates of water use for processing and cooling in the oil refinery sector

7.1 Step 1: estimating volumes

The oil refinery sector is a substantial user of water both for processing and cooling. At present there are two BREFs available to obtain water use coefficients for processing and cooling — the BREF from 2003 and draft BREF from 2012 that is available for consultation rounds. In addition, there is information on annual production capacity from the Oil and Gas Journal that we can use to estimate the water use (intake), water consumption (loss) and water return. A parallel approach is adopted to calculate water use for cooling, based on CO_2 emissions for the E-PRTR facilities.

CONCAWE has provided useful global information on water use, consumption and return, based on a comprehensive survey of its members, making it likely that estimates of both volumes and spatial distribution are accurate enough for the water accounting.

Table A10.16 Water use coefficients for the oil and gas refining sector (NACE 23)

Type of use	Average m³/T throughput/annum	Source
Freshwater usage	0.62	BREF 2003 p399
Freshwater process effluent	0.53	BREF 2003 p399
Freshwater process effluent	0.44	BREF 2012 p217
Freshwater effluent from cooling	0.08	BREF 2012 p217

Information on the capacity of throughput volumes and the geographical location of the major European oil refineries is obtained from the following websites ([106]):

- http://en.wikipedia.org/wiki/List_of_oil_refineries#Europe
- http://abarrelfull.wikidot.com/european-refineries

Using this information makes it possible to derive the following information on water use for processing in the oil refining sector — which is also mapped to the facility.

([106]) This source tells a positive message of the usefulness of Wikipedia, which is a much mobilised source for the accounts and the very negative message that such essential information must be collected and can be collected out of established data collection processes.

Table A10.17 Global overview of total production (throughput), Water Use (intake), Water Consumption (loss) and Water Return for processing in oil refining facilities in the EU-27

Total throughput Mtonnes/Year	**Water Intake** hm^3/year	**Water Consumption** hm^3/year	**Water Return** hm^3/year
776.08	481.17	69.85	411.32

Again, the selection of the cooling technology and water source is based on the GIS buffer analysis.

Table A10.18 Global overview of Water Use (intake), Water Consumption (loss) and Water Return for cooling oil refining facilities in the EU-27, for both brackish and fresh surface water sources

Item	value	Unit
Number of facilities using brackish river water	57	count
Number of facilities using fresh river water	34	count
Water Use (intake) - Brackish	8 907	hm^3/y
Water Use (intake) - Fresh	616	hm^3/y
Water Consumption (loss) - Brackish	0	hm^3/y
Water Consumption (loss) - Fresh	431	hm^3/y
Water Return - Brackish	8 907	hm^3/y
Water Return - Fresh	185	hm^3/y

7.2 Step 2: spatial disaggregation

The oil refinery water uses are mapped out to the identified oil refineries across Europe — as the oil refining production georeferences were made available thanks to mobilised sources. The only uncertainty is the cooling technology and water resource — this is selected on the basis of the GIS buffer analysis.

7.3 Suggestions for improvement

CONCAWE has a report that provides an overview of water use in the oil refining sector — because of confidentiality reasons they are not able to publish the water use at the facility level. However, they would be prepared to test the validity of the mapped water use. Hence, as for pulp and paper, deepening cooperation with the professional organisation is the only way to go forward, possibly exploring to which level the accuracy of reported data could be transmitted.

In the case of oil, there is no significant distribution of site sizes: they should all fall into the 'large' stratum and hence be individually explored.

This deepened cooperation should go in parallel with establishing a list of sites and negotiating with Eionet NRC, but this is likely to be fruitless if not first supported by reasonable proxy values obtained from the professional organisation.

8 Estimates of water use for processing and cooling in the iron and steel sector

8.1 Step 1: country aggregates estimation

Iron and steel are important products that are widely used. The production of crude steel in the European Union stood at 176 880 Ktonnes in 2011 (EUROFER), equivalent to about 12 % of the world production. EUROFER represents 100 % of the steel production in the EU. In the EU about two thirds of the crude steel is produced via the blast furnace route at 40 sites and one third is produced in 246 electric arc furnaces. The classic blast furnace route is the most complex taking place in large industrial complexes, known as integrated steelworks. Integrated steelworks are characterised by networks of interdependent material and energy flows between the various production units (sinter plants, pelletisation plants, coke oven plants, blast furnace and basic oxygen steel-making plants). The electric arc furnaces use smelting of iron-containing materials, such as scrap. Data on water usage, quality and optimised use are only starting to be collected (Worldsteel Survey ([107])): 50 % to 90 % of the water is recycled or reused, depending on the production step.

The methodology consisted of combining the information on average water use and consumption of the World Steel Association with reported production levels at the national level to estimate annual water use for the manufacturing industry of iron and steel (Table A10.19 and following). The total average intake is estimated at 28.6 m³/tonne crude steel, total discharge of water is 25.3 m³/tonne for integrated steelworks, and an average consumption and discharge of 28.1 m³ and 26.5 m³ per tonne produced steel for electric arc furnaces. Some plants are working in closed loop (World Steel Association, 2011). Total water intake for the iron and steel manufacturing industry in the EU-27 is estimated to be 5 024.3 hm³ per year. Net consumption levels are much lower of recycling and reuse, averaging 466.5 hm³ per year, which is nevertheless a very significant volume.

The estimates must be carried out by category of plants, and inside a category by process or per intermediate product.

Integrated Steelworks

Water is used for processing and cooling. There is a wide range of water consumption rates depending on technology choice, the availability of water and the legal framework. From the Reference Document on BAT in the iron and steel industry (IPPC, 2001), water used in the different production units of an integrated steelworks varies strongly depending on the techniques used (Table A10.19), with the coke oven plants and the blast furnace plants being the highest water-using steps.

Table A10.19 Water use in different production units

Main Production Steps	Source in BREF	Min. water use	Max. water use	Unit
Sinter plants	Table 4.1	0.01	0.35	m³/t sinter
Pelletisation plants	Table 5.1	0.11	1.5	m³/t pellets
Coke oven plants	Table 6.2	0.8	10	m³/t liquid steel
Blast furnaces	Table 7.1	0,8	50	m³/t pig iron
Basic oxygen steel making and casting	Table 8.2	0.4	5	m³/t liquid steel

Water use is first estimated in the different production steps and then the overall water use and water consumption are estimated.

1. Sinter plants: Sintering products may also be added to the furnace. Sintering is a process in which solid wastes are combined into a porous mass that can then be added to the blast furnace. These wastes include iron ore fines, pollution control dust, coke breeze, water treatment plant sludge and flux. Sintering plants help reduce solid waste by combusting waste products and capturing trace iron present in the mixture. Sintering plants are not used at all steel production facilities.

Table A10.20 Water use in sinter plants

Sinter production	KTonnes	Water use (min.) m³ per year hm³ per year	Water use (max.) hm³ per year
	1 999		m³ per year
Austria	2 720	0.027	9.520
Belgium	9 828	0.098	34.398

([107]) DG Env-EIP Water efficiency, 2011 (de Lamberterie, 2011).

Sinter production	KTonnes	Water use (min.) m³ per year hm³ per year	Water use (max.) hm³ per year
Germany	25 451	0.255	89.079
Spain	6 200	0.062	21.700
France	18 385	0.184	64.348
Italy	12 800	0.128	44.800
the Netherlands	4 143	0.041	14.501
Poland	400	0.004	1.400
Sweden	960	0.010	3.360
United Kingdom	13 602	0.136	47.607
Together		0.945	330.713

Source: Wirtschaftsvereinigung Stahl (Ed.): Statistisches Jahrbuch der Stahlindustrie 1995. Verlag Stahleisen GmbH, Dusseldorf, 360.

Table A10.21 Water use in the pelletising step

Period of Pellet production: 2000–2010	Production MTonnes per year	Water use (min.) hm³ per year	Water use (max.) m³ per year
Europe	80	8.800	120.000

Source: Trust fund on iron ore information - Iron ore market 2008-2010, Geneva, June 2009, United Nations Conference on Trade and Development.

2. Coke making: Coke is a solid carbon fuel and carbon source used to melt and reduce iron ore. Coke production begins with pulverised, bituminous coal. The coal is fed into a coke oven which is sealed and heated to very high temperatures for 14 to 36 hours. Coke is produced in batch processes, with multiple coke ovens operating simultaneously.

Table A10.22 Water use in the coke-making step

Country	Production Tonnes/annum	Water use (min.) hm^3/annum	Water use (max.) m^3/annum
Poland	8 791 000	7.033	87.910
Germany	8 297 000	6.638	82.970
United Kingdom	4 364 000	3.491	43.640
Czech Republic	3 613 000	2.890	36.130
Belgium	2 856 000	2.285	28.560
Spain	2 662 000	2.130	26.620
the Netherlands	2 343 000	1.874	23.430
Slovakia	1 917 000	1.534	19.170
Romania	1 891 000	1.513	18.910
Austria	1 437 000	1.150	14.370
Sweden	1 402 000	1.122	14.020
Finland	894 000	0.715	8.940
Bulgaria	771 000	0.617	7.710
Hungary	653 000	0.522	6.530
Portugal	67 000	0.054	0.670
Estonia	37 000	0.030	0.370
Together	**41 995 000**	**33.598**	**419.95**

Source: Energy Statistics Database, United Nations Statistics Division, 2005 (UNSD, 2013), recomputed as hm^3 and rounded off for readability.

3. Blast furnace: During iron making, iron ore, coke, heated air and limestone or other fluxes are fed into a blast furnace.

Table A10.23 Water use in the iron-making step

Pig iron/Blast furnace iron	Total 2011 KTonnes	Water use (min.) hm³ per year	Water use (max.) hm³ per year
Austria	5 815	4.7	290.8
Belgium	4 725	3.8	236.3
Czech Republic	437	3.3	206.9
Finland	2 600	2.1	130
France	9 698	7.8	484.9
Germany	27 795	22.2	1390
Hungary	1 315	1.1	65.8
Italy	9 824	7.9	491.2
The Netherlands	5 943	4.8	297.1
Poland	3 975	3.2	198.7
Romania	1 555	1.2	77.8
Slovakia	3 346	2.7	167.3
Spain	3 540	2.8	177
Sweden	3 240	2.6	162
United Kingdom	6 625	5.3	331.3
Together	**90 433**	**75.5**	**4 707.1**
All EU	94 134		

Source: World Steel Association, Statistics, 2011.

4. Basic oxygen steel making and casting: Molten iron from the blast furnace is sent to a basic oxide furnace, which is used for the final refinement of the iron into steel. High purity oxygen is blown into the furnace and combusts carbon and silicon in the molten iron. The basic oxide furnace (BOF) is fed with fluxes to remove any final impurities. Alloy materials may be added to enhance the characteristics of the steel.

Table A10.24 Water use in basic oxygen steel making and casting (BOF technology)

Country	Total crude steel production Ktonnes	BOF KTonnes	Water use (min.) hm³ per year	Water use (max.) hm³ per year
Austria	7 474	6 801.34	2.7	34.0
Belgium	8 026	6 019.5	2.4	30.1
Bulgaria	835	509.35	0.2	2.5
Czech Republic	5 583	5 080.53	2.0	25.4
Finland	3 986	2 830.06	1.1	14.2
France	15 780	9 941.4	4.0	49.7
Germany	44 284	30 555.96	12.2	152.8
Greece	1 934	0	0.0	0.0
Hungary	1 732	1 454.88	0.6	7.3
Italy	28 726	11 490.4	4.6	57.5
Latvia	568	522.56	0.2	2.6
Luxembourg	2 521	0	0.0	0.0
the Netherlands	6 937	6 798.26	2.7	34.0
Poland	8 779	5 179.61	2.1	25.9
Portugal	1 351	0	0.0	0.0
Romania	3 645	2 223.45	0.9	11.1
Slovakia	4 242	3 902.64	1.6	19.5
Slovenia	665	0	0.0	0.0
Spain	15 504	3 720.96	1.5	18.6
Sweden	4 829	3 332.01	1.3	16.7
United Kingdom	9 478	7 582.4	3.0	37.9
EU-27	**176 880**	**107 945.3**	43.2	539.8

Source: EUROFER; split between BOF and EAF based on Steelmap (2005).

Total water use and net water consumption in the steel and iron industry

As the sources of the different production steps are from different years, we take the average figures from the World Steel Association for calculating the total water intake, water consumption (i.e. losses), and water return for the entire process.

Table A10.25 Total water use and net water consumption in the EU-27 for BOF

Country	BOF KTonnes	Mean water intake hm³ per year	Mean water return m³ per year	Mean net water consumption m³ per year
Austria	6 801.34	194.5	172.1	22.4
Belgium	6 019.5	172.2	152.3	19.9
Bulgaria	509.35	14.6	12.9	1.7
Czech Republic	5 080.53	145.3	128.5	16.8
Finland	2 830.06	80.9	71.6	9.3
France	9 941.4	284.3	251.5	32.8
Germany	30 555.96	873.9	773.1	100.8
Greece	0	0	0.0	0.0
Hungary	1 454.88	41.6	36.8	4.8
Italy	11 490.4	328.6	290.7	37.9
Latvia	522.56	14.9	13.2	1.7
Luxembourg	0	0	0.0	0.0
the Netherlands	6 798.26	194.4	172.0	22.4
Poland	5 179.61	148.1	131.0	17.1
Portugal	0	0	0.0	0.0
Romania	2 223.45	63.6	56.3	7.3
Slovakia	3 902.64	111.6	98.7	12.9
Slovenia	0	0	0.0	0.0
Spain	3 720.96	106.4	94.1	12.3
Sweden	3 332.01	95.3	84.3	11.0
United Kingdom	7 582.4	216.9	191.8	25.0
EU-27	**107 945.3**	**3 087.1**	**2 730.9**	**356.1**

Electric arc furnaces

The electric arc furnaces have an average consumption and discharge of 28.1 m³ and 26.5 m³ per tonne produced steel (World Steel Association, precise years of averaging not provided).

Table A10.26 Total water use and net water consumption in the EU-27 for EAF

Country	EAF KTonnes/y	Mean water intake hm³/y	Mean water return hm³/y	Mean net water consumption hm³/y
Austria	672.66	18.9	17.8	1.1
Belgium	2 006.5	56.4	53.2	3.2
Bulgaria	325.65	9.2	8.6	0.1
Czech Republic	502.47	14.1	13.3	0.8
Finland	1 155.94	32.5	30.6	1.8
France	5 838.6	164.1	154.7	9.3
Germany	13 728.04	385.8	363.8	22.0
Greece	1 934	54.3	51.3	3.1
Hungary	277.12	7.8	7.3	0.4
Italy	17 235.6	484.3	456.7	27.6
Latvia	45.44	1.3	1.2	0.1
Luxembourg	2521	70.8	66.8	4.0
the Netherlands	138.74	3.9	3.7	0.2
Poland	3 599.39	101.1	95.4	5.8
Portugal	1 351	38.0	35.8	2.2
Romania	1 421.55	39.9	37.7	2.3
Slovakia	339.36	9.5	9.0	0.5

Slovenia	665	18.7	17.6	1.1
Spain	11 783.04	331.1	312.3	18.9
Sweden	1 496.99	42.1	39.7	2.4
United Kingdom	1 895.6	53.3	50.2	3.0
EU-27	**68 934.69**	**1 937.0**	**1 826.7**	**109.8**

Source: EUROFER; split between BOF and EAF based on Steelmap (2005).

See http://www.eurofer.org/index.php/eng/Facts-Figures/European-Steel-Map online.

8.2 Consolidation

The figures are very different and were cautiously entered into the databases. As indicators of orders of magnitude of the importance of the branch, an indicative synthesis is presented, collating the values from the previous tables.

Table A10.27 Collating figures from the different tables in relation to the steel industry

Production type	Water use (min.) hm³ per year [1]	Water use (max.) m³ per year [2]	Mean (as ([1]+[2])/2	Net consumption	Source table
Pelletising	8.8	120	64.400		Table A10.21
Sinter	0.945	330.712	165.828		Table A10.20
Coke	33.598	419.95	226.774		Table A10.22
Pig iron	75.5	4707.1	2 391.300		Table A10.23
BOF	43.2	539.8	291.500		Table A10.24
Together	153.243	5 997.562	3 075.402		
EAF			1 937.000	109.8	Table A10.26
BOF			3 087.100	356.1	Table A10.25
Together	306.5	.1231 995.11	1174.904 11 1465.900	Over 465.9	

The important figures are on the one hand the range of uncertainty. However, the total requirements for abstraction are huge: in a range of 3 billion m^3 to 12 billion m^3 a year. Similarly, the rather certain net uses (water not returned) are in the range of 500 hm^3, making this branch a significant net consumer, all the more plants are very concentrated.

This is why it is extremely important to reach much better figures and assess over those volumes if all are withdrawn from freshwater or if some significant contribution of seawater (in cooling) is of relevance.

8.3 Step 2: spatial disaggregation

The spatialisation of the coking industry is the same as for pulp and paper — national water use spread equally to the E-PRTR facilities.

8.4 Suggestions for improvement

Iron and steel is now a well organised branch at the professional level, and several data were obtained from EUROFER, as the delivered information suggests many other sources had to be considered as well, suggesting a lack of EU-level sufficient information sources. Hence, most data are quite aggregated and have demanded many disaggregation efforts for limited spatial results and local accuracy.

In the case of iron and steel, there is a quite significant distribution of site sizes: they should mostly fall into the ‘large’ and ‘medium’ strata and hence be individually explored, others being completed by statistical coefficients.

This deepened cooperation should go in parallel with establishing a list of sites to sort out by stratum and negotiating with Eionet NRC, but this is likely to be fruitless if not first supported by reasonable proxy values obtained from the professional organisation.

9 *Estimates of water use for processing and cooling in the textile sector*

9.1 Step 1: estimations per country

The textile industry is one of the longest and most complicated industrial chains in the manufacturing industry. It is a fragmented and heterogeneous sector dominated by small and medium-sized enterprises (SMEs). The textile industry is composed of a wide number of sub-sectors covering the entire production cycle from the production of raw materials to semi-processed and final products. The BREF focuses on the sub-sectors that involve wet processes: wool scouring, textile finishing and the carpet sector. In the BREF Textile the following water uses are mentioned. The water use differs a great deal depending on the techniques or processes being used; this is the way the volumes are estimated.

Table A10.28 Water consumption rates in the textile industry

Type of production	Min. water consumption $m^3/T/annum$	Max. water consumption $m^3/T/annum$	Mean water consumption $m^3/T/annum$	Source of information
Wool scouring	10	15	12.5	Source: VITO, 1998, BAT for the Textile Industry
Textile finishing	40	180	110	Table 3.10 of BREF (very variable per process)
Carpets	30	55	42.5	BREF

Wool scouring sector

Wool scouring involves the use of hot water and detergents to remove soil, vegetable impurities, grease and other contaminants from fibres. Wool scouring typically uses water and alkali, although scouring with an organic solvent is also possible. Water consumption is estimated to be 10 to 15 m^3 per tonne greasy wool per year, although lower values were observed in the surveyed companies in the BREF. Net specific consumption can be reduced by installing a grease and dirt recovery loop, through which water is recycled to the scouring bowls. Net water consumption varies widely over different companies depending on the rate of water recycling techniques being installed. No figures are found on the general percentage of water that is recycled in the sector.

Table A10.29 Water consumption for wool scouring per country

Country	Greasy wool (2009) tonnes	Min. water consumption hm^3/y	Max. water consumption hm^3/y
Austria	256	0.0026	0.0038
Belgium	215	0.0022	0.0032
Bulgaria	7 353	0.0735	0.1103
Croatia	659	0.0066	0.0099
Czech Republic	312	0.0031	0.0047
Cyprus	160	0.0016	0.0024
Denmark	180	0.0018	0.0027
Estonia	134	0.0013	0.0020
Finland	95	0.0010	0.0014
France	8 646	0.0865	0.1297
Germany	12 500	0.1250	0.1875
Greece	7 420	0.0742	0.1113
Hungary	4 444	0.0444	0.0667
Ireland	13 711	0.1371	0.2057
Italy	9 071	0.0907	0.1361
Latvia	42	0.0004	0.0006
Lithuania	80	0.0008	0.0012
Luxembourg	15	0.0002	0.0002
Malta	26	0.0003	0.0004
the Netherlands	3 036	0.0304	0.0455
Poland	968	0.0097	0.0145
Portugal	6 980	0.0698	0.1047
Romania	18 038	0.1804	0.2706
Slovakia	820	0.0082	0.0123
Slovenia	190	0.0019	0.0029
Spain	27 049	0.2705	0.4057
Sweden	148	0.0015	0.0022
United Kingdom	65 393	0.6539	0.9809
EU-27	187 941	1.8794	2.8191

Source: FAOSTAT.

Textile finishing

The 'finishing processes' can take place at different stages of the production process (i.e. on fabric, yarn, loose fibre, etc.), the sequence of treatments being variable and dependent on the requirements of the final user. The water consumption varies strongly between mills depending on used processes and techniques. The majority of EU textile finishing companies are SMEs. No specific production data per Member State are found ([108]).

Considering the likelihood of limited impact at the water balance level and the cost in resource to collect the information, this search was abandoned.

Carpets

Tufted carpets account for 66 % of EU production. Assumptions for estimating volumes are that average carpet density is 1 821 kg/m³ and that carpet thickness is 0,025 m to turn areas of products into water uses (ratios are per tonne of carpet).

Table A10.30 Water consumption for tufted carpets production per country

Tufted Carpets	Production (area) 2010 m² per year	Production (volume) 2010 m³ per year	Production (weight) 2010 Tonnes per year	Water consumption (min.) hm³ per year	Water consumption (max.) hm³ per year
Austria	8 500 000,00	212 500	386 856	11.61	21.28
Belgium	163 000 000,00	4 075 000	7 418 538	222.56	408.02
Denmark	11 300 000,00	282 500	514 291	15.43	28.29
France	14 300 000,00	357 500	650 829	19.52	35.80
Germany	49 200 000,00	1 230 000	2 239 215	67.18	123.16
Greece	4 200 000,00	105 000	191 153	5.73	10.51
Ireland	1 600 000,00	40 000	72 820	2.18	4.01
Italy	19 000 000,00	475 000	864 738	25.94	47.56
the Netherlands	101 300 000,00	2 532 500	4 610 416	138.31	253.57
Portugal	2 800 000,00	70 000	127 435	3.82	7.01
Spain	6 000 000,00	150 000	273 075	8.19	15.02
UK	72 300 000,00	1 807 500	3 290 554	98.72	180.98
Total	**453 500 000**	**1133750011 337 500 20639919**	**20 639 920**	**619.20**	**1 135.20**

Source m² production: Intercontuft, 2011 and ENco, 2001.

9.2 Consolidation

Volumes are not very significant, except the highly local carpet production and do not deserve consolidation.

9.3 Step 2: spatial disaggregation

Textile-related volumes are concentrated in 2 countries: Belgium and the Netherlands, for example, consume together close to 60 % of the total EU estimate of tufted carpet production. This means that spatial disaggregation is a key issue in this branch.

9.4 Suggestions for improvement

([108]) CIRFS (European Man Made Fibres Association) and EURATEX might have data available, which could be used after checking.

The textile industry has undergone many substantial changes in the past decade, with lots of outsourcing. There is no substantial statistical basis: according to the indications above, carpet making should be rather a 'large' stratum issue, whereas many other activities would share between 'medium' and 'small'; the total of the latter not being necessarily negligible.

Significant data identification should be carried out with national authorities through Eionet.

10 Estimates of water use for processing in the non-metal mineral products

The non-metal mineral industry is a heterogeneous cluster of rather different branches that process massive quantities of products and have quite different water demands and uses. For practical issues, they are lumped together in this section.

10.1 Step 1: estimations per country and per branch

Cement industry

The European Union cement industry consists of some 250 plants operated by 64 groups (though a large part involves cement mills only). In the European Union (EU-27), total cement production in 2005 is estimated at 242.1 million tonnes, representing 10.5 % of world production. Cement production in the EU is dominated by Spain, at over 19 % of the EU total, followed by Italy and Germany (Materials technology publications, 2008). About 90 % of Europe's cement production is from dry process kilns, a further 7.5 % of production is accounted for by semi-dry and semi-wet process kilns, with the remainder coming from wet process kilns (2.5 %). When renewed, the wet and semi-wet processes are expected to be converted in dry processes (BREF). Cement is a finely ground, non-metallic, inorganic powder. Cement is the basic material for building and civil engineering that breaks down into 27 different cement types.

1. Average water usage per tonne throughput

Water is primarily used in cement production for cooling heavy equipment and exhaust gases. In most cases the water consumed is not potable water. For the wet process a typical water consumption of 100–600 litres per tonne clinker is reported (BREF). The average water consumption of a modern dry cement plant is between 100 and 200 litres per tonne of clinker (Cement Australia, 2013).

2. Production levels and total water usage

This is a **list of countries by cement production** in 2005 mostly based on 'index mundi' accessed in September 2008 (source is Wikipedia (2013)). The calculated total production of clinkers for the European Member States is 194 million tonnes of which 10 % are wet and semi-wet/semi-dry process kilns (water consumption equals on average 350 l/tonne clinker) and 90 % dry process kilns (average water consumption equals 150 l/tone of clinker). Total estimated water consumption hence reaches 32.926 hm^3 per year. We divided this total amount proportionally over the Member States, assuming that every country has a similar percentage of wet and dry process kilns, in the absence of more accurate information.

Table A10.31 Estimation of water uses and consumption by cement industry

Rank (world)	Country	In million Tonnes Cement	In million Tonnes Clinker	Mean water consumption in hm^3 / y
45	Austria	4.7	3.76	0.639
39	Belgium	7	5.6	0.952
71	Bulgaria	2.1	1.68	0.286
75	Cyprus	1.8	1.44	0.245
52	Czech Republic	3.9	3.12	0.530
68	Denmark	2.2	1.76	0.299
87	Finland	1.3	1.04	0.177

Rank (world)	Country	In million Tonnes Cement	In million Tonnes Clinker	Mean water consumption in hm³ / y
19	France	21	16.8	2.856
15	Germany	30	24	4.080
23	Greece	15	12	2.040
56	Hungary	3.5	2.8	0.476
49	Ireland	4	3.2	0.544
9	Italy	46	36.8	6.256
66	the Netherlands	2.4	1.92	0.326
27	Poland	12.6	10.08	1.714
35	Portugal	9	7.2	1.224
38	Romania	7	5.6	0.952
57	Slovakia	3.4	2.72	0.462
88	Slovenia	1.2	0.96	0.163
7	Spain	50	40	6.800
64	Sweden	2.6	2.08	0.354
29	United Kingdom	11.4	9.12	1.550
	Total	**242.1**	**193.68**	**32.926**

Comment: Assumption: 1 kg Clinker = 1,25 kg Cement.

Lime industry

In 2003, there were approximately 211 installations producing lime in the EU-27. Belgium, Germany, Spain, France and Italy are the largest producers of lime in EU-27, which account for 20 % of the world lime production. The structure of the lime industry varies between Member States. It is mostly characterised by SMEs. Recently there has been a growing trend towards concentrations. The total production in the EU-27 in 2007 was 28.4 million tonnes.

1. Average water usage per tonne throughput

The raw materials from the quarry may contain small amounts of clay and sand. In this case, the limestone is washed with water. The water demand for washing is 0.5 to 2 m³ per tonne of raw material. It is usually taken from surface waters or from lowering of the groundwater during excavation. Other common sources for washing water are rainfall and wells. Fewer than 10 % of the lime plants wash the limestone (BREF, 2010). Limestone may be cut, crushed or pulverised and chemically altered. 'Burning' (calcination) converts them into **quicklime** (calcium oxide (CaO)) and, through subsequent addition of water, into **hydrated lime** (calcium hydroxide ($Ca(OH)_2$)). For hydration the quantity of water added is about twice the stoichiometric amount required for the hydration reaction (56 g quicklime + 18 g water = 74 g hydrated lime). If we use the reaction and the fact that twice the amount of water is used, the hydration process consumes approximately 0.64 m³ of water per tonne of quicklime.

2. Production levels and water estimates

When the consultant's report was drafted (June 2012), no production data that split the production of quicklime and hydrated lime could be found, and since the specific water consumptions are quite different, the consultant was not able at that time to estimate the water consumption per country or in total. This could be investigated further, especially if data from EuLA could be obtained.

In the meantime, the estimates of lumped production per country are reported.

Table A10.32 Production of quicklime and hydrated lime, including dead-burned dolomite in EU countries

Rank	Country	In thousand metric tonnes
7	Germany	6 000
8	Italy	6 000
12	France	3 500

Rank	Country	In thousand metric tonnes
15	Spain	2 000
17	Belgium	2 000
18	Poland	1 950
21	Bulgaria	1 500
22	United Kingdom	1 500
24	Czech Republic	1 200
25	Slovakia	1 080
31	Croatia	600
32	Sweden	600
33	Austria	500
34	Finland	500
35	Hungary	500
	Total	29 430

Source: United States Geological Survey (USGS) Minerals Resources Program from http://www.indexmundi.com/minerals/?product=quicklime&graph=production

Glass

The glass industry within the EU is extremely diverse, both in products made and the manufacturing techniques employed. In this sector there is an increasing use of recycled glass.

The glass industry as a whole is not a major consumer of water. Water is used mainly for cleaning, batch humidification and cooling, and can be readily treated or reused. Actual water consumption may vary much according to local conditions. Therefore, no specific numbers on water consumption are given in the BREF for most glass types. Water can be taken either from the mains supply or from natural sources.

The continuous filament glass fibre sector typically needs large amounts of water for cooling (semi-closed circuits), and also needs significant amounts of water in coating preparation and wash down in the forming/winding area. The total water consumption is estimated at between 4 and 20 m^3 per tonne finished product; cooling system losses account for around 20 % of this figure.

For mineral wool production, the overall water consumption is 3 to 10 m^3/tonne of product for stone wool.

No volumes can be proposed for the time being; this industry is not well suited to estimation from proxy information.

10.2 Consolidation

Apart from cement, data is too scarce to provide even estimates of water uses.

10.3 Step 2: spatial disaggregation

For the branches for which volumes were assessed, apportionment according to sites was carried out, on an equality of site basis. The remarks previously made on the poor accuracy of such an approach, which cannot be improved for the time being, apply as well.

10.4 Suggestions for improvement

Industries belong to rather different branches and knowledge is still too scarce to anticipate possible size distribution. Regarding the cement industry, there are necessarily quite large sites that should be candidates for being addressed as individual sites.

Specific investigations have to be carried out to better address size distribution of plants across Europe.

11 Estimates of water use for processing in the food industry

11.1 Generic issues

The food and drinks manufacturing (FDM) sector produces both finished products destined for consumption and intermediate products destined for further processing. Water consumption is one of the key environmental issues for the FDM sector. Most of the water that is not used as an ingredient ultimately appears in the wastewater stream. Water has many different uses, e.g. for cooling and cleaning, as a raw material, as process water, e.g. for washing raw materials, intermediates and products, for cooking, dissolving and for transportation, as auxiliary water. The quality of the water needed depends on the specific use. In the sector as a whole, about 66 % of the total freshwater used is of drinking water quality. In some sectors such as dairies, soft drinks and breweries, up to 98 % of the freshwater used is of drinking water quality. Some of the water is re-used, e.g. washing water of a higher production step for washing the raw materials. Some activities may require more than drinking water standards for their production (all that have wet fermentation processes, for example).

Although the overall water use can be much higher for some sectors than their actual water consumption, there is an increasing share of re-use; however, because of a lack of consistent data, no specific sector percentages are given. Data on water consumption vary, not only with the type of process and how it is operated, but also with the size of the operation.

This last point would deserve a stratified approach, but current data is too scarce for such an approach to be envisaged.

11.2 Step 1: Estimations per branch

In combining the water use coefficients from the BREF with some production data, the average water use in the different subsectors can be estimated.

As reported, the total volume is somehow significant, compared to branches analysed in previous sections.

Table A10.33 Water consumption for different sub-sectors

Sub-sector of activity	NACE code	Min. $m^3/t/y$	Max. $m^3/t/y$	mean $m^3/t/y$	Reference in BREF	Production (Eurostat 1999) in KT	Average water (hm^3/y)
Meat and Poultry	15.13	2	20	11	p.148	11 445	125.90
Breweries	15.96	0.4 (per hl)	1 (per hl)	0.7 (per hl)	p.200	28 030 (*)	196.2
Sugar	15.83	1.56	3.21	2.39	p.198	16 700	39.91
Fish canning	15.20			15	p.156	4 908	73.62
Fruit canning	15.33	2.5	4	3.25	p.162	15 485	50.33
Together							**485.96**

Source: Water use coefficients from Reference document of Best Available Techniques in the Food, Drink and Milk Industries (2006) (http://eippcb.jrc.es/reference/fdm.html).

Note: (*) Unit for production in breweries is hl (10 hl assimilated to 1 tonne).

For the above sectors, more recent or slightly different data based on per country production are proposed.

1. Breweries

Table A10.34 Water used for beer production (NACE 15.96), per country

Country	Beer prod. (1 000 hl/annum)	Min. water abstraction (hm^3/y)	Max. water abstraction (hm^3/y)	Min. use of water	Max. use of water	Avg. water use/y	Avg. use of water
Austria	8905	4.35	7.95	3.31	3.81	6.15	3.56
Belgium	8 680	4.24	7.75	3.23	3.72	6.00	3.47
Bulgaria	4 872	2.38	4.35	1.81	2.09	3.37	1.95
Cyprus	379	0.19	0.34	0.14	0.16	0.26	0.15
Czech Republic	16 190	7.92	14.46	6.02	6.93	11.19	6.48
Denmark	3 946	1.93	3.52	1.47	1.69	2.73	1.58
Estonia	1 164	0.57	1.04	0.43	0.50	0.80	0.47
Finland	4 682	2.29	4.18	1.74	2.00	3.24	1.87
France	20 014	9.79	17.87	7.45	8.57	13.83	8.01
Germany	89 853	43.94	80.24	33.43	38.46	62.09	35.94
Greece	4 329	2.12	3.87	1.61	1.85	2.99	1.73
Hungary	6 500	3.18	5.80	2.42	2.78	4.49	2.60
Ireland	4 832	2.36	4.31	1.80	2.07	3.34	1.93
Italy	16 855	8.24	15.05	6.27	7.21	11.65	6.74
Latvia	1 423.4	0.70	1.27	0.53	0.61	0.98	0.57
Lithuania	2 583.9	1.26	2.31	0.96	1.11	1.79	1.03
Luxembourg	409.1	0.20	0.37	0.15	0.18	0.28	0.16
the Netherlands	199.2	0.10	0.18	0.07	0.09	0.14	0.08
Poland	12 035.6	5.89	10.75	4.48	5.15	8.32	4.81
Portugal	34 317	16.78	30.65	12.77	14.69	23.71	13.73
Romania	6 100	2.98	5.45	2.27	2.61	4.22	2.44
Slovakia	17 600	8.61	15.72	6.55	7.53	12.16	7.04
Slovenia	4 050	1.98	3.62	1.51	1.73	2.80	1.62
Spain	1 740	0.85	1.55	0.65	0.74	1.20	0.70
Sweden	35 774.5	17.49	31.95	13.31	15.31	24.72	14.31
United Kingdom	4 884	2.39	4.36	1.82	2.09	3.37	1.95
Together	312 317.7	152.72	278.90	116.18	133.67	215.81	124.93

Source: Brauer bund, Beer statistics 2010.

Comment: decimals kept to avoid zeros for small producers. 1 000 hl assimilated to 100 tonnes or 0.1 hm^3.

The total production obtained from detailed source is consistent with the Eurostat reported data, 31 000 tonnes in the detailed estimate and 28 000 in the Eurostat statistics. The estimated water consumption is for 25 % of the beer itself.

2. Sugar industry

EU-27 sugar production varies between approximately 19 and 20 million tonnes per year. Sugar is produced in all Member States of the EU-25 except Estonia, Cyprus, Luxembourg and Malta. Germany, France and Poland are the largest producers, accounting for half of EU-27 sugar production, followed by Italy and the United Kingdom. The efficiency of sugar production varies significantly across Member States. The water requirement for fluming is about 500–800 % of the amount of beet. For washing, 150–200 % of the amount of beet is needed and for a single stone catcher 70–100 % water is needed. The mechanical clarified water is re-used for fluming and washing, thus only 25–30 % beet-based industrial water needs to be added during the last rinsing of the beets. As the sugar beets are 75 % water they are a net producer of water; therefore, compared to the overall water use, the water consumption is relatively low (BREF).

Table A10.35 Water consumption in relation to sugar beets sugar production

Country	Production 2005/2006 in tonnes	Mean water consumption hm^3/y
Austria	488 932	1.17
Belgium	925 266	2.21
Bulgaria	2 105	0.01

Czech Republic	558 879	1.34
Denmark	475 000	1.14
Finland	179 000	0.43
France	4 140 243	9.90
Germany	4 040 625	9.66
Greece	310	0.74
Hungary	490 836	1.17
Ireland	205 160	0.49
Italy	1 804 422	4.31
Latvia	71 019	0.17
Lithuania	92 000	0.22
the Netherlands	976 148	2.33
Poland	2 053 975	4.91
Portugal	37 239	0.09
Slovakia	263 767	0.63
Slovenia	46 920	0.11
Spain	1 083 000	2.59
Sweden	406 000	0.97
United Kingdom	1 341 015	3.21
Total	**19 681 861**	**48**

Water for sugar processing is however abstracted in large quantity over a short period of time (sugar beets are collected in late autumn) and at very concentrated places.

11.3 Consolidation

The available data could only give a flavour of water susceptible of being required by food production, in the range of certainly more than 500 hm^3/year, on the average. The data provided makes it very difficult to address the spatial distribution and the source of water use. In most cases, however, following the simple rule proposed, those industries that require tap water quality or better (e.g. breweries) are likely to abstract from groundwater.

11.4 Step 2: spatial disaggregation

For both the beer and sugar industries the production is provided at the national rather than the facility level, so as for most other sectors the water use can only be divided on an equal basis to E-PRTR facilities. The sugar industry facilities are on the whole well represented in the E-PRTR database — as these are generally large industrial concerns; however, in the brewery sector there can be a large number of small facilities that do not feature in the E-PRTR database.

11.5 Suggestions for improvement

Food and caning industries are exerted in a very large number of plants, the size of which depends largely on the industrial policy of the country. The milk industry for example is significantly processed in large plants in Germany and in smaller ones in France. Both numbers are very distributed in sizes and spatial location highly depending (for the largest, hence having significant local demand) on agricultural specialisation. For example, sugar beet processing plants are all large and placed in the centre of crop fields to minimise beet transport costs.

In this case, more than in other industrial branches, the making of a statistical basis is a critical issue. The three categories are likely to be existing and the strata with smaller plants not necessarily those where the total of abstraction is expected to be the smaller.

12 Final storage of data in computation databases

This is part of the Nopolu process and not detailed in this report. Refer to manual.

13 *Generic conclusions*

13.1 The facts

The current data sets that can be mobilised at the European level share the same insufficiencies as source of volumes for the accounting process:

- Only two sources of information have some ambition in making a list of facilities having a certain degree of completeness for the largest facilities:
    - E-PRTR, but this list is limited to the EU (and hence misses non-EU albeit EEA countries); moreover, its design and poor rate of populating makes it a poor source of information that does not link well with national data sources;
    - Platts database for energy, with major inconvenient of overlap with E-PRTR without linkage, being private (and hence copyrighted) and often insufficiently QA.
- There are many branch-oriented data sources, as from professional associations, that have at their disposal high quality aggregated data, but that are largely semi-private and cannot as such be the future source of information for the accounts;
- None of the different EU data sources have (or even attempt to collect) water volume values and as a result the source of provisioning is not addressed either for individual facilities (even the largest ones);
- None of the EU data sources have (or only envisage as optional information) possible surrogates (e.g. volume of activity, number of employees) that could provide a possible substitute to estimate water uses thanks to commonly agreed technical constants (e.g. BREF document);
- There is no consistent and comprehensive reference population of activities. By comprehensive, understood as a basis for a stratified approach, that could be even as aggregated as the E-PRTR list for the largest (albeit completed for some branches) and completed for smaller by aggregates at a certain NUTs level, for example.
- As a conclusion, the obtained industrial volumes (both cooling and manufacturing and their sources of abstraction) are highly questionable despite efforts of reconstruction carried out. They should not be understood as a description of the situation but as rough spatialised estimates and indicative 'seed' for proof of concept of the accounts and future developments.

13.2 Proposals

The development of information in relation with cooling and manufacturing can be developed in a consistent, affordable and constructive way only if it results from the meeting of political will by countries and technical support by both statistical institutions and professional associations.

A possible and effective way to start the process that should necessarily end up with a significant revision of data collection procedure is a joint Commission — EEA – Eurostat, Eionet (NRCs statistics and industry or having competency in this matter) and professional organisations.

The aim of this meeting (early 2013) should be:

- Agree on the principle of stratified approach ([109]) of industrial and energy water-related information as voluntary complements and harmonisation of current official reporting;
- Define the principle of a stratified approach so that the key secrecy fields would be preserved;
- Define a road map, exemplified by volunteer countries where a water uses information system is already advanced;
- Foresee the integration of the revised dataflow in the AMP of the next EEA strategy.

([109]) Rationales have been given; immediate reasons are minimise the burden for EU institutions and countries whilst maximising the obtained information and preserve industrial secrecy as much as possible.

Annex 11 Turbining waters

1 Place in the process

Turbining is a special use of water in which only energy is taken. In reality, a very small fraction of water can be evaporated because of the aerosols produced (depends on the technique implemented).

In most cases, water is returned next to the abstraction point; however, in some large plants, return can be done in different recipients, hence changing substantially the water resource.

Another developing category is pumping — turbining, in which water is pumped back into reservoirs during periods when electricity is cheap and then turbined back at peak hours.

In many rivers systems, like the Rhone, a large share of total discharge is turbined several times through low-head dams.

In the event the water wearing indicator should be considered (see §3.3 , p. 48 main report), turbining volumes should be considered.

2 Current situation

The turbined volumes have not been considered as such in this work, because there is no suitable data. In some cases, this results in an over estimate of resource in catchments where substantial diversions occur.

Annex 12 Irrigation water data collection

1 Place in the process

1.1 Data destination

Agriculture is an essential driving force in the management of water use. Especially in southern European countries, irrigation is an important element of agricultural production and agricultural water use has a substantial share in total water use since a large deal of the irrigated water is evapotranspired.

However, irrigation is not the single source of water; the 'rain-fed' agriculture is carried out over the largest part of agricultural areas. In the water accounts, it is computed from evapotranspiration (taken from climatic data) applied to crop areas. These latter volumes are not clearly identified in the accounting tables, but can be extracted in the Nopolu *System2* used for making the accounts and can be of significant meaning in further data usage.

1.2 Meeting stratification principles

As with any water use category, agricultural water abstraction could be dealt with thanks to stratification. This has not been the case because data sources were too maladapted, but the principle is the following:

- Stratum L: large irrigation systems, having a specific water supply (e.g. fed by irrigation canal), where direct data should be obtained.
- Stratum M: seems not practical, all important should be dealt with L and disseminated as S. **An improvement could be to assume a spatial defined stratum of lumped users where the density of hillside dams is significant.**
- Stratum S: irrigated crop area at FEC level (taken or disaggregated from statistics) and computed with a per ha technical coefficient, preferably spatially defined.

2 Assessment of agriculture activities water volumes

2.1 Natural soil water

The understanding of SEEAW regarding agriculture is that only abstracted water (e.g. passing through economy) participates in the accounts. Rain-fed agriculture is implicitly considered from the 'natural' evapotranspiration from cropland. Hence, rain-fed agriculture is not part of I/O tables.

By contrast, ancillary information derived from the accounting application can be considered as the amount of precipitation that falls onto agricultural fields. The excess of water, e.g. the part that is not used by the crop, is recorded as a return flow to the environment from soil. Both values can be extracted thanks to appropriate queries, albeit not reported in the tables. It is important to record this flow for several reasons: it shows, for example, the relative contribution of rain-fed and irrigated agriculture for food production. In addition, considering the importance of rain-fed agriculture worldwide (more than 60 % of all food production in the world is produced under rain-fed conditions), this information can be used to assess the efficiency of rain-fed agriculture (e.g. crop production per volume of water used) and to formulate water policies.

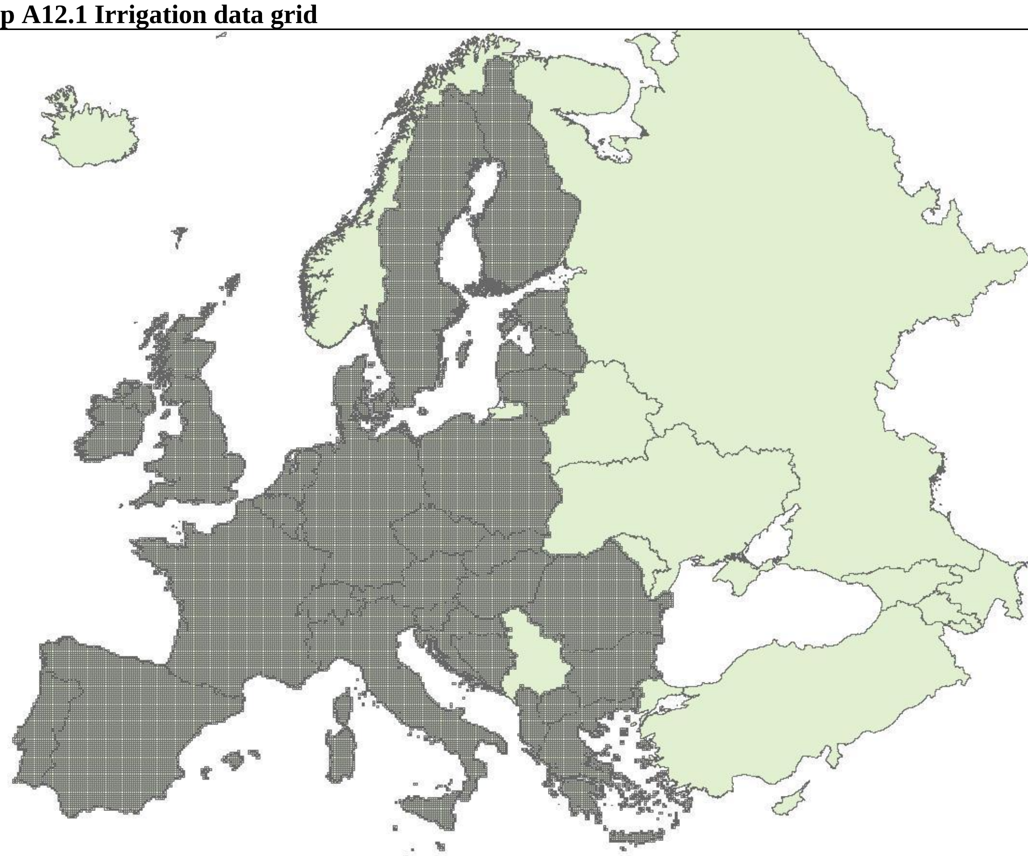
2.2 Abstractions for irrigation

Irrigation volumes were assessed by JRC and the agriculture abstraction information is based on data sets obtained from JRC, compiled and set in the Nopolu *System2* geodatabase EU_IRGA.mdb.

The process comprises several steps to transform source raster into time data sets, then restructuring the volumes and correcting them with additional data sources and when possible share between surface and groundwater. The steps are rather complicated and require hypothesis since data is poorly documented.

This database does not cover all EEA areas and does not fully cover the accounting domain, as shown in the source gridding.

Map A12.1 Irrigation data grid



Source: JRC, mapped by Pöyry.

The source JRC database comprises, for most of the grid cells, monthly values covering all the years from 1984 until 2008. The spatial resolution is 10 by 10 km, congruent with the LEAC kilometre gridding; the database can hence be processed with MS Access based on ID of cells. The processed data have been allocated per FEC and stored in table IRGAmonthly.

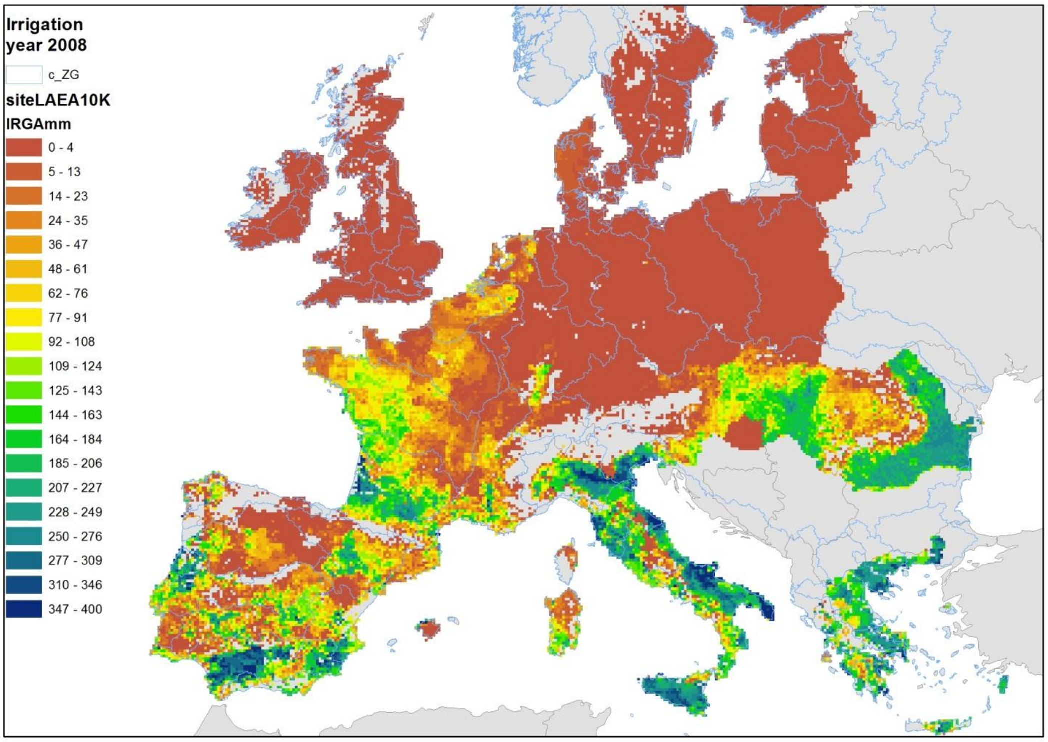
Table A12.1 Source JRC table (sample)

IRGAmonthly

OBJECTID	SiteID	YY	MM	CropArea	avgIRGA	DateValue
3364561	10852	1984	1	2097	0	01/01/1984
8473542	10852	1984	2	2097	0	01/02/1984
6560414	10852	1984	3	2097	0	01/03/1984
3798672	10852	1984	4	2097	3.81	01/04/1984

This table shows that a certain crop area for a cell (field 'site'), total area of which is 10 000 ha (10 km × 10 km), has received a certain water blade, in the table 3.81 mm in April 1984. The actual extent of data is represented in Map A12.2.

Map A12.2 Map of irrigation for year



Source: JRC database, reprocessed by Pöyry.

This source of information has been systematically reprocessed by the consultant (Pöyry) that aggregated data from the source rasters to FECs using a simplified transfer matrix: each source cell has been in totality allocated to the FEC into which the centre of the cell was falling. This method is not very accurate at the FEC level but sufficient, considering the large uncertainty in data, for further aggregation at the sub-basin level.

Each of the last 10-year monthly data sets are calculated and stored in a specific database EU_IRGA_ZHYD.mdb, under table IRGA_monthly_Zhyd, which sample structure is displayed in Table A12.2. The assumption is that the water blade (as from table IRGAmonthly) applies to the whole area of the cell, and is then used to compute a volume from mm * crop area.

Table A12.2 Transformed source JRC irrigation table (sample)

IRGA_monthly_ZHYD

ZHYD	YY	MM	IRGAmm	CropHa	CropIrrigł	VolMm3	VolcMm3
A020001377	2008	5	0	12363	3898	0	0
A020001377	2008	6	27.92	12363	3898	0.3459571	0.3459571
A020001377	2008	7	0	12363	3898	0	0
A020001377	2008	8	0	12363	3898	0	0
A020001377	2008	9	0	12363	3898	0	0
A020001377	2008	10	0	12363	3898	0	0

An ancillary source, from JRC as well, was used to estimate the irrigated areas. This source has been allocated to FECs as well and is used to check the consistency of the water blade when irrigation volumes are recomputed for only the actual irrigated areas.

The important information is that volumes are computed the following way:

1. compute volumes from water blade * crop area (source 1) and aggregate to relevant item;
2. allocate computed volume to irrigated area in the same aggregate and deduce water blade of irrigated area in the item;
3. if required, correct the volumes and reprocess from 1.

When reallocating the gridded crop areas to the FECs, once computed, data is transferred into the final tables, the structure of which is reported and displayed in Figure A12.1 below. For practical reasons, the table displayed refers to districts, but could apply to any other entity in ECRINS.

Figure A12.1 Structure of final irrigation table

IRGA_ZG

Nom du champ	Type de données	
ZG	Texte	basin code
NameEntity	Texte	basin name
YY	Numérique	year ref
IRGAmm	Numérique	mm of irrigated water (JRC)
Irrigha	Numérique	irrigated area ha
CropHa	Numérique	crop area in ha
SurfGridha	Numérique	square grid area ha
SurfShapeha	Numérique	Entity area ha
Volrefhm3	Numérique	reference volumes for irrigation computed from JRC in hm3
valref	Numérique	Bibliography values for volume for irrigation in hm3
coefRec	Numérique	bias coeficient application
IrrigEff	Numérique	Irrigation efficiency
SurfWaterPerc	Numérique	percentage of surface water
GroundWaterPerc	Numérique	percentage of ground water
ImportWaterPerc	Numérique	percentage of imported water
RecyclingWaterPerc	Numérique	percentage of recycle water
SprinklingPerc	Numérique	percentage of sprinkling irrigation system
DrippingkPerc	Numérique	percentage of Dripping irrigation system
SurfacePerc	Numérique	percentage of Surface irrigation system
OtherPerc	Numérique	percentage of Others irrigation system

Source: Nopolu *System2* application.

This table contains many important fields required by the accounting procedure that are not part of the source data.

2.3 Other sources of irrigation volumes

Once data exploded at FEC level and was integrated into the final table IRGA_ZG, different aggregates (e.g. per country or per district) were carried out and compared to reference volumes taken from local statistics, showing large differences in some areas. This comparison presents significant differences, leading to the introduction of a correction factor that applies globally to the basin concerned, as shown in Table A12.4.

The explored information sources are listed in the next table A 12.3.

Table A12.3 Source of irrigation correction data found and used

Country	*Report / source*
Austria	*AT_Abschlussbericht EU-Grant Bewõsserung 2010.pdf*

Bulgaria	*BG_Final Report Water Use.pdf*
Cyprus	*CY Final irrigation report.pdf*
Czech Republic	*CZ_Final report.pdf*
Germany	*DE_Final Technical Report 40701 2008 001-2008 129.pdf*
Denmark	*DK_Water for Irrigation in Water .pdf*
Greece	*GR_Report_Volume_of _Water.pdf*
Italy	*IT_REPORT_MARSALa_vers_27_07.pdf*
Lithuania	*LT_Water project_final report_rev2.doc*
Latvia	*LV_Annex_1_LV.doc*
Latvia	*LV_Annex_2_LV.doc*
Latvia	*LV_Final_report_LV.doc*
Malta	*MA_Irrigation Report_120310.pdf*
Norway	*NO_BIOFORSK RAPPORT nr 174 _2.12.2009_ til ESTAT.pdf*
Norway	*NO_Final report Irrigation ESTAT.pdf*
Portugal	*PT_Pilot studies on estimating the volume of water used for irrigation - PT.pdf*
Slovenia	*SI_Pilot studies on estimating the volume of water used for irrigation - final report.pdf*
France	*FR_ IRRIGATION DURABLE ([110])*
Spain	*SP_ The development of irrigated agriculture in twentieth-century Spain: a case study of the Ebro basin, and other Ebro basin related reports*

The exploiting of data allowed in many cases (not in all) establishing correction factors and breakdown between water sources, generally at the district level. No more detailed data could be found before presenting results to countries.

The final collated data, that are quite scarce, are reported in Table A12.4. In this table the correcting coefficient is source volume / vol. corrected, where volumes are present.

Table A12.4 Values of irrigation correction data found and used

ZG	Name Entity	YY	*Irrig. km²*	Crop km²	Vol. JRC hm³	Vol. corrected hm³	Coef. Cor.	Ground Water	IrrigEff
WFD0000001	Loire, Brittany and Vendee coastal waters	2008	*4 259*	78 598	332	505	66 %		65 %
WFD0000008	Adour, Garonne, Dordogne, Charente and coastal waters	2008	*6 357*	42 758	1 037	1 032	101 %		65 %
WFD0000024	Seine	2008	*1 067*	47 351	44	116	38 %		65 %
WFD0000082	Rhone and Coastal Mediterranean	2008	*2 829*	30 419	188	3 009	6 %		65 %
WFD0000092	Corsica	2008	*81*	623	1	43	2 %	21 %	65 %
WFD0000095	East Aegean Region Basin District	2008	*1 159*	2 271	265		37 %		65 %
WFD0000096	Sicily	2008	*1 434*	9 554	353	990	36 %	68 %	65 %
WFD0000097	West Aegean Region Basin District	2008	*1 113*	1 826	264	*	37 %		65 %
WFD0000100	Thessalia	2008	*2 289*	3 693	345	*	37 %		65 %
WFD0000101	Western Macedonia	2008	*1 580*	2 964	288	*	37 %		65 %

([110]) Rapport établi par: Jean-Didier LEVY Ingénieur général du génie rural, des eaux et des forêts Michel BERTIN Ingénieur général du génie rural, des eaux et des forêts Bernard COMBES Ingénieur général du génie rural, des eaux et des forêts Josy MAZODIER Ingénieur général du génie rural, des eaux et des forêts Alain ROUX Ingénieur général du génie rural, des eaux et des forêts.

ZG	Name Entity	YY	*Irrig. km²*	Crop km²	Vol. JRC hm³	Vol. corrected hm³	Coef. Cor.	Ground Water	IrrigEff
WFD0000102	Central Macedonia	2008	*688*	2 589	169	*	37 %		65 %
WFD0000103	Epirus	2008	*311*	606	71	*	37 %		65 %
WFD0000104	Western Sterea Ellada	2008	*424*	758	81	*	37 %		65 %
WFD0000105	Western Peloponnese	2008	*182*	1 077	34	*	37 %		65 %
WFD0000106	Eastern Sterea Ellada	2008	*826*	2 120	195	*	37 %		65 %
WFD0000107	Northern Peloponnese	2008	*496*	1 473	100	*	37 %		65 %
WFD0000108	Eastern Peloponnese	2008	*468*	1 309	119	*	37 %		65 %
WFD0000109	Attica	2008	*53*	281	16	*	37 %		65 %
WFD0000111	Crete	2008	*498*	2 056	104	301	35 %	97 %	65 %
WFD0000112	Aegean Islands	2008	*23*	329	4	*	37 %		65 %

Source: Nopolu database.

Comment: corrections are from obtained volumes (reported in Vol. corrected column) or derived from general source mentioned by * (hence only Coeff. cor. is populated).

The groundwater to surface water read as surface = 100 – groundwater percentage. Where not populated, a 50/50 ratio is imposed by default.

One of the important outcomes of this process is the ratio of irrigated to total crop area. Data for the district (or countries) with no correction are not reported as district, because they are too numerous. In total, 9 344 km² are irrigated over 152 639 km² of total crop area, making an average irrigation percentage area of 6 %. In some basins, by contrast up to 62 % of the crop area is irrigated, making local demand in water high.

2.4 Summary of abstracted volumes

After applying the correction factors to the source data, the aggregation of the FECs per country yields the final volumes (the computations are carried out with FEC level data, the aggregation is to facilitate the assessments. Table is presented by country because the list of districts is much longer and data are in the electronic tables of the accounts.

Table A12.5 Summary of yearly (average) of volumes (hm³) abstracted for irrigation per country

CTY	No of years	Input in mm (source) (1) = (5)/(3)	Input in mm (corrected) (2) =((5)+(6))/(3)	Irrigated area km² (3)	Total crop area km² (4)	Vol. from JRC data hm³ (5)	Supplement of volume (correction) hm³ (6)
AL	25	273	273	0.7	2.1	0.2	-
AT	25	25	25	277.9	17 582.4	6.9	-
BE	25	44	44	27.4	11 838.4	1.2	-
BG	25	116	116	47.0	2 149.5	5.5	-
BY	25	0	0	-	411.7	-	-
CH	25	13	13	133.2	7 467.5	1.8	-
CY	0						
CZ	25	0	0	154.3	32 741.3	0.0	-
DE	25	3	3	2 168.2	125 929.1	5.9	-
DK	25	7	7	1 909.5	17 375.2	13.4	-
EE	25	0	0	-	9 374.6	-	-
ES	25	127	127	29 756.3	177 611.0	3 784.6	-
FI	25	0	0	-	16 398.7	-	-
FR	25	108	309	15 442.9	222 437.2	1 665.0	3 102.7
GR	25	203	554	10 021.9	23 171.9	2 033.2	3 520.4
HR	25	30	30	0.7	392.8	0.0	-
HU	25	158	158	640.2	49 604.1	101.3	-

CTY	No of years	Input in mm (source) (1) = (5)/(3)	Input in mm (corrected) (2) =((5)+(6))/(3)	Irrigated area km² (3)	Total crop area km² (4)	Vol. from JRC data hm³ (5)	Supplement of volume (correction) hm³ (6)
IE	25	0	0	-	32 768.9	-	-
IS	25	0	0	-	-	-	-
IT	25	205	233	22 846.7	90 158.4	4 688.3	637.0
LI	25	0	0	0.8	30.4	-	-
LT	25	0	0	-	29 150.7	-	-
LU	25	0	0	5.2	1 187.7	-	-
LV	25	0	0	-	21 212.7	-	-
MD	25	214	214	37.6	348.2	8.0	-
MK	25	162	162	2.9	12.2	0.5	-
MT	0	0	0	-	-	-	-
NL	25	34	34	628.9	16 464.4	21.3	-
NO	25	0	0	0.1	2.7	-	-
PL	25	0	0	331.6	158 679.3	-	-
PT	25	117	117	2 056.1	23 648.1	241.1	-
RO	25	214	214	3 886.5	108 409.3	831.4	-
RS	25	182	182	2.3	1 368.7	0.4	-
RU	25	0	0	0.4	707.8	-	-
SE	25	0	0	451.3	20 211.7	-	-
SI	25	118	118	24.5	4 441.4	2.9	-
SK	25	94	94	1 075.0	16 174.3	100.7	-
TR	25	292	292	75.0	140.4	21.9	-
UA	25	97	97	3.7	762.0	0.4	-
UK	25	0	0	1 442.2	115 963.5	0.6	-
Together				93 451.0	1 356 330.4 ##########	13 536.3 13,536.3	7 260.1 7,260.1

Source: recomputed from Pöyry data sets by EEA.

Comments: greyed cell lines to mark the incorporated changes in volumes.

Volumes correction is a rather tricky issue since the values are related to sub-basins, provinces or districts. In this table these corrections have been aggregated at the country level to make reading possible. The correction factor makes a global change of + 60 % in volumes, reaching 3-fold in some countries.

However, the analysis of water blade over irrigated areas seem quite elevated (at the country level) and too small in other cases, suggesting quite significant uncertainties in the source data.

2.5 Returns

In the absence of other data and considering the rather poor value of irrigation volumes obtained that did not deserve investing resources in the searching for data, a systematic efficiency of 65 % was used. This means that for 1 m^3 of irrigated water, 350 litres are returning to soil and are not evaporated.

2.6 Comments on the volumes used in computations

During the consultation of Member States, only Spain expressed strong reservations indicating that, according to their data (unfortunately not provided), the irrigations volumes for Spain ([111]) are underestimated by a several fold factor (24 000/3 598 = 6.67). This underlines that in Spain,

([111]) Extract from the note sent on behalf of Spain: 'Indica que ha utilizado la base de datos de demanda mensual elaborada por el JRC. La cifra global recogida en el documento para España, 3598 hm^3/año es muy inferior a la estimación actualmente disponible, que asciende a 24000 hm^3/año. Por tanto, la fuente de información utilizada parece incompleta.'

24 000/82 000 = 29 % of annual resource is on the average used for irrigation (both figures from the Spanish comments, second being the run-off).

However, in the absence of more utilisable information, correction cannot be undertaken. Hence, in Table A12.5 above, Spain is not corrected, data coming only from the JRC source.

Were Spanish values to be incorporated, this would more than double the total corrected figure in Table A12.5, demonstrating the importance of accurate information and the current level of uncertainty.

3 Suggestions for data collection improvement

Irrigation is a very important source of water abstraction and consumption; as a rule of thumb, more than 75 % of abstracted water is eventually evaporated. Moreover, irrigation needs are in many countries deeply rooted in the weather: dry years when resource is scarce demand more water for irrigation than wet years when water is abundant. The pressure on resource is hence all the more sensible when that resource is poor.

There is indeed some baseline demand; since many irrigation activities are driven by the weather, values per year (disaggregated at the monthly level) are needed. In the current case, the dilemma is that time-disaggregated data are significantly erroneous, whereas available and more accurate information is seemingly lumped and averaged.

It is suggested that a stratified approach be developed jointly with the JRC, in which:

- large systems are identified, with the help of Eionet, and based on a common specification;
- infrastructures in relation to agriculture are inserted into ECRINS, and related to resources (to better document the change in reserve);

specific information on weather dependent technical coefficients (if data on volumes is not available) is collected with countries' support at, that is, sub-unit level, for example, to correct current estimates in close relationship with Eionet NRC.