Summary document on the applied methods for data gap filling

1. Introduction

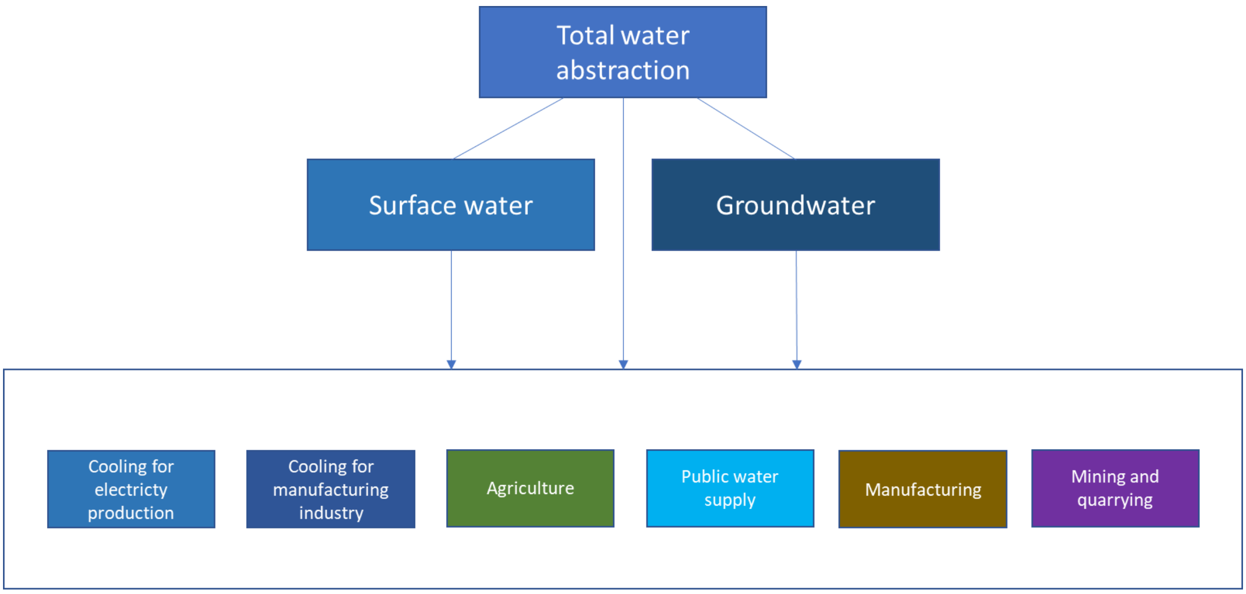
Water abstraction by source and sector is an important indicator to address properly on pressure of socio-economic water demand as well as the state of surface water and groundwater. The EEA indicator for water abstraction by source and sectors aims to address the quantitative changes of water abstraction from surface and groundwater resources for different economic activities. The current data has been compiled mainly from the data reported by the EIONET Member Countries to WISE SoE water quantity as well as data reported to Eurostat and OECD. In addition, the indicator has used data from the EEA water accounts database, which was published in 2019 (Development of water abstraction in Europe since the 1990s)

Nevertheless the European database still holds large gaps on certain sectors sucs as agriculture, manufacturing industry, public water supply, cooling for electricity production and many others. This shotr document aims to provide short clarifications on the data gap filling exercises performed when the underlying dataset for developing the indicator on water abstraction by source and sectors.

2. Schema of data availability on water abstraction by source and sectors

The European databases (WISE SoE, Eurostat and OECD) hold the data on water abstraction by source and sectors at different spatial and temporal scales. There is no fully homogenous spatial and temporal data covering all EIONET Member countries. Some report the data on water abstraction from surface water and groundwater, while some others report the data only on total water abstraction (Figure 1). Similarly, water abstraction by sectors also present different spatio-temporal coverage with large gaps among the countries.

Figure 1 - Schema of data availability on water abstraction by source and sectors



The fllowing sections provide short clarification on how simple statistical methods have been employed in filling the gaps in the database. In addition, a short statistics can be found in the excel file the level of gap filling per country and economic sector.

3. Water abstraction from surface and groundwater

In many cases, the current database contains data only total water abstraction by economic sectors without specifying the source of water (e.g. surface water and groundwater). Quantifying the proportional share of water abstraction from surface water and groundwater the following method has been implemented;

Step 1 – Estimate the annual ratio of water abstraction between groundwater and surface water based on available data on water abstraction from surface and groundwater reported by the countries

Step 2- Apply average nearest neighbour based on latest available three years to fill the gaps in ratio of water abstraction by groundwater and surface water

Step 3- Estimate water abstraction of groundwater and surface water for those cases data reported only on total water abstraction based on the estimated coefficient in the Step 2.

4. Water abstraction for agriculture

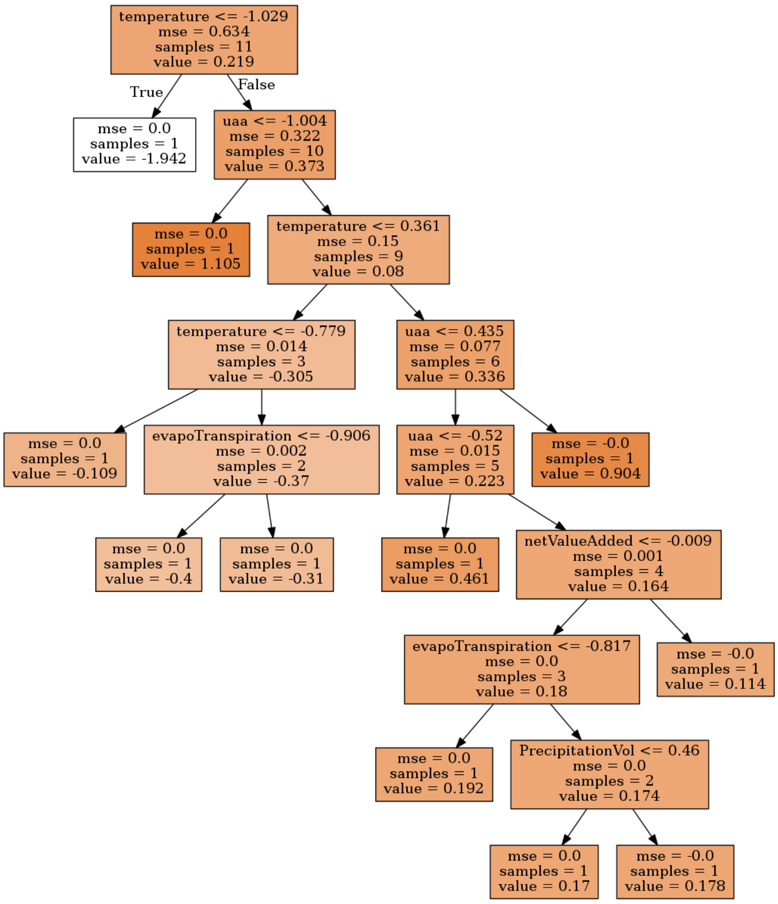
Data sources

- The original data has been obtained from WISE SoE 3 and Eurostat env_wat_abs + national statistical office websites
- Then the gaps have been filled from the EEA water accounts database at the country level developed by the ETC in 2017.
- Utilisied agricultural area of Eurostat database ((apro_acs_a_and_h) and [TAG00025] have been used to calculate annual water abstraction per ha by dividing the total water abstraction (only reported data) with the total utilised agricultural area.

Gap filling

- Gap filling has been performed by multiplying water abstraction from surface and groundwater per ha with total UUA.
- For few cases, machine learning (ML) algorithms have been employed for testing the possiblity of using multi-parametric regressional analyses in data gap filling based on the random forest tree of which one example is illustrated below;

Figure 2 – Example of random forest tree method



The ML algorithms have used the following variables;

- Annual total precipitation at the country level
- Annual total actual evapotranspiration
- Annual mean temperature
- Utilised agricultural area
- Net value added generated from agricultural sector
- Population

The first test with the ML has been performed by estimating the known values from the reported data with the 10 % of deviation thresholds. The second step is to ask support from the EIONET Member countries to cross-check the estimated values with the national data for those records highlighted in red in the excel sheet "-1- Agriculture".

The ML algorithms will be expanded over the other sectors as well in the coming period.

Uncertainities

- There is large implementation of the data gap filling based on the water used density per ha. In many cases, UUA would not be directly related to the area subject to irrigation. The assumption was to adopt one-to-one relation between water abstraction for agriculture and UUA, which might not be the case in the reality.

- Water abstraction from groundwater has a number of external factors e.g. abundance of surface water availability would reduce the water abstraction from the groundwater. This could not be adressed in the data gap filling. Similarly, energy price would have also impact on the choice of water source, as water abstraction from groundwater may require using the energy.

- Due to above and many other uncertainities, the gap filling would not provide consistent overlapping with national data. Nevertheless, the gap filling with such large assumption aims only to provide a European overview.

5. Total public water supply

Data sources

The source data for public water supply has been obtained from the WISE SoE3 and the Eurostat env_wat_abs. Data gaps further has been filled by the data from the national websites where the respective data is available.

Gap filling

- Population data (tps000001) and population connected to public water supply (env_wat_pop) have been obtained from the Eurostat database
- Number of population connected to public water supply has baeen estimated based on the tps000001 and env_wat_pop.
- Total annual water supply per capita has been estimated by comparing annual water abstraction for public water supply divided by population connected to public water supply.
- Average nearest neighbour based on latest available three years has been applied to fill the gaps in the dataset on total annual water supply per capita.
- Then the annual total public water supply has been estimated by multiplying water abstraction per capita with the total population connected to public water supply for those years no data is available

Uncertainities

- Serbian GW/SW ratio has been used for Montenegro and Kosovo (under United Nations Security Council Resolution 1244/99)

6. Manufacturing

Data sources

WISE SoE water quantity and Eurostat (env_wat_abs) data are the main source of dataset. Then OECD water abstraction and Aquastat databases have been used as complementary to add the data where available. Later National statistical offices and KNOMEA databases have been also used. The source of data is provided with the link on the respective data in the dataset.

Gap filling

It was adopted that there is one-to-one relation between water abstraction and production in industry. Based on that assumption, gap filling on total water abstraction for manufacturing industry has been performed by using the Eurostat data of Production in industry - annual data [sts_inpr_a] and [sts_inpr_a].

The [sts_inpr_a] is the data indicating percentage change in industrial production compared to previous year. The same percentage of changes in industrial production has been applied to the change in water abstraction for manufacturing at the first phase for gap filling. Then [sts_inpr_a] data which is indicating the index change (2010=100 index) in manufacturing industry production was also used with the same principle to fill the remaining gaps where such baseline data is available.

Water abstraction from groundwater and surface water

As any record on water abstraction from surface water and groundwater is not available for some countries (e.g. Italy, Iceland) groundwater/surface water seperation has not been estimated. Therefore, it is suggested to use the field of "surface and groundwater" in the "9-Pivot" for Eurpean overview in the context of water abstraction for manufacturing inducatry.

Uncertainities

Adoption of one-to-one relation between water abstraction for manufacturing industry and production in manufacturing industry is a very rough approach which would result in large unquantifiable uncertainities. For example, in many cases industrial plants might change production or processing technologies which would improve the water use efficiency per unit production.

7. Water abstraction for electricity cooling

Data sources

The primary source of data has been obtained from WISE SoE 3 and Eurostat env_wat_abs.

Complementary data has been obtained from the EEA Water accounts database which was published by the EEA in 2019. Similarly, OECD freshwater abstractions database has also been visited to compile the data from open sources.

Gap filling

As a first step, the Eurostat data on "Production of electricity and derived heat by type of fuel (nrg_bal_peh)" and data on water abstraction for cooling in electricity production reported to WISE SoE and Eurostat env_wat_abs have been used to calculate average water abstraction (m3) per kWh of electricity production from combustion plants.

At the second step, the long-term annual average of water abstraction for electricity cooling per kWh has been calculated. The average ratio (coefficient) between water abstraction for cooling has been multiplied with electricity production from combustion plants as an estimate of total annual water abstraction for electricity -cooling for missing data.

Uncertainities

The method applied for data gap filling based on total water abstraction and electricity production from combustion plants is adopting one-to-one relation between energy production and water demand for per KWh electricity produced. That might not always be the case. Power plants might change their cooling technology over the time.

Another uncertainty is about separation of water abstraction between groundwater and surface water. Using the average nearest neighbour of three consecutive year would mislead the actual volume of water either from surface or groundwater resources. The time series based on the reported data present large ranges between groundwater and surface water over the years.

8. Cooling water in manufacturing, mining and quarrying

Despite large data gaps on cooling for manufacturing industry as well as mining, quarrying and construction sectors, no gap filling has been performed for those sectors due to higher potential uncertainties.

9. Baseline periods

In the database, three baseline periods have been identified for comparative analyses and assessment across the countries and regions; 2000, 2010 and 2019. In order to obtain a comparable and homogenous data set for those three periods, the nearest data to those periods have been added into the time series (Figure 2).

Figure 2- Developing the baseline for comparative analyses

Country Name	EcoCode	Source	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2000 baseline
Spain	Agriculture	surface and groundwater	23 797.9	23 700.0	20 279.4	17 000.0	21 504.6	24 116.0	23 762.1	23 413.5	25 010.9	26 325.2	23 688.0	23688.0
Finland	Agriculture	surface and groundwater	20.0	20.0	20.0	28.9	50.0	50.0	50.0	50.0	50.0	50.0	50.0	50.0
France	Agriculture	surface and groundwater	4 884.0	4 900.0	4 919.0	4 949.0	4 971.0	4 023.9	3 536.4	3 350.0	3 561.0	3 181.0	4 871.9	4871.9
Croatia	Agriculture	surface and groundwater	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	7.0	30.0	30.0
Hungary	Agriculture	surface and groundwater	439.5	452.4	949.0	945.0	711.0	662.1	455.6	407.5	407.2	441.5	720.7	720.7
Ireland	Agriculture	surface and groundwater	0.0	0.0	0.0	0.0	179.0	0.0	0.0	0.0	0.0	0.0	0.0	179.0
Iceland	Agriculture	surface and groundwater	0.0	0.0	0.0	0.0	0.0	0.0	70.0	70.0	70.0	70.0	70.0	70.0
Italy	Agriculture	surface and groundwater	23 231.9	23 223.3	23 137.6	23 137.6	23 137.6	23 120.5	23 120.5	25 376.0	20 865.0	20 430.1	20 010.0	20010.0
Lithuania	Agriculture	surface and groundwater	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.1	1.1
Luxembourg	Agriculture	surface and groundwater	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2	0.0	0.2
Latvia	Agriculture	surface and groundwater	122.2	143.4	135.5	74.6	62.8	100.0	54.9	58.6	53.4	50.0	47.4	47.4
North Macedonia	Agriculture	surface and groundwater	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	383.0	383.0
Malta	Agriculture	surface and groundwater	18.2	19.2	18.2	17.7	17.7	18.7	17.6	16.7	16.7	16.8	15.0	15.0
Netherlands	Agriculture	surface and groundwater	189.0	184.0	184.0	220.8	241.4	260.0	230.0	90.0	53.0	76.0	59.8	59.8
Norway	Agriculture	surface and groundwater	210.4	210.4	210.4	210.4	210.4	271.1	228.1	762.7	696.4	769.3	80.0	80.0
Poland	Agriculture	surface and groundwater	1 693.7	1 518.7	1 369.3	1 392.8	1 237.8	1 176.8	1 057.5	1 082.9	999.2	1 045.4	1 060.6	1060.6
Portugal	Agriculture	surface and groundwater	3 991.0	5 100.0	6 016.4	6 479.3	10 000.0	10 000.0	6 829.0	6 834.3	8 754.6	5 862.9	4 752.1	4752.1
Romania	Agriculture	surface and groundwater	6 990.0	2 309.0	2 794.0	2 980.0	2 520.0	1 910.0	2 320.0	1 030.0	1 300.0	1 027.0	940.0	940.0
Serbia	Agriculture	surface and groundwater	802.0	802.0	802.0	802.0	802.0	802.0	802.0	802.0	802.0	802.0	838.4	838.4
Sweden	Agriculture	surface and groundwater	160.0	172.0	174.0	174.0	176.0	150.0	150.0	150.0	150.0	150.0	150.0	150.0
Slovenia	Agriculture	surface and groundwater	3.9	2.9	3.7	3.5	5.5	4.8	0.2	5.9	6.2	6.4	6.6	6.6
Slovakia	Agriculture	surface and groundwater	321.0	91.0	152.0	192.0	138.0	96.0	75.0	67.0	60.4	37.6	91.3	91.3
Turkey	Agriculture	surface and groundwater	17 842.0	27 394.1	27 394.1	27 394.1	27 394.1	29 112.0	25 676.0	26 980.0	28 760.0	27 025.0	32 907.0	32907.0

10. References:

Use of Freshwater Resources (ETC/ICM, 2016)

Results and lessons from implementing the Water Assets Accounts in the EEA area (EEA, 2013)

Reference system and resources datasets (Poyry, 2012)

Uses and supply (Poyry, 2012)