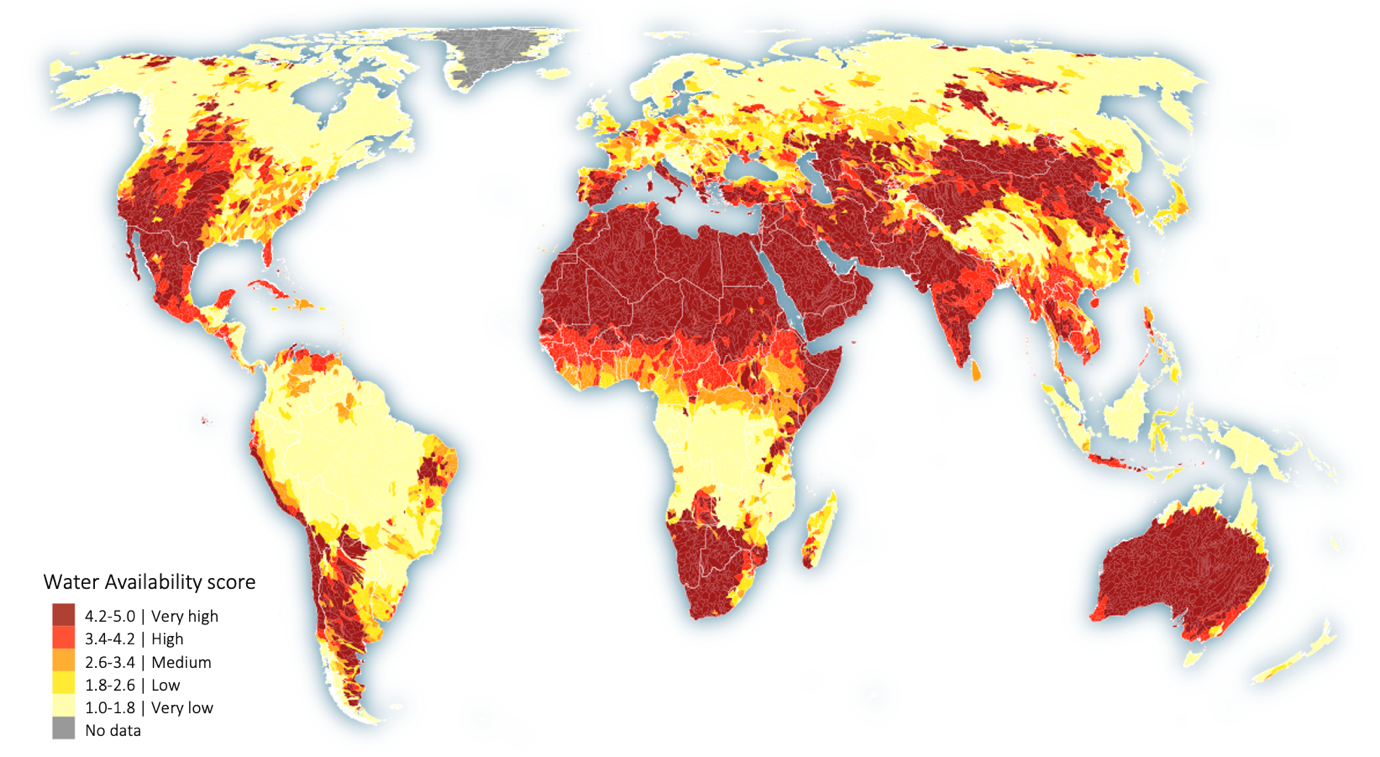
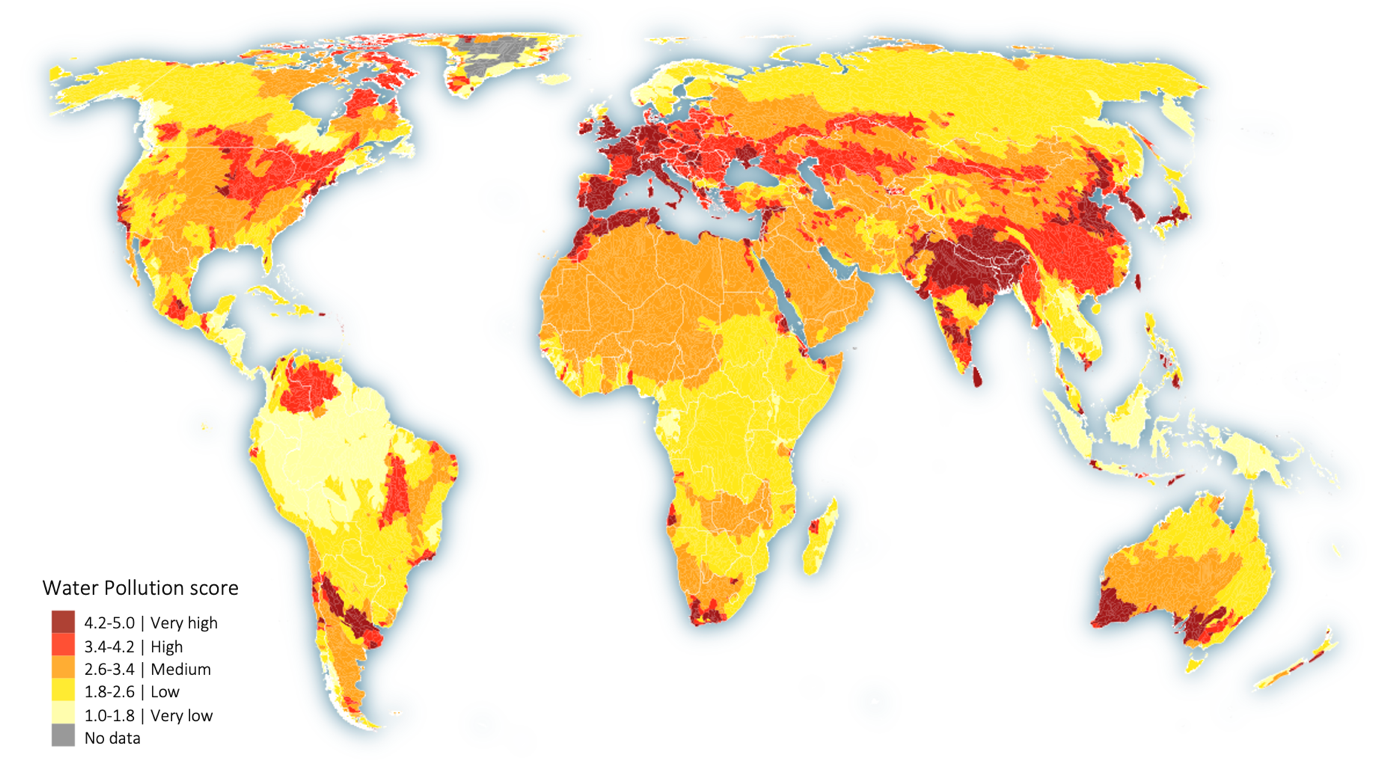
Documentation

# State of Nature layers for Water Availability and Water Pollution to support SBTN Step 1: Assess and Step 2: Interpret & Prioritize

Version 1.1 (June 2024)

Rafael Camargo, Liam Bailey, Sara Walker, Elizabeth Saccoccia, Richard McDowell, Allen Townsend, Ariane Laporte-Bisquit, Samantha McCraine, & Varsha Vijay. (2024). State of Nature layers for Water Availability and Water Pollution to support SBTN Step 1: Assess and Step 2: Interpret & Prioritize (Version 1.1) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7797979>





# Introduction

There are multiple well-recognized and peer-reviewed global datasets that can be used to assess water availability and water pollution. Each of these datasets are based on different inputs, modeling approaches, assumptions, and limitations. Therefore, in SBTN Step 1: Assess and Step 2: Interpret & Prioritize, companies are required to consult different global datasets for a robust and comprehensive State of Nature (SoN) assessment for water availability and water pollution.

To streamline this process, WWF, the World Resources Institute (WRI), and SBTN worked together to develop two ready-to-use unified layers of SoN – one for water availability and one for water pollution – in line with the [Technical Guidance for Steps 1: Assess and Step 2: Interpret & Prioritize](https://sciencebasedtargetsnetwork.org/resources/) (July 2024). The main outputs contain the maximum values of Water Availability and of Water Pollution as well as the individual indicators' values. This information is available at different spatial resolutions, thus in two data formats: 1) a shapefile with values at HydroBasins (Pfafstetter level 6); and 2) an excel file with values at sub-national divisions (Adm1) and national divisions (Adm0). These datasets and complete documentation are publicly available for download from <https://doi.org/10.5281/zenodo.7797979> . See the data’s columns and description in page 5.

These unified layers will make it easier for companies to implement a robust approach, and they will lead to more aligned and comparable results between companies. For transparency and reproducibility, the code is publicly available at <https://github.com/rafaexx/sbtn-SoN-water>

# How were these layers prepared?

# 1) Shapefile with values at HydroBasins (Pfafstetter level 6)

We used the best available geospatial datasets for water availability and pollution, which are publicly available (as of July 2024), and which has global coverage.

For the SoN for water availability, we considered the following datasets:

* Baseline Water Stress ([Kuzma et al. 2023](https://www.wri.org/research/aqueduct-40-updated-decision-relevant-global-water-risk-indicators)), data available [here](https://www.wri.org/data/aqueduct-global-maps-40-data)
* Water Depletion ([Brauman et al. 2016](https://doi.org/10.12952/journal.elementa.000083)), data available [here](http://www.earthstat.org/water-depletion-watergap3-basins/)
* Blue water scarcity ([Mekonnen & Hoekstra 2016](https://doi.org/10.1126/sciadv.1500323)), data upon request to the authors

For the SoN for water pollution[[1]](#footnote-1), we considered the following datasets:

* Coastal Eutrophication Potential ([Kuzma et al. 2023](https://www.wri.org/research/aqueduct-40-updated-decision-relevant-global-water-risk-indicators)), data available [here](https://www.wri.org/data/aqueduct-global-maps-40-data)
* Nitrate-Nitrite Concentration ([Damania et al. 2019](http://hdl.handle.net/10986/32245)), data available [here](https://wbwaterdata.org/dataset/global-nitrate-nitrite-surface-water)
* Periphyton Growth Potential ([McDowell et al. 2020](https://doi.org/10.1038/s41598-020-60279-w)), data available [here](https://figshare.com/s/e44fe7319585183d888f)

In general, we performed the same processing steps for all datasets:

1. Harmonize spatial resolution to [HydroSHEDS HydroBasins](https://www.hydrosheds.org/products/hydrobasins) Pfafstetter level 6. In the case of Baseline Water Stress and Coastal Eutrophication Potential datasets no harmonization was needed as these are already at such resolution. In the case of Blue Water Reduction Target, we transposed the original data from Pfafstetter level 5 to level 6, without altering the data. For the other datasets, that are originally gridded data (raster format), we compute the area-weighted median for the HydroBasins Pfafstetter level 6.
2. Classify datasets to a common range as reclassifying raw values to 1-5 values, where 0 (zero) was used for cells or features with no data. See the thresholds used for each dataset on the next pages. These were dataset’s authors definitions and/or recommendations.

Lastly, we put all data together and identified the maximum value between the classified datasets, separately, for Water Availability and for Water Pollution. See the final data’s columns and description in page 5.

Below we briefly describe the datasets used and how values were reclassified to 1-5 values.

**Baseline Water Stress** ([Kuzma et al. 2023](https://www.wri.org/research/aqueduct-40-updated-decision-relevant-global-water-risk-indicators)), data available [here](https://www.wri.org/data/aqueduct-global-maps-40-data)

The WRI’s Baseline Water Stress measures the ratio of total surface and groundwater withdrawals to available renewable water. This indicator is based on model outputs from PCR-GLOBWB 2 to compute average monthly values, for the period 1960-2019, then to produce regression values for the year 2019 (baseline). Raw values were classified as follows:

|  |  |  |
| --- | --- | --- |
| Raw value | Label | New value |
| ≤ 0.1 | Low (<10%) | 1 |
| > 0.1 – 0.2 | Low - Medium (10-20%) | 2 |
| > 0.2 – 0.4 | Medium - High (20-40%) | 3 |
| > 0.4 – 0.8 | High (40-80%) | 4 |
| > 0.8 | Extremely High (>80%) | 5 |
|  | Arid and Low Water Use | 5 |

**Water Depletion** ([Brauman et al. 2016](https://doi.org/10.12952/journal.elementa.000083)), data available [here](http://www.earthstat.org/water-depletion-watergap3-basins/)

Water Depletion measures the ratio of surface and ground water consumptive use to available renewable water. This indicator is based on model outputs from WaterGAP3 to compute average annual and monthly values, for the period 1971-2000, and to map seasonal depletion and dry-year depletion. Raw values were classified as follows:

|  |  |  |
| --- | --- | --- |
| Raw value | Label | New value |
| 1 | ≤ 5% annual depletion | 1 |
| 2 | > 5 – 25% annual depletion | 2 |
| 5 | > 25 – 75% annual depletion and at least 3 out of 30 years had at least one month with monthly depletion ratio >75% (Dry-Year) | 3 |
| 6 | > 25 – 75% annual depletion and at least one month every year the monthly depletion ratio is >75% (Seasonal) | 4 |
| 7 | > 75 – 100% annual depletion | 5 |
| 8 | > 100% annual depletion | 5 |

**Blue water scarcity** ([Mekonnen & Hoekstra 2016](https://doi.org/10.1126/sciadv.1500323)), data upon request to the authors

Blue water scarcity measures the ratio of the blue water footprint to the total blue water availability. This indicator is based on the global standard for water footprint assessment to compute average monthly values and an annual average value (10-year average for the period 1996-2005). Raw values were classified as follows:

|  |  |  |
| --- | --- | --- |
| Raw value | Label | New value |
| ≤ 0.2 | Very low | 1 |
| > 0.2 – 1 | Low | 2 |
| > 1 – 2 | Moderate | 3 |
| > 2 – 5 | Significant | 4 |
| > 5 | Severe | 5 |

**Coastal Eutrophication Potential** ([Kuzma et al. 2023](https://www.wri.org/research/aqueduct-40-updated-decision-relevant-global-water-risk-indicators)), data available [here](https://www.wri.org/data/aqueduct-global-maps-40-data)

Coastal eutrophication potential measures the potential for riverine loadings of nitrogen (N), phosphorus (P), and silica (Si) to stimulate harmful algal blooms in coastal waters. Raw values were classified as follows:

|  |  |  |
| --- | --- | --- |
| Raw value | Label | New value |
| ≤ -5 | Low | 1 |
| > -5 – 0 | Low - Medium | 2 |
| > 0 – 1 | Medium - High | 3 |
| > 1 – 5 | High | 4 |
| > 5 | Extremely High | 5 |

**Nitrate-Nitrite Concentration** ([Damania et al. 2019](http://hdl.handle.net/10986/32245)), data available [here](https://wbwaterdata.org/dataset/global-nitrate-nitrite-surface-water)

Nitrate-Nitrite Concentration is based on a combination of monitoring data and a machine learning prediction model. Here we used the average predicted values between 2006 and 2010, i.e. the last 5 years of available data. Raw values were classified as follows:

|  |  |  |
| --- | --- | --- |
| Raw value | Label | New value |
| ≤ 0.4 | Very Low Concentration | 1 |
| > 0.4 – 0.8 | Low Concentration | 2 |
| > 0.8 – 1.2 | Moderate Concentration | 3 |
| > 1.2 – 1.6 | High Concentration | 4 |
| > 1.6 | Very High Concentration | 5 |

**Periphyton Growth Potential** ([McDowell et al. 2020](https://doi.org/10.1038/s41598-020-60279-w)), data available [here](https://figshare.com/s/e44fe7319585183d888f)

Periphyton Growth Potential is based on global model of dissolved and total nitrogen (N) and phosphorus (P) concentrations and ratios to determine which nutrient is limiting periphyton proliferation during the growing season. Raw values were classified as follows, i.e., a similar classification as in the original publication, however, using more thresholds:

|  |  |  |
| --- | --- | --- |
| Raw value | Label | New value |
| N:P ratio < 7 and N ≤ 0.4 | N-limited growth acceptable | 1 |
| N:P ratio < 7 and N < 0.4 – 0.8 | N-limited growth acceptable | 2 |
| N:P ratio < 7 and N < 0.8 – 1.2 | N-limited growth undesirable | 3 |
| N:P ratio < 7 and N < 1.2 – 1.6 | N-limited growth undesirable | 4 |
| N:P ratio < 7 and N > 1.6 | N-limited growth undesirable | 5 |
| N:P ratio ≥ 7 and P ≤ 0.023 | P-limited growth acceptable | 1 |
| N:P ratio ≥ 7 and P < 0.023 – 0.046 | P-limited growth acceptable | 2 |
| N:P ratio ≥ 7 and P < 0.046 – 0.100 | P-limited growth undesirable | 3 |
| N:P ratio ≥ 7 and P < 0.100 – 0.150 | P-limited growth undesirable | 4 |
| N:P ratio ≥ 7 and P > 0.150 | P-limited growth undesirable | 5 |

# 

Columns and description of the data available at <https://doi.org/10.5281/zenodo.7797979>

|  |  |
| --- | --- |
| **HYBAS\_ID** | id from the [HydroSHEDS HydroBasins](https://www.hydrosheds.org/products/hydrobasins) Level 6 (h6) |
| **WMOBB\_id** | id from the [WMO Basins and Sub-Basins](https://panda.maps.arcgis.com/home/item.html?id=be4b6f13121b4670ad8f006bc1908e14) |
| **WMOBB\_name** | Name of the basin in which the h6 is within |
| **region** | Name of the region in which the h6 is within |
| **wa\_max** | Maximum value between water availability layers: **bws\_n**, **wdp\_n**, **wsb\_n** |
| **bws\_raw** | Raw value of Baseline Water Stress (as in the original dataset) |
| **bws\_n** | Variable **bws\_raw** classified to 1-5 values |
| **bws\_label** | Label of Baseline Water Stress classes (based on the original dataset) |
| **wdp\_raw** | Median of Water Depletion pixel values within h6 (previously classified to 1-5) |
| **wdp\_n** | Variable **wdp\_raw** classified to 1-5 values |
| **wdp\_label** | Label of Water Depletion classes (based on the original dataset) |
| **wsb\_raw** | Median of Blue Water Scarcity pixel values within h6 |
| **wsb\_n** | Variable **wsb\_raw** classified to 1-5 values |
| **wsb\_label** | Label of Blue Water Scarcity classes (based on the original dataset) |
| **wp\_max** | Maximum value between water pollution layers: **cep\_n**, **nox\_n**, **pgp\_n** |
| **cep\_raw** | Raw value of Coastal Eutrophication Potential (as in the original dataset) |
| **cep\_n** | Variable **cep\_raw** classified to 1-5 values |
| **cep\_label** | Label of Coastal Eutrophication Potential classes (based on the original dataset) |
| **nox\_raw** | Median of Nitrate-Nitrite Concentration pixel values within h6 |
| **nox\_n** | Variable **nox\_raw** classified to 1-5 values |
| **nox\_label** | Label of Nitrate-Nitrite Concentration classes |
| **tnc\_raw** | Median of Total Nitrogen Concentration pixel values within h6 |
| **tpc\_raw** | Median of Total Phosphorus Concentration pixel values within h6 |
| **pgp\_n** | Variable **tnc\_raw** and **tpc\_raw** classified to 1-5 values |
| **pgp\_label** | Label of Periphyton Growth Potential classes (based on the original dataset) |

# 2) Excel file with values at sub-national divisions and national divisions

The administrative level aggregations are based on the shapefile of HydroBasins Pfafstetter level 6 (described above) and [Global Administrative Areas (GADM) data, version 4.1](https://gadm.org/download_world.html), level 0 (national) and level 1 (sub-national). Additionally, we used gridded data on water demand of different sectors and cropland extent to generate sector-specific aggregations, to better represent sectors’ impacts. We applied a weighted average approach ([Gassert et al. 2015](https://wriorg.s3.amazonaws.com/s3fs-public/Aqueduct_Global_Maps_2.1-Constructing_Decicion-Relevant_Global_Water_Risk_Indicators_final_0.pdf)), where basins with more water demand/impact have a higher influence over the final score of the administrative region.

This approach was repeated for all six individual indicators. Lastly, we put all data together and identified the maximum value between the datasets, separately, for Water Availability and for Water Pollution, and per sector. The aggregations were organized as follow:

|  |  |
| --- | --- |
| **Suffix** | **Description** |
| \_n | Aggregation at administrative level, without weighting |
| \_dom | Aggregation at administrative level weighted by domestic water demand\* |
| \_ind | Aggregation at administrative level weighted by industrial water demand\* |
| \_irr | Aggregation at administrative level weighted by irrigation water demand\* |
| \_liv | Aggregation at administrative level weighted by livestock water demand\* |
| \_tot | Aggregation at administrative level weighted by total water demand\* |
| \_agr | Aggregation at administrative level weighted by cropland extent\*\* |

\* Data for 'water demand' differs depending on the indicator aggregated: for Baseline Water Stress, WRI used water withdrawal from PCR-GLOBWB 2; for Water Depletion, we used water consumption from WaterGAP 3; and for Blue Water Scarcity, we used blue water footprint from Water Footprint Network.

\*\* For the water pollution indicators, we used cropland extend from [Potapov et al. 2021](https://www.nature.com/articles/s43016-021-00429-z).

Note: Aggregations for Baseline Water Stress were not recomputed. These are as the ones available in the [Aqueduct 4.0 Country Rankings](https://www.wri.org/data/aqueduct-40-country-rankings), but with values rescaled from 0-5 to 1-5, to be aligned with the other indicators.

1. Note that, for the moment, water pollution datasets are limited to nutrient pollution. [↑](#footnote-ref-1)