

Funded by the European Union within the
7th Framework Programme, Grant Agreement 603378.
Duration: February 1st, 2014 – January 31th, 2018



Deliverable 7.2: MARS Suite of Tools II

D7.2-1: Scenario Analysis Tool (SAT)

D7.2-2: Bayesian Belief Networks

Lead beneficiary: IGB, Germany (D7.2-1), Deltares, Netherlands (D7.2-2)

Authors: Multiple authors, see inside for details

Due date of deliverable: **Month 45**

Actual submission date: **Month 48**

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Content

MARS deliverable 7.2 contains two parts: "D7.2-1: Scenario Analysis Tool (SAT)" and "D.7.2-2: Bayesian Belief Networks".

Summary

D7.2-1: Scenario Analysis Tool (SAT)

Report on data, scientific methods and tool implementation

The reduction of some dominating stressors in many surface waters of Europe over the last decades revealed the impact of multiple, presumably minor, but jointly acting stressors. The MARS scenario analysis tool (SAT) addresses the type of interactions between selected main stressors and their current and future impact on aquatic ecosystems at the European scale. The resolution of model results is limited to the FEC level (Functional Elementary Catchments, with a mean spatial resolution of 58 km²), but the European wide application opens long gradients and increases the number of relevant stressors, thus potentially allowing to identify stressor-response relationships which are often concealed at smaller scales.

The MARS Scenario Analysis Tool (SAT) is an online tool to visualize and analyse multi-stressor conditions in European rivers. With 6.13 Mio. km², the model extent covers EU-27 countries, EFTA states and hydrological connected areas (e.g. of Ukraine/Danube or Russia/Baltic Sea). The tool operates at the level of 104,300 hydrological sub-catchments, resembling spatial units similar to the 'water bodies' delineated by the European countries for the surface water management according to the WFD. The SAT provides a harmonized European-wide assessment, comparing geo-climatic regions under different anthropogenic stress, with an emphasis on aggregation levels larger than 1,000 km² and mean conditions over a ten year period (due to the underlying data and model features). It also offers a detailed overview of stressor conditions and potential impact on the ecological status across Europe. The tool predicts the effect of selected mitigation measures and targets users working on EU legislation, river basin managers, and scientists interested in multi-stressor conditions in a broad context.

D7.2-2: Bayesian Belief Networks: Linking abiotic and biotic data

In this deliverable, the predictive use of Bayesian Belief Networks (BBNs) is presented for several case studies for rivers and lakes in Europe. The construction of this BBNs is based on the results of MARS WP 4, in which causal relationships are constructed according to the DPSIR- approach, ensuring causal relationships between causes for deterioration, pressures, state variables and biota. In WP4, these relationships were subsequently statistically tested with large). The aim of Deliverable 7.2-2 was to combine abiotic and biotic models for river basin management planning. In this work package, BBNs have been used for the coupling of these models.

In this report we have developed predictive BBN models for five case studies catchments across Europe to explore the effects of future scenarios on biological responses and ecological status of water bodies. The case studies cover many dimensions of the MARS project, such as:

- Three regions of Europe (North, Central, South), with case studies from Finland (Lepsämäenjoki), Denmark (Odense), The Netherlands (Regge and Dinkel), Portugal (Sorraia), and Norway (Vansjø);
- The two water categories: rivers and lakes;
- The three story lines: Techno, Fragmented and Consensus world that have been used in MARS work package 4.2;
- Various stressor types: Total P, Total N, hydrology, hydromorphological alterations, temperature, etcetera;
- Biological indicators: chlorophyll a in rivers and lakes, cyanobacteria in lakes, macrophytes, macroinvertebrates, fish, and total ecological status of the water body.

For all case studies, the BBN method enabled the coupling of abiotic and biotic models, and facilitated predictions of biological responses under the different future storylines. Therefore, BBNs had a clear additional value compared to the abiotic process-based catchment models (MARS work package 4). Below, the main results are presented for the case studies.

Funded by the European Union within the
7th Framework Programme, Grant Agreement 603378.
Duration: February 1st, 2014 – January 31th, 2018



Deliverable 7.2-1: MARS Suite of Tools II

Scenario Analysis Tool (SAT)

Report on data, scientific methods and tool implementation

Lead beneficiary: **Markus Venohr (IGB)**

Contributors (in alphabetic order): Sebastian Birk (UDE), Vanessa Bremerich (IGB), Andreas Gericke (IGB), Lidija Globevnik (UL), Maja Koprivšek (UL), Judith Mahnkopf (IGB), Yiannis Panagopoulos (NTUA), Luka Snoj (UL), Marta Faneca Sánchez (DELTARES), Kostas Stefanidis (NTUA), Frederiek Sperna Weiland (DELTARES).

Due date of deliverable: **Month 45, October 2017**

Actual submission date: **Month 48, January 2018**

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Content

1.	General concept of the Scenario Analysis Tool	7
2.	Considered input data and models	12
2.1	PCR-GlobWB	12
2.2	MONERIS	21
3.	Joint Geo-Database, data analysis and post-processing	47
3.1	Hydrological indicators.....	47
3.2	Geodata-Base, effects and thresholds.....	55
4.	Implementation of the Scenario Analysis Tool	60
5.	General conclusions	79
	References.....	83
	Appendix	88

List of figures

Figure 1: Principle concept of a simplified DPSIR approach used for the scenario analysis tool. 9

Figure 2: Conceptual model of the data flow to assess the impacts of multiple stressors on a European scale..... 10

Figure 3: The “challenges space” to be spanned by SSPs (source: O’Neill et al., 2013; based on Kriegler et al. 2012, Fig. 3), divided into five “domains” with one SSP located within each domain, represented by a star 14

Figure 4: PCR-GLOBWB model structure, adapted from Van Beek et al. (2011). Symbols’ definitions are as follows: Precip, precipitation; Evap, evaporation; T , temperature; S1, first soil layer; S2, second soil layer; S3, third soil layer; S4, groundwater reservoir; Q_{channel}, total runoff; QDR, direct runoff; QIF, intermediate flow; QBF, base flow; and Inf, water flow from the river channel to the groundwater reservoir..... 16

Figure 5: Monthly average discharge large European discharge simulated with PCR-GLOWB (blue) compared with observations from the GRDC (orange). 18

Figure 6: Monthly average discharge changes for large European discharge simulated with PCR-GLOWB based on historic data (blue) together with discharges for the different RCPs and future time horizons..... 20

Figure 7: Correlation between P-content in soils and measured WSP in soil samples of Germany and Swiss according to Pöthig, Behrendt, Opitz, & Furrer, 2010 and, Pöthig (unpublished data). 23

Figure 8: Available P-balance on country (left) and the accomplished time series (right). 24

Figure 9: P-accumulation on agricultural land per country in the period from 1950 and 2014. 25

Figure 10: Some of the main input datasets for the application of MONERIS: (a) reclassified land use (b) mean soil loss per FEC in t/ha (c) Nitrogen Surplus per FEC in kg/ha*a (d) degree of phosphorus saturation per FEC in % (e) population density for the year 2010 (f) Inhabitants connected to WWTP and sewer system for the year 2010 [%] 27

Figure 11: Plot of water temperature and air temperature (time period = 2001-2010) for a FEC (ID = 35402) situated in Germany. 29

Figure 12: Dependency of connection rate to population density for 175 NUT2 units in 16 European (AT, BG, CY, CZ, DE, EL, FR, HU, IE, LV, NL, PL, RO, SI, SK, TR) countries (left). Corrected connection rates at different CSCT compared to the base model. 32

Figure 13: Comparison of inhabitants connected to WWTP derived from country statistics (years 2006-2011) and the UWWTD inventory (year 2012). 33

Figure 14: Structure of the MONERIS model showing the external framework, catchment characteristics, pathways, and surface waters. 37

Figure 15: Response of the new approach for P retention and remobilisation to hydraulic load (HL), left: large rivers - mean HL = 10000 m/yr, mean slope = 1 %), middle: small lowland brook - mean HL = 500 m/yr, mean slope = 1 %), right: small hilly brook - mean HL = 1000 m/yr, mean slope = 15 %). 41

Figure 16: Comparison of the accumulative share of FECs on the total area and their share on the total emissions, for nitrogen and phosphorus in the entire MARS modelling extent. 41

Figure 17: Some of the main results delivered by MONERIS: (a) specific nitrogen emissions in kg/ha (b) specific phosphorus emissions in kg/km² (c) nitrogen retention in % (d) phosphorus retention in % (e) DIN concentrations in mg/l (f) phosphorus concentrations in mg/l..... 44

Figure 18: Comparison of mean modelled and observed TN and TP loads for the years 2001-2010. Categories are characterised by catchment size, deviation between modelled and observed run-off (provided by PCR-GlobWB), and the number of monitoring years from GRDC-EWA. For attribution of categories see Table 8. 45

Figure 19: From left to right: Example showing a) calculated centroids of FECs polygons, b) calculated centroids of PCR-GLOBWB model raster cells, c) buffer zones created based upon the FEC centroids and d) the PCR-GLOBWB centroid cells that fall within the created buffer zones.. 50

Figure 20: Top Left: Mean Annual Flow (m³/s) of the 10-y period 2001-2010 under the natural scenario (historic climate with no water abstractions). Top Right: Alteration of baseline (2001-2010 with abstractions) mean annual flow from the natural conditions, expressed as mean annual flow (baseline) / mean annual flow (natural). Ratios below unity indicate decrease in flow due to abstractions, values equal to unity show no alteration and values above unity show increase from the natural conditions. Reduction is depicted with lighter colours and is observed mostly in parts of the Mediterranean countries. Increase in flows (ratio >1) is found in Central Europe, with very large parts all across the continent and especially in the North remaining unaltered with respect to annual river flows (ratio = 1). Middle and Bottom: Alteration of mean annual flow under four future scenarios (S). S1: RCP4.5 - SSP2 for 2026-2035, S2: RCP4.5 - SSP2 for 2056-2065, S3: RCP8.5 – SSP5 for 2026-2035 and S4: RCP8.5 – SSP5 for 2056-2065..... 52

Figure 21 Top Left: Base Flow Index of the 10-y period 2001-2010 under the natural scenario (historic climate with no water abstractions). Base flow Index is defined as the 7-d minimum flow/mean annual flow of the year. Top Right: Alteration of the base flow index from the natural conditions, expressed as BFI (baseline) / BFI (natural). Ratios below unity indicate decrease in

BFI due to abstractions, values equal to unity show no alteration and values above unity show increase from the natural conditions. Reduction is depicted on the right map with lighter colours and is observed mostly in parts of the Mediterranean countries. Increase in BFI (ratio >1) is found in Central Europe, with very large parts all across the continent and especially in the North remaining unaltered. Middle and Bottom: Alteration of BFI under four future scenarios (S). S1: RCP4.5 - SSP2 for 2026-2035, S2: RCP4.5 - SSP2 for 2056-2065, S3: RCP8.5 – SSP5 for 2026-2035 and S4: RCP8.5 – SSP5 for 2056-2065. 53

Figure 22: Principle workflow and data exchange for calibrating and learning the SAT-approach and for the application at scenario conditions. 60

Figure 23: Structure of the Bayesian Belief Network to estimate the probability to reach a good ecological status, shown for the example of broad river type “RT8-Highland and glacial rivers”, using the Software package GENIE. 64

Figure 24: Share of individual stressor indicator combinations per FEC on all FECs in the MARS modelling extent (left). Share of FECs with the same number of active stressor indicators. Both figures refer to current conditions 2001-2010. 66

Figure 25: Probability to reach a good ecological status derived from the trained Bayesian Belief Network for the ten different broad River Types 67

Figure 26: a) mean annual precipitation in 2001-2010, b) difference between the mean annual precipitation for Storyline2 in 2026-2035 and the mean annual precipitation in 2001-2010. 70

Figure 27: a) Current percentage share of agricultural areas on the FEC area, and b) percentage changes in agricultural areas between current conditions and Storyline 2 in 2026-2035. 70

Figure 28: N-surplus on agricultural areas for a) the year 2009 and under b) assumption of applied mitigation measures in $\text{kg ha}^{-1} \text{yr}^{-1}$ 70

Figure 29: Change of the DIN concentration at FEC outlet for Storyline 1, 2026-2035, with and without mitigation measures, compared to the mean concentration at current conditions, 2001-2010. 71

Figure 30: Change of the TP concentration at FEC outlet for Storyline 1, 2026-2035, with and without mitigation measures, compared to the mean concentration at current conditions, 2001-2010. 71

Figure 31: Probability to reach a good ecological status for a) current conditions (2001-2010), b) Storyline 2, 2026-2035, no mitigation measures, and c) Storyline 2, 2026-2035, with mitigation measures. Maps e) and f) show the difference of maps b) minus a) and c) minus a),

respectively. Diagram d) displays the calculated probabilities to reach a good ecological state derived from a), b), and c). 72

Figure 32: Start window after initial start of the SAT. 74

Figure 33: The navigation menu. 75

Figure 34: The results panel. 76

Figure-A 1: Aggregated broad river types derived for the modelling extent in MARS, on basis of ETC/ICM(2015).....90

Figure-A 2: Share of correctly predicted ecological status in % of all FECs of broad river Type derived from the trained Bayesian Belief Network. 91

Figure-A 3: Mean specific total nitrogen and total phosphorus emissions for the years 2001-2010 per country in the MARS modelling extent..... 93

List of tables

Table 1: Performance statistics for river discharges simulated with PCR-GLOBWB compared to GRDC measured discharges. 16

Table 2: Brief overview of the main input datasets used for the application of MONERIS 21

Table 3: Relationships between MONERIS classes and attributes AQUIFER_TYPE and LITHO5 of the IHME map. 22

Table 4: Completed national statistic on inhabitants connected to sewer systems in the years 2000-2011 based on reported data from (EUROSTAT). Colour code: coloured cells indicate countries, for which the same connection rate was assumed; red numbers indicate countries to which connection rates were transferred. No-coloured cells with red figures indicate Years, for which connection rates were calculated from linear interpolations. 31

Table 6: Mean WWTP size specific effluent nutrient concentrations derived from the UWWTD inventory, calculate as discharge weighted mean from all WWTP of a specific size class. 34

Table 7: Considered soil-type specific TP concentrations for the calculation of emissions via tile drainage and groundwater. 38

Table 8: Share of both nitrogen and phosphorus emissions from different land-use types and via considered pathways, area specific emission for nitrogen in kg/ha and for phosphorus in kg/km², numbers in brackets represent the share on the total nitrogen or phosphorus emissions. WSA = water surface area. 42

Table 9: Comparison of mean modelled and observed TN and TP loads distinguished by categories. 45

Table 10: Description of broad river types for which thresholds were defined for the scenario analysis tool 56

Table 11: Variables predicting general ecological state with the highest explanatory power (%).	57
Table 12: Thresholds for six stressor indicators by broad river types derived with classifier algorithm C4.5.	58
Table 13: Classification accuracy and selected classification scheme parameters used to derive thresholds for selected pressure variables.	59
Table 14: List of considered stressor indicators.....	61
Table 15: Broad river type specific thresholds for the six considered stressor indicators. RT1- very large rivers, RT2 – Lowland brooks, RT3 – Lowland streams and rivers, RT4 – Mountains, siliceous, brooks, RT5 – mountains, siliceous, streams and rivers, RT6 – mountains, calcareous, brooks, RT7 – Mountainous, calcareous, stream and rivers, RT8 – Highland and glacial rivers, RT9 – Mediterranean perennial rivers, RT10 – Mediterranean temporary or very small rivers (brooks).	63
Table 16: Share of FECs with allocated river type, ecological status information and the share of FEC with a good or high reported status per broad river type (BRT). ¹⁾ Values in brackets describe the percentage share on all 104,300 FECs of the entire MARS extent. ²⁾ Values in brackets describe the percentage share of all FECs allocated to the respective BRT. ³⁾ Values in brackets describe the percentage share on all FEC with reported status per BRT.....	64
Table 17: Mean euclidean strength of influence derived using software „Genie“	65
Table 18: Ten most frequent stressor indicator combination in the MARS modelling extent. X = active, - = inactive.	66
Table 19: Parameters and mean changes considered for the scenario modelling.	68
Table-A 1: Land use classification used within MONERIS.....	88
Table-A 2: Land use classification of the GLCC data used within MONERIS for scenario calculations.....	89
Table-A 3: Country specific data sources on population density, used to supplement population data from EUROSTAT.	90
Table-A 4: The list of candidates stressor indicators as variables to estimate the probability to reach a GoodOrBetter ecological status.....	90
Table-A 5: Completed national statistic on inhabitants connected to sewer systems in the years 2000- 2011 based on reported data from (EUROSTAT). Colour code: coloured cells indicate countries, for which the same connection rate was assumed; red numbers indicate countries to which connection rates were transferred. No-coloured cells with red figures indicate Years, for which connection rates were calculated from linear interpolation. Blue coloured figures indicate that values had to be adapted due to data inconsistencies.	92

1. GENERAL CONCEPT OF THE SCENARIO ANALYSIS TOOL

Lead: Markus Venohr, Vanessa Bremerich, Judith Mahnkopf

Introduction

The reduction of some dominating stressors in many surface waters of Europe over the last decades revealed the impact of multiple, presumably minor, but jointly acting stressors. The MARS scenario analysis tool (SAT) addresses the type of interactions between selected main stressors and their current and future impact on aquatic ecosystems at the European scale. The resolution of model results is limited to the FEC level (Functional Elementary Catchments, with a mean spatial resolution of 58 km²), but the European wide application opens long gradients and increases the number of relevant stressors, thus potentially allowing to identify stressor-response relationships which are often concealed at smaller scales.

Scope of the SAT

The MARS Scenario Analysis Tool (SAT) is an online tool to visualize and analyse multi-stressor conditions in European rivers. With 6.13 Mio. km², the model extent covers EU-27 countries, EFTA states and hydrological connected areas (e.g. of Ukraine/Danube or Russia/Baltic Sea). The tool operates at the level of 104,300 hydrological sub-catchments, resembling spatial units similar to the ‘water bodies’ delineated by the European countries for the surface water management according to the WFD. The SAT provides a harmonized European-wide assessment, comparing geo-climatic regions under different anthropogenic stress, with an emphasis on aggregation levels larger than 1,000 km² and mean conditions over a ten year period (due to the underlying data and model features). It also offers a detailed overview of stressor conditions and potential impact on the ecological status across Europe. The tool predicts the effect of selected mitigation measures and targets users working on EU legislation, river basin managers, and scientists interested in multi-stressor conditions in a broad context.

Principle approach

In agreement to the overall concept of the MARS project we followed the Driver, Pressure, State, Impact, and Response (DPSIR) concept as the central approach for the SAT. However, the DPSIR approach had to be simplified as a consistent driver, pressure, and state cascade could not be derived on a European scale and was substituted by the collective term “stressor indicators” (Figure 1). Further, the impact assessment on biodiversity measures could only be done indirectly. Instead, we used the ecological state

reported by the European countries for the second assessment of the EU-WFD (Table 15). Based on the results in Chapter 3.2, we developed a Bayesian Belief Network to estimate the probability to reach a good ecological status under a specific combination of active stressor indicators.

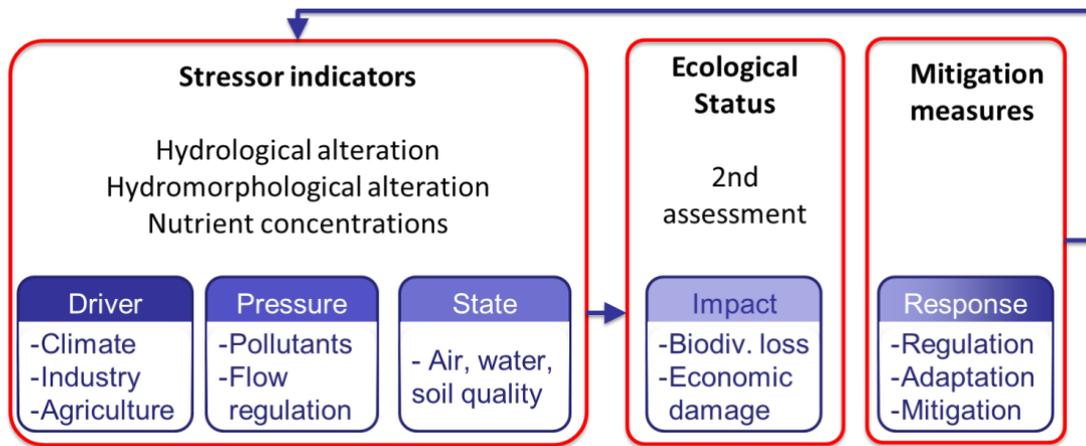


Figure 1: Principle concept of a simplified DPSIR approach used for the scenario analysis tool.

The backbone of the SAT is a combination of the models PCR-GlobWB (Chapter 2.1) and MONERIS (Chapter 2.2) and is linked with the MARS geo-database (Chapter 3.2). PCR-GlobWB provides information on daily water balances for near-natural (i.e. no reservoirs, no water abstraction or addition) current and future conditions. These data are used to analyse hydrological alterations (Chapter 3.1), and as input data for MONERIS which in turn quantifies nitrogen and phosphorus emissions to surface waters, in-stream retention, and resulting loads and concentrations.

Results of both models, together with an additional extended data collection on various catchment parameters and a complete data set of the ecological status reported by the EU member states feed into the MARS geo-database. The outputs links climate, water availability, nutrient fluxes and management options by quantifying and evaluating multi-stressor conditions and the aquatic response.

Machine learning approaches (Chapter 3.2) have been used to identify major stressor indicators with the strongest power to explain the ecological state reported for the 2nd assessment of the EU-WFD. Six dominant stressor indicators were identified and thresholds, i.e. tipping points describing the impact of a stressor indicator on the reported ecological state, were derived using regression tree analysis. A stressor is considered active if the threshold is exceeded, and inactive if the value remains below the threshold. As thresholds for active stressors vary considerably between different river types, the analysis was conducted for different broad river types (BRT) (ETC/ICM, 2015). The stressor indicators and derived thresholds are used in Bayesian Belief Networks to derive probabilities for a FEC to reach a good or high

ecological status. This allowed estimating probabilities to reach a good ecological status not only in regions outside the EU, considered in this modelling task, but also under future conditions.

All data have been collated and models been applied for the period 2001-2010. This period was used to conduct the statistical analysis, identify stressor indicators and relevant thresholds and finally to train the Bayesian belief network for estimating the probability to reach a good ecological status. All analyses were conducted on FEC level.

Considered scenarios and mitigation measures

In addition to the current state conditions we derived consistent data sets for two future periods (2026-2035 and 2056-2065, Figure 2) for two of the three story lines 1) Techno world and 2) Consensus world (Figure 2, Table 18). The MARS storylines are described in detail in “Report on the MARS scenarios of future changes in drivers and pressures with respect to Europe’s water resources” (MARS Deliverable 2.1).

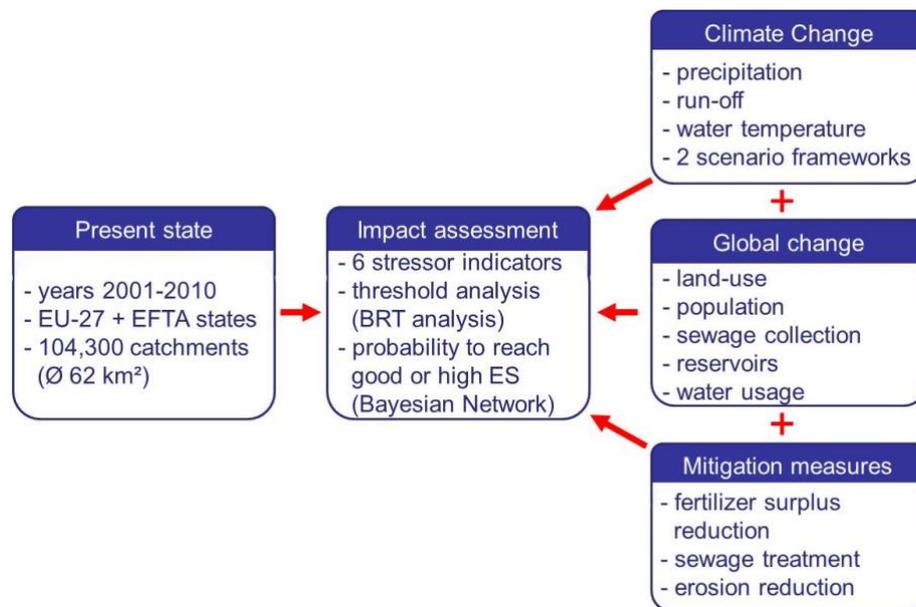


Figure 2: Conceptual model of the data flow to assess the impacts of multiple stressors on a European scale.

MARS Storyline 1 “Techno World” is based on the representative concentration pathway RCP 4.5 (moderate change) in combination with the Shared Socioeconomic Pathways SSP2. MARS Storyline 2 “Consensus World” is based on RCP 8.5 (largest changes) in combination with the Shared Socioeconomic Pathways SSP5.

Within these scenario frameworks we modelled and considered changes in climatic conditions (e.g. precipitation, air temperature), runoff, water temperature, construction of new reservoirs, land use, population, and collection and treatment rates. The data bases, assumptions, back ground and changes

are described in more detail throughout the respective Chapters 2.1, 2.2, 3.1 and 4. Additionally, mitigation measures (e.g. N surplus reduction, improved waste water collection and treatment, see Figure 2 and Chapter 2.2) to reduce nutrient emissions to and concentrations in surface waters were derived and applied in combination with the scenario frameworks. The type and extent of measures selected on basis of feasibility, rather than from a reduction need to improve the ecological state. This is indicated in Figure 1 by a missing arrow between status and response. Nevertheless, the effect of these measures on the nutrient concentrations and subsequently on the probability to reach a good ecological status was quantified and assessed. Changing stressor indicator values modelled for scenario conditions are translated by the trained Bayesian Belief Network to future probabilities to reach a good or high ecological status.

2. CONSIDERED INPUT DATA AND MODELS

2.1 PCR-GLOBWB

Lead: Frederiek Sperna Weiland, Marta Faneca Sánchez

Input data

Historical meteorological data

Global daily fields of precipitation and temperature were collected from the **ERA-40** reanalysis dataset. The ERA-40 dataset is obtained with a numerical weather prediction system that includes assimilation of meteorological observations. The prediction system has a horizontal resolution of ca. 125 km and a vertical resolution of 60 levels (Uppala et al., 2005). The ERA-40 datasets slightly overestimates precipitation globally, but underestimates precipitation in the Danube basin (Troccoli and Kalberg, 2004; Sperna Weiland et al., 2010). Inter-annual variability is relatively high because the observational data, especially satellite data, included in the system vary over time (Hagemann et al., 2005). The ERA-40 dataset is available for the period 1958-2001

For the period 1979-2010 the ERA-Interim dataset (Dee and Uppala, 2009) is used. This dataset supersedes the ERA-40 reanalysis, and includes several improvements to the numerical weather prediction system. The horizontal resolution has been increased from T159 to T255, the model physics have been improved, radiance information is used for bias-correction and better data sources are utilized for wave height, radiance, and ozone profiles. Nevertheless, a strong correlation exists between the ERA-40 and ERA-Interim reanalysis datasets. Not only is the ERA-Interim system an evolution of the existing ERA-40 system, up to 2001 the boundary forcing of the ERA-Interim system has been taken from the ERA-40 system. For the present analysis the ERA-Interim time-series of precipitation and have been extracted from the ERA-Interim runs.

The monthly amounts of precipitation and monthly average temperature have been retrieved from the monthly **CRU** TS 3.21 dataset (Harris et al., 2014). The CRU data is based on station data that is interpolated as a function of longitude, latitude and elevation above sea level using the thin-plate spline method. The monthly data have been downscaled to a daily time-step using the ERA-40 and for the later period the ERA-Interim re-analysis dataset. The final data have been corrected for snow-undercatch as suggested by Fiedler and Döll (2007).

The resulting daily grids have been disaggregated to the PCR-GLOBWB model resolution of 0.08° (~ 10 km × 10 km at the equator) using nearest neighbour interpolation.

Climate scenarios and data

Future precipitation and temperature datasets have been taken from the datasets belonging to the 5th assessment report of the Intergovernmental Panel on Climate Change (IPCC; IPCC, 2014). Focus is on a short-term (2026-2035) and mid-term (2056-2065) future time-horizon. The precipitation and temperature projections are based on representative concentration pathways (RCPs) that belong to pre-defined emission scenarios (Van Vuuren et al. 2013). There are 4 RCPs available. From modest to extreme:

- RCP 2.6: In this pathway the radiative forcing peaks around 2050 after which there is a modest decline towards 2100 due to a declining use of oil and an overall decrease in energy use;
- RCP 4.5: In this pathway the radiative forcing stabilizes before 2100 due to the introduction of technologies and strategies that reduce greenhouse gas emissions;
- RCP 6.0: Here a stabilization, due to the introduction of technologies for greenhouse gas emissions, is reached after 2100;
- RCP 8.5: In this pathway there is a continuously increasing radiative forcing.

For this project **RCP4.5** (moderate change) and **RCP8.5** (largest changes) were used.

For the calculation of future changes we have used datasets from global climate models (GCMs) that were part of the international inter-sectoral impact model inter-comparison project (ISI-MIP). The ISI-MIP project developed future projections that were later used as input for the IPCC 5th assessment report. The ISI-MIP data portal can be found at: <https://esg.pik-potsdam.de/search/isimip-ft/>.

Within the ISI-MIP project GCM datasets were corrected using the EU-WATCH dataset (Weedon et al., 2011). This dataset is constructed from the ERA-40 re-analysis dataset (ECMWF; Uppala et al., 2005) corrected with the CRU dataset (Mitchell and Jones, 2005).

Running a scenario throughout the full modelling chain described in this project requires quite some calculation time and therefore we have restricted ourselves to the use of one single GCM, namely **GFDL-ESM2M** (developed by the Geophysical Fluid Dynamics Laboratory for NOAA; Dunne et al., 2013).

For the current situation, PCRGLOBWB was forced with a combination of the CRU TS2.1 and ERA-40 / ERA-Interim datasets (CRU-ERA). This is a different dataset than the one used by ISIMIP for the historical and time-period and the bias-correction of the GCM datasets. However, as PCRGLOBWB has been calibrated for the CRU-ERA dataset, we used this one for the current situation. Therefore monthly mean changes derived from the ISIMIP datasets have been applied to the historical CRU-ERA datasets to construct future precipitation, temperature and evaporation time-series.

Socio-economic scenarios

Future socio-economic changes are described by the Shared Socioeconomic Pathways (SSPs). The SSP scenarios serve as a tool for integrated analysis of future climate impacts, vulnerability, adaptation and mitigation. There are 5 SSP scenarios, which all differ in challenges on adaptation and mitigation. Each SSP is a narrative of possible future socio-economic developments. In this report we consider SSP2 (in combination with RCP4.5) and SSP5 (in combination with RCP8.5).

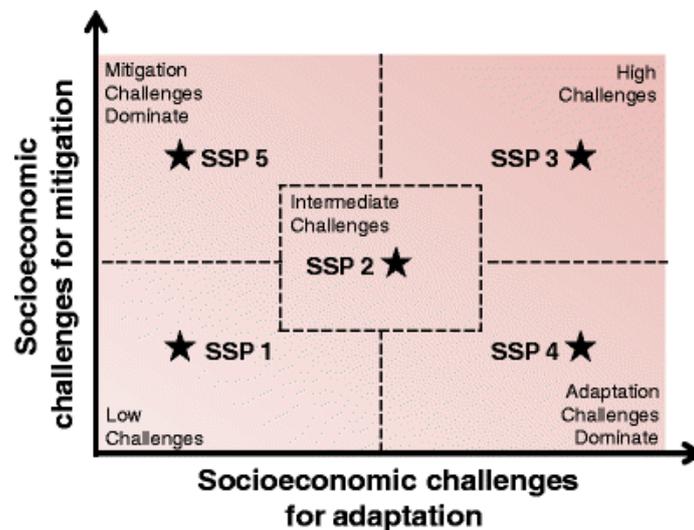


Figure 3: The “challenges space” to be spanned by SSPs (source: O’Neill et al., 2013; based on Kriegler et al. 2012, Fig. 3), divided into five “domains” with one SSP located within each domain, represented by a star

In SSP5 the challenges for mitigation are high due to a lack of climate policy and high emissions whereas at the same time there are factors that reduce the mitigative capacity of the society, such as rapid population increases, large heterogeneity between different groups within the society, lack of political will or limited financial resources (O’Neill et al., 2013). In SSP2 both the mitigative and adaptation challenges are intermediate. This SSP can be seen as a continuation of the current trends, it is the SSP with medium population growth, medium economic growth and medium technological change. Adaptation challenges are a result of the exposure to more extreme changes and the resilient capacity of the society. Narratives of future changes belonging to these pathways have been sketched.

With integrated assessment models changes in and absolute quantitative data of amongst other future population, urbanization, income, energy use, land use and agricultural production can be estimated for the different SSPs. Data for this project have been estimated with the integrated assessment model IMAGE of the Netherlands Environmental Assessment Agency (PBL; van der Esch et al., 2017; Stehfest et al., 2014). The model outcomes provided information to estimate:

- Future irrigation efficiency and irrigated area

- Future domestic and industrial water demand
- Limitations for groundwater abstractions

Reservoirs

In addition to the above SSP information that is included in the model simulations we have added hydropower reservoirs that are planned to be constructed or under construction (see the assessment of Zarfl et al., 2015). As the differentiation in time for the construction of these reservoirs is difficult to assess we have included all planned reservoirs for both future time horizons.

Methods

Model description: PCR-GLOBWB

The global hydrological model PCR-GLOBWB was employed in this study for the simulation of river flows and runoff for Europe (Van Beek and Bierkens, 2009; Van Beek et al., 2011; Lopez Lopez et al., 2016). PCR-GLOBWB is essentially a leaky-bucket type of model applied on a cell-by-cell basis. PCR-GLOBWB is coded in the PCRaster-Python environment (Wesseling et al., 1996; Karszenberg et al., 2010). A spatial resolution of 0.08° (~ 10 km × 10 km at the equator) and a daily temporal resolution were used in this study. A schematic representation of PCR-GLOBWB is given in Figure 4. For each time step and cell, PCR-GLOBWB calculates the water balance components, including the water storage in three vertical soil layers (0–5, 5–30 and 30–150 cm) and one underlying groundwater reservoir, as well as the water exchange between the layers (percolation, capillary rise) and between the top layer and the atmosphere (rainfall, evapotranspiration and snowmelt). Sub-grid variability is taken into account considering the variations of elevation, land cover, vegetation and soil. The total runoff of a cell consists of direct runoff (saturation excess surface runoff), non-infiltrating melt water, interflow (lateral drainage from the soil profile) and base flow (groundwater runoff from the lowest linear reservoir). The simulated runoff is routed along the river network based on the Simulated Topological Networks (STN30; Vörösmarty et al., 2000).

The model has not been calibrated. Model parameterization is based on global datasets of soil properties, vegetation and geological information. This parameterization together with model evaluation details are presented in Van Beek et al. (2011) and Sutanudjaja et al. (2014).

Water abstraction and consumptive water use (domestic, industrial, livestock, irrigation) and reservoir management are included the model and have been derived from global datasets (Wada and Bierkens, 2014). Future changes are based on the Shared Socioeconomic Pathways defined by the IPCC (Van

Vuuren et al., 2013). Abstractions for industry, irrigation, domestic use and livestock can be abstracted from the surface and groundwater depending on availability and pre-defined abstraction limits.

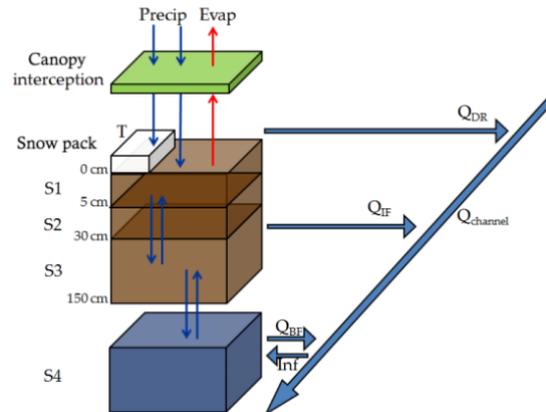


Figure 4: PCR-GLOBWB model structure, adapted from Van Beek et al. (2011). Symbols' definitions are as follows: Precip, precipitation; Evap, evaporation; T, temperature; S1, first soil layer; S2, second soil layer; S3, third soil layer; S4, groundwater reservoir; $Q_{channel}$, total runoff; Q_{DR} , direct runoff; Q_{IF} , intermediate flow; Q_{BF} , base flow; and Inf , water flow from the river channel to the groundwater reservoir.

Results

Validation

Global validation of PCR-GLOBWB has been peer-reviewed published in Van Beek et al. (2011) and Sutanudjaja et al. (2014). Here we show the validation results for a number of large European rivers. Discharge observations have been obtained from the Global Runoff Data Centre (GRDC). Table 1 presents a number of performance measures and Figure 5 displays the monthly average simulated discharge (blue) together with the observed discharge (orange).

Table 1: Performance statistics for river discharges simulated with PCR-GLOBWB compared to GRDC measured discharges.

Location	NRMSE (-)	AbsPercBias s (%)	Nash-Sutcliffe (-)	correlation (-)	meanAbsoluteError (m ³ /s)	KGE (-)
Venlo Meuse	0.42	51.94	0.39	0.84	150.73	0.49
Kienstock Danube	0.19	14.29	0.71	0.85	268.58	0.76
Rheinhalle Basel Rhine	0.20	18.73	0.61	0.88	198.58	0.82
Saultbrenaz Rhone	0.32	22.54	0.33	0.68	105.33	0.66
Wittenberg Elbe	0.48	58.77	-0.15	0.59	211.23	0.31
Ruse Danube	0.23	23.19	0.32	0.75	1350.86	0.66

The performance for the river Rhine and Danube at Kienstock is very good. Although for the Rhine the model overestimates the observed discharge, yet the variation in time is represented very well (correlation of 0.88). For both the Elbe and Meuse the performance is less, especially the base flow is highly overestimated for both rivers. The Meuse is a relatively small basin where sub-daily processes are relevant while these are not considered in PCR-GlobWB that runs on a daily time-step, furthermore the basin contains a number of weirs upstream of Venlo for which the regulation is not included in the model. For the Elbe the correlation between simulated and observed flows is lower than for the other basins, this in combination with the overestimation of the observed discharge indicates poor performance for this basin, which could as well result from too high precipitation amounts in the forcing datasets and underestimations of the demand in the Elbe basin. In the Rhone the discharge peaks are highly overestimated, this is most likely a result of the presences of many weirs and reservoirs that are used for electricity generation. These have a dampening effect on discharge peaks and apparently the reservoirs are not represented well in the model, nor is their regulation.

The above gives an overview of the performance for a selection of rivers spread over Europe. In general it has been reported before that global hydrological models tend to overestimate discharges and runoff in relatively dry river basins (Sperna Weiland et al., 2015). In general the model performs best in basins such as the Rhine where more data is available from ground observations for assimilation in the meteorological re-analysis dataset and the global datasets of land use, soil type and geohydrology used for the model parameterization.

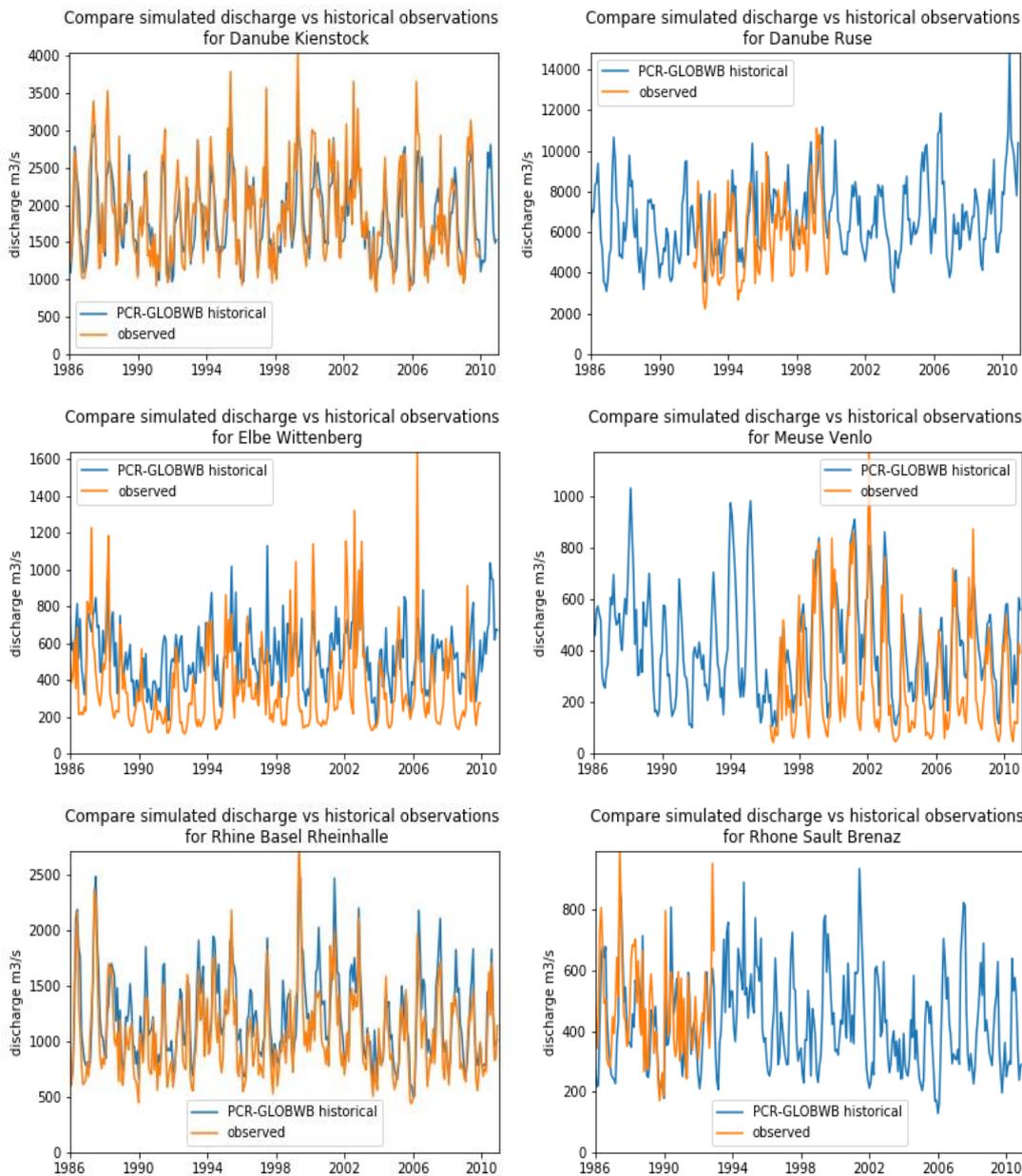


Figure 5: Monthly average discharge large European discharge simulated with PCR-GLOWB (blue) compared with observations from the GRDC (orange).

Future scenarios

PCR-GLOBWB has been run for RCP4.5 in combination with SSP2 and RCP8.5 in combination with SPP5 for the future time-horizons 2026 and 2056. Discharge regimes have been calculated from a ten year window around these future time horizons.

In the Danube discharge will likely increase. Increases are largest for late winter / early spring this is a result of increased temperatures leading to earlier snowmelt and more precipitation falling as snow

instead of rain. This also causes a decrease in intra-annual variability. For the Rhine where the discharge regime is also influenced by snow fall during winter we see similar changes. Increases in late winter / spring discharge and slight discharge decrease in summer. Similar changes can be expected for similar snow influenced basins.

For the Rhone changes are largest and larger than expected. This is mainly caused by strong increases in precipitation. For the Ebro flow during the wet months, January to May, will likely increase whereas low flows show little changes. For the Elbe an increase in river discharge is projected throughout the year. Here the differences between the scenarios are much smaller than the change from the current situation.

For the Meuse changes are smaller. The winter and early spring months will likely become wetter and summer discharge may increase slightly as well, although RCP8.5 projects small decreases for 2056.

These are some examples for major river basins in Europe that form a representation of similar basins. For the Ebro basin in Spain we would, as well as for other basins in the Mediterranean, have expected small decreases in flow (see Sperna Weiland et al. 2010) here it should be noted that due to computational constraints only one GCM has been applied and herewith not the full range of changes or the uncertainty therein is considered.

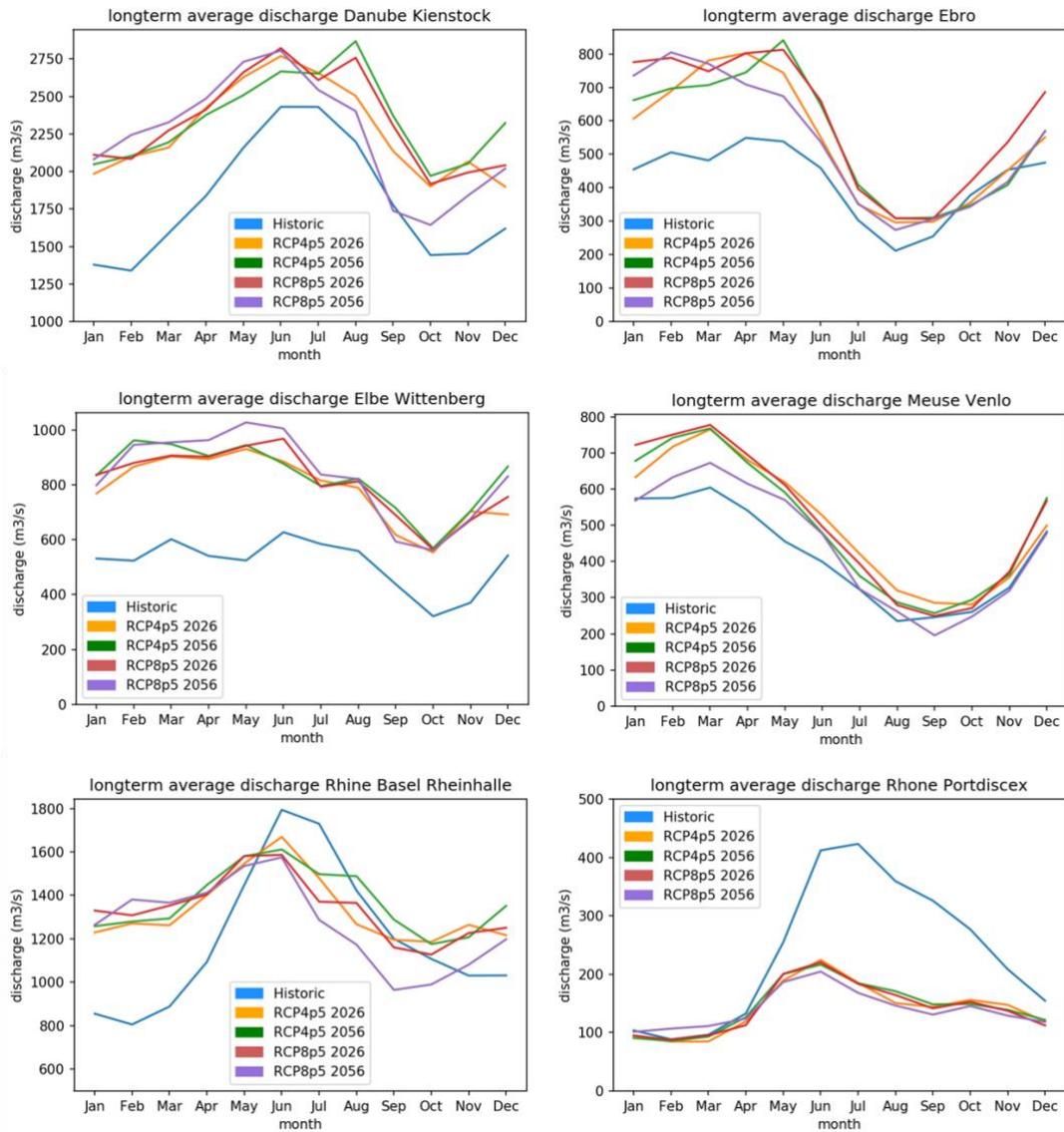


Figure 6: Monthly average discharge changes for large European discharge simulated with PCR-GLOWB based on historic data (blue) together with discharges for the different RCPs and future time horizons.

2.2 MONERIS

Lead: Markus Venohr, Andreas Gericke, Roshni Arora, Judith Mahnkopf

Input data

All data processing and analysis were conducted on basis of 104,300 hydrological sub-catchments (Functional Elementary Catchments, FEC) as spatial modelling units, with a mean spatial resolution of 58 km² (ECRINS, MARS geo-database). MONERIS requires a variety of input data comprising information on hydro-climatic, geo-physical, administrative-demographic conditions. All input data were derived on the maximum spatial resolution and subsequently aggregated as area weighted mean to the FEC level. Table 2 gives an overview of the main input datasets used for the application of MONERIS.

Table 2: Brief overview of the main input datasets used for the application of MONERIS

<i>Dataset</i>	<i>Data sources</i>
<i>Modelling units (FEC), River basin districts</i>	ECRINS / MARS geo-database
<i>river network</i>	ECRINS
<i>water surface areas</i>	ECRINS / Venohr et al., 2011
<i>Hydrology and Climate</i>	
<i>Precipitation</i>	Model results PCR-GlobWB
<i>Evapotranspiration</i>	Model results PCR-GlobWB
<i>Runoff</i>	Model results PCR-GlobWB
<i>water temperature</i>	own calculations
<i>Land use and Land management</i>	
<i>topography – slope / height</i>	EU-DEM; ASTER
<i>land use</i>	Corine 2012, ESA 2010 and UC Louvain; ECRINS; GLCC
<i>soil loss, C-Factor</i>	own calculations
<i>N Surplus on agricultural areas</i>	Venohr et al. 2018
<i>P accumulation</i>	own calculations, OECD, EUROSTAT
<i>soil type</i>	HWSD;ESDB; FAO
<i>tile drained areas</i>	Feick et al. 2005
<i>Hydrogeology</i>	IHME 1500
<i>Other data</i>	
<i>atmospheric deposition of N</i>	EMEP MSC-W
<i>atmospheric deposition of P</i>	Behrendt et al. 2002
<i>solar radiation</i>	EUMETSAT - CM SAF
<i>Population</i>	various sources, see Table-A 3
<i>Sewage collection & treatment</i>	own calculations, Eurostat, Table 4, Table-A 5
<i>WWTPs</i>	own calculations, UWWTD
<i>industrial point sources</i>	E-PRTR

Land use information was derived from Corine 2012 (mostly EU countries), GlobCorine 2009 (e.g. EFTA countries or Switzerland), and ECRINS Lakes v1.1 (merged with the Corine data) with 100m resolution for the entire modelling extent. This mosaicked raster dataset was reclassified to meet the requirements of

MONERIS (see Figure 10 and Table-A 1) and used for all land-use related calculations (e.g. soil loss, nitrogen surplus, or degree of phosphorus saturation). We created a digital elevation model (DEM) by combining the EU-DEM and ASTER G-DEM, in accordance to the process described for land-use data. The DEM was intersected with land use to attribute arable land with different slope classes and the FEC with mean height and slope information. For agricultural land (mean for arable and grassland), the share of tile drained areas was calculated using the dataset by Feick et al. 2005. For each FEC, we calculated the area of sandy, silty, loamy and clayey soils as well as fens and bogs from the European Soil Database (ESDB) and the Harmonized World Soil Database (HWSD). These data were further used to calculate the mean soil nitrogen and clay content.

The soil loss map is based on the universal soil loss equation (USLE, Equation 1) which considers topography (L, S), soil properties (K), land use and land cover (C), and rainfall erosivity (R).

$$\text{Equation 1: } E = L \times S \times K \times C \times R$$

- With:
- L = slope length factor, calculated according to Fuchs 2010
 - S = slope factors, calculated according to Nearing 1997
 - K = soil properties, calculated after Gericke 2015 using the Harmonized World Soil Database (HWSD) and the European Soil Database (ESDB)
 - R = rainfall erosivity, estimated from long-term mean annual precipitation (1961 to 2007) and taken from Gericke 2015
 - C = land cover factor, distinguished for 60 land use classes after Gericke 2015 using the land use mosaic grid described above

Information on hydro-geology were reclassified from the International Hydrogeological Map of Europe (IHME 1500) according to Table 3. After translation, the area of all hydrogeological classes per FEC as input for MONERIS were derived.

Table 3: Relationships between MONERIS classes and attributes AQUIFER_TYPE and LITHO5 of the IHME map.

<i>MONERIS class</i>	<i>AQUIFER_TYPE (IHME)</i>	<i>LITHO5 (IHME)</i>
<i>high porosity, consolidated</i>	Highly productive fissured aquifers	all
	Low and moderately productive fissured aquifers	all
	Locally aquiferous rocks, porous or fissured	(Partly) consolidated
<i>Impermeable, consolidated</i>	Practically non-aquiferous rocks, porous or fissured	(Partly) consolidated
	Low and moderately productive porous aquifers	All
<i>deep groundwater, unconsolidated</i>	Locally aquiferous rocks, porous or fissured	Unconsolidated
	Practically non-aquiferous rocks, porous or fissured	Unconsolidated
<i>shallow groundwater, unconsolidated</i>	Highly productive porous aquifers	all

Nitrogen balances (see Figure 10) on agricultural areas are a key input dataset for MONERIS and were derived using a novel approach by Venohr et al. 2018 as distributed data in the required spatial resolution was unavailable. We disaggregated national N balances reported by the countries by distinguishing three input compounds: atmospheric deposition, organic and mineral fertilizer application. For each compound, individual use efficiency was applied and the resulting distribution was calibrated against the reported country-wide means.

Together with nitrogen agricultural soils are usually fertilized with phosphorus. In contrast to nitrogen, phosphorus easily adsorbs or associates to small particles (e.g. loam, silt but also organic and calcareous material). This leads to an immobilisation and accumulation of phosphorus in soils. The plant-available phosphorus is determined by applying approaches such as P-CAL or Olsen. At the same P content, the plant-available share can vary considerably depending on soil type. The amount of plant-available phosphorus depends on the share of sorption partners occupied by phosphorus on all available sorption partners, also expressed as degree of phosphorus saturation (DPS). Unfortunately, assessment approaches like P-CAL or Olsen cannot directly be used to derive the P loss risks from soils by surface runoff. Here, water soluble phosphorus (WSP), which can directly be derived from P content in soils, is a much better descriptor.

WSP was calculated as weighted mean (Equation 2) per 500 m grid cell according to Pöthig, Behrendt, Opitz, & Furrer 2010 and Pöthig (unpublished data) by distinguishing sand and loamy-silty soils (Figure 7). WSP values calculated by Equation 2 were limited to a maximum of 60 mg/kg, as the range of observed WSP did not exceed this value.

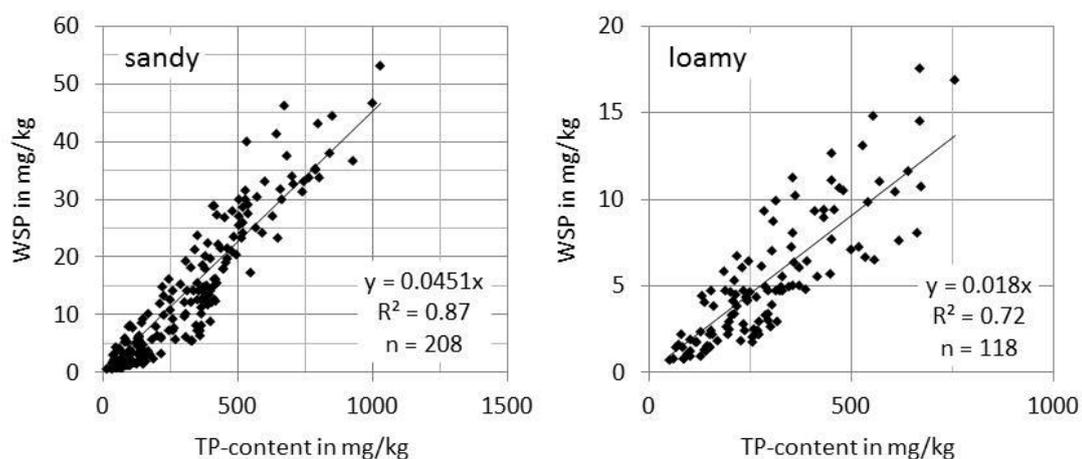


Figure 7: Correlation between P-content in soils and measured WSP in soil samples of Germany and Swiss according to Pöthig, Behrendt, Opitz, & Furrer, 2010 and, Pöthig (unpublished data).

Equation 2:
$$WSP = \frac{((P\text{-content} \times 0.0451 \times Sand) + ([P\text{-content}] \times 0.018 \times [Clay]) + ([P\text{-content}] \times 0.018 \times [Silt]))}{([Silt] + [Sand] + [Clay])}$$

With: WSP = water soluble phosphorus, mg/kg
 Sand = share of sand fraction in soils, in %
 Clay = share of clay fraction in soils, in %
 Silt = share of silt fraction in soils, in %
 P-content = Phosphorus content in upper 30 cm soil layer, in mg/kg

As a prerequisite for this we had to derive the spatially distributed P content in agricultural soils. This was done on basis of country wide P-accumulation, to calibrate the total P content and using the N-surplus described above to derive the spatial distribution of applied fertilizers. This approach was developed, tested and calibrated for agricultural soils in Germany first (not shown) and subsequently transferred to European data.

In a first step country wide P balance data on agricultural areas were collected from EUROSTAT, and area corrected as described before (see Figure 8). Longest time series ranged from 1985 to 2014, whereas the shortest time series only covered data after 2004. To estimate the P-accumulation also fertilisation from earlier years had to be considered. From a reconstruction of historic nutrient balances in central Europe (Gadegast & Venohr, in prep) we know, that intensive fertilisation already took place in the 1960ies and in many European countries found its maximum in the 1980ies. From this we derived following rules of thumb:

- 1) P-balances in 1960 equal the earliest reported available value per country (between 1985 and 2004)
- 2) In 1950, P-balances were 10 % of the values in 1975 (for this year P balances in all countries were positive, but not at their maximum)
- 3) In 1980, P-balances were 20 % higher than in 1960. These values were corrected for Estonia and Hungary, to ensure, that P-accumulation in all years remained positive.

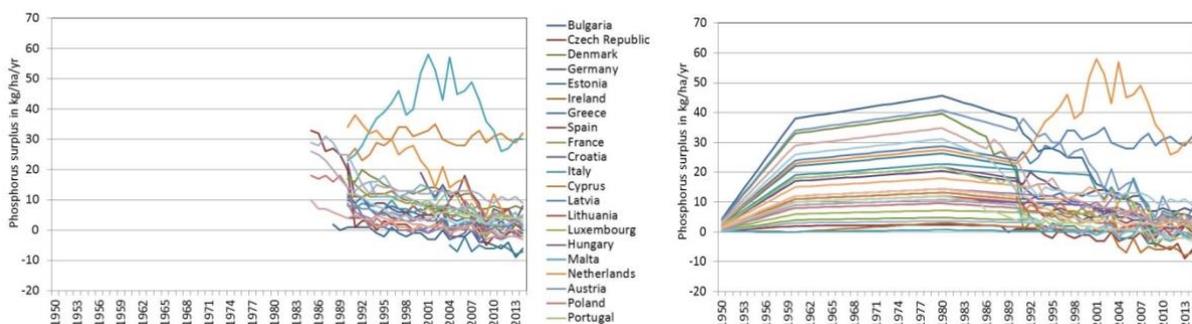


Figure 8: Available P-balance on country (left) and the accomplished time series (right).

The P-accumulation was calculated as the accumulative sum of P-balances over the years (Figure 9).

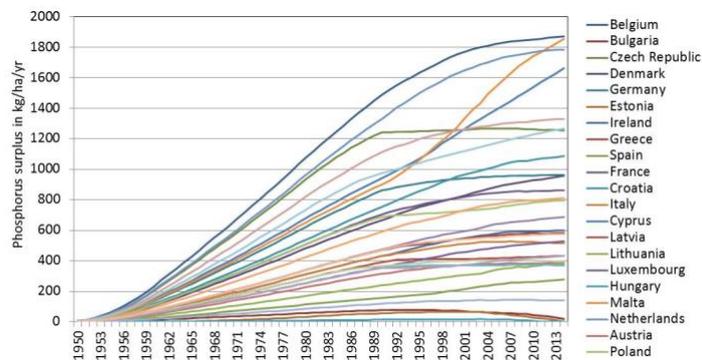


Figure 9: P-accumulation on agricultural land per country in the period from 1950 and 2014.

The P accumulation was distributed following the approach for N-surplus by Venohr et al. (2018), however without taking atmospheric deposition into account, for which spatially distributed data was unavailable.

The P-content was derived from the bulk density information in the LUCAS physical top soil information map (Ballabio, Panagos, & Monatanarella, 2016). First the soil weight of the top 30 cm soil layer (ploughing horizon) was calculated (Equation 3).

$$\text{Equation 3: } \text{Soil weight} = \text{BulkDensity} \times \text{LayerDepth} \times \text{UCF}$$

With: soil weight = soil weight of the top 30 cm soil layer, kg/ha
 Bulk density = Bulk density, in g/cm³
 LayerDepth = 30 cm
 UCF = unit correction factor (g/cm² → kg/ha) = 100000

By dividing the corrected and spatially distributed P accumulation by the derived soil weight the mean P content in top soils was estimated (Equation 4).

$$\text{Equation 4: } \text{P_content} = \frac{[\text{P}_{acc}]}{[\text{Soil weight}] \times 1000000}$$

With: P-content = Phosphorus content in upper 30 cm soil layer, in mg/kg
 P-acc = P-accumulation, in kg/ha
 Soil weight = soil weight of the top 30 cm soil layer, kg/ha

DPS was estimated considering the soil type information by LUCAS and considering the transformation function published by Pöthig, Behrendt, Opitz, & Furrer (2010) (see Figure 10). P-concentrations in surface run-off was finally calculated according to Vadas et al. (2005), which was corrected on basis of findings by Fischer et al. (2016), to eliminate effects originating from different soil to water ratios used by Vadas et al. (2005) (Equation 5).

$$\text{Equation 5: } \mathbf{PconcSR} = \left(\frac{11.2 * \mathbf{WSP_arable} + 66.9}{1000} \right) \times \mathbf{WSP_corr}$$

With: PconSR = P-concentration in surface run-off, in mg/l
WSP = water soluble phosphorus, mg/kg
WSP_corr = WSP correction factor, without unit)

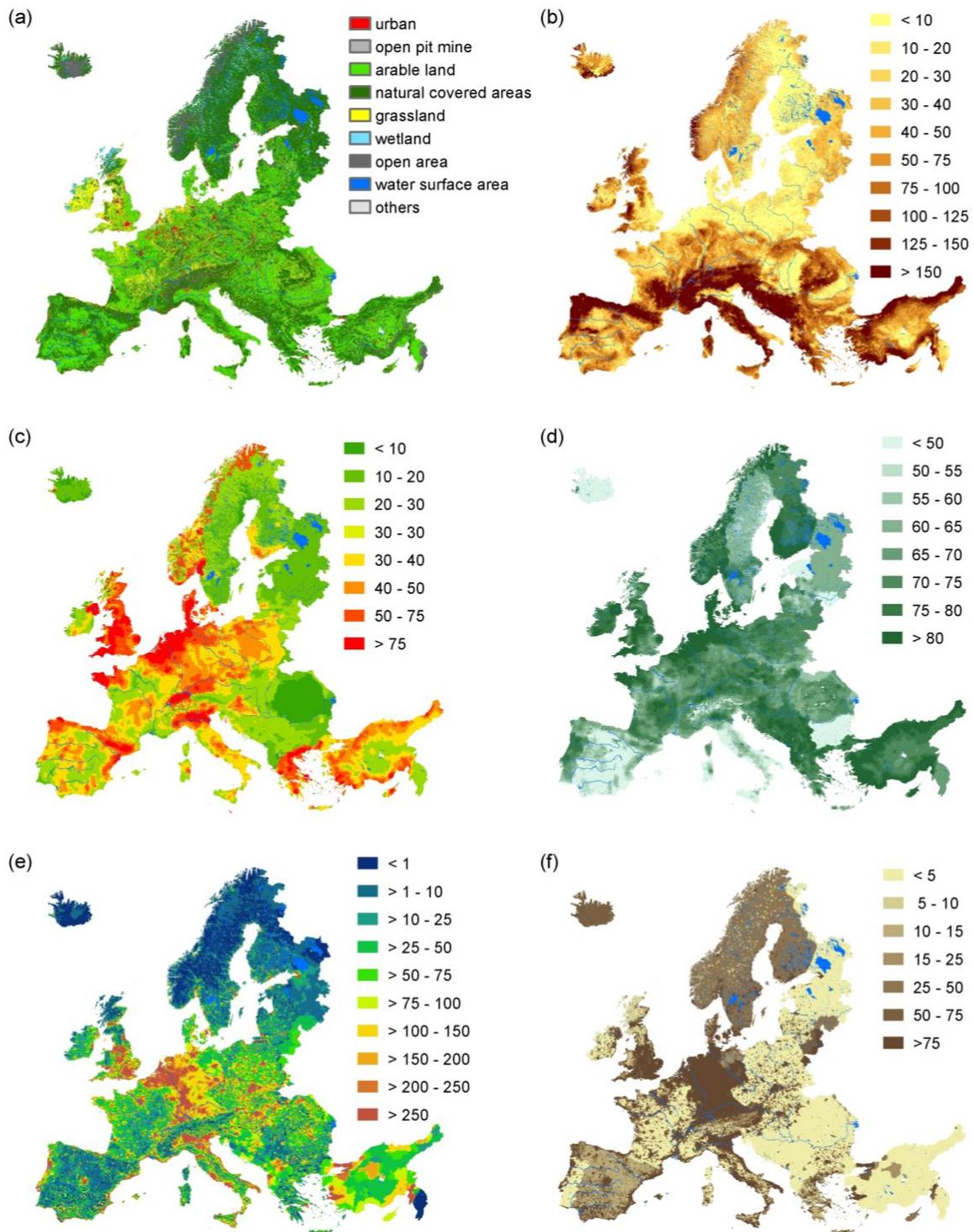


Figure 10: Some of the main input datasets for the application of MONERIS: (a) reclassified land use (b) mean soil loss per FEC in t/ha (c) Nitrogen Surplus per FEC in kg/ha*a (d) degree of phosphorus saturation per FEC in % (e) population density for the year 2010 (f) Inhabitants connected to WWTP and sewer system for the year 2010 [%]

Hydrology, climate, and water temperature

Climate and hydrological data (precipitation, evapo-transpiration, water balance) PCR-GlobWB were provided as ascii files with a resolution of $0.83^{\circ} \times 0.83^{\circ}$ as monthly means (m/month) for the current period 2001 to 2010 and the future scenarios periods 2026-2035 and 2056-2065. The pre-processing included a series of work steps, such as, conversion to GeoTIFFs, re-projection and resampling to a resolution of $500\text{m} \times 500\text{m}$ and application of the “zonal statistics” tool implemented in ESRI-ArcGIS to calculate FEC specific mean precipitation/runoff/evapotranspiration/air temperature weighted means.

For very small FECs and for coastal FECs not covered by the PCR-GlobWB grids, ArcGIS did not deliver weighted mean values. Here, we used the information on the hydrological topology, i.e. the information of the next down-stream FEC, given in the ECRINS data set. Like this precipitation and specific run-off was transferred to neighbouring FECs. PCR-GlobWB provides data on water balances, which can be negative if evapo-transpiration exceeds precipitation. According to PCR-GlobWB negative values water balances were calculated for most grids/FECs with dominating shares of water surface areas. However, MONERIS requires information on the generated run-off per FEC and negative values had to be substituted by small positive values. As this change only had a minor effect on the total run-off of a river basin, a counter balancing of the added run-off by subtracting it from other FEC run-off values was renounced. Both corrections were conducted using the software package R.

Deriving scenario values for precipitation and specific run-off required a further correction. The grids for current and future conditions provided by PCR-GlobWB showed some small dislocations, which could not be removed by ArcGIS functions. These dislocations partly led to erroneous and un-realistic high changes between current and future conditions, when values of different grid cells formed the basis for the zonal statistics per FEC. To eliminate these inconsistencies we calculated the mean change for groups of FECs. These were FECs of a similar elevation in a sub-unit or small river basins. Rather than using the individual changes per FEC the mean changes per FEC group was applied to the spatial pattern of current precipitation and specific run-off.

Statistical approaches, most commonly linear regression, offer simple and efficient means to predict water temperature of rivers and have been used widely (Webb et al., 2003; Benyahya et al., 2007). Air temperature is generally used as the independent variable in regression analysis since it is proxy for the net changes in heat flux affecting the water surface (Webb et al., 2003). Water-air temperature regression models have been applied successfully at several time scales such as daily, weekly and monthly and annual means (Webb and Walling, 1993; Erickson and Stefan, 1996, 2000; Webb and Nobilis, 1997). At smaller time scales and at annual scales, air and stream temperature correlations are typically weak

and use of linear regression to predict water temperature might not be the best choice (Mohseni and Stefan, 1999; Erickson and Stefan, 2000). At monthly scales, however, strong water-air temperature relationships are usually observed (Figure 1) and therefore, linear regression yields satisfactory results (Erickson and Stefan, 1996; Webb et al., 2003).

For this study, linear regression analysis was carried out at a monthly scale to derive coefficients for water temperature extrapolation. Monitored water temperature data were checked for erroneous (missing, lower than 0, higher than 35 °C) values, converted to monthly means (if otherwise) and trimmed for the desired time period (2001-2010). The FECs with data containing ≥ 80 water temperature values were retained (a complete series has 120 values). For these FECs and air temperature data, data gaps were filled with average monthly values for that particular month and station. Linear regressions between air temperature and water temperature were then applied within these FECs to obtain the regression coefficients (all statistically significant, $P < 0.05$). The air temperature coefficient ranged from -0.08 to 1.3 and the average r^2 was 0.86 (S.D. = 0.09; $n = 2,056$).

Extrapolation of air-water temperature coefficients to FECs without data with less than <80 values was done on the basis of similarity in elevation and hydro region. Hence, as a next step, the FECs were assigned hydro region classes. Originally, the hydro region classification for Europe consists of seven classes. For our calculations, the classes CB_CON and NOR were split in two sub regions each, France, Belgium and Luxembourg assigned to new group CB_CON_2 and Norway and Iceland formed the new group NOR_2; in total resulting in nine classes.

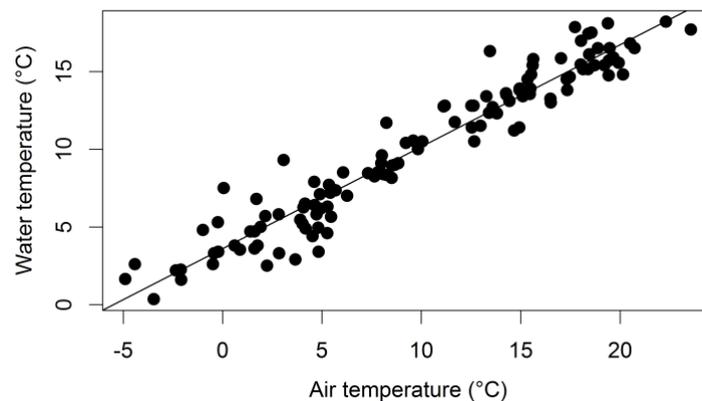


Figure 11: Plot of water temperature and air temperature (time period = 2001-2010) for a FEC (ID = 35402) situated in Germany.

Population, collection and connection rates, WWTPs and industrial point sources

Annual population data for 2001 to 2010 were collected from a variety of different sources like European data from EUROSTAT on a NUTS 3 level and supplemented with country wide data (see Table-A 2). Single population numbers were joined to available shapefiles for administrative regions. Gridded population data (1 km² spatial resolution) were available for 2006 from GISCO/Eurostat. This dataset was used for the spatial description of population density and was modulated between years of the modelling period with the relative changes derived from the annual times series on administrative level.

Further data sets required by MONERIS are the share of households connected to sewer systems and the share connected to sewer system and a waste water treatment plant. This data was available from EUROSTAT as country wide means for most of the country contributing to the MARS modelling extent. Incomplete time series were filled by linear interpolation between years. For countries with missing data we transferred the percentage share from countries with similar geo-political conditions (Table 1).

The country wide connection rate had subsequently to be broken down to FEC level. To do this connection rates to sewers given for 175 Nut-2 units (out of 1373 NUT-2 units in Europe) between 2002 and 2010 were used to build a simple population density driven model, following to steps. In a first step we used a sigmoidal function to develop the base distribution of connections rates (Figure 12, Equation 6). In a second step the calculated spatial distributed raw connection rates had to be corrected to fit the reported national wide mean connection rates (Figure 12, Equation 7, Table 4).

$$\text{Equation 6: } \text{ConRateToSewers} = \frac{1}{1.001 + 1 \cdot 10^{-8} \text{PopDens}^{4.5}} \cdot 100$$

With: PopDens = population density, in habitant / km²

$$\text{Equation 7: } \text{ConRate_FEC_corr} = \left(\frac{\text{ConRate_FEC}}{100} \right)^{\text{CSCT}} \cdot 100$$

With: CSCT = country specific correction term

Table 4: Completed national statistic on inhabitants connected to sewer systems in the years 2000-2011 based on reported data from (EUROSTAT). Colour code: coloured cells indicate countries, for which the same connection rate was assumed; red numbers indicate countries to which connection rates were transferred. No-coloured cells with red figures indicate Years, for which connection rates were calculated from linear interpolations.

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2001-2005	2006-2010
Albania	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7
Andorra	87	88	89	90	91	92	94	94	94	96	98	98	90	96
Austria	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Belarus	67	67	67	72	68	68	67	100	100	100	100	100	69	95
Belgium	65	69	69	72	73	74	88	97	96	97	96	97	71	95
BosniaHerzegovina	8.6	8.6	8.6	8.6	8.6	8.6	8.9	22	26	30	35	39	8.6	27
Bulgaria	70	71	71	72	72	72	72	73	74	75	77	82	71	75
Croatia	8.6	8.6	8.6	8.6	8.6	8.6	8.9	22	26	30	35	39	8.6	27
Cyprus	100	100	100	100	100	100	100	100	100	100	100	100	100	100
CzechRepublic	66	68	72	73	74	75	74	76	78	78	79	81	72	78
Denmark	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Estland	67	67	67	72	68	68	67	100	100	100	100	100	69	95
Finland	100	100	100	100	100	100	100	100	100	100	100	100	100	100
France	95	96	96	96	96	97	97	98	98	99	99	100	96	99
Germany	97	97	97	98	98	97	99	100	99	100	100	100	98	100
Gerorgia	26	27	27	30	36	42	43	46	46	50	52	56	32	49
Great Britain	98	100	100	100	100	100	99	98	97	97	100	100	100	98
Greece	80	81	81	82	83	84	84	85	86	87	87	88	82	86
Hungary	52	54	56	58	59	61	63	67	68	69	72	73	57	69
Iceland	39	39	56	56	56	96	97	99	100	100	73	73	61	90
Ireland	85	86	87	88	89	90	91	92	93	94	95	96	88	94
Isle of Man	98	100	100	100	100	100	99	98	97	97	100	100	100	98
Italy	95	96	96	96	96	97	97	98	98	99	99	100	96	99
Kosovo	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6
Lebanon	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Lettland	67	67	67	72	68	68	67	100	100	100	100	100	69	95
Lichtenstein	65	69	69	72	73	74	88	97	96	97	96	97	71	95
Lithuania	67	67	67	72	68	68	67	100	100	100	100	100	69	95
Luxembourg	100	100	100	100	100	100	99	99	99	99	99	100	100	99
Makedonia	5	6	6	6	6	6.5	7	7	7	7	7.5	7.7	6.1	7.2
Malta	16	16	18	18	15	15	11	10	17	23	30	37	16	21
Moldowa	21	23	24	26	28	29	31	32	32	32	34	42	26	34
Montenegro	8.6	8.6	8.6	8.6	8.6	8.6	8.9	22	26	30	35	39	8.6	27
Netherlands	100	100	100	100	100	100	100	100	100	100	100	100	100	100
Norway	94	93	92	94	94	96	96	96	96	97	97	97	94	96
Poland	54	55	82	84	85	60	61	62	63	64	65	66	73	64
Portugal	74	75	76	77	78	79	80	81	82	83	84	85	77	83
Romania	21	23	24	26	28	29	31	32	32	32	34	42	26	34
Russia	67	67	67	72	68	68	67	100	100	100	100	100	69	95
Serbia	8.6	8.6	8.6	8.6	8.6	8.6	8.9	22	26	30	35	39	8.6	27
Slovakia	98	98	98	98	98	98	98	99	98	99	99	98	98	98
Slovenia	71	72	73	75	84	85	87	88	88	89	88	90	78	88
Spain	87	88	89	90	91	92	94	94	94	96	98	98	90	96
Sweden	99	99	99	100	100	100	100	100	100	100	100	100	100	100
Switzerland	98	99	99	99	99	99	99	99	99	99	99	99	99	99
Syria	5	5	5	5	5	5	5	5	5	5	5	5	5	5
Turkey	26	27	27	30	36	42	43	46	46	50	52	56	32	49
Ukraine	21	23	24	26	28	29	31	32	32	32	34	42	26	34

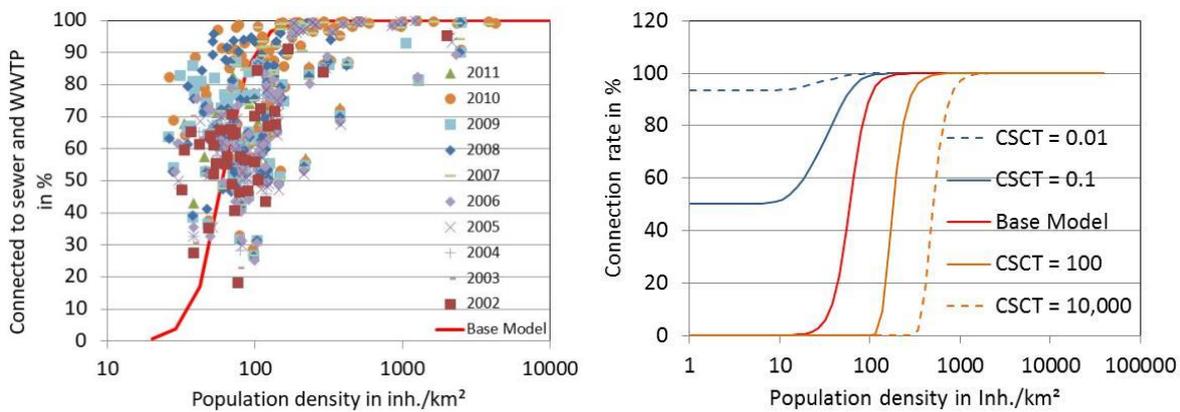


Figure 12: Dependency of connection rate to population density for 175 NUT2 units in 16 European (AT, BG, CY, CZ, DE, EL, FR, HU, IE, LV, NL, PL, RO, SI, SK, TR) countries (left). Corrected connection rates at different CSCT compared to the base model.

The same procedure was applied to derive and correct connection rates for inhabitants connected to sewer and WWTP (Table-A 5). The share inhabitants connected to sewer and WWTP must by definition always be lower than the rate of inhabitants connected to sewers (but not necessarily to a WWTP). For some countries this prerequisite partly or in total was not given and had to be corrected, by setting the share of inhabitants connected to sewers and WWTP to the connection rate to sewers only. Further in some countries the reported connection rate to WWTP was expressed as shares on the number of inhabitant connected to sewers. This was translated into shares on the total population, accordingly.

The derived connection rates to sewers and WWTP do not consider if and where inhabitants from several FEC are connected to the same WWTP, i.e. in which FEC WWTP effluents are discharged. This information was taken from the UWWTD inventory (Waterbase-UWWTD: The European Topic Centre on Inland, Coastal and Marine waters. Version 5, date of delivery (date sent to the Data Service): 18/02/2015.; E-PRTR-database: European Commission Directorate-general for Environment), containing information on the location of a WWTP and the treatment capacity. We used version 5 of the UWWTD inventory (reporting year 2012) as basis for our analysis. But, in particular for Finland, in former versions WWTPs were reported including inhabitant capacity, which was missing in the following versions. Therefore, for all WWTP named in version 5, we checked, if in former version capacities have been reported and included these in the used inventory, assuming that the total amount of waste water from connected household does not decline from past to present.

Treatment capacity was assumed to reflect the actual number of connected inhabitants. Comparing the number of inhabitants connect to sewer and WWTP as described above and the UWWTD inventory showed that the latter in average is 1.7 times higher than the number of inhabitants derived from the country statistics (Figure 13).

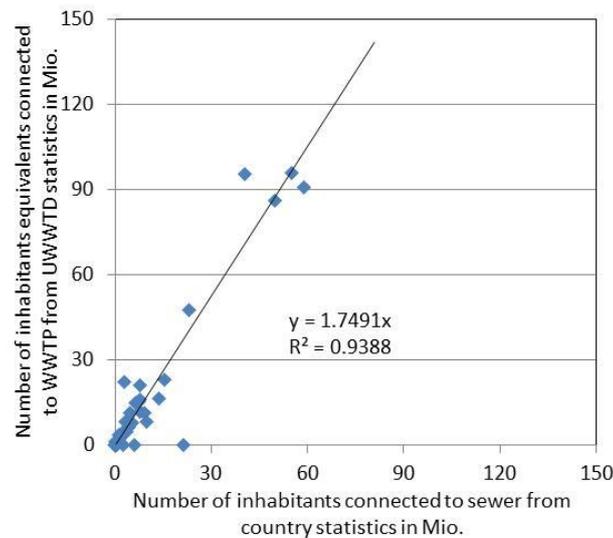


Figure 13: Comparison of inhabitants connected to WWTP derived from country statistics (years 2006-2011) and the UWWTD inventory (year 2012).

This difference can be explained by water from industries and sealed urban areas treated in WWTP but not considered as inhabitants in country statistics. In countries for which no WWTP inventory was available, the share of connected inhabitants per FEC was increased by a factor of 1.7 and WWTPs added to the inventory accordingly. Further, for countries where the reported inventory included a smaller share of inhabitants than such derived from the country statistics, the WWTP inventory was corrected accordingly.

For 4534 WWTP reported Inhabitant capacities as well as TN and TP loads were available from the UWWTD inventory. Discharges originating from inhabitants, precipitation and industry were in general not reported and had to be estimated. Inhabitant specific water consumption was set to 100 l/Inh/day. For the WWTP inventory also discharges from industry and sealed urban areas had to be included and have been estimated under consideration FEC based mean long-term precipitation (Equation 8). Here, we the ratio of precipitation and water consumption based discharge was estimated from the ratio described above.

$$\text{Equation 8: } WWTP_Q = \left(\frac{1}{1.7} \text{Inhabs } 100 \cdot 365 + \left(1 - \frac{1}{1.7} \right) \text{Inhabs } 100 \frac{PP_{LT}}{800} 365 \right)$$

With: $WWTP_Q$ = annual discharge from a WWTP, in l/yr

Inhabs = reported inhabitant equivalent, in heads

PP_{LT} = long-term mean precipitation, in mm/yr

Additionally, for the WWTPs without reported TN and TP effluent concentrations had to be derived. This was also done on basis of the 4534 WWTPs with a complete dataset from the UWWTD inventory. Mean effluent concentrations per WWTP size class were derived from the reported nutrient loads and the calculated discharges.

Table 5: Mean WWTP size specific effluent nutrient concentrations derived from the UWWTD inventory, calculate as discharge weighted mean from all WWTP of a specific size class.

Reported WWTP size class	Reported WWTP			Share on total discharge		discharge weighted mean concentration	
	connected inhabitants In thousand	count #	share %	TN %	TP %	TP mg/l	TN mg/l
1	< 2	47	1	0.2	0.4	19.6	118.1
2	2 – 5	718	16	2.1	5.7	7.6	35.2
3	5 – 10	778	17	3.2	7.8	4.6	24.2
4	10 – 50	1938	43	22.0	25.3	1.8	19.5
5	50 – 100	558	12	16.4	14.8	1.2	17.3
6	100 – 500	433	10	31.0	25.6	1.0	15.3
7	>500	62	1	25.1	20.5	1.1	17.8
Total		4534	100			1.4	41.4

Other data

The total deposition of reduced nitrogen and oxidized nitrogen was calculated as the average monthly deposition (kg km^{-2}) per FEC using the Co-operative Programme for Monitoring and Evaluation of the Long-Range Transmission of Air Pollutants in Europe (EMEP/MSC-W model version rv4.5, data for 2001-2010). The deposition rate of phosphorus, which depends on the land use of the observed area, lies between 0.3 and 3.0 $\text{kg P}/(\text{ha}\cdot\text{a})$. Behrendt et al. (2002b) derived an average value of 0.37 $\text{kg P}/(\text{ha}\cdot\text{a})$ for European catchment areas. This value was defined as constant for the whole calculation period and modelling extent. The surface incoming shortwave radiation (SIS) was calculated as the average monthly mean per FEC in W/m^2 from EUMETSAT’s Satellite Application Facility on Climate Monitoring (CM SAF).

Scenarios

As shown in Figure 2 and Table 18 changes in various input data were considered for the two future periods (2026-2035 and 2056-2065) for two different story lines 1) Techno world and 2) Consensus world. For deriving future conditions on FEC level we used in most cases data sources different from those used for current conditions. In order to derive scenario conditions consistent and comparable to current

conditions, we, in most case, derived changes based on the scenario data sources and applied the relative (percentage) changes to the data set derived to run and validate the model under current conditions. Hydro-climatic data were derived using model results from PCR-GlobWB (see Chapter 2.1). For MONERIS additional changes have been considered. To estimate water temperature for the four scenarios periods, the respective air temperature provided by PCR-GlobWB data was used. Here, we used the regression coefficients derived for current conditions, assuming unchanged air-water temperature relations. Here, the same approach as described in MONERIS – input data was applied.

Future land use was available as GLCC maps from the SCENES project (according to MARS deliverable 2.1). As SCENES data do not provide the same land-use classes (Table-A 1) and spatial resolution as the Corine data set, used for our modelling task, we applied the relative land use changes (between current and future situations) to the current land on FEC level use derived from Corine. Population data have been extracted from the IIASA (International Institute for Applied Systems and Analysis) SSP database (MARS deliverable 2.1) and transferred to FEC level. Demographic changes had further to be considered for collection and treatment of waste water. Her in principle it was assumed that new additional inhabitants (population increase) is connected to a sewer systems and a WWTP. In turn for a population decrease it was assumed that first households without a connection to collection and treatment are abandoned. This basically represent a land-urban migration as found in many regions. Additionally, to this an overall increase in waste water collection was assumed. For Storyline 1 around 2030 and 2060 a general increase of sewage collection and treatment by 10% and 20%, respectively, was assumed (Equation 9). For Storyline 2 an additional increase of the collection rate in rural areas was assumed (Equation 10).

$$\text{Equation 9: } Conn_SL1_{2030/60} = Conn_{2010} + a$$

$$\text{Equation 10: } Conn_SL2_{2030/60} = Conn_SL1_{2030/60} + (100\% - Conn_SL1_{2030/60}) \cdot b$$

With: $Conn_SL1_{2030/60}$ = Share of collected and treater waste water in 2030 for Storyline 1, in %

$Conn_{2010}$ = Share of collected and treater waste water under current conditions, in %

a, b = increase factor, 2030 = 10 %, 2060 = 20 %

Further, a general reduction of TN and TP effluent concentrations for all treatment plants sizes by 25%, with unchanged domestic water consumption was assumed. This reduction was assumed for both periods and storylines.

Three different mitigation measures were considered, to be applied in addition to the changes assumed for the storylines:

- 1) During the last years several measures throughout the European countries have been put in place to **reduce nitrogen surplus**. These have partly been successful and in many places an increased use efficiency of applied fertilizer and a reduction of N surplus have been reported. Nonetheless, mostly in areas with high livestock densities, N surplus, mostly caused by manure is still far too high and does not meet European regulations. In short, the excess in N surplus can be to a large share explained by too much und insufficiently distributed manure. As during the last decade no clear reduction in livestock densities can be found in most European countries the mitigation measures addresses an improved distribution of manure, rather than a reduction of produced manure. The measure assumes that manure can be dried and efficiently transported and distributed across Europe. Further we assume that the dried manure replaces 80 % of the currently applied mineral fertilizer and can be applied with the same use efficiency as mineral fertilizer. Like this in no places N surplus will increase, but the current hot-spot regions will see a massive reduction in n surplus. For Romania, Moldova and Ukraine, however, N surplus was set to 30 kg/ha with a UE of 75%, as intensification in agricultural production was assumed.
- 2) An implementation of **riparian buffer stripes** along 75 % of all rivers and brooks was assumed. The retention efficiency of the buffer stripes was estimated according to Venohr et al. (2011), in dependence from their width. As soil loss is dependent, among other factors, by the slope of an area, we assumed a buffer width of 2-5m, 5-10m, and 10-20m for slopes <2%, 2-4%, and >4%, respectively. Here, the mean slope of arable land was used for class selection. Although different retention efficiencies are known for dissolved and particulate material, we applied for this large scale modelling task fixed retention rates for nitrogen and phosphorus of 30%, 50%, and 80% for buffer with of 2-5m, 5-10m, and 10-20m, respectively. Further effects of buffer strips like, temperature regulation, carbon source or habitat have not been considered.
- 3) A further reduction of **WWTP N and P effluent concentrations** by 50 % related to current conditions (see above and Chapter 2.2 - Input data).

Methods

MONERIS (MOdelling Nutrient Emissions in River Systems, Venohr et al. 2011) is a process-oriented model, for quantifying nutrient fluxes, i.e. emissions from the catchment to the surface waters, in-stream retention and resulting loads/concentrations in surface waters. The model considers 7 different

emissions pathways: atmospheric deposition on surface waters, surface runoff, erosion tile drainages, interflow-groundwater, urban systems and point sources. All calculations are conducted on basis of hydrological sub-catchments. In the case of the MARS project we used the Functional elementary Catchments (FECs) as described below. The model is designed to work on annual basis, however, emissions can also be disaggregated to monthly values and in-stream retention and transport can be calculated on monthly basis. While point source emissions are discharged directly to surface waters, the modelling of diffuse emissions needs contemplating pathway-specific nutrient transport and retention processes. After entering surface waters nutrients are further retained, transported and transformed in surface waters. In contrast to these model-inherent approaches, water balances, soils-loss, or nutrient balances are required as input data.

A main principle during the model development was to establish and calibrate the different modules addressing pathways-specific processes independently to minimize the risk of a factitious accuracy of a potentially over-parameterised model. The basic principles of the individual pathway modelling approaches are described in the following.

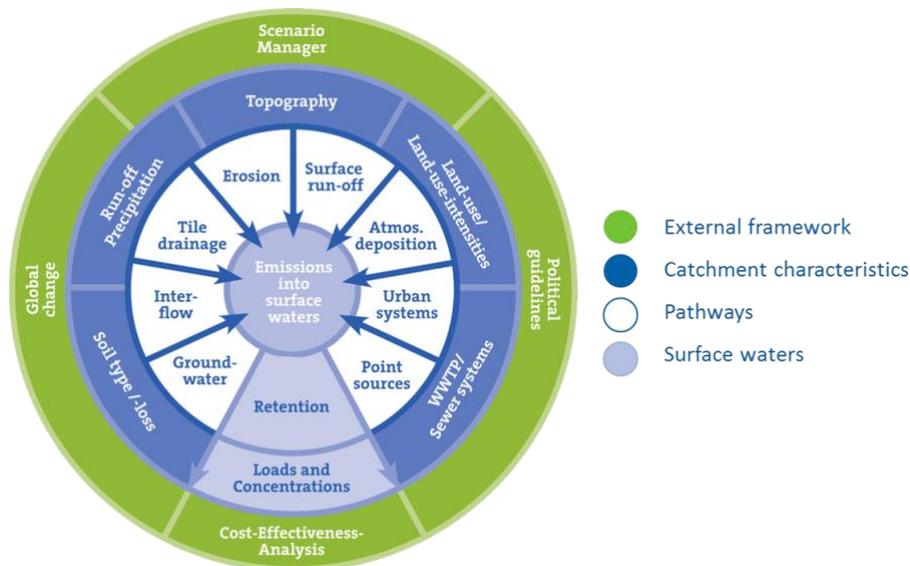


Figure 14: Structure of the MONERIS model showing the external framework, catchment characteristics, pathways, and surface waters.

Atmospheric deposition und water surface areas: Deposition on water surface areas is considered as direct input and derived from N deposition maps. An equivalent data source for P deposition is currently not available. We assumed a constant P deposition of 0.37 kg P/(ha-a) (Behrendt et al. 2002) for all surface waters in Europe. Whereas larger lakes and rivers are shown as polygons of a known size, in large-scale GIS maps, smaller rivers are often only considered as line elements and are commonly not attributed

with area or width information. *To estimate the water surface of these smaller river sections the approach by Venohr et al. (2005, 2006) was used. This calculates a mean river width per FEC on basis of specific run-off, slope and catchment size. Subsequently, the length of the river section per FEC is derived from the GIS map, corrected by factor to reduce scale-dependent generalisation effects, and multiplied with the calculated width. The total water surface area further considered areas from the larger lakes and river given in the map. This approach was applied separately for main river (MR) and tributaries (TRIB) of all FECs.*

Erosion: Surface run-off is calculated as function of the mean annual run-off on basis of a flow disaggregation approach by Carl et al. (2008). Here it is assumed that all areas except from urban areas and water surface areas contribute to surface run-off generation. Nitrogen concentrations in surface run-off are derived from atmospheric deposition and N-surplus and on for phosphorus from the phosphorus saturation of soils.

Sediment and nutrient emissions via erosion are calculated based on the universal soil loss equation (USLE). Soil loss is considered separately for arable land of different slope-class, and for land use types grass land, forest, glaciers and open areas. The share of contributing areas (sediment delivery ratio), with a slope of more than 1.5 %, is calculated with an empirical equation calibrated by Behrendt et al. (2002) based on a 25 m resolution digital elevation model. For phosphorus a fixed enrichment ration of 1.86 according to Wilke and Schaub (1996) assumed.

Surface run-off: Surface run-off is calculated as function of the mean annual run-off on basis of a flow disaggregation approach by Carl et al. (2008). Here it is assumed that all areas except from urban areas and water surface areas contribute to surface run-off generation. Nitrogen concentrations in surface run-off are derived from atmospheric deposition and N-surplus and on for phosphorus from the water extractable phosphorus in agricultural soils.

P concentrations in surface run-off are calculated according to Vadas et al. (2005), including a correction factor derived by Fischer et al. (2016), to eliminate effects originating from different soil to water ratios used by Vadas et al. (2005) in Brazil and Europe.

Tile drainages: Tile drained agricultural soils can be a major source for nutrient emissions. Nutrient concentrations are calculated on basis Nitrogen surplus, under consideration of plant up-take (fixation) under grass land and denitrification in the saturated soil zone (for further details see Heidecke et al. 2015). Phosphorus concentrations are considered as soil dependent values as shown in Table 6. Discharge via tile drainages is calculated as percentage share of monthly precipitation according to Hirt et al. 2009.

Table 6: Considered soil-type specific TP concentrations for the calculation of emissions via tile drainage and groundwater.

	TP concentrations (mg/l) in	
	tile drainage	groundwater
<i>Sandy soils</i>	0.2	0.1
<i>Loamy soils</i>	0.06	0.03
<i>Fen</i>	0.3	0.1
<i>Bog</i>	2.0	0.2
<i>Forest</i>		0.02

Interflow-groundwater: Groundwater discharge per FEC is calculated as residual from the total run-off minus all other flow components considered for the pathways described above. Phosphorus concentrations are considered as described in Table 6. Nitrogen concentrations are calculated similar as for tile drainages, but additionally considering groundwater retention time. Here, N-surplus is considered as mean value during the ground water retention period.

Urban systems: For urban systems basically four different emissions pathways are considered: sealed urban areas connected to separate or combined sewer systems and households connected to waste water treatment plants or to decentralised treatment systems. Sealed urban areas are derived from the urban area given in the land-use map, corrected by population densities. The share of areas connected areas was taken from country wide statistics, again, disaggregated by population density information (see input data). The share of used combined or separated sewer systems was estimated based on an approach by Behrendt et al. 2002, assuming that due to the lower construction work combined sewers are predominately in place under consolidated rock, whereas separate sewers are preferred in presence of un-consolidated undergrounds. Waste water treatment plants were considered as inventory containing information on connected population, discharge and effluent concentrations. From the country wide statistics on connect rates we derived the share of households connected to decentralised treatment plants as residual. For these a fixed small retention capacity was assumed (TN: 10%, TP; 7%). Further it was distinguished whether effluents are discharged via a pipe directly to surface water or via a soil-groundwater discharge. For the latter case additional retention processes as considered for the groundwater pathways were used.

In-stream nutrient retention: The phosphorus retention approach in MONERIS was originally developed to describe mean annual net retention in surface waters. As MONERIS was applied on a monthly basis, the processes of sedimentation and (possible) remobilisation had to be considered. Additionally, the former approach was unable to differentiate between retention conditions in mountainous and lowland rivers, which is also addressed by the new approach. Sedimentation depends on flow velocity and particle size, whereas remobilisation is controlled by sheer stress. Both flow velocity and sheer stress express the power which the flowing water provides to transport sediments. At lower flows also finer sediments are

deposited whereas finest particles may still be remobilised. Looking at the actual processes in real-time sedimentation and remobilisation are therefore no functions of absolute flow, but of changes in flow. At higher flows large particle sediment and also larger particles are remobilised. Consequently, sedimentation and remobilisation can occur at any flow condition but are relevant for different particle sizes.

Looking at mean monthly conditions, the co-existing sedimentation and remobilisation can be simplified to a net retention. The basic assumption of the new monthly approach is that in rivers, on a long-term perspective, sedimentation and remobilisation are (almost) balanced. “Balanced” does not necessarily mean that there is no net-retention of sediments on a mid- or long-term perspective. In drier years net-retention (due to sedimentation) can occur, whereas in wetter years (not even extreme floods) sediments even of several years can be remobilised. This leads to the assumption that for each river stretch an equilibrium between sedimentation and remobilisation can be assumed at mean runoff, depending on the distribution of particle size. With the new monthly approach sedimentation is calculated as a steady process with changing rates and considers monthly runoff, water-surface area and slope. Remobilisation is only considered if mean long-term run-off is exceeded. MONERIS as a static, empirical model does not provide information on sediment pools and P content in sediments. So, remobilisation rates were coupled to the P emissions and incoming loads from headwater catchments and receiving water, respectively.

A further improvement is the slope-dependent calculation of both processes. Basically, for both processes, hydraulic load (i.e. runoff divided by water-surface area) was considered as major control. For identical flow conditions and river widths, the flow velocity is higher in mountainous rivers, given their smaller depth and cross-section area compared to lowland rivers. Therefore, sedimentation and remobilisation rates decrease with increasing river slope. Figure 15 conceptually shows the resulting sedimentation and remobilisation rates for changing hydraulic loads, for a lowland and mountainous river.

Nitrogen retention follows an approach by Venohr (2006) and Venohr et al. (2011). For nitrogen denitrification as main and permanent retention process was considered. Here, the water temperature as indicator for the activity of denitrifying bacteria and the ration of water volume per sediment area (expressed as hydraulic load = annual mean run-off divided by water surface area) are used as central drivers.

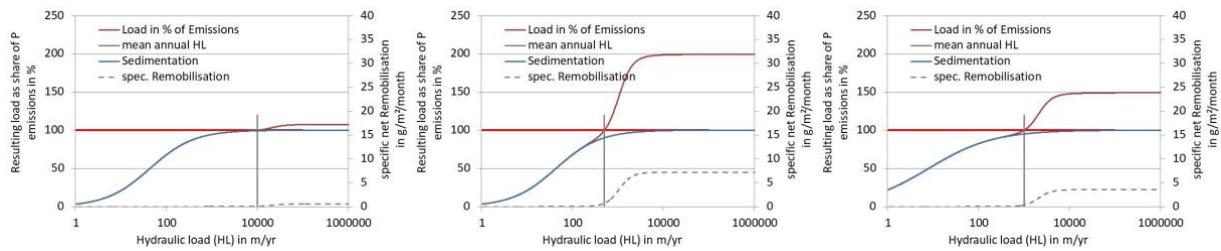


Figure 15: Response of the new approach for P retention and remobilisation to hydraulic load (HL), left: large rivers - mean HL = 10000 m/yr, mean slope = 1 %), middle: small lowland brook - mean HL = 500 m/yr, mean slope = 1 %), right: small hilly brook - mean HL = 1000 m/yr, mean slope = 15 %).

Results

MONERIS delivers three main results: a) nutrient emissions per pathway, land-use, FEC and month, b) in-stream retention of nutrients in main rivers and tributaries and main rivers per month and FEC, and c) resulting nutrient loads and concentrations at the outlet of a FEC per month. As the scenario analysis tool only considered the modelled concentrations in surface waters, this chapter will only briefly describe the general results and conclude on the validation of the modelled loads and concentrations.

As a mean for 2001 – 2010 in total $6629 \cdot 10^3$ tons nitrogen and $292 \cdot 10^3$ tons phosphorus per year are emitted to surface waters in the entire modelling extent. This equals a mean area specific emission of $11 \text{ kg ha}^{-1}\text{yr}^{-1}$ and $48 \text{ kg km}^{-2}\text{yr}^{-1}$. However, these values vary considerably among pathways and source areas (Table 7) and countries (Figure-A 3). Comparing emissions from different land-uses types, arable land with areas share of 33% contributes 47 % and 35% of total nitrogen and total phosphorus emissions, respectively. This ratio becomes even more disbalanced, when looking at the individual share of FECs on the total emissions (Figure 17). According to Figure 16 20 % of the total area contributes 50% and 58% of the total nitrogen and total phosphorus emissions.

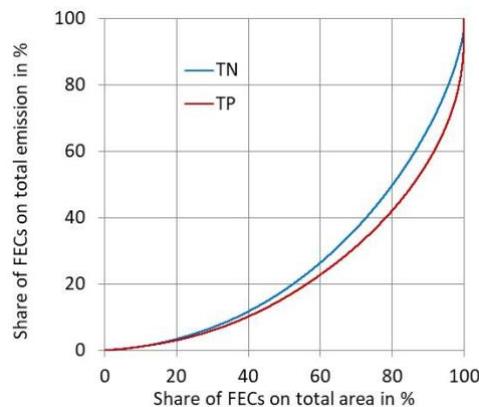


Figure 16: Comparison of the accumulative share of FECs on the total area and their share on the total emissions, for nitrogen and phosphorus in the entire MARS modelling extent.

Nutrient retention was modelled as function of the hydraulic load and water temperature (TN) and on basis of mean slope and hydraulic load (TP). Hydraulic load is defined as the ratio between run-off divided by water surface area. In general, with increasing water surface area and decreasing run-off TN and TP retention increases. Additionally, TN retention is modulated by varying water temperature (increasing retention with increasing temperature), whereas TP retention in general decrease with increasing slope and flow velocity. These patterns are well described by the spatial distribution of the mean TN and TP retention in surface waters of a FEC (Figure 17). In general TP retention is considerably lower than TN retention, but can reach values of 30% in presence of lakes.

Applying the calculated retention to the nutrient emissions and adding a routing function to this, the resulting loads in surface waters are calculated. Here, it is important to note, that retention in surface waters is calculated in a cascade-like way, meaning that the load leaving a FEC is subject to retention in the main river of the next downstream FEC.

The model validation was done by comparing calculated and observed run-off and concentrations, obtained from the European Water Archive of the Global Runoff Data Centre (GRDC-EWA). As the present modelling task was done in context of the MARS scenario analysis tool (SAT) development, and as the SAT works on the mean annual conditions during the period 2001-2010 we restricted the load comparison to corresponding data. A central input data for MONERIS was the modelled run-off provided by PCR-GlobWB. For the validation of the modelled run-off please see Chapter 2.1). When comparing all modelled and observed loads a high deviation with a poor statistical agreement was found.

Table 7: Share of both nitrogen and phosphorus emissions from different land-use types and via considered pathways, area specific emission for nitrogen in kg/ha and for phosphorus in kg/km², numbers in brackets represent the share on the total nitrogen or phosphorus emissions. WSA = water surface area.

Land use	WSA	Arable land	Grassland	Forest	Urban area	Other Areas	Total
Area in km ²	210965	2052587	411781	2769989	219559	469221	6134102
area share in %	3	33	7	45	4	8	100
Nitrogen emission in kg ha⁻¹ yr⁻¹ (share on total emission in %)							
atmospheric deposition	6 (2)						0.2 (2)
surface run-off		7 (2)	1 (0.7)	0.5 (2)		0.4 (0,3)	0.6 (6)
Erosion		0.39 (1)	0.03 (0.02)	0.05 (0.2)		0.004 (0.003)	0.2 (1)
tile drainages		2 (7)	0.5 (0.3)				0.8 (8)
groundwater		12 (36)	18 (11)	3 (13)	6,4 (2)	7 (5)	7 (67)
urban systems from this: sewer systems					19 (6)	1 (1)	0.8 (7)
					2 (0.7)		

<i>DCTP</i>					17 (5)		
point sources					27 (9)		0.9 (9)
Total (N)	6 (2)	15 (47)	19 (12)	4 (16)	52 (17)	9 (6)	11 (100)
Phosphorus emission in kg km⁻² yr⁻¹ (share on total emission in %)							
atmospheric deposition	36 (3)						1 (3)
surface run-off		2 (2)	5 (1)	4 (4)		10 (2)	4 (8)
Erosion		34 (24)	2 (0.3)	3 (3)		0.3 (0.04)	13 (27)
tile drainages		3 (2)	5 (0.8)				1 (3)
Groundwater		11 (7)	20 (3)	5 (5)	31 (2)	27 (4)	10 (21)
urban systems					188 (14)	9 (1)	7 (15)
<i>from this:</i>							
<i>sewer systems</i>					54 (4)		
<i>DCTP</i>					134 (10)		
point sources					303 (23)		11 (23)
Total (P)	36 (3)	49 (35)	33 (5)	13 (12)	521 (39)	45 (7)	48 (100)

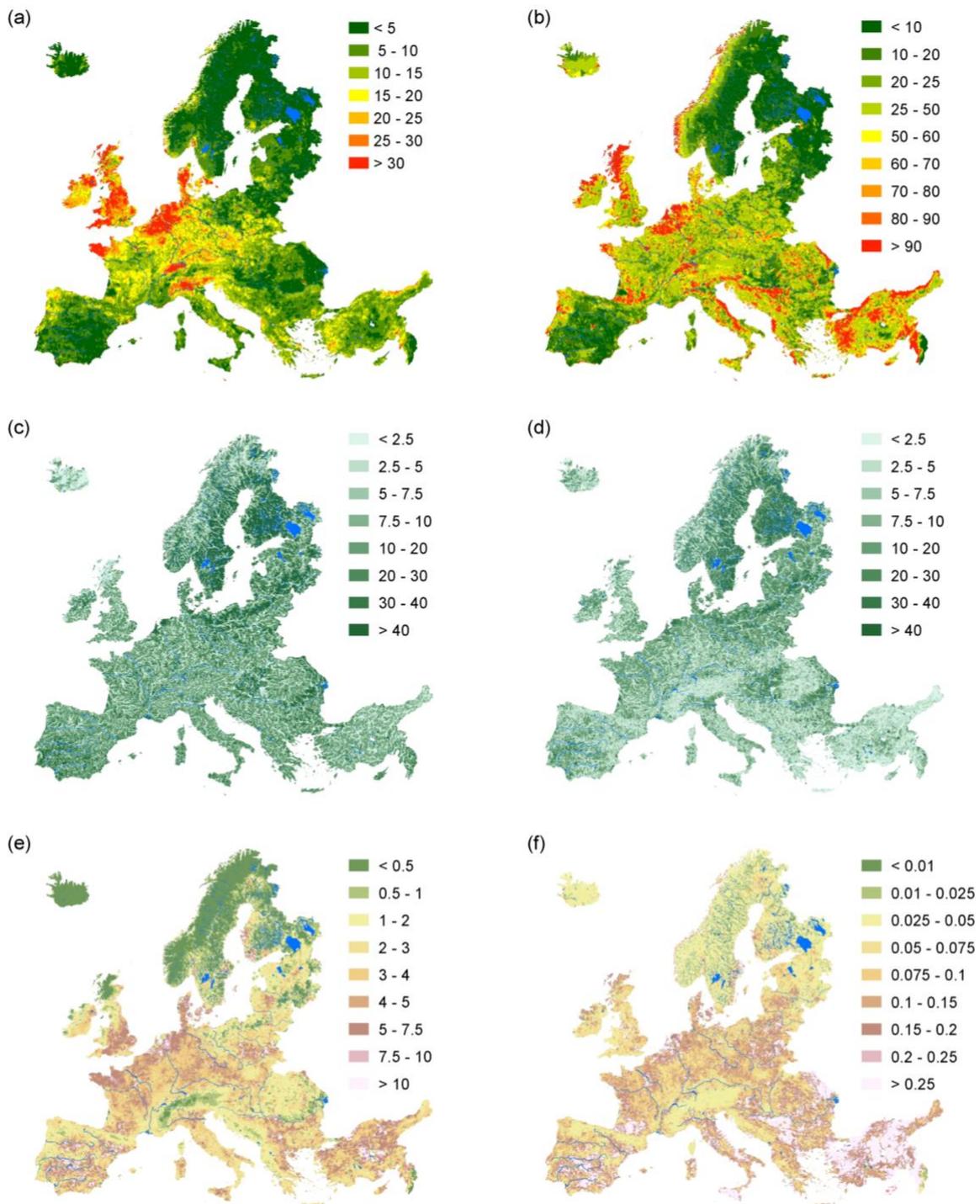


Figure 17: Some of the main results delivered by MONERIS: (a) specific nitrogen emissions in kg/ha (b) specific phosphorus emissions in kg/km² (c) nitrogen retention in % (d) phosphorus retention in % (e) DIN concentrations in mg/l (f) phosphorus concentrations in mg/l

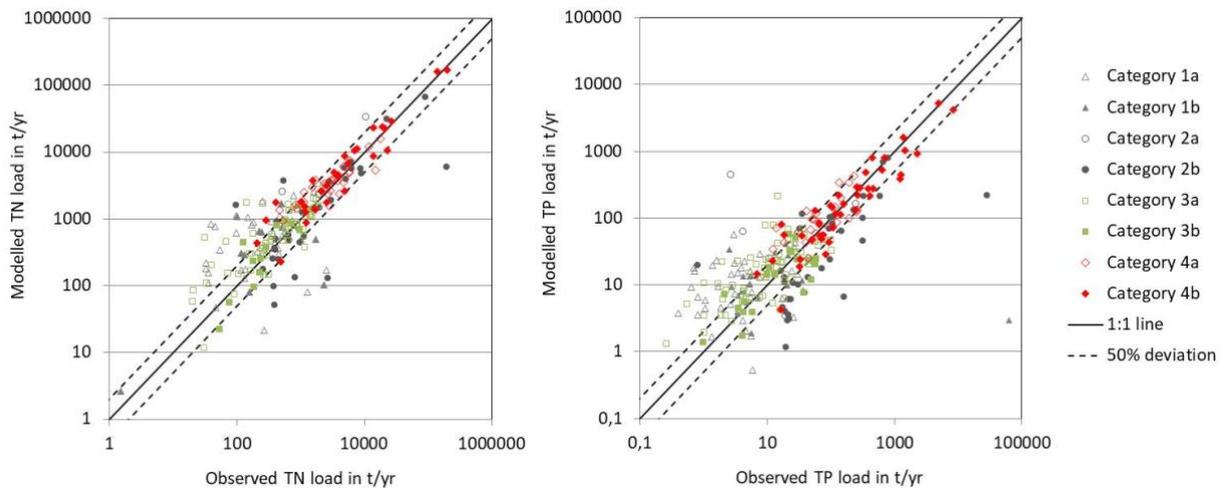


Figure 18: Comparison of mean modelled and observed TN and TP loads for the years 2001-2010. Categories are characterised by catchment size, deviation between modelled and observed run-off (provided by PCR-GlobWB), and the number of monitoring years from GRDC-EWA. For attribution of categories see Table 8.

Table 8: Comparison of mean modelled and observed TN and TP loads distinguished by categories.

Category	Catchment area in km ²	deviation mod.-obs. run-off in %	Available monitoring years (TN)	regression coefficient	r ²	mean abs. deviation in %	n
Total Nitrogen							
1a	< 1000	> 50	< 6	0.30	0.06	264	36
1b	< 1000	> 50	≥6	-0.05	0.00	2034	12
2a	> 1000	> 50	< 6	3.49	0.97	198	5
2b	> 1000	> 50	≥6	0.15	0.19	96	31
3a	< 1000	< 50	< 6	0.90	0.52	92	54
3b	< 1000	< 50	≥6	1.16	0.76	237	13
4a	> 1000	< 50	< 6	0.68	0.75	45	26
4b	> 1000	< 50	≥6	0.91	0.97	58	29
	> 1000	all	≥6	0.59	0.63	77	61
	all	< 50	≥6	0.92	0.97	115	43
	all	all	all	0.60	0.64	242	206
Total Phosphorus							
1a	< 1000	> 50	< 6	0.78	0.54	440	39
1b	< 1000	> 50	≥6	0.00	0.04	2896	21
2a	> 1000	> 50	< 6	0.06	0.00	3684	4
2b	> 1000	> 50	≥6	0.01	0.03	5640	34
3a	< 1000	< 50	< 6	0.41	0.08	191	63
3b	< 1000	< 50	≥6	0.35	0.16	63	19
4a	> 1000	< 50	<6	0.85	0.40	86	22
4b	> 1000	< 50	≥6	0.61	0.83	52	41
	> 1000	all	≥6	0.07	0.09	2592	76
	all	< 50	≥6	0.61	0.84	55	61
	all	all	all	0.01	0.02	1254	243

Consequently, we distinguished different categories of monitoring data and FECs. Monitoring data were not completely available for the entire modelling period. We assumed that time series covering less than 6 years are not representative for the modelling period and should not be used for the comparison. Nutrient emissions and resulting loads are strongly driven by run-off. In order to have a sound data base for assessing the model performance of the water quality model MONERIS from its error caused by input data we distinguished monitoring stations with a deviation in modelled (PCR-GlobWB) and observed run-off smaller/higher than 50%. Lastly, the SAT focuses on a European wide assessment of stressors and does not aim to provide individual results on local scale. Consequently, we split the available data set in monitoring stations with a catchment smaller (larger than 1,000 km²). The combination of these attributes formed 6 different categories of monitoring stations (see Table 8 and Figure 18). In general, for catchments larger 1,000 km² or less than 50 % deviation for run-off (both with more than 6 years data) provide a reasonably good agreement between observed and modelled loads. Underestimate in TP loads (for Q-dev. < 50%) is mostly caused by two monitoring stations. Stricter assumptions for the maximum acceptable run-off deviation would in general lead to a considerable improvement of the statistical agreement between modelled and observed loads but would also reduce the data set significantly.

3. JOINT GEO-DATABASE, DATA ANALYSIS AND POST-PROCESSING

3.1 HYDROLOGICAL INDICATORS

Lead: Yiannis Panagopoulos, Kostas Stefanidis

Towards the need for a comprehensive large-scale hydrologic stress analysis in MARS, we have tried to assess the hydrologic alteration of European rivers. Hydrologic alteration is attributed to water abstractions for the satisfaction of urban, industrial and agricultural needs and expresses the hydrologic stress of rivers and streams, namely the disturbance or deviation of their hydrologic regime from the ideal undisturbed or natural conditions.

In MARS Deliverable 5.1 related to WP5 we had conducted a European scale analysis of hydrologic data at the resolution of the Functional Elementary Catchment (FEC) (<http://www.eea.europa.eu/data-and-maps/data/european-catchments-and-rivers-network>). Simulated daily time-series of river flows from the PCR-GLOBWB global model (Van Beek et al., 2011, Sutanudjaja et al., 2014) were used based on a hypothetical near-natural scenario where water abstractions from water bodies did not exist and an anthropogenic scenario with water abstractions occurring. The latter practically represented the reality. Many hydrologic indicators expressing the characteristics of the rivers' hydrologic regime were calculated for all FECs with the Indicators of Hydrologic Alteration (IHA) methodology and software package and the deviations of the indicators' values between the two scenarios were used as proxy metrics of hydrologic alteration or hydrologic stress of rivers.

For the needs of WP7 and the present deliverable in particular, we followed the same procedure to assess the hydrologic alteration of rivers at the FEC level due to future climate and socio-economic changes. We have chosen two future scenarios that give rise to four future scenario runs, namely: RCP4.5 in combination with SSP2 for the future time-horizons 2026-2035 and 2056-2065 and RCP8.5 in combination with SPP5 for the future time-horizons 2026-2035 and 2056-2065 as well (see Chapter 2 above). Daily discharge time-series have been calculated with the PCR-GLOBWB global model again after changing climate and management input according to each scenario. We used as basis the natural scenario results to compare our new findings and estimate the hydrologic alteration under future conditions.

This chapter describes briefly the Indicators of Hydrologic Alteration (IHA) approach (Richter et al., 1996), used to address hydrologic stress in Europe at the FEC level. The methodology also describes how we associated simulated hydrologic data to the thousands of FECs across Europe. The results are presented

and discussed through maps which show the change in future river characteristics across Europe from the near-natural scenario.

Input data

The implementation of the proposed IHA method requires time series of daily flows for at least 10-15 years. This means that a detailed dataset covering the whole Europe and containing the needed information is crucial for the objectives of this work. Data from gauged sites that meet the above requirements are rare. But given that we look for future data, observations cannot meet this requirement anyway. Thus, we take advantage of modelled hydrology data simulated by the large scale hydrology model PCR-GlobWB (Van Beek et al., 2011, Sutanudjaja et al., 2014).

The model includes an online water demand scheme to estimate irrigation water requirement. Briefly, this scheme separately parameterizes two different irrigated crop groups: paddy and non-paddy, aggregated from 26 crop classes from the MIRCA2000 dataset (Portmann et al., 2010) that accounts for various growing season lengths under different regional practices and climatic conditions. Other sectoral water demands include livestock, industry and households. Data for those sectors were obtained from several sources for the historic simulations (see MARS Deliverable 5.1) and have been developed for future conditions based on the MARS storylines (see Chapter 2.1 above).

PCR-GlobWB simulations have been performed for a historic 10-y period under both a naturalized (no abstraction) run and an anthropogenic run (with human influence). Here the main characteristics of these scenarios are listed.

Near-natural (no abstraction) scenario:

- A grid-cell in this scenario constitutes up to three land cover classes: short vegetation, tall vegetation and surface water bodies.
- Basically, the parameters for the first two land cover classes are based on the Global Land Cover Characteristics Data Base Version 2.0 (GLCC 2.0, http://edc2.usgs.gov/glcc/globe_int.php).
- Fractions of land cover classes are assumed to be fixed throughout the entire model simulation (e.g. no deforestation), i.e. there is no land use/cover change. For this scenario, only natural surface water bodies, e.g. rivers, wetland and lakes, are considered. Reservoirs (dam constructions) are not simulated.
- No water demand was simulated, and, therefore, no water abstraction.

Anthropogenic scenario:

- A grid-cell in this scenario constitutes up to five land cover classes: short natural vegetation, tall natural vegetation, surface water bodies (including reservoirs), as well as two classes of irrigated crop types: paddy and non-paddy types.
- The parameters for the first two classes (natural, non-irrigated land types) are basically based on the GLCC 2.0.
- For this scenario, areal extents of fractions of all land cover classes change on yearly basis, particularly due to expansion of irrigated areas and progressive construction of dams/reservoirs. Therefore, land use/cover change is simulated.
- To parameterize the reservoirs, the GRanD dataset was used.
- Water demand is simulated and, therefore, water abstraction is also simulated.

The result of the two scenarios was two data sets of daily discharges for a ten year period (2001-2010). The purpose of the natural scenario dataset is to simulate the hydrologic conditions in Europe under a status of minimal anthropogenic pressures on water from abstractions and land use modifications. By implementing the IHA method for analysing these two hydrologic datasets we have compared the results and derived a degree of alteration between the baseline conditions and the near-natural scenario expressing proxies of pressures related to hydrologic alteration (MARS Deliverable 5.1).

Here, based on IHA methodology we analysed the four hydrologic datasets that represent four future scenario runs. These runs are based on two RCPs combined with a single GCM, the GFDL-ESM2M (see Chapter 2):

- RCP 4.5 (moderate change): In this pathway the radiative forcing stabilizes before 2100 due to the introduction of technologies and strategies that reduce greenhouse gas emissions;
- RCP 8.5 (largest changes): In this pathway there is a continuously increasing radiative forcing,

and two SSPs (O'Neill et al., 2013):

- SSP2 (in combination with RCP4.5), where mitigation and adaptation challenges are intermediate, thus, it can be seen as a continuation of the current trends (medium population growth, economic growth and technological change);
- SSP5 (in combination with RCP8.5), where the challenges for mitigation are high due to a lack of climate policy and high emissions whereas at the same time there are factors that reduce the mitigation capacity of the society, such as rapid population increases, large heterogeneity between different groups within the society, lack of political will or limited financial resources.

Narratives of future changes belonging to these pathways have been sketched. As described in Chapter 2.1, quantitative data for future irrigation efficiency and irrigated area, domestic and industrial water demand, and limitations for groundwater abstractions were estimated to feed the PCR-GlobWB to produce future predictions of river discharges across Europe.

Data allocation to FECs

We assigned representatively the gridded hydrologic data produced by the PCR-GlobWB model to the 104,334 Functional Elementary Catchments (FECs). At first, the centroids of the FEC polygons were calculated and a new shapefile was produced as shown in Figure 19(left). The share of the upstream area that corresponds to each FEC was matched with the FEC centroid objects. Next, a new shapefile of the centroids of the modelled raster cells was created adding the upstream area for each cell (see Figure 19). For each FEC's centroid a buffer with a 15 km radius was created (see Figure 19) and then intersected with the PCR-GlobWB centroid shapefile to identify which cell centroids fall within the buffer area of each FEC's centroid. This resulted into having several grid points in one FEC buffer (see Figure 19). Then, for each case (FEC buffer) we selected the cell centroid for which the absolute difference between the FEC's upstream area and Grid cell's upstream area was the minimum. This allowed us to minimize the number of cases where a grid cell with a large upstream area was wrongly assigned to a FEC with minimal influence from the upstream area (e.g. a small tributary).

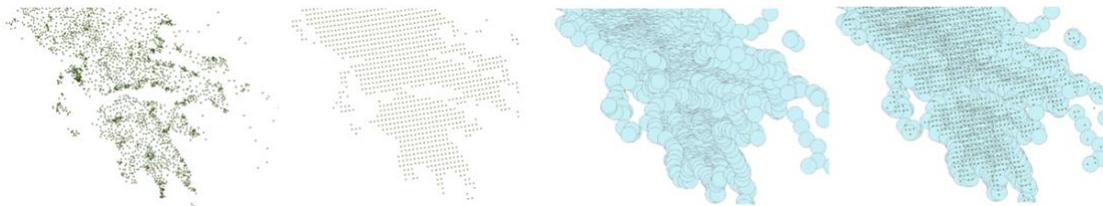


Figure 19: From left to right: Example showing a) calculated centroids of FECs polygons, b) calculated centroids of PCR-GLOBWB model raster cells, c) buffer zones created based upon the FEC centroids and d) the PCR-GLOBWB centroid cells that fall within the created buffer zones.

Indicators of Hydrologic Alteration

The Indicators of Hydrologic Alteration (IHA) was originally proposed by Richter et al. (1996, 1997, 1998; Poff et al., 1997) to assess the degree of hydrologic alteration caused by human intervention on rivers. The method is based on the calculation of several hydrologic parameters that characterize the intra- and inter-annual variability in water conditions, including the magnitude, frequency, duration, timing and rate of change of flows or water levels (Richter et al., 1996). Apart from their ability to reflect human-induced changes, the parameters have ecological relevance (Richter et al., 1997). Other researchers propose a smaller set of hydrologic parameters after identifying those that are redundant and inadequate. The calculation of the hydrologic parameters is computed with the use of a free software tool developed by

The Nature Conservancy, called the Indicators of Hydrologic Alteration (IHA, 2009). The parameters are considered as representative of crucial relationships between flow and ecological functions and are categorized in the major components of flow, all considered as ecologically important: low flows, high flows and floods. It should be mentioned that some of the hydrologic parameters calculated are inter-correlated, resulting in considerable information redundancy (Gao et al., 2009). Moreover, due to the short length of time-series, some other parameters such as the ones related to floods cannot be considered reliable. Therefore, to increase the efficiency of the analysis at a large-scale (multiple sites with flow time-series) there is a need to reduce the number of indices to be used to those which are adequate to provide a comprehensive overall determination of the hydrologic alteration. Based on our experience and the analysis of the actual data we chose the most two comprehensive and reliable parameters to express hydrologic alteration the mean annual flow and the base flow index.

Results

As described before, the Indicators of Hydrologic Alteration (IHA) had been already calculated for two datasets of simulated daily discharges. One dataset is modelled assuming zero water abstractions and natural type of land uses (near-natural scenario), and the other dataset is obtained through baseline model runs (anthropogenic scenario) of the PCR-GlobWB model. In order to assess the deviation of the baseline hydrologic conditions from the near-natural scenario we have calculated the ratios between the values of the indicators for the anthropogenic scenario and the values of the indicators for the near-natural scenario (anthropogenic scenario over near-natural scenario). Focusing on ratios, if the value for a certain indicator is 1, it means that there is no alteration between the “anthropogenic” model run and the “near-natural” model run. If the ratio is above 1, then the value of the hydrologic indicator for the anthropogenic is greater than the near-natural scenario. This was repeated four more times, each one for each future scenario run and alteration for each FEC is calculated as the ratio *future scenario X / near-natural scenario*.

It should be noted that the ‘pressure’ on water is not expressed consistently by above unity numbers (or the opposite) but depends on the nature of the indicator analysed. However, for both our indicators selected to show the alteration results (mean annual flow and base flow index) a ‘negative’ alteration is shown by ratios below unity. Thus, a below ‘1’ value for alteration in the mean annual flow shows that the 10-y mean annual flow of a river under a future scenario has been reduced from the natural scenario or the base flow index, which expresses the magnitude of base flow over the total river flow, has been decreased as well. The following GIS European maps at the FEC level (Figure 20 and Figure 21) show the alteration results for both parameters under all scenarios simulated.

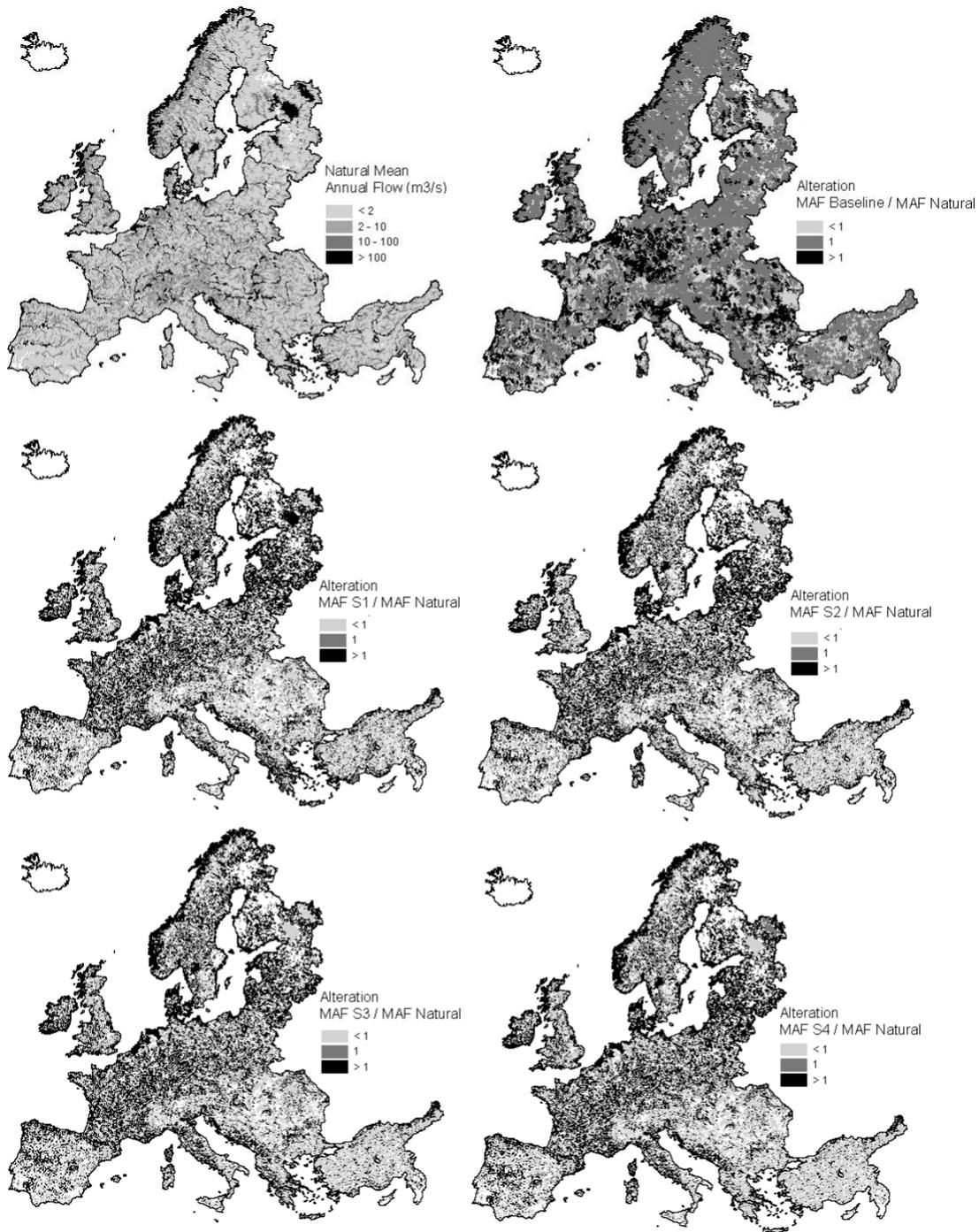


Figure 20: Top Left: Mean Annual Flow (m^3/s) of the 10-y period 2001-2010 under the natural scenario (historic climate with no water abstractions). Top Right: Alteration of baseline (2001-2010 with abstractions) mean annual flow from the natural conditions, expressed as mean annual flow (baseline) / mean annual flow (natural). Ratios below unity indicate decrease in flow due to abstractions, values equal to unity show no alteration and values above unity show increase from the natural conditions. Reduction is depicted with lighter colours and is observed mostly in parts of the Mediterranean countries. Increase in flows (ratio >1) is found in Central Europe, with very large parts all across the continent and especially in the North remaining unaltered with respect to annual river flows (ratio = 1). Middle and Bottom: Alteration of mean annual flow under four future scenarios (S). S1: RCP4.5 - SSP2 for 2026-2035, S2: RCP4.5 - SSP2 for 2056-2065, S3: RCP8.5 - SSP5 for 2026-2035 and S4: RCP8.5 - SSP5 for 2056-2065.

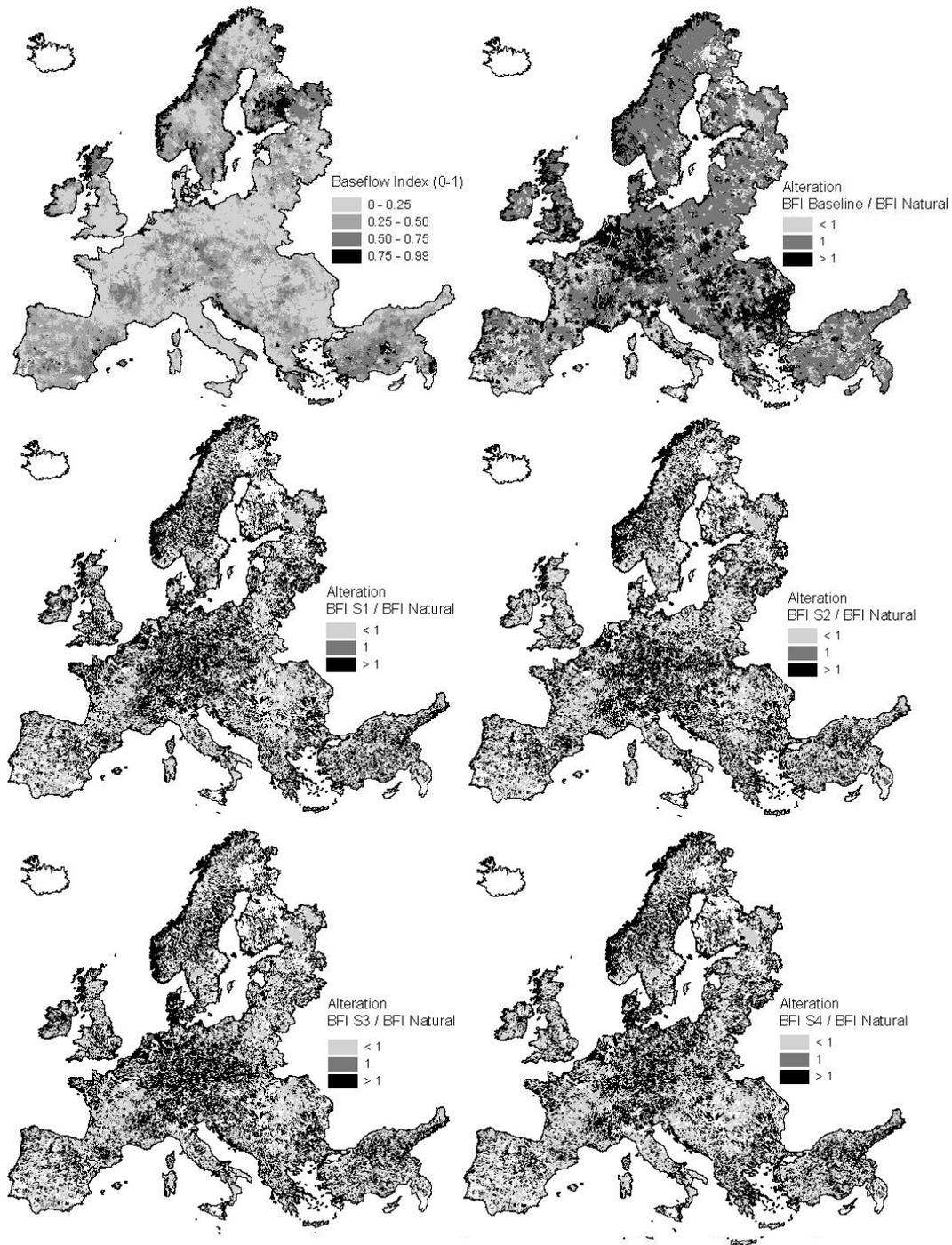


Figure 21 Top Left: Base Flow Index of the 10-y period 2001-2010 under the natural scenario (historic climate with no water abstractions). Base flow Index is defined as the 7-d minimum flow/mean annual flow of the year. Top Right: Alteration of the base flow index from the natural conditions, expressed as BFI (baseline) / BFI (natural). Ratios below unity indicate decrease in BFI due to abstractions, values equal to unity show no alteration and values above unity show increase from the natural conditions. Reduction is depicted on the right map with lighter colours and is observed mostly in parts of the Mediterranean countries. Increase in BFI (ratio >1) is found in Central Europe, with very large parts all across the continent and especially in the North remaining unaltered. Middle and Bottom: Alteration of BFI under four future scenarios (S). S1: RCP4.5 - SSP2 for 2026-2035, S2: RCP4.5 - SSP2 for 2056-2065, S3: RCP8.5 - SSP5 for 2026-2035 and S4: RCP8.5 - SSP5 for 2056-2065.

A clear reduction of annual river flows was indicated for parts of all the Mediterranean countries (Greece, Turkey, Spain, France, and Italy) due to water abstractions occurring in the anthropogenic (baseline) scenario. However, the average flow conditions are not influenced for most of the rest of Europe, while for central Europe the baseline mean annual flow is even increased due to abstractions. This can only be interpreted by water diversions or water returns to rivers from industrial activities via point sources.

The base flow index map also shows that alteration of rivers' hydrology from the natural conditions due to the baseline anthropogenic activities is clearer in Southern Europe where agricultural water use predominates. Light colours of BFI alteration (<1) are observed in parts of all the Mediterranean countries but not much in the rest of Europe where the majority of FECs seem to remain unaltered (ratio = 1). The reduced significance of base flow is certainly attributed to the main water use of the Mediterranean countries which is agricultural. Irrigation water in these countries highly depends on groundwater resources, abstractions lower groundwater storage and subsequently, groundwater contribution to streamflow is reduced.

By studying the middle and bottom parts of the Figures we see that future scenarios cause almost everywhere either a positive or negative alteration. This is expected due to the climate change that was included in the drivers of hydrologic alteration in contrast to the baseline case where only water management was responsible for the changes in the rivers' hydrologic regimes.

Under both RCPs both mean annual flows and base flow indices show that in the majority of the Eastern and Southern Europe water availability deteriorates (ratios <1) in 2030 and 2060. In the mid-century scenarios (2060) Southern countries such as Greece seem to suffer even more as more areas appear with light colours for the mean annual flows (ratios <1). However, studying the left and right maps of the middle and bottom parts of the figures we see that changes from early- to mid-century are not significant for the majority of Europe with the trend in hydrologic alteration being the same (either positive or negative alteration). This shows that under the two scenarios the changes occur rapidly in the century but possibly need more decades until 2100 to become more severe.

The FEC level calculations allow investigating possible high alterations all across Europe and to this end the produced GIS alteration layers are valuable. We have to mention that the assessment of hydrologic alteration in this work was based on simulated data from a global model with modelling uncertainties and simplifications which are unambiguously transferred to the present results. The short length of the simulated flow time-series does not allow us to focus more on flood events and their hydrologic characteristics as small and large floods are defined in the method based on a 2-y and a 10-y interval respectively, certainly not in line with the 10-y length of the available data.

3.2 GEODATA-BASE, EFFECTS AND THRESHOLDS

Lead: Lidija Globevnik

Input data

The list of candidate stressor indicators have been developed in work package 5.1 (see MARS del. 5) and calculated (modelled) for each main drain (river section) of an elementary functional catchment (FEC). The first group of variables that represent stressor indicators related to excessive nutrient concentrations and loads due to pollution coming from either point or diffuse sources. The second and third group of variables represent hydrological and morphological alterations. Hydrological alterations (water flow changes due abstraction, flow diversion, reservoirs, impounding, and concentrated water releases) are represented as ratio between present flow (altered) and semi-natural flow. Morphological alterations (natural riverine habitats losses, disturbed longitudinal connectivity, reduced lateral connection to floodplains due to levelling of the riverbed and consolidation of the banks and the bottom) are represented by flood plain and FEC land uses. The list of candidate stressor indicators as predicting variables selected for the scenario analysis tool is given in Table-A 4.

Methods

The final selection of stressor indicators is based on the following starting points and conditions:

- to reduce the final number of stressor indicators as much as possible,
- to keep a representative set of stressor indicators comprising the most complete picture of related drivers and pressures
- to use the same set of stressor indicators for all river FECs in Europe
- to define thresholds for each broad river type used in the model setting (rivers types grouped by three catchment size classes, three altitude classes and on group of Mediterranean rivers what gives ten (10) broad river types (BRT); see Table 9
- to include those stressor indicators that can be modelled and assessed for future scenarios
- to exclude auto correlated stressor indicators and to reduce redundancy
- to obtain acceptable level of modelling accuracy for selected non-probabilistic regression analysis.

From the list of candidate predicting variables we selected the first set of weakly redundant variables and those with high expectation of explanatory power. Table 10 shows ranking of variables according to their relative influence (%) in each broad river type for 12 variables.

Alteration of selected hydrological parameter represent situations when current hydrological parameter (e.g. mean annual flow) is lower or larger than semi-natural (ratio between current and semi-natural

hydrological indicator is lower or larger than one (1) respectively). We analysed correlations between the alteration of selected hydrological parameters with five (5) ecological status to check if unilateral relations exist between them, since this a required condition for the development of a scenario tool. To satisfy this prerequisite we divided hydrological alterations into two conditions, one representing an increase, the other a decrease of stressor indicators compared to semi-natural state. The alteration indicator is then defined as difference between the calculated ratio and one (1). Negative values present decrease of hydrological parameter compared to semi-natural state, whereas positive values present an increase.

Table 9: Description of broad river types for which thresholds were defined for the scenario analysis tool

BRT	Broad River Type Name	Altitude (m.a.s.l.)	Catchment size (km ²)	Geology	Broad Type Rivers (ETC/ICM 2015)
1	Very large rivers	any	> 10,000	any	1
2	Lowland Brooks	≤ 200	≤ 100	any	2-7 if catchment size ≤ 100
3	Lowland Stream and rivers	≤ 200	100 – 10,000	any	2-7 if catchment size between 100 - 10000
4	Mountains, Siliceous, Brooks	200 - 800	≤ 100	siliceous and organic	8, 9, 12 if catchment size ≤ 100
5	Mountains, Siliceous, Streams and rivers	200 - 800	100 – 10,000	siliceous and organic	8, 9, 12 if catchment size between 100 - 10000
6	Mountains, Calcareous, Brooks	200 - 800	≤ 100	calcareous/mixed and organic	10, 11, 13 if catchment size ≤ 100
7	Mountains, Calcareous, Streams and rivers	200 - 800	100 – 10,000	calcareous/mixed and organic	10, 11, 13 if catchment size between 100 - 10000
8	Highland and glacial rivers	> 800 > 200 glacial	< 10,000	any	14, 15, 16
9	Mediterranean perennial rivers	< 800	100 -10,000	any	17, 18
10	Mediterranean temporary or very small rivers (brooks)	< 800	≤ 100 <1,000 (temporary)	any	19, 20

The analysis showed that both, decrease and increase of mean annual flow negatively correlates with ecological status deterioration. In opposite, the decrease of a base flow index negatively correlates with ecological status deterioration, whereas the increase does not show any positive correlation. In the

further analysis we take into account only “negative” alteration of base flow index as an indicator and normalised its values: value “0” presents no alteration (difference between ratio and 1 is null), the value “1” presents 95-percentilis of all negative values in the dataset. The normalisation of negative values of mean annual flow alteration indicator follows the same rules as for negative base flow index alterations. The positive values of mean annual flow alteration indicator between null and 2.5 are normalised between “0” and “1”, whereas all larger than 2.5 are set to “1”.

Results

By applying multivariate analysis we calculated the explanatory power of stressor indicators on the reported ecological state. As the scenario analysis tool aims to deliver results on European wide conditions (including EFTA and other neighbouring countries) we were limited to input data available for the entire modelling extent. This for example made it necessary referring morphological stressor indicators to the land-use in a FEC and not only to such in riparian areas, as e.g. Copernicus data was not available for all FECs. The explanatory power of stressor indicators on land uses (lu_urb, lu_for, lu_agr), nutrients (nu_din, nu_tp) and two hydrological alterations, mean annual flow alteration (hy_maf) and base flow index (hy_basef) have the largest explanatory power, either to explain ecological status of all rivers in one data set (column “all”) or averaged by broad river types (last column, Table 10).

Table 10: Variables predicting general ecological state with the highest explanatory power (%).

	all	BRT 1	BRT 2	BRT 3	BRT 4	BRT 5	BRT 6	BRT 7	BRT 8	BRT 9	BRT 10	average by BRT
<i>nu_din</i>	16.6	18	12.7	17.1	17	19.3	19.4	20	11.1	12.1	8.1	15.5
<i>nu_tp</i>	12.3	18.2	8.5	11.1	8.1	13.8	8.5	13	16.2	16.7	9.1	12.3
<i>hy_basef</i>	11.4	9.6	12.2	12.8	7.7	13.5	8.1	12.2	12.1	11.4	7.1	10.7
<i>lu_urb</i>	13.5	5.8	21.6	9.5	15.5	5.5	14.2	9.1	12.3	6.3	16.2	11.6
<i>hy_maf</i>	10.6	10.7	9.3	11.7	5.7	8.8	3.3	11.4	10.2	11.2	4.3	8.7
<i>lu_for</i>	6.7	6.9	7.2	6.4	12.5	8.6	21.6	7.3	8.1	7	4.2	9.0
<i>lu_agr</i>	7.1	7.9	8.7	8.1	4.6	5	6.4	6.8	6.7	7.1	34.9	9.6
<i>hy_lp</i>	3.6	4.3	3	3.4	1.7	3	3	3	3.6	5.9	2.9	3.4
<i>hy_hp</i>	2.4	8.5	2.5	2.8	2	2.7	4	4.3	1	2.5	1.1	3.1
<i>hy_ld</i>	2.8	6.3	2.3	2.6	2.5	2.8	1.2	2.7	3.5	5	2.9	3.2
<i>hy_hd</i>	1.6	3.9	1.8	2.2	2.2	1.6	3.1	1.3	0.8	2.2	0.4	2.0

Legend: *nu_din*: concentration of dissolved nitrogen in water; *nu_tp*: total phosphorus concentration in water; *hy_basef*: alteration of base flow index (ratio between present base flow index and semi-natural base flow index); *lu_f_urb*: share of urban land use in FEC; *hy_maf*: alteration of mean annual flow (ratio between present mean annual flow and semi-natural mean annual flow); *lu_f_for*: share of forest in FEC, *lu_f_agr*: share of agricultural land use in FEC; *hy_lp*: alteration of low pulse threshold; *hy_hp*: alteration of high pulse threshold; *hy_ld*: alteration of extreme low flow duration; *hy_hd*: alteration of high flow duration;

Based on their explanatory power we selected a final set of six (6) variables for which we define thresholds between good or high ecological status and less than good ecological status:

nu_din:	nitrate, nitrite + ammonium concentration in water (mg/l N)
nu_tp:	total phosphorus concentration in water (mg/l P)
hy_basef:	increased base flow index – positive, normalised (%)
hy_maf:	mean annual flow alterations – absolute normalised value (%)
lu_urb:	share of urban area in FEC (%)
lu_agr:	share of agricultural area in FEC (%)

Table 11: Thresholds for six stressor indicators by broad river types derived with classifier algorithm C4.5.

	BRT 1	BRT 2	BRT 3	BRT 4	BRT 5
<i>nu_din</i>	2.0	0.42	2.26	2.43	2.8
<i>nu_tp</i>	0.10	0.03	0.34	0.12	0.04
<i>lu_urb</i>	5.17	6.24	1.78	1.9	1.1
<i>lu_agr</i>	5.29	1.9	24.8	16.2	16.2
<i>hy_maf</i>	12.2	0.01	15.5	3.2	0.6
<i>hy_basef</i>	5.1	0.02	3.1	20	1.2
	BRT 6	BRT 7	BRT 8	BRT 9	BRT 10
<i>nu_din</i>	2.0	3.1	2.07	2.67	2.76
<i>nu_tp</i>	0.05	0.14	0.14	0.12	0.12
<i>lu_urb</i>	6.9	3.3	2.0	3.76	13.4
<i>lu_agr</i>	27.8	37.1	15.5	12.1	24.6
<i>hy_maf</i>	20	1.8	20	3.8	5.9
<i>hy_basef</i>	4.6	2.2	4.1	8.3	6.3

Thresholds for selected six variables are defined using classification approach of data mining with algorithm C4.5 (Quinlan, 1993). It is a statistic classifier algorithm that uses info gain ratio for variable (feature) selection and to construct the decision tree and is widely used because of its quick classification and high precision (Sharma et al., 2013). Evaluation of results (validation) is done with cross validation methods (10 folds), where learning is first done on test data (randomly chosen 10% of instances). To avoid over fitting the supervised class balancing with weightings method (re-weighting the instances in the data so that each class has the same total weight) and validated with SMOTE method (Synthetic Minority Oversampling Technique according to Nitesh et al., 2002). To capture important structural information and to decrease error as much as possible, we used pruning (removal of decision nodes that do not provide any significant additional information). The confidence factor used for pruning (smaller values incur more pruning) is 0.25 and or 0.05. The minimum number of instances per node (leaf) varied between 10 and 100.

Thresholds are given in Table 11. The classification accuracy and other parameters used in classification (selected classification scheme parameters) are given in Table 12.

Table 12: Classification accuracy and selected classification scheme parameters used to derive thresholds for selected pressure variables.

	BRT 1	BRT 2	BRT 3	BRT 4	BRT 5
<i>Number of instances</i>	1397	5471	11259	3458	4372
<i>Confidence factor</i>	0.25	0.25	0.25	0.25	0.25
<i>Pruning*</i>	50	50	100	50	50
<i>Correctly classified (%)</i>	72.7	60.6	70.6	66.9	69.9
<i>Kappa statistics</i>	0.44	0.22	0.20	0.34	0.40
	BRT 6	BRT 7	BRT 8	BRT 9	BRT 10
<i>Number of instances</i>	2240	2576	2761	3486	3170
<i>Confidence factor</i>	0.25	0.25	0.25	0.25	0.25
<i>Pruning*</i>	50	25	50	50	50
<i>Correctly classified (%)</i>	64.9	64.4	63.5	65.0	66.4
<i>Kappa statistics</i>	0.40	0.29	0.27	0.30	0.33

* Number of minimum instances in a leaf

The highest classification accuracy is obtained for datasets representing BRT 1 (large rivers), but accuracies for the other types do not fall below 60%. The prediction performance of classifiers (measured by Kappa statistics (inter-rater agreement) is moderate for majority of broad river types.

4. IMPLEMENTATION OF THE SCENARIO ANALYSIS TOOL

Lead: Markus Venohr, Judith Mahnkopf, Andreas Gericke

General approach

The joint application of the above described models (Figure 22) and the consistent analysis and assessment of the results led to the identification and consideration of in total six stressor indicators (Table 13) and thresholds determining whether they have a significant impact on the ecological status. As a last step for the development of the SAT, the combination of active or inactive stressor indicators (see Chapter 1) had to be translated into the probability to reach a good ecological status. This was done using a Bayesian Belief Network trained for current conditions. The assessment of changing probabilities under future conditions was eventually done by feeding changed stressor indicator values into the trained Bayesian Belief Networks. The technical steps are described in more detail below.

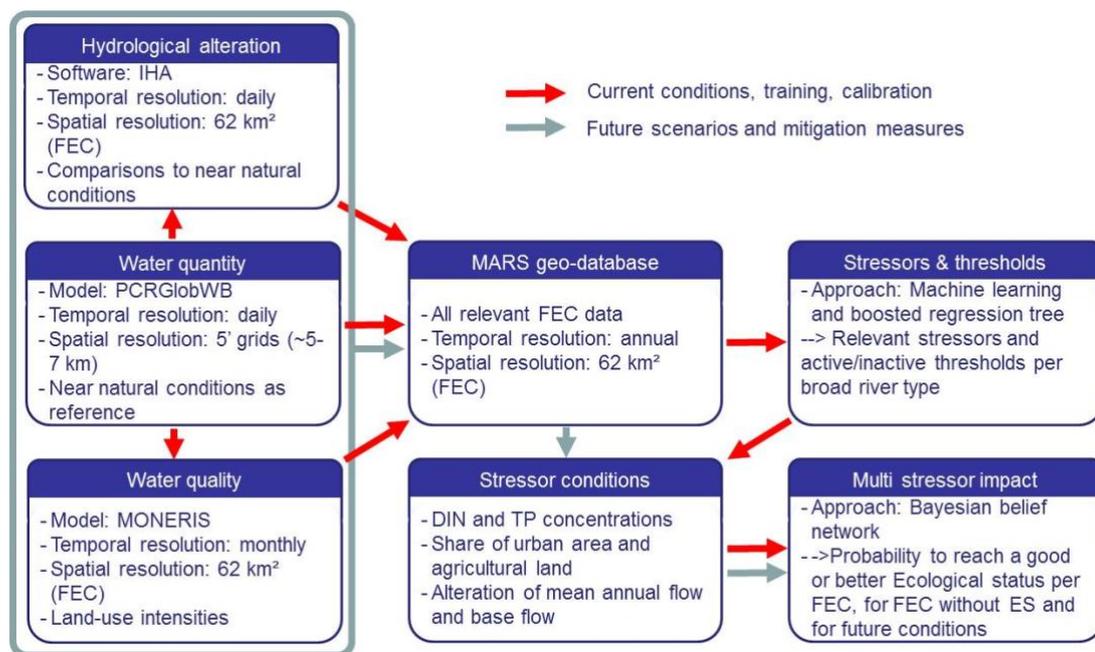


Figure 22: Principle workflow and data exchange for calibrating and learning the SAT-approach and for the application at scenario conditions.

Estimating the probability to reach a good ecological status

As stated before the impact assessment of stressor indicators was derived on the basis of the reported ecological state. Here we distinguished whether a good or high status is reached or not. These two conditions were compared to the distribution and combination of the six active or inactive stressor

indicators (Table 13). A stressor is considered active if the threshold is exceeded, and inactive if the value remains below the threshold. These stressor indicators were shown to explain the major share or the reported ecological status (see Chapter 3.2, Table 15).

Table 13: List of considered stressor indicators.

Name	Description	Units
<i>nu_din</i>	Concentration of dissolved inorganic nitrogen (DIN) in main river at outlet of sub-catchment by MONERIS	mg/l
<i>nu_tp</i>	Concentration of total phosphorus (TP) in main river at outlet of a sub-catchment modelled by MONERIS	mg/l
<i>lu_urb</i>	Area share of urban area per sub-catchment derived from land-use maps (CORINE & GlobCorine)	%
<i>lu_agr</i>	Area share of urban area per sub-catchment derived from land-use maps (CORINE & GlobCorine)	%
<i>hy_maf</i>	Change of mean annual flow (maf) between near natural conditions and current/scenario conditions.	%
<i>hy_basef</i>	Change of base flow index (basef) between near natural conditions and current/scenario conditions. Base flow index is the ratio between 7-day minimum flow divided by mean annual flow. Only positive changes (increasing base flow indices were considered)	%

As described above the six stressor indicators represent three different categories of stressors:

nu_din and nu_tp: the nutrient concentration represent an integral of the entire catchment. Beyond the land-use and land-use intensities (population density, N surplus, extent of waste water treatment) it is also majorly impacted by water availability and, to a lesser extent, by the size and distribution of larger lakes for which increased retention rates can be assumed.

lu_urb and lu_agr: These shares of land-use-class-areas on the FEC-area is an indicator for the local hydro-morphological alteration. It strictly only represent the local conditions in a single FEC. Any hydro-morphological alterations in up-stream catchments (Hinterland FECs) are not reflected by our analysis. During the selection process we tested these two hydro-morphological alteration stressors in two variants: a) share of urban area on the riparian area (according to Copernicus, 2015) along the main river of a FEC and b) share of urban area on the total FEC area. Our analysis showed that the explanatory strength of the two variants does not significantly change. As Copernicus data is not available for the entire MARS model extent, we decided to consider the area share on the total FEC area.

hy_maf and hy_basef: These stressor reflecting hydrological alterations are derived from GlobWB results for two different conditions: a) daily values current conditions between 2001 and 2010 and b) daily values

derived from current climatic conditions but assuming a complete absence of water abstractions and reservoirs. The latter version was assumed to represent near natural conditions. For both conditions the IHA software was applied (see Chapter 3.1) and more than 25 hydrological parameters were derived. By comparing the parameter of both conditions the stressor indicators were generated. Throughout the analysis two most explanatory stressor indicators were identified and selected for the final tool development.

All six stressor indicators more or less strictly reflect an alteration of natural conditions. For *hy_maf* and *hy_basef* the alteration has directly be derived while for *lu_urb* and *lu_agr* alteration is implemented as any presence of the two anthropogenic induced land-use classes implement a change from natural conditions. For nutrient concentrations (*nu_din* and *nu_tp*) the alteration aspect is less clear. However, nutrient concentrations under near natural conditions were much lower than present concentrations and did, according to Gadegast et al. (2014) and Hirt et al. (2013), vary much less between river basins than under current conditions. Here, current, anthropogenic altered concentrations in FECs are commonly bias by a much lower and less variable near natural concentration.

Hydro-bio-geo-chemical conditions vary considerably across Europe. Accordingly, thresholds were derived separately for different water body types. Our statistical approaches - thresholds but also the later described training of the Bayesian Belief Network – required a data set containing a representative combination of stressor indicator and status conditions. A larger sub-set of FECs for each water body type was also required to ensure long gradients of values. With this background we used a set of 10 different aggregated broad water body types (BRT, ETC/ICM 2015) (Figure-A 1) to conduct the analysis.

By applying regression tree analysis, thresholds for each of the six stressor indicators were determined and tested with cross validation methods (Table 14). The derived thresholds have to be interpreted carefully and cannot as such directly be seen as thresholds outside the here derived context. For e.g. RT2 (lowland brooks) the lowest share of FECs with a good or high ecological status was reported. Lowland areas are often used for intensive agricultural productions, so in many FECs of RT2 stressor indicators like *lu_agr*, *nu_din* or *nu_tp* are often on a high level. Consequently, there is a high probability for a multiple stressors situation co-prohibiting a good ecological status. This further means that only at a very low value of an individual stressor indicator, the statistical probability to reach a good ecological status can be presumed. In turn, *lu_urb* is often very on low level in RT2, as here, larger cities are less common. Consequently, only very high shares of urban areas lead to a singular impact on the ecological status, which results in a high threshold. In spite of this peculiar interpretation of the thresholds, the approach is able to describe co-limitations and multiple stressor conditions, as will be shown in the following.

Table 14: Broad river type specific thresholds for the six considered stressor indicators. RT1- very large rivers, RT2 – Lowland brooks, RT3 – Lowland streams and rivers, RT4 – Mountains, siliceous, brooks, RT5 – mountains, siliceous, streams and rivers, RT6 – mountains, calcareous, brooks, RT7 – Mountainous, calcareous, stream and rivers, RT8 – Highland and glacial rivers, RT9 – Mediterranean perennial rivers, RT10 – Mediterranean temporary or very small rivers (brooks).

BRT	lu_urb in %	lu_agr in %	hy_maf in %	hy_basef in %	nu_din in mg/l	nu_tp in mg/l
1	5.17	5.29	12.20	5.10	2.00	0.10
2	6.24	1.9	0.01	0.02	0.42	0.03
3	1.78	24.8	15.50	3.10	2.26	0.34
4	1.9	16.2	3.20	20.00	2.43	0.12
5	1.1	16.2	0.60	1.20	2.80	0.04
6	6.9	27.8	20.00	4.60	2.00	0.05
7	3.3	37.1	1.80	2.20	3.10	0.14
8	2.0	15.5	20.00	4.10	2.07	0.14
9	3.76	12.1	3.80	8.30	2.67	0.12
10	13.4	24.6	5.90	6.30	2.76	0.12

The BBNs were developed using the software Genie (licensed from the University of Pittsburgh), in a simple parallel structure without any co-influences between the stressor indicators (Figure 23). This was possible as statistical analysis revealed no or only very weak correlations between the stressor indicators (Chapter 3.2).

Each of the 10 BRT related datasets containing information on the six stressor indicators and the reported ecological state per FEC was randomly split into 4 subsets of equal size. Two subsets were merged to train and the remaining two were used to validate the BBNs. The combination of subsets was iterated in all possible combinations allowing 4 training-validation cycles. This process showed for all BRT types that the share of correctly predicted ecological status and the explanatory strength of the stressor indicators did not change or depend on the combination of subsets. This allowed finally training the BBN with the complete BRT data set.

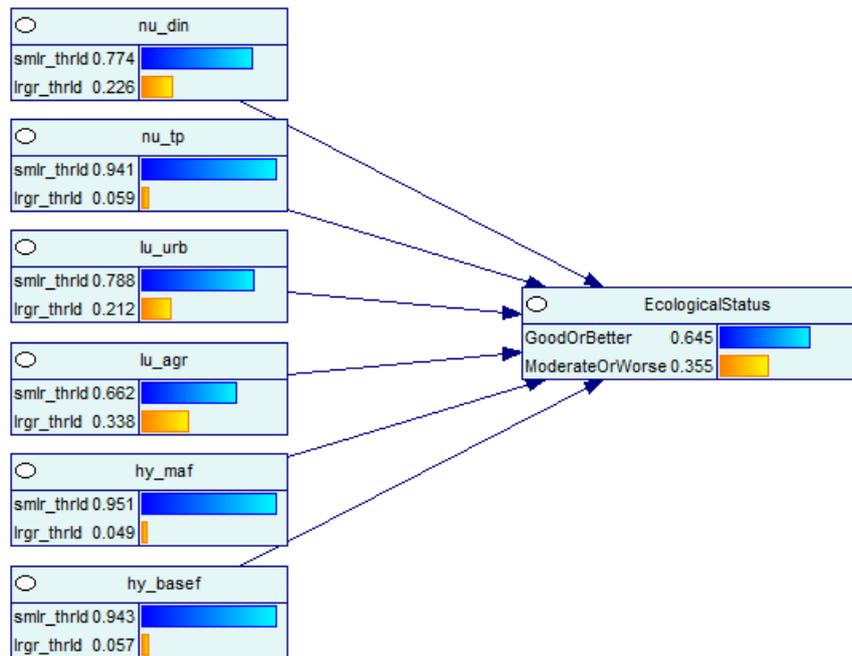


Figure 23: Structure of the Bayesian Belief Network to estimate the probability to reach a good ecological status, shown for the example of broad river type “RT8-Highland and glacial rivers”, using the Software package GENIE.

The trained BBNs using the data subset in mean predicted 69 % of the reported ecological status correctly. The lowest and highest share of correctly predicted ecological status was 61 % (RT4) and 81 % (RT1). For all other RTs these values range between 65% and 70% (Figure-A 2).

Table 15: Share of FECs with allocated river type, ecological status information and the share of FEC with a good or high reported status per broad river type (BRT). ¹⁾ Values in brackets describe the percentage share on all 104,300 FECs of the entire MARS extent. ²⁾ Values in brackets describe the percentage share of all FECs allocated to the respective BRT. ³⁾ Values in brackets describe the percentage share on all FEC with reported status per BRT.

BRT	Number of FECs with allocated river type ¹⁾	FECs with reported status ²⁾	FECs with good or high status ³⁾
1	4045 (4)	2830 (70)	599 (21)
2	16083 (15)	6733 (42)	2147 (32)
3	23065 (22)	15695 (68)	4484 (29)
4	8632 (8)	3721 (43)	1629 (44)
5	7958 (8)	5400 (68)	1947 (36)
6	4976 (5)	2438 (49)	1169 (48)
7	4830 (5)	3327 (69)	1073 (32)
8	11840 (11)	4226 (36)	2744 (65)
9	7017 (7)	4459 (64)	1705 (38)
10	11852 (11)	3409 (29)	1722 (51)
All	100298 (96)	52238 (52)	19219 (37)

The explanatory share of each stressor indicator on the ecological status was tested by comparing the mean euclidean influence of strength calculated by GENIE. This comparison showed a trend of nu_din, nu_tp and hy_basef being the most explanatory stressors for RT1-RT5, whereas lu_urb and lu_agr tend to be stronger explanatory stressors for RT6-RT10. In none in the river types hy_maf was identified as strongest indicator. However, in general the explanatory share of all stressor indicators was quite balanced in all RT, and none of the stressor indicators completely failed in any of the river types. This indicates multi-stressor situations caused or explained by the selected stressor indicators.

Table 16: Mean euclidean strength of influence derived using software „Genie“

BRT	lu_urb	lu_agr	hy_maf	hy_basef	nu_din	nu_tp
1	13.42	22.18	21.80	31.58	30.60	22.99
2	21.97	19.78	18.70	16.77	22.78	22.09
3	14.03	16.30	11.58	16.91	14.57	20.60
4	32.0	23.91	24.73	25.66	27.80	14.84
5	16.70	12.23	14.40	14.34	24.41	23.13
6	29.41	21.63	19.17	22.25	22.19	13.61
7	28.12	25.72	21.65	17.72	22.00	20.32
8	25.99	19.98	20.60	21.41	19.70	24.78
9	23.93	29.79	18.22	23.33	19.70	23.81
10	24.59	28.35	19.23	18.25	14.66	18.03

From six stressors being active or inactive, 64 stressor combinations result. All 64 combinations were found for the FECs in the MARS modelling extent. However, with a highly skewed distribution, the 10 most frequent combinations represent 55 % of all FECs (Figure 24). Interestingly, the two most frequent combinations were found for combinations where only hy_maf is active and no stressor is active (Table 17). This is surprising, as hy_maf is the stressor indicator with the lowest strength of influence, but it also shows that a ubiquitous presence of a stressor indicator does not equal its potential to impact the ecological status. In total the combinations of 1, 2 or 3 jointly acting stressor indicators represent the major share of all FECs. This share significantly decreases for 5 or 6 jointly acting stressors (Figure 24).

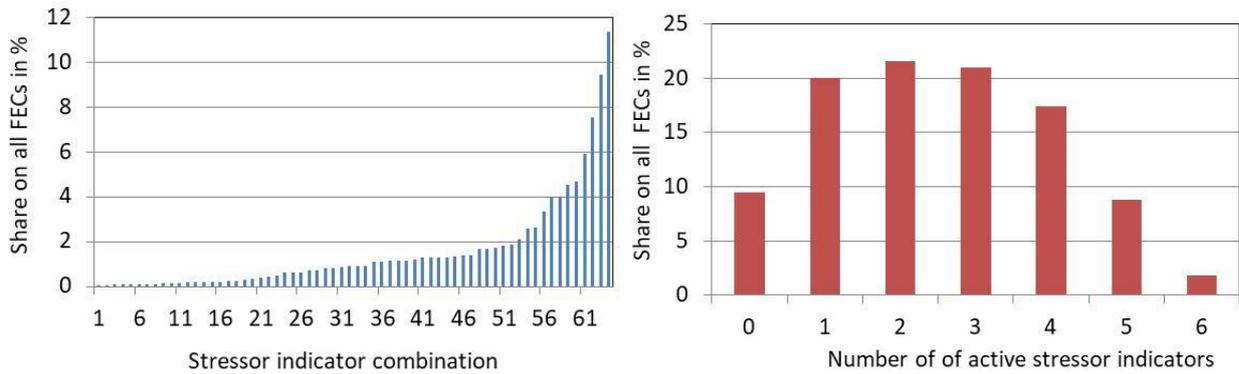


Figure 24: Share of individual stressor indicator combinations per FEC on all FECs in the MARS modelling extent (left). Share of FECs with the same number of active stressor indicators. Both figures refer to current conditions 2001-2010.

Table 17: Ten most frequent stressor indicator combination in the MARS modelling extent. X = active, - = inactive.

No. of FECs	Share on all FECs in %	lu_urb	lu_agr	hy_maf	hy_basef	nu_din	nu_tp
2624	2.6	-	x	x	-	-	x
3354	3.3	x	X	x	-	-	-
4014	4.0	-	X	-	-	X	x
4017	4.0	-	-	x	x	-	-
4548	4.5	-	X	-	-	-	-
4716	4.7	-	X	x	x	x	X
5959	5.9	-	X	x	-	-	-
7546	7.5	-	X	x	-	x	X
9470	9.4	-	-	-	-	-	-
11392	11.4	-	-	x	-	-	-

Table 15 shows the share of FECs for which a good or high ecological status was reported. The trained Bayesian Belief Network provides a very good agreement between the reported status and the mean probability to reach a good or high status ecological (“GoodOrBetter”) calculated for the FECs of a BRT (Figure 25).

For each stressor combination the software “Genie” compares the share of GoodOrBetter or ModeratOrWorse conditions to derive the probabilities. The software Genie only allocates probabilities higher than 50% to a GoodOrBetter ecological status. In particular for BRTs with a low share of FECs with a good or high reported status, the predicted share of GoodOrBetter conditions tends to be underrepresented. Here, only for very specific stressor indicator combinations, likely with few active stressors, a GoodOrBetter status will be dominating. As a consequence for such BRTs the predicted share of FECs with a GoodOrBetter status will underestimate the reported share. This aspect is important for

the interpretation of the predicted probability to reach a good ecological status and suggests only comparing relative changes of the percentage probabilities.

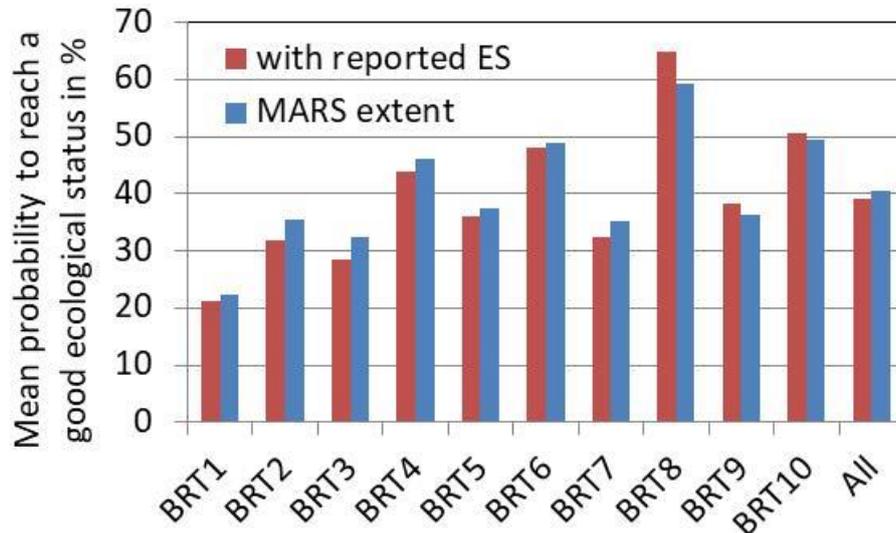


Figure 25: Probability to reach a good ecological status derived from the trained Bayesian Belief Network for the ten different broad River Types

Effects of considered scenarios and management options

The SAT compares current conditions calculated for the years 2001-2010 to the two future scenarios, described as MARS story lines shown in Table 18, for the two periods 2026-2035 and 2056-2065. The calculations of the scenarios are split-up into two principle components: a) exogenous factors (climate, demography, and land-use changes) and b) endogenous factors (e.g. mitigation measures, planned and conducted on local or country level). Consequently, changes in the selected six stressor indicators and in the probability to reach a good or better ecological status are calculated separately for the exogenous factors for both story lines and periods. The determined changes in active and inactive stressors as well as the resulting changes to reach a good ecological status can be visualised and assessed in the SAT. In the following a concluding overview on the general changes (

Table 18) and some exemplary results on changing effects are given.

The mean changes in precipitation and run-off in the entire modelling extent are with less than 10 % relatively small. However a strong spatial pattern (varying throughout the different climate scenarios) can be found, in general indicating, as shown for Storyline 2 in Figure 26, decreasing precipitation in southern Europe and increasing precipitating in central and northern Europe.

Land-use changes are also highly skewed. Whereas the European wide mean change in agricultural land is almost negligible (Table 18), regionally considerable changes of more than $\pm 25\%$ were presumed. Here, a general pattern is less clear, but overall an decrease in agricultural areas for central Europe, a slight increase in southern Europe and only some local increases in northern Europe were derived (Figure 27).

Table 18: Parameters and mean changes considered for the scenario modelling.

	Storyline 1: Techno world	Storyline2: Consensus world
Modelling period	2026-2035 and 2056-2065	
Climate/global change scenario	RCP8.5, SSP5	RCP4.5, SSP2
Precipitation change in %	2030: mean 8 2060: mean: 9	2030: mean 9 2060: mean: 8
Reservoirs Water abstraction	Increase of reservoir area by 0.084 % of land area	
	2030: 2.1 % 2060: 2.2%	2030: 0.7 % 2060: 2.1 %
Land-use change in %	AGR: 2030: -2, 2060: -3 URB: 2030: 1, 2060: 2	AGR: 2030: -3, 2060: -3 URB: 2030: 1, 2060: 1
Population change in %	2030: mean: 10 2060: mean: 13	2030: mean: 6 2060: mean: 9
Sewage collection	2030: 2010+10 %-points 2060: 2010+20 %-points	2030: 2010+10%-points+10 % of diff. 2060: 2010+20%-points+20 % of diff.
Sewage treatment	Current run-off concentrations of treatment plants discharges are reduced by 25 %, unchanged domestic water consumption	

The assumptions for Storyline 2 for an increased sewage collection shows in particular in rural areas an apparent effect and increasing collection rates. Here collection rates may have increased at maximum from current 0 % to 36 %. However, this effect does not necessarily have a strong effect on the nutrient emission and concentrations, as in sparsely populated area the increased collection rate only effects a very limited number of households.

In addition to the changes describe above and in

Table 18, three mitigation measures were applied. Whereas the reduction of N-surplus clearly aims at the reduction of Din concentrations I surface waters, the other to mitigation measures are more relevant for changes in TP concentrations. The proposed mitigation measure to dry and distribute manure leads from current mean European wide N-surplus of 40 kg ha⁻¹yr⁻¹ to a reduction by 40% to a resulting mean N surplus of 24 kg ha⁻¹yr⁻¹. As shown in Figure 28, the reduction again shows a distinct spatial pattern, causing the strongest reduction in areas with currently highest N-surplus.

The resulting changes in DIN concentrations based on global changes and in addition considering the mitigation measures is shown in Figure 29 and Figure 30. In general, increasing precipitation and run-off, decreasing agricultural areas and an improved sewage collection and treatment already causes a general reduction of high nutrient concentrations in favour of an increase in lower concentrations for TN and TP. The additional application of the mitigation measures fosters this development, in particular for nitrogen.

A question, the SAT aims to answer is, how these changes together effect the probability to reach a good ecological status. As the changes of the other considered stressor indicators are more ambiguous in their spatial distribution, direction and extent (see Chapters 3.1 and 3.2) the spatial pattern of active stressor and the joint impact on the ecological status more complex. The interplay of scenario changes is given in Figure 31 on the example of Storyline 2 and for the period 2026-2035. Changed stressor indicators cause increasing probabilities to reach a good ecological status in particular for central Europe, whereas probabilities decrease for southern Europe. When additionally applying the mitigation measures, no further improvement can be found in central Europe, but in many regions of southern Europe probabilities. In southern Europe a further improvement will considerably decrease WWTP effluents and negative effects from an increase in agricultural areas is partly compensated by the implementation of buffer strips. In central Europe the decrease of DIN concentrations does not led to a further increase of probabilities. Three possible causes seem to come together here: 1) other stressors are still active, 2) N-surplus changes are not sufficient reduce nu_din below threshold, 3) the stressor indicator combination has not been sufficiently represented for training the BBN. In particular the last aspect is crucial for the interpretation of the results and requires further research.

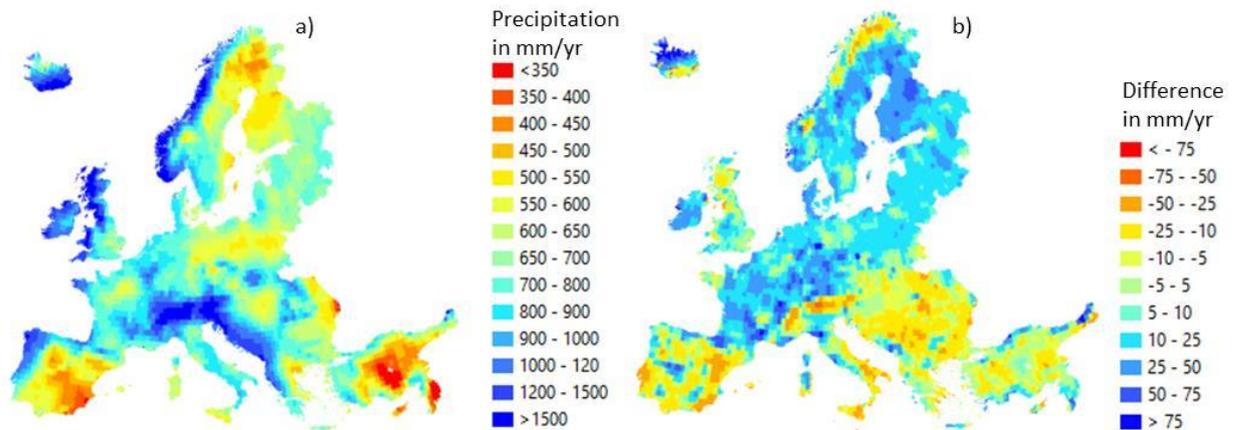


Figure 26: a) mean annual precipitation in 2001-2010, b) difference between the mean annual precipitation for Storyline 2 in 2026-2035 and the mean annual precipitation in 2001-2010.

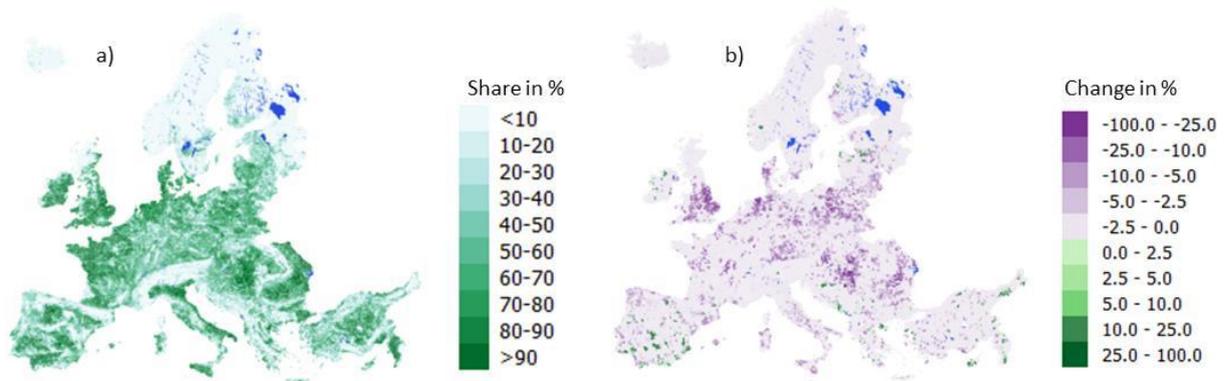


Figure 27: a) Current percentage share of agricultural areas on the FEC area, and b) percentage changes in agricultural areas between current conditions and Storyline 2 in 2026-2035.

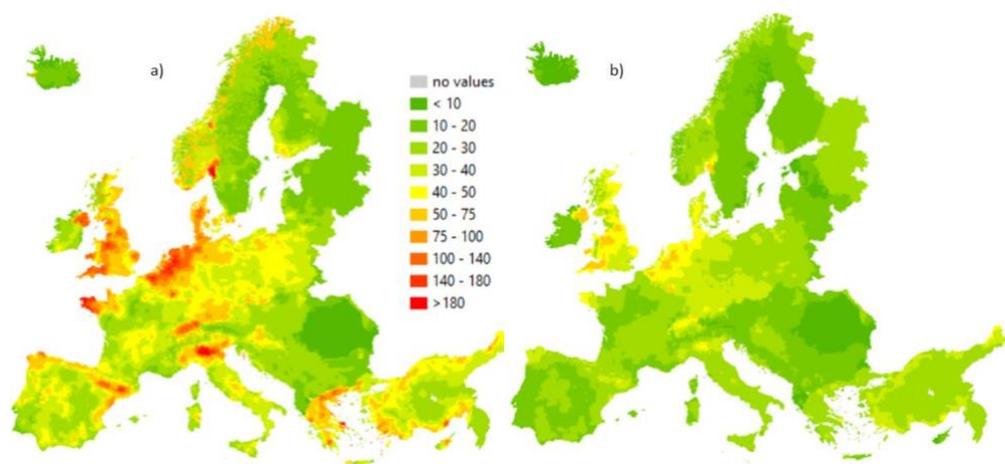


Figure 28: N-surplus on agricultural areas for a) the year 2009 and under b) assumption of applied mitigation measures in $\text{kg ha}^{-1} \text{yr}^{-1}$.

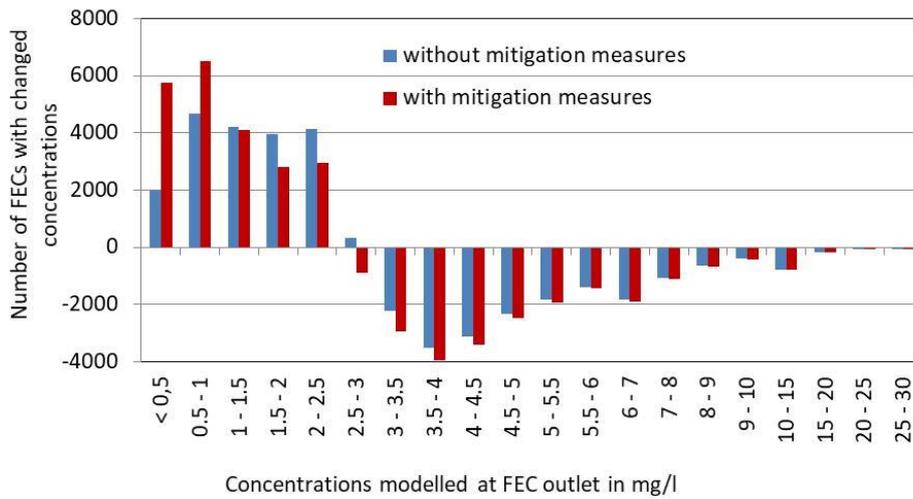


Figure 29: Change of the DIN concentration at FEC outlet for Storyline 1, 2026-2035, with and without mitigation measures, compared to the mean concentration at current conditions, 2001-2010.

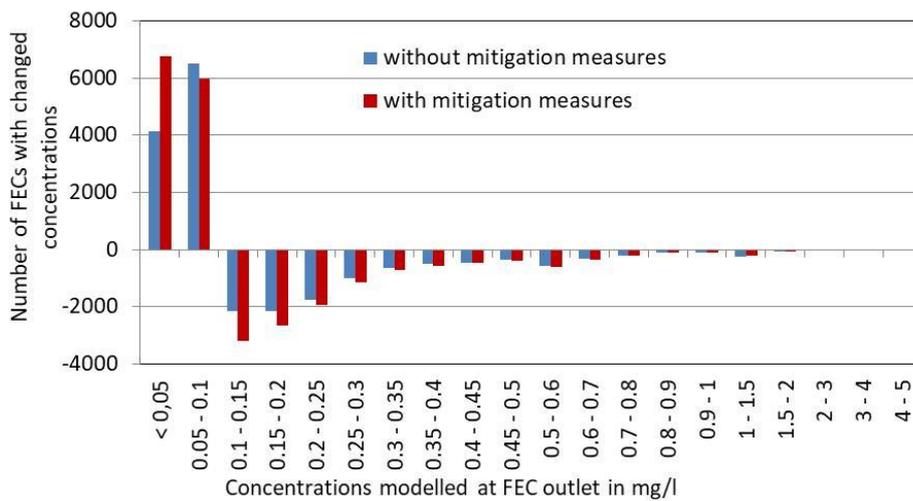


Figure 30: Change of the TP concentration at FEC outlet for Storyline 1, 2026-2035, with and without mitigation measures, compared to the mean concentration at current conditions, 2001-2010.

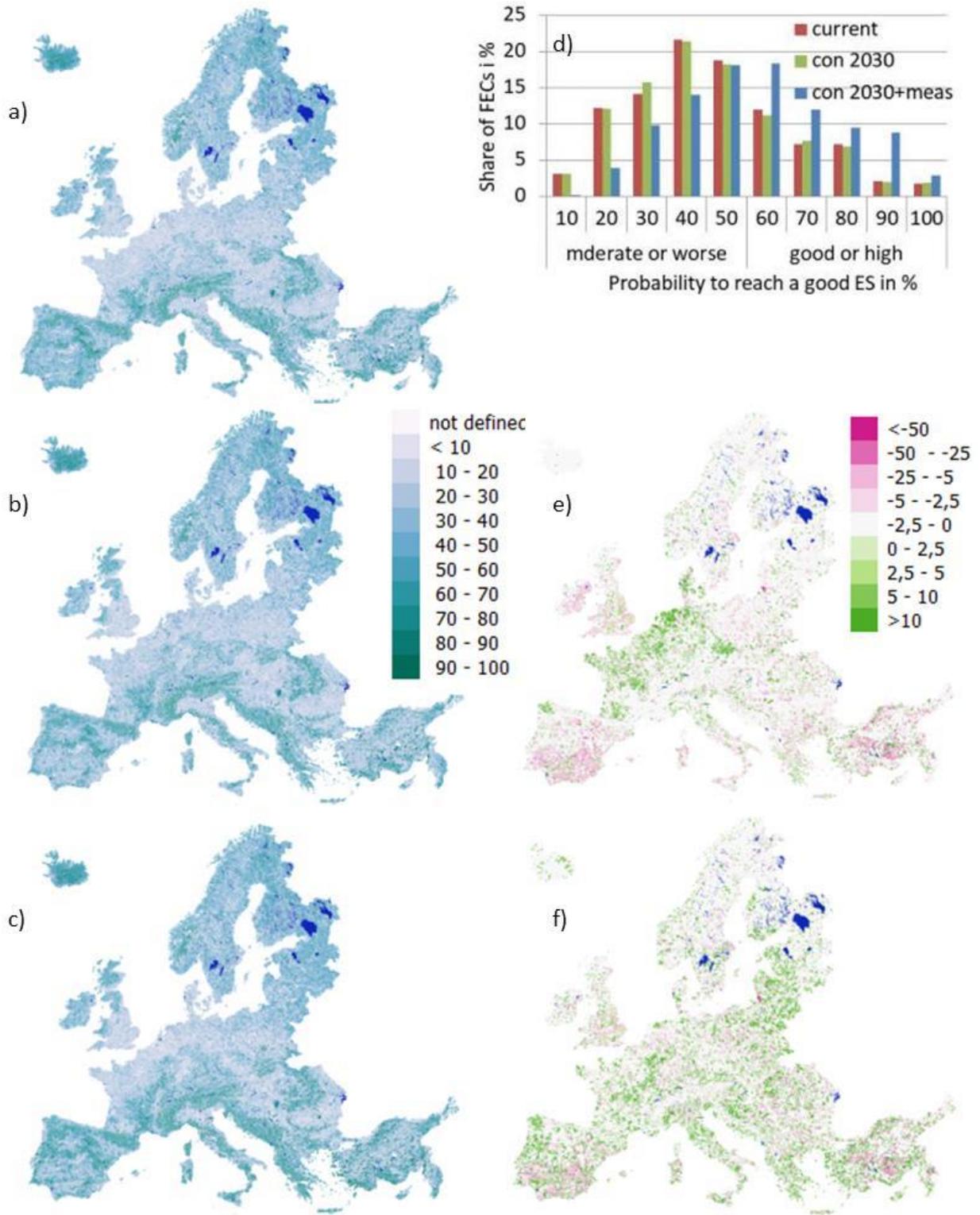


Figure 31: Probability to reach a good ecological status for a) current conditions (2001-2010), b) Storyline 2, 2026-2035, no mitigation measures, and c) Storyline 2, 2026-2035, with mitigation measures. Maps e) and f) show the difference of maps b) minus a) and c) minus a), respectively. Diagram d) displays the calculated probabilities to reach a good ecological state derived from a), b), and c).

Technical set-up

In order to fulfil the EC funding requirements, the scenario analysis tool was developed as open-source software which includes its dependencies. All material published in the SAT are published under a GNU GPLv3 unless otherwise stated (see <https://choosealicense.com/licenses/gpl-3.0/> for details).

The SAT is an interactive web app realised in the programming language R (<https://www.r-project.org/>) with the Shiny library (<http://shiny.rstudio.com>). The tool is expected to work with common operating systems and browsers. The details on diagrams, content, and results are described in the next section.

The user interface contains three core elements, each based on an individual technical approach:

a) Map views

Maps and legends are hosted on a GeoServer (<http://geoserver.igb-berlin.de/geoserver/web/>).

The Java-based GeoServer server software (<http://geoserver.org/>) allows users (**not** the SAT user) to create, edit, and share spatial data. This open-source software supports the Web Map Service (WMS) standard to provide maps in a variety of output formats.

The map visualisation and interaction in the SAT is realized with Leaflet (<http://leafletjs.com/>), an open-source JavaScript library to compose interactive maps. It displays the spatial data and enables the user to zoom and pan the map. According to the current user settings and extent, the SAT downloads the required geometry, styles, and legend from the GeoServer using the WMS capabilities integrated in Leaflet (except for scenario changes).

b) Data analysis and plots

The data analysis and handling in R relies on these packages/libraries (<https://cran.r-project.org/web/packages/>):

- data.table: fast data storage and access
- rgdal: reading geo-spatial data
- ggplot2: creating plots

After reading the spatial and FEC data as well as the thresholds from different files, the SAT stores all data in memory and populates the selection controls for the user interface. User actions trigger events in Shiny which are used to change the current data subset and to update the map view and plot accordingly.

c) Hierarchical multi stress analysis

We adapted the KRONA tool (Ondov et al. 2011), developed to visualize and analyse hierarchies and abundances in meta-genomic data.

The interactive visualisation of multi-stressor conditions in a transparent and easy-to-read way was a major challenge. In order to allow a hierarchical ordering of the extent and combination in

different spatial sub-units, we decided to use the KRONA widget (<https://github.com/marbl/Krona/wiki>), which was originally used in meta-genomics. With KRONA, hierarchical data can be explored as zoomable pie charts. This is realised by a variant of radial, space-filling displays and interactive polar-coordinate zooming. The HTML5 and JavaScript implementation enables fully interactive charts in any modern Web browser. For embedding the KRONA widget in the SAT, the interface between JavaScript, HTML and R-shiny required various adjustments.

How to use

The SAT can be accessed via the freshwater information platform (FIP: www.freshwaterplatform.eu) under Tool/MARS SCENARIO TOOL, using all commonly available browsers.

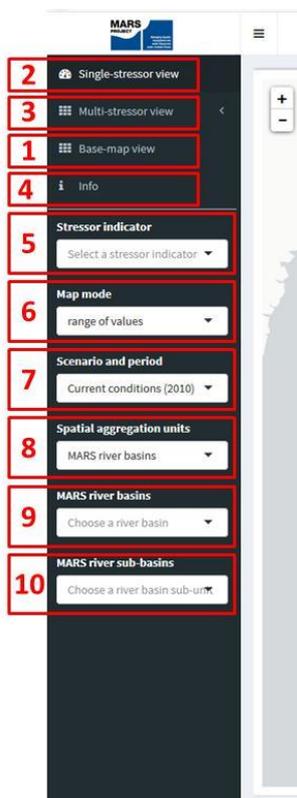
When starting the SAT a small box in the bottom-right corner informs you about the progress in loading maps and background information, the loading can take 10 to 30 seconds depending on the transfer rate of the internet connection. The start window is shown in Figure x, the main elements and their function are described below:



Figure 32: Start window after initial start of the SAT.

- 1) The tool was developed in the context of the FP7 EU-Project MARS. Clicking on the logo will lead you to the project web-site giving more information on the overall goals and results of the project.
- 2) This button will open the navigation menu. Further details are given in Figure 2.
- 3) Clicking on this button will open an URL containing all settings of your current view of the SAT. This URL can be used to share a specific result with colleagues or to save it for a later continuation of the work.
- 4) Main map view. Here all maps are visualised.
- 5) By pressing “+” or “-”, the map will zoom-in and zoom-out. This can also be done by using your mouse wheel, CTRL “+” or CTRL “-“. You can pan the map by clicking and holding the left mouse button.
- 6) You can open the Results panel by clicking on the “v” button. You can pan the Results panel by clicking and holding the left mouse button. Further details are given in Figure qq.
- 7) Short information on the settings in the navigation menu.

After clicking on the 3 bars above the map (2), the navigation menu appears and offers the user many selection and setting options described below:



- 1) Base-map view enables you to have a look on 4 different datasets, each visualized in a map, which are the basis of scenario analysis.
- 2) The default selection of the scenario analysis tool is the single-stressor view. By selecting a single stressor in dropdown-menu (5) the respective map and results are depicted.
- 3) Click on Multi-stressor view and choose between »Map«, that leads you to the visualized number of active stressors and the probability of good ecological status, or »Explorer«, that opens the Krona tool which presents you the data in multi-layered pie charts (see Figure 31). You can use the multi-stressor »map« view in the same way like the single-stressor view.
- 4) The »Info« menu option opens an information page on the tool. Here you can access a manual and background information on the SAT.

Figure 33: The navigation menu.

- 5) The »Info« menu option opens an information page on the tool. Here you can access a manual and background information on the SAT.
- 6) This dropdown menu allows you to select one of 6 different stressor indicators.
- 7) By selecting a »map mode« you can switch two different visualization types. »Range of values« is the default setting that shows you the selected stressor indicator in its corresponding units. In contrast to this you can switch to »stressor status« that just shows you FECs in which the selected stressor is active or not active depending on its defined threshold.
- 8) »Scenario and period« allow you to decide whether you want to depict the selected stressor indicator in current conditions based on the year 2010 or in the two scenarios consensus world and techno world in the years 2030 and 2060.
- 9) By clicking on »Spatial aggregation units« you can change how the results are displayed - on a geographical or political level. MARS river basins or countries are available for selection.
- 10) Here you can choose a river basin / country by using the dropdown list or typing in the name of the basin / county.
- 11) If you choose river basins as spatial aggregation unit (8) and select a river basin (9) you can zoom further in river sub-basins by using the dropdown list or typing in the name of the sub-basin.

Single- and multi-stressor »map« view allows you to get detailed information about the data that is depicted in the map. To open the Results panel, click on the arrow to unfold it.

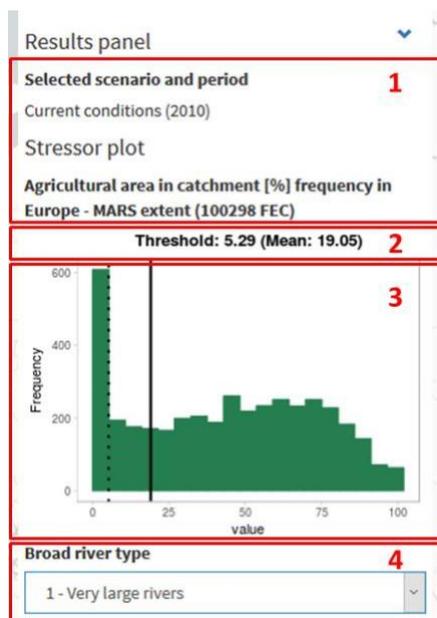


Figure 34: The results panel.

- 1) The first part of the Results panel summarizes your current selection and describes the plot below including the number of analysed FECs (MARS extent).
- 2) In the second part you receive information about the threshold / mean threshold of all broad river types that must not be exceeded in order to achieve a good ecological status.
- 3) The plot in the third part of the results panel displays the frequency of FECs and its value ranges in a column diagram. The vertical line in the plot shows the threshold and the dotted vertical line shows the mean threshold of all broad river types.

- You can select one of 10 broad river types of your current selection to have a closer look on this broad river type. By selecting a broad river type, the threshold and the plot are adjusted.

Multi-stressor view »Explorer« / Krona tool

The Krona tool (see Ondov et al. 2011) allows hierarchical data to be explored with zooming in multi-layered pie charts. Beside the »map« view it is the second possibility to explore MARS river basins or countries and their proportion of stressor indicators under current or future conditions.

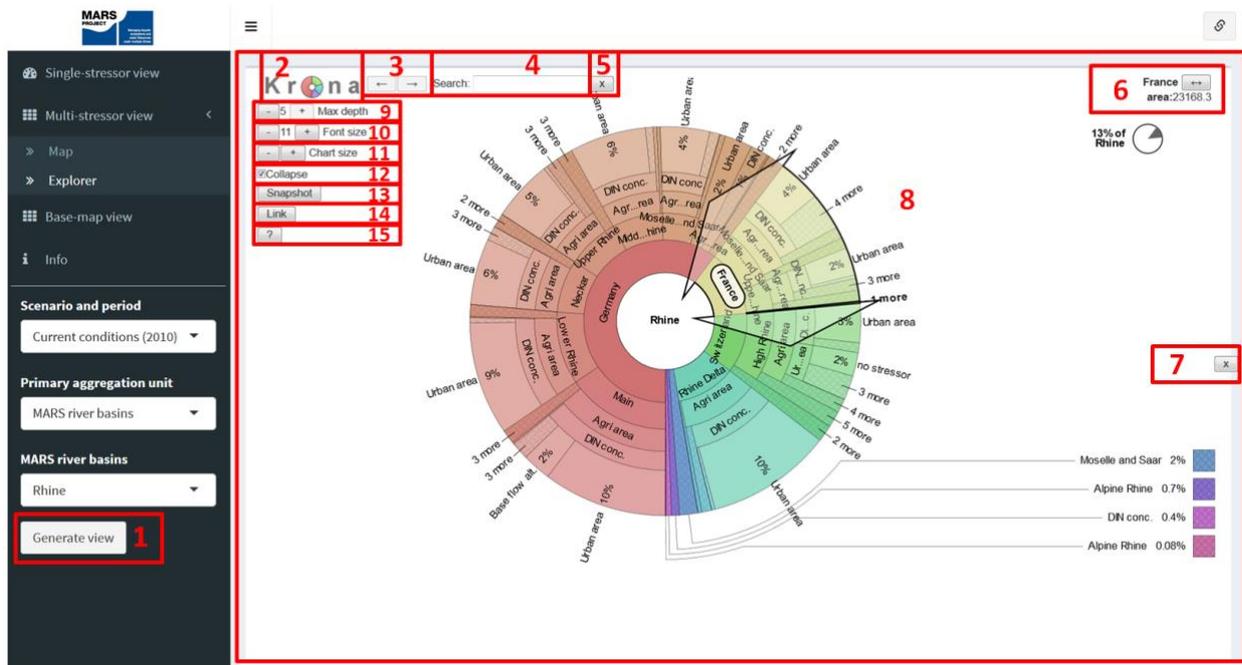


Figure 31: The Multi-stressor view »Explorer« / Krona tool.

- The multi-layered pie-charts are created by clicking on »Generate view«. This requires that you select previously the scenario and period, the primary aggregation unit and a specific MARS river basin or country.
- The Krona tool was developed by Ondov et al. 2011 and is available open source. By clicking on the logo you will be forwarded to the tool website and source code.
- The navigation arrows allow you to undo a selection ← or to restore a selection → inside the tool.
- Here you can search for specific words or terms used in the tool. Wedges that contain hidden matches will also be highlighted.
- The »X« button deletes your search.
- By clicking on a wedge inside of the pie charts, you can unfold your selected wedge. The same applies for double-clicking on a wedge.

- 7) Some MARS river basins or countries include too many wedges for a complete labelling. In this case a supporting legend appears that you can hide by clicking on the »x« button.
- 8) Main Krona view. Here all stressors are depicted in multi-layered pie charts. You can unfold wedges on different levels by double-clicking on a wedge. To zoom in on a wedge briefly and then return to the current view, click and hold on a wedge.
- 9) You can reduce the number of displayed layer by clicking on the »-« button or increase the maximum depth by clicking on the »+« button.
- 10) Further customise your view by change the font size. Smaller font sizes (- button) will be harder to read than larger font sizes (+ button), but will allow smaller wedges to be shown and can reduce label crowding.
- 11) Improve the legibility by changing the size of the pie charts.
- 12) You can use this checkbox to simplify the chart by collapsing "redundant" wedges that are entirely composed of another wedge.
- 13) Click on the Snapshot button to open a new window with the current view rendered in SVG (Scalable Vector Graphics) format.
- 14) This shows a link to the current view that can be copied for bookmarking or sharing.
- 15) The Help button opens the developers' general manual of the Krona tool in a new tab / window.

5. GENERAL CONCLUSIONS

Lead: Markus Venohr, Lidija Globevnik, Yiannis Panagopoulos, Sebastian Birk

Data and models

- 1) A first European wide (EU, EFTA states and neighbouring countries) joint model application on a spatial resolution of 58 km² (ECRINS FECs, average size) and for monthly time steps build the core of the developed scenario analysis tool.
- 2) Modelling of near natural and current run-off conditions allowed deriving the alteration of hydrological parameters.
- 3) Current condition (2001 – 2010), two storylines (based on RCP8.5, SSP5 and RCP4.5, SSP2) and two future periods (2026 – 2035, 2056 – 2065) constitute reference conditions for the scenario frameworks to assess the spatial distribution, extent and combination of stressor indicators on FEC level.
- 4) Additional to the future scenario frameworks a set of three mitigations measures (N-surplus reduction, riparian buffer strips, and improve waste water treatment) were applied and assessed.
- 5) A candidate list of more than 25 stressor indicators (Table-A 4) were tested regarding their explanatory power to describe the ecological status reported by the EU countries for the 2nd EU Water Framework Directive management cycle.
- 6) Six selected stressor indicators (DIN and TP concentration in rivers, share of urban and agricultural land, mean annual flow alteration and base flow alteration) were shown to explain the majority of the reported ecological status.
- 7) For all selected stressor indicators river type-specific thresholds identifying when the ecological status most probably deteriorates from good to moderate were derived, using machine learning techniques.
- 8) Bayesian Belief Networks (BBN) were trained to estimate the probability to reach good ecological status under given river type-specific stressor combinations.
- 9) An online browser-based tool was developed to visualise and analyse all stressor conditions and their probabilities under current and future conditions.

Key results

- 1) Results of both applied models, PCR-GlobWB and MONERIS, show a reliable spatial and temporal distribution in water quantity and water quality, respectively. However, at smaller scales and for individual monitoring stations high deviations to observed run-off and nutrient loads were found, which suggests to restrict the interpretation of model results to catchments larger 1,000 km².

- 2) Hydrologic stress of European rivers can be comprehensively expressed through the calculation of the ratio (alteration) of hydrologic indicators derived from time-series of daily river discharge occurring in a certain scenario and in a near-natural scenario without water abstractions, and time-series of discharge occurring in future scenarios of existing water abstractions to cover needs.
- 3) Mapping the hydrologic alteration on Europe's geographic background, trends in hydrologic stress were indicated across a north-south gradient.
- 4) From the hydrologic indicators calculated, the mean annual flow and base flow index were the most informative showing that Southern Europe is always more hydrologically stressed than the rest of Europe. Specifically, water availability in rivers decreases in large parts of the Mediterranean countries, mostly because of water abstractions for agriculture. This is the case both in the baseline scenario and in future scenarios.
- 5) In the rest of Europe, especially the Northern part, temporal flow variations are much less pronounced, and the natural hydrologic conditions are preserved in the anthropogenic (baseline) scenario. However, future scenarios indicate positive or negative changes from the natural conditions which seem to alternate even among neighbouring FECs. The combination of future climate and water management governs the direction of change in each FEC. However, these local-based rules do not seem to change from 2030 to 2060.
- 6) Both storylines show that, on average, river discharges are likely to increase compared to the current situation, being the highest increase during the high flow season, and the lowest increase during the low flow season. However, for several rivers discharge during the low flow season will decrease for both storylines.
- 7) The differences between the two simulated storylines are smaller than the differences between the current and the future situation.
- 8) Nutrient emission modelled to calculate in-stream nutrient concentrations originate to 59 % (TN) from agricultural areas and to 39 % (TP) from urban areas and to 35 % (TP) from arable land. In total, 20 % of area contribute 50 % (TN) and 58 % (TP) of the total nutrient emissions.
- 9) In 49 % of the FECs at least three stressors act jointly, only for 9 % of the FECs no active stressor was calculated. Ten out of 64 stressor indicator combinations represent conditions in 55 % of all FECs.
- 10) In 50 % (nu_din) and 40 % (nu_tp) of the FEC thresholds for nutrient concentrations are exceeded.
- 11) In 26 % (lu_urb) and 60 % (lu_agr) of the FEC thresholds for urban and agricultural land use are exceeded, respectively.

- 12) In 23 % (hy_maf) and 15 % (hy_basef) of the FEC thresholds for the hydrological indicators are exceed.
- 13) Nutrient concentrations and land use were found to be the stressor indicators contributing the strongest influence on the derived status using the BBN.
- 14) Based on the reported ecological status, for 32 % of the FECs a high or good ecological status was derived. With an agreement of 69 % this share was reproduced by the BBN.
- 15) For all scenarios without mitigation measures the mean probability to reach a good ecological status decreases by 3% to 5% on average, with distinct spatial differences. Here, for Broad River Types 1 to 5 and 10 average decreasing probabilities of up to -3 %-points were derived, whereas for Broad River Types 6 to 9 average increasing probabilities of up to 6 %-points were calculated. These mean type-specific values are partly much higher for the individual FEC belonging to a Broad River Type.
- 16) In principle, the proposed mitigation measures help to foster positive trends and to curb negative effects caused by global change. When derived as a Broad River Type mean, the effect of global change on the probabilities was in general stronger than such of the mitigation measures. Assessed on a FEC level it strongly depends on the local conditions and can hardly be generalised.

Constraints

- 1) Historical and future discharges are simulated with the global hydrological model PCR-GLOBWB. Global hydrological models have a relatively coarse resolution compared to catchment-specific models and therefore the degree of aggregation of local processes is relatively high. This can result in a lower performance that could have been reached using local catchment scale hydrological models, yet for a European scale assessment the homogeneity that can be reached using a single model is very valuable.
- 2) River flows simulated by the global hydrological model show deviations from observed flows due to simplified approximations of water use / management and aggregation of hydrological processes. Basins with more information and wetter conditions are simulated better than basins with less information and dryer conditions. The projected directions of change are more reliable.
- 3) Due to computational constrains of the full modelling chain, data from only one GCM has been considered providing a possible future. Uncertainties within GCMs and consequently differences between GCMs can be large (Sperna Weiland et al., 2012) and this may have influenced the magnitude of the projected hydrological changes considered in this tool.
- 4) Collection and treatment of waste water is based on assumptions and discharge concentrations had to be estimated for many waste water treatment plants. A more comprehensive data set

would help quantifying emission from this important pathway. Further, more reliable complete and differentiated input data for tile drained areas, atmospheric P deposition or already implemented buffer strips would be valuable input data for this modelling task.

- 5) The BBN showed the tendency to underestimate high or good ecological status for broad river types already featuring a high share of FECs in moderate or worse conditions. A normal distribution of reported status by defining new sub-groups could help to derive more consistent or stable probability tables.
- 6) A further in-depth analysis of the limited effect of mitigation measures on the probability tables could help identifying most effective measure combinations in hot-spot regions.

REFERENCES

- ASTER G-DEM: ASTER Global Digital Elevation Model via <http://reverb.echo.nasa.gov/reverb/> (last accessed: 12.03.2018).
- Beck, H. E., A. I. J. M. van Dijk, A. de Roo (2015): Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydrometeorology* 16(4), 1478–1501.
- Beck, H. E., A. I. J. M. van Dijk, D. G. Miralles, R. A. M. de Jeu, L. A. Bruijnzeel, T. R. McVicar, J. Schellekens (2013): Global patterns in baseflow index and recession based on streamflow observations from 3394 catchments. *Water Resources Research* 49(12), 7843–7863.
- Behrendt, H., R. Dannowski, D. Deumlich et al. (2002): Investigation on the quantity of diffuse entries in the rivers of the catchment area of the Odra and the Pomeranian Bay to develop decision facilities for an integrated approach on waters protection (Phase III). Institute of Freshwater Ecology and Inland Fisheries, Berlin, Final report, 271.
- Carl, P., K. Gerlinger, K. K. Hattermann, V. Krysanova, C. Schilling, H. Behrendt (2008): Regularity-based functional streamflow disaggregation: II. Extended demonstration. *Water Resources Research* 44, W03426.
- CLC (2012): Corine Land Cover, Version 18.5.1 via <http://land.copernicus.eu/pan-european/corine-land-cover/clc-2012> (last accessed: 12.03.2018).
- Copernicus (2015): Riparian Zones Delineation via <http://land.copernicus.eu/local/riparian-zones/riparian-zones-delineation/view> (last accessed: 12.03.2018).
- Dee, D. P., S. M. Uppala (2009): Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Quarterly Journal of the Royal Meteorological Society* 135, 1830-1841.
- Dunne, J. P., J. G. John, E. Shevliakova, R. J. Stouffer et al. (2013): GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part II: Carbon System Formulation and Baseline Simulation Characteristics. *Journal of Climates* 26, 2247–2267.
- Crouzet, P., W. Simonazzi (2012): EEA Catchments and Rivers Network System - ECRINS v 1.1. Technical report, European Environment Agency, Copenhagen, Technical Report No 7/2012.
- Esch, S. V. D., B. ten Brink, E. Stehfest, M. Bakkenes, A. Sewell, A. Bouwman, J. Meijer, H. Westhoek, M. V. D. Berg (2017): Exploring future changes in land use and land condition and the impacts on food, water, climate change and biodiversity: Scenarios for the UNCCD Global Land Outlook. PBL, The Hague, Policy Report No 2076, 116.
- EMEP MSC-W (2014): Meteorological Synthesizing Centre – West modelled air concentrations and depositions via http://webdab.emep.int/Unified_Model_Results/ (last accessed: 12.03.2018).
- ESDB - The European Soil Database distribution version 2.0. European Commission and the European Soil Bureau Network, CD-ROM, EUR 19945 EN, 2004.
- ETC/ICM (2015): European Freshwater Ecosystem Assessment: Cross-walk between the Water Framework Directive and Habitats Directive types, status and pressures. ETC/ICM, Magdeburg, Technical Report No 2/2015, 95.
- EU-DEM: Digital Elevation Model over Europe via <http://www.eea.europa.eu/data-and-maps/data/eu-dem-tab-european-data> (last accessed: 12.03.2018).
- EUMETSAT CM SAF (2016): Meteorological Satellite Second Generation (MSG): The Satellite Application Facility on Climate Monitoring via <http://www.cmsaf.eu> (last accessed: 12.03.2018).
- EUROSTAT (2016): Ihr Schlüssel zur europäischen Statistik via <http://ec.europa.eu/eurostat/de/data/database> (last accessed: 15.09.2016).
- Erickson, T. R., H. G. Stefan (1996): Correlation of Oklahoma Stream Temperatures with Air Temperatures. Project Report No 398, University of Minnesota, St Anthony Falls Laboratory: Minneapolis.
- Erickson, T. R., H. G. Stefan (2000): Linear air/water temperature correlations for streams during open water periods. *Journal of Hydrologic Engineering, American Society of Civil Engineers* 5(3), 317–321.

- Feick, S., S. Seibert, P. Döll (2005): A Digital Global Map of Artificially Drained Agricultural Areas. Frankfurt Hydrology Paper 4, Frankfurt, Germany.
- Fiedler, K., P. Döll (2007): Global modelling continental water storage changes – sensitivity to different climate datasets. *Advances Geologic Sciences* 11, 63–68.
- Fischer, P., R. Pöthig, B. Gücker, M. Venohr (2016): Estimation of the degree of soil P saturation from Brazilian Mehlich-1 P data and field investigations on P losses from agricultural sites in Minas Gerais. *Water Science and Technology* 74(3), 691-697.
- Gadegast, M., Hirt, U., Venohr, M. (2014): Changes in Waste Water Disposal for Central European River Catchments and Its Nutrient Impacts on Surface Waters for the Period 1878–1939. *Water, Air and Soil Pollution* 225, 1914.
- Gao, Y., R. M. Vogel, C. N. Kroll, N. L. R. Poff, J. D. Olden (2009): Development of representative indicators of hydrologic alteration. *Journal of Hydrology* 374, 136-147.
- Gericke, A. (2015): Soil loss estimation and empirical relationships for sediment delivery ratios of European river catchments. *International Journal of River Basin Management* 13, 179-202.
- GISCO/Eurostat (2006): Gridded population data for 2006 via <http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/population-distribution-demography/geostat> (last accessed: 12.03.2018).
- GlobCorine (2010): ESA 2010 and UCLouvain via <https://doi.pangaea.de/10.1594/PANGAEA.778363?format=html#download> (last accessed: 12.03.2018).
- GRDC-EWA (2017): European Water Archive (EWA) of the Global Runoff Data Centre (GRDC) via http://www.bafg.de/GRDC/EN/04_spcldtbss/42_EWA/ewa_node.html;jsessionid=A6B4EF1AAEC50C900C5D3CC2C74ADBB2.live21304 (last accessed: 12.03.2018).
- Hagemann, S., K. Arpe, L. Bengtsson (2005): Validation of the hydrological cycle of ERA-40. Max Planck Institute, Hamburg, Germany, Reports on Earth System Science No 24.
- Harris, I., P. D. Jones, T. J. Osborn, D. H. Lister (2014): Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *International Journal of Climatology* 34, 623–642.
- Heidecke, C., U. Hirt, P. Kreins, P. Kuhr, R. Kunkel, J. Mahnkopf, M. Schott, B. Tetzlaff, M. Venohr, A. Wagner, W. Wendland (2015): Endbericht zum Forschungsprojekt "Entwicklung eines Instrumentes für ein flussgebietsweites Nährstoffmanagement in der Flussgebietseinheit Weser" AGRUM+-Weser. Johann Heinrich von Thünen-Institut, Braunschweig, Germany, Thünen Report No 21, 380.
- Hirt, U., A. Wetzig, M. Venohr (2009): How does tile drainage react on precipitation? Results of landscapes in the Northern Hemisphere. Leibnitz-Institute of Freshwater Ecology and Inland Fisheries, ICID Conference Proceedings, 1-6.
- Hirt, U., J. Mahnkopf, M. Gadegast, L. Czudowski, U. Mischke, C. Heidecke, G. Schernewski, M. Venohr (2013): Reference conditions for rivers of the German Baltic Sea catchment: Reconstructing nutrient regimes using the model MONERIS. *Regional Environmental Change* 14, 1123-1138.
- HWSD - Harmonized World Soil database version 1.2 via <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/> (last accessed: 12.03.2018).
- IHME 1500 - International Hydrogeological Map of Europe via https://www.bgr.bund.de/EN/Themen/Wasser/Projekte/laufend/Beratung/Ihme1500/ihme1500_projektbeschr_en.html (last accessed: 12.03.2018).
- LUCAS 2009 TOPSOIL data, European Commission, Joint Research Centre via <http://esdac.jrc.ec.europa.eu/> (last accessed: 12.03.2018).
- Karssenber, D., O. Schmitz, P. Salamon, K. de Jong, M. F. Bierkens (2010): A software framework for construction of process-based stochastic spatio-temporal models and data assimilation. *Environmental Modelling and Software* 25, 489–502.
- López López, P., N. Wanders, J. Schellekens, L. J. Renzullo, E. H. Sutanudjaja, M. F. P. Bierkens (2016): Improved large-scale hydrological modelling through the assimilation of streamflow and

- downscaled satellite soil moisture observations. *Hydrology and Earth System Science* 20, 3059-3076.
- MARS Deliverable 5.1 (2017): Five Reports on stressor classification and effects at the European scale. <http://mars-project.eu/index.php/deliverables.html>
- MARS geo-database, MARS Deliverable 5.1.1. (2017): <http://fgg-web.fgg.uni-lj.si/~mars/>
- Mitchell, T. D., P. D. Jones (2005): An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology* 25, 693–712.
- Mohseni, O., H. G. Stefan (1999): Stream temperature/air temperature relationship: A physical interpretation. *Journal of Hydrology* 218, 128–141.
- Nearing, M. A. (1997): A single, continuous function for slope steepness influence on soil loss. *Soil Science Society of America Journal* 61, 917-919.
- Chawla, N. V., K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer (2002): Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research* 16, 321-357.
- Ondov, B. D., N. H. Bergman, A. M. Phillippy (2011): Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics* 12(1), 385.
- O'Neill, B. C., E. Kriegler, K. Riahi, K. L. Ebi, S. Hallegatte, T. R. Carter, R. Mathur, D. P. van Vuuren (2014): A new scenario framework for climate change research: The concept of shared socioeconomic pathways. *Climatic Change* 122 (2), 387–400.
- Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegard, B. D. Richter, R. E. Sparks, J. C. Stromberg (1997): The natural flow regime. A paradigm for river conservation and restoration. *BioScience* 47, 769-784.
- Portmann, F. T., S. Siebert, P. Döll (2010): MIRCA2000—global monthly irrigated and rainfed crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological modeling. *Global Biogeochemical Cycles* 24, GB1011.
- Pöthig, R., H. Behrendt, D. Opitz, G. Furrer (2010): A universal method to assess the potential of phosphorus loss from soil to aquatic ecosystems. *Environmental Science and Pollution Research* 17(2), 497–504.
- Quinlan, R. (1993): C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Francisco, USA.
- Richter, B. D., J. V. Baumgartner, J. Powell, D. P. Braun (1996): A method for assessing hydrological alteration within ecosystems. *Conservation Biology* 10, 1163-1174.
- Richter, B. D., J. V. Baumgartner, R. Wigington, D. P. Braun (1997): How much water does a river need? *Freshwater Biology* 37, 231-249.
- Richter, B. D., J. V. Baumgartner, D. P. Braun, J. Powell (1998): A spatial assessment of hydrologic alteration within a river network. *Regulated Rivers: Research and Management* 14, 329-340.
- Sharma, S., J. Agrawal, S. Sharma (2013): Classification Through Machine Learning Technique: C4.5 Algorithm based on Various Entropies. *International Journal of Computer Applications* 82, 20-27.
- Sperna Weiland, F.C., L. P. H. van Beek, J. C. J. Kwadijk, M. F. P. Bierkens (2010): The ability of a GCM-forced hydrological model to reproduce global discharge variability. *Hydrology and Earth System Sciences* 14, 1595-1621.
- Sperna Weiland, F. C., J. A. Vrugt, L. P. H. van Beek, A. H. Weerts, M. F. P. Bierkens (2015): Significant uncertainty in global scale hydrological modeling from precipitation data errors. *Journal of Hydrology* 529, 1095-1115.
- Stehfest, E., D. P. van Vuuren, L. Bouwman, T. Kram, R. Alkemade, M. Bakkenes, H. Biemans, A. Bouwman, M. den Elzen, J. Janse, P. Lucas, J. van Minnen, C. Müller, A. Prins (2014): Integrated Assessment of Global Environmental Change with Model description and policy applications IMAGE 3.0. PBL, The Hague, Policy Report No 735, 336.

- Sutanudjaja, E. H., L. P. van Beek, Y. Wada, D. Wisser, I. E. de Graaf, M. W. Straatsma, M. F. Bierkens (2014): Development and validation of PCR-GLOBWB 2.0: A 5 arc min resolution global hydrology and water resources model. EGU General Assembly 16, 9993.
- Sutanudjaja, E., L. Van Beek, S. de Jong, F. Van Geer, M. Bierkens (2014): Calibrating a large-extent high-resolution coupled groundwater-land surface model using soil moisture and discharge data. *Water Resources Research* 50, 687–705.
- The Nature Conservancy (2009): Indicators of Hydrologic Alteration, Version 7.1, User's Manual.
- Troccoli, A., P. Kallberg (2004): Precipitation correction in the ERA40 reanalysis. ECMWF, Reading, UK, ERA-40 Project Report Series 13, 06.
- Uppala, S. M., P. W. Kallberg, A. J. Simmons et al. (2006): The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society* 131, 2961–3012.
- Vadas, P. A., P. J. A. Kleinman, A. N. Sharpley, B. L. Turner (2005): Relating soil phosphorus to dissolved phosphorus in runoff: A single extraction coefficient for water quality modeling. *Journal of Environmental Quality* 34(2), 572-580.
- Van Beek, L. P. H., M. F. P. Bierkens (2009): The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification. Department of Physical Geography, Utrecht University, the Netherlands, Technical Report. available at <http://vanbeek.geo.uu.nl/suppinfo/vanbeekbierkens2009.pdf> (last access: 28 February 2011).
- Van Beek, L. P. H., Y. Wada, M. F. P. Bierkens (2011): Global monthly water stress: I. Water balance and water availability. *Water Resources Research* 47, W07517.
- Venohr, M., J. Mahnkopf, A. Gericke (2018): Distributed nitrogen surplus derived from European national statistics. *Hydrological Processes*, submitted.
- Venohr, M., M. Gadegast (2018): Modelling nutrient emissions and resulting concentrations for the period around 1880s in Central Europe. In preparation.
- Venohr, M., U. Hirt, J. Hofmann, D. Opitz, A. Gericke, A. Wetzig, S. Natho, F. Neumann, J. Hürdler, M. Matranga, J. Mahnkopf, M. Gadegast, H. Behrendt (2011): Modelling of Nutrient Emissions in River Systems - MONERIS - Methods and Background. *International Review of Hydrobiology* 96(5), 435-483.
- Venohr, M. (2006): Modellierung der Einflüsse von Temperatur, Abfluss und Hydromorphologie auf die Stickstoffretention in Flusssystemen. *Weißensee Verlag, Berliner Beiträge zur Ökologie* 14, 193.
- Venohr, M., I. Donohue, S. Fogelberg, K. Irvine, H. Behrendt (2005): Nitrogen retention in a river system under consideration of the river morphology and occurrence of lakes. *Water Science and Technology* 51(3-4), 19-29.
- Vörösmarty, C. J., B. M. Fekete, M. Meybeck, R. Lammers (2000): A simulated topological network representing the global system of rivers at 30 min spatial resolution (STN-30). *Global Biogeochemical Cycles* 14, 599–621.
- van Vuuren, D. P., E. Kriegler, B. C. O'Neill, K. L. Ebi, K. Riahi, T. R. Carter, J. Edmonds, S. Hallegatte, T. Kram, R. Mathur, H. Winkler (2013): A new scenario framework for Climate Change Research: Scenario matrix architecture. *Climatic Change* 122, 373-386.
- Wada, Y., M. F. P. Bierkens (2014): Sustainability of global water use: Past reconstruction and future projections. *Environmental Resource Letters* 9, 10.
- Webb, B. W., F. Nobilis (1997): A long-term perspective on the nature of the air–water temperature relationship: A case study. *Hydrological Processes* 11, 137–147.
- Webb, B. W., D. E. Walling (1993): Longer-term water temperature behaviour in an upland stream. *Hydrological Processes* 7, 19–32.
- Webb, B. W., P. D. Clack, D. E. Walling (2003): Water-air temperature relationships in a Devon river system and the role of flow. *Hydrological Processes* 17, 3069–3084.
- Weedon, G. P., S. Gomes, P. Viterbo, W. J. Shuttleworth, E. Blyth, H. Österle, J. C. Adam, N. Bellouin, O. Boucher, M. Best (2011): Creation of the WATCH forcing data and its use to assess global and

regional reference crop evaporation over land during the twentieth century. *Journal of Hydrometeorology* 12(5), 823-848.

Wesseling, C. G., D. J. Karssenbergh, P. A. Burrough, W. Deursen (1996): Integrating dynamic environmental models in GIS: The development of a dynamic modelling language. *Transactions in GIS* 1, 40–48.

Wilke, B., D. Schaub (1996): Phosphatanreicherung bei Bodenerosion. *Mitteilungen DBG* 79, 435-438.

Zarfl, C., A. E. Lumsdon, J. Berlekamp, L. Tydecks, K. Tockner (2015): A global boom in hydropower dam construction. *Aquatic Sciences* 77, 161.

APPENDIX

Table-A 1: Land use classification used within MONERIS.

Dataset Original	Class Original	Code Original	Code MONERIS	Class MONERIS
Corine	Continuous urban fabric	111	1	Urban Area
Corine	Discontinuous urban fabric	112	1	Urban Area
Corine	Industrial or commercial units	121	1	Urban Area
Corine	Road and rail networks and associated land	122	1	Urban Area
Corine	Port areas	123	1	Urban Area
Corine	Airports	124	1	Urban Area
Corine	Mineral extraction sites	131	2	Open pit mine
Corine	Dump sites	132	2	Open pit mine
Corine	Construction sites	133	2	Open pit mine
Corine	Green urban areas	141	1	Urban Area
Corine	Sport and leisure facilities	142	1	Urban Area
Corine	Non-irrigated arable land	211	3	Arable land
Corine	Permanently irrigated land	212	3	Arable land
Corine	Rice fields	213	3	Arable land
Corine	Vineyards	221	3	Arable land
Corine	Fruit trees and berry plantations	222	3	Arable land
Corine	Olive groves	223	3	Arable land
Corine	Pastures	231	5	Grassland
Corine	Annual crops associated with permanent crops	241	3	Arable land
Corine	Complex cultivation patterns	242	3	Arable land
Corine	Land principally occupied by agriculture	243	3	Arable land
Corine	Agro-forestry areas	244	3	Arable land
Corine	Broad-leaved forest	311	4	Natural covered areas
Corine	Coniferous forest	312	4	Natural covered areas
Corine	Mixed forest	313	4	Natural covered areas
Corine	Natural grasslands	321	4	Natural covered areas
Corine	Moors and heathland	322	4	Natural covered areas
Corine	Sclerophyllous vegetation	323	4	Natural covered areas
Corine	Transitional woodland-shrub	324	4	Natural covered areas
Corine	Beaches	331	7	Open area
Corine	Bare rocks	332	7	Open area
Corine	Sparsely vegetated areas	333	7	Open area
Corine	Burnt areas	334	7	Open area
Corine	Glaciers and perpetual snow	335	7	Open area
Corine	Inland marshes	411	6	Wetland
Corine	Peat bogs	412	6	Wetland
Corine	Salt marshes	421	6	Wetland
Corine	Salines	422	6	Wetland
Corine	Intertidal flats	423	6	Wetland
Corine	Water courses	511	8	Water surface area
Corine	Water bodies	512	8	Water surface area
Corine	Coastal lagoons	521	8	Water surface area
Corine	Estuaries	522	8	Water surface area
Corine	Sea and ocean	523	8	Water surface area
GlobCorine	Urban and associated areas	10	1	Urban Area
GlobCorine	Rainfed cropland	20	3	Arable land
GlobCorine	Irrigated cropland	30	3	Arable land
GlobCorine	Forest	40	4	Natural covered areas
GlobCorine	Heathland and sclerophyllous vegetation	50	4	Natural covered areas
GlobCorine	Grassland	60	5	Grassland
GlobCorine	Sparsely vegetated area	70	7	Open area
GlobCorine	Vegetated low-lying areas on regularly flooded soil	80	6	Wetland
GlobCorine	Bare areas	90	7	Open area

<i>GlobCorine</i>	Complex cropland	100	3	Arable land
<i>GlobCorine</i>	Mosaic cropland / natural vegetation	110	3	Arable land
<i>GlobCorine</i>	Mosaic of natural (herbaceous, shrub, tree) vegetation	120	4	Natural covered areas
<i>GlobCorine</i>	Water bodies	200	8	Water surface area
<i>GlobCorine</i>	Permanent snow and ice	210	7	Open area
<i>GlobCorine</i>	No data	230	9	Other areas
<i>ECRINS</i>	Lakes		8	Water surface area

Table-A 2: Land use classification of the GLCC data used within MONERIS for scenario calculations.

MONERIS classes	GLCC classes
Urban areas	Urban
Arable Land	Arable land (IR>=80% and IR<80%)
Natural covered areas	Forest, other natural vegetation
Grassland	Grazing Land
Wetland	Wetland
Open Area	Barren Land, Snow and Ice, Set aside

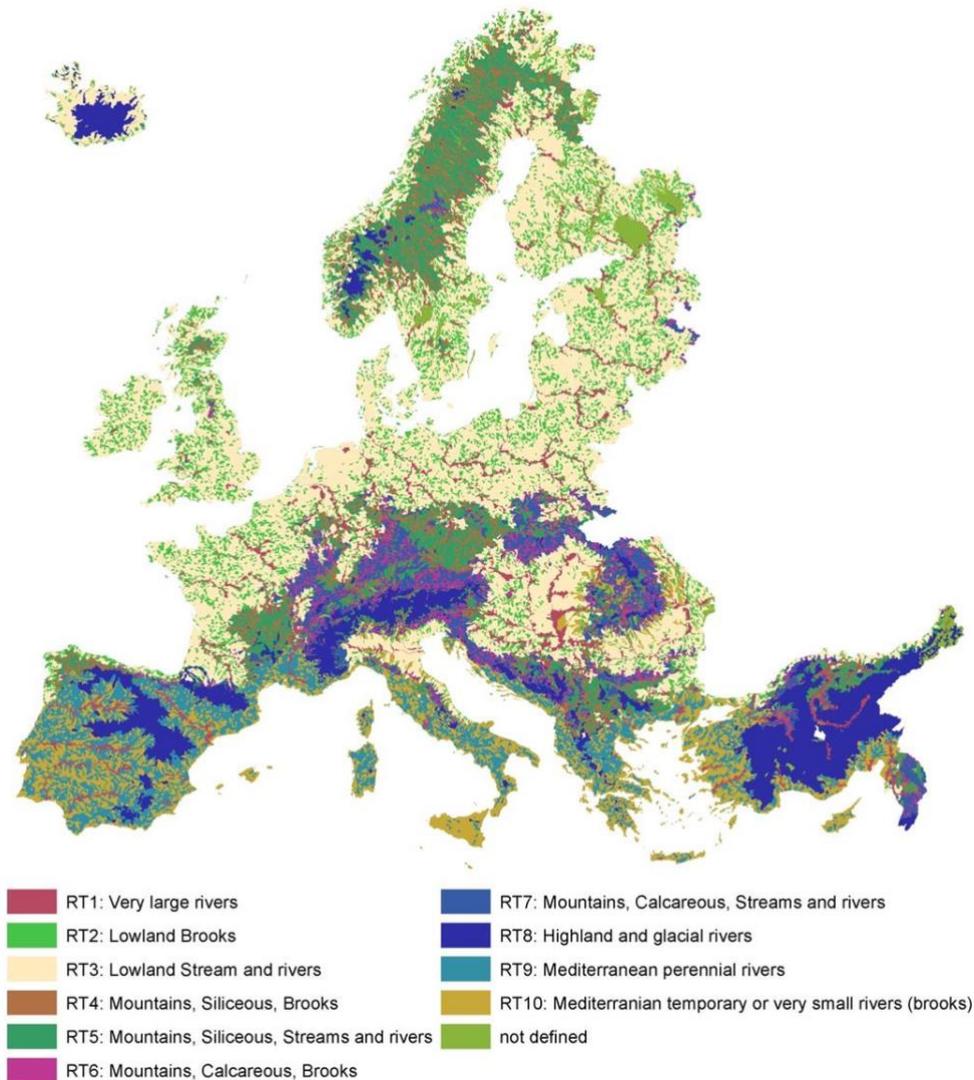


Figure-A 1: Aggregated broad river types derived for the modelling extent in MARS, on basis of ETC/ICM (2015).

Table-A 3: Country specific data sources on population density, used to supplement population data from EUROSTAT.

Country	Source
SRB	http://www.citypopulation.de/
IS	https://www.statice.is/statistics/population/
TUR	EUROSTAT; http://www.turkstat.gov.tr/UstMenu.do?metod=temelist
AND	http://www.estadistica.ad/serveiestudis/web/banc_dades4.asp?tipus_grafic=&check=0&bGrafic=&formules=inici&any1=01/01/2001&any2=01/01/2011&codi_divisio=9&lang=4&codi_subtemes=8&codi_tema=2&chkseries=
IM	https://www.gov.im/media/1355784/2016-isle-of-man-census-report.pdf
DK	http://www.dst.dk/en/Statistik/emner/befolkning-og-valg
BLR	http://www.belstat.gov.by/en/ofitsialnaya-statistika/social-sector/demografiya_2/osnovnye-pokazateli-za-period-s- -po- gody_3/population-size-by-regions-and-minsk-city/
NOR	https://www.ssb.no/a/english/kortnavn/fobhoved_en/tab-2012-06-21-03-en.html
GER	Statistisches Bundesamt, Wiesbaden 2016; http://www.geodatenzentrum.de/geodaten/gdz_rahmen.gdz_div?gdz_spr=deu&gdz_akt_zeile=5&gdz_anz_zeile=1&gdz_unt_zeile=0&gdz_user_id=0
ALB	http://open.data.al/sq/lajme/lajm/lang/sq/id/669/Popullsia-ne-Shtetin-Shqiptar-1870-2011
HR	http://www.ksh.hu/nepszamlalas/tables_regional_00
KO	http://ask.rks-gov.net/en/kosovo-agency-of-statistics/social/population-and-housing-census
MNE	http://www.monstat.org/eng/page.php?id=234&pageid=48
NL	https://www.cbs.nl/en-gb/society/population ; http://www.citypopulation.de/
RU	http://www.citypopulation.de/ ; http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/en/figures/population/
UKR	http://www.citypopulation.de/

Table-A 4: The list of candidates stressor indicators as variables to estimate the probability to reach a GoodOrBetter ecological status.

Attribute code	unit	Description	comment	Data source
<i>m_zhyd</i>	code	functional elementary catchment (FEC)		MARSgeoDB
<i>m_hyreg5c</i>		broad hydro-region: NOR, CB, ALP, EC, MED		MARSgeoDB
<i>m_btype10size</i>		broad river type, grouped by size criteria: BT1- BT10		MARSgeoDB
<i>m_wfdgroup</i>	natural, HMA	modification of river water body		WFD 2016
<i>lu_urb</i>	%	share of urban land use in FEC from Corine Land Cover 2012 (CLC codes 111, 112, 121, 122, 123, 124, 141, 142)		CLC2012
<i>lu_agr</i>	%	share of agricultural land use in FEC from Corine Land Cover 2012 (CLC codes 211, 212, 213, 221, 222, 223, 241, 242, 243, 244)		CLC2012
<i>lu_for</i>	%	share of forest land use in FEC from Corine Land Cover 2012 (CLC codes 311, 312, 313, 324)		CLC2012
<i>lu_r_urb</i>	%	share of urban land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 1 code 1)		Copernicus LC/LU
<i>lu_r_agr</i>	%	share of agricultural land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 1 code 2)		Copernicus LC/LU
<i>lu_r_for</i>	%	share of forest land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 1 code 3)		Copernicus LC/LU
<i>lu_r_fort</i>	%	share of transitional forest land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 2 code 34)		Copernicus LC/LU
<i>lu_r_ford</i>	%	share of broadleaved (deciduous) forest land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 2 code 31)		Copernicus LC/LU

<i>lu_r_forc</i>	%	share of coniferous forest land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 2 code 32)	Copernicus LC/LU
<i>lu_r_form</i>	%	share of mixed forest land use in riparian zone (strip along rivers where Copernicus MAES LC/LU data are available) (LC/LU level 2 code 33)	Copernicus LC/LU
<i>nu_din</i>	mg/l N	nitrate, nitrite + ammonium concentration in water	MONERIS
<i>nu_tp</i>	mg/l P	total phosphorus concentration in water	MONERIS
<i>nu_tpero</i>	kg/km ² /yr	total phosphorus emissions from erosion per FEC	MONERIS
<i>nu_tpemi</i>	kg/km ² /yr	total phosphorus emissions per FEC	MONERIS
<i>nu_tnemi</i>	kg/ha/yr	total nitrogen emissions per FEC	MONERIS
<i>nu_nbal</i>	t/km ² /yr	nitrogen surplus per FEC	MONERIS
<i>nu_psat</i>	%	degree of phosphorus saturation on agricultural land	MONERIS
<i>eco_gen5c</i>	-	general ecological status in 5 text classes	WFD 2016
<i>hy_maf(_abs_norm)</i>	%	mean annual flow alterations - normalised value	hy_maf: ratio between present mean annual flow and semi-natural mean annual flow PCR-GlobWB, IHA
<i>hy_basef(_abs_norm)</i>	%	base flow index alteration - normalised	hy_basef: ratio between present base flow index and semi-natural base flow index PCR-GlobWB, IHA
<i>hy_basef(_pos_norm)</i>	%	increased base flow index - normalised	hy_basef: ration between present base flow index and semi-natural base flow index PCR-GlobWB, IHA
<i>hy_basef(_neg_norm)</i>	%	decreased base flow index - normalised	hy_basef: ration between present base flow index and semi-natural base flow index PCR-GlobWB, IHA

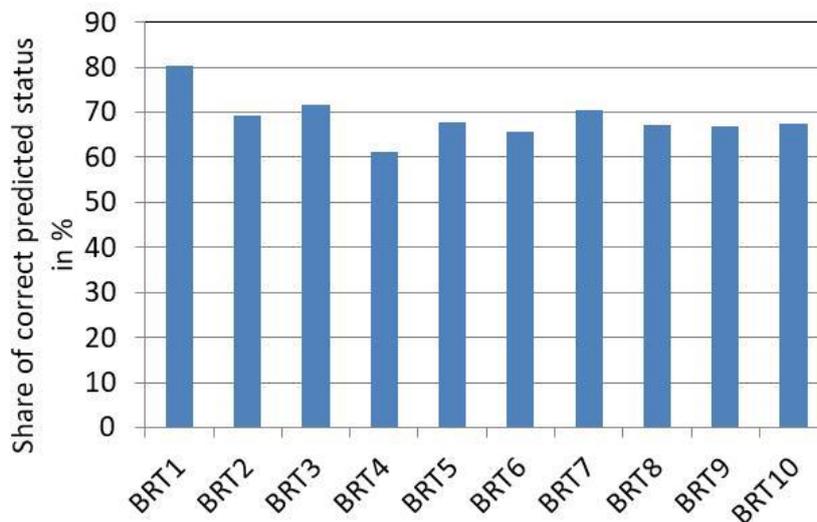


Figure-A 2: Share of correctly predicted ecological status in % of all FECs of broad river Type derived from the trained Bayesian Belief Network.

Table-A 5: Completed national statistic on inhabitants connected to sewer systems in the years 2000-2011 based on reported data from (EUROSTAT). Colour code: coloured cells indicate countries, for which the same connection rate was assumed; red numbers indicate countries to which connection rates were transferred. No-coloured cells with red figures indicate Years, for which connection rates were calculated from linear interpolation. Blue coloured figures indicate that values had to be adapted due to data inconsistencies.

Country	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2001-2005	2006-2010
Albania	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3.0	3	3
Andorra	70.4	71.3	73	74	74	74	77	74	78	81	80	81	73	79
Austria	85.4	86	86	88.9	89	90	91.8	92	92.7	94	93.9	95	88	93
Belarus	65.2	65.2	65.2	70.2	66.1	65.5	64.7	64.1	61.3	66.5	64.3	67.8	66	65
Belgium	51.4	54.7	55.0	58.7	60.9	62.2	75.1	85.0	84.3	85.6	86.5	88.6	58	84
BosniaHerzegovina	2.7	2.7	2.6	2.6	2.6	2.4	2.5	6.4	8.1	9.7	11.2	12.9	3	9
Bulgaria	66.7	67.9	68.4	68.6	68.7	69	69.4	69.7	70	70.4	70.6	74.1	69	71
Croatia	8.45	8.45	8.45	8.45	8.45	8.45	8.75	21.6	25.7	29.8	33.9		9	26
Cyprus	14.3	15.9	18.3	23	28.4	29.8	34	37	40	44	47	51	23	42
CzechRepublic	48.1	48.9	54.4	55.0	55.5	57.8	57.6	59.3	63.3	63.5	65.3	67.1	54	63
Denmark	87.8	88	88.3	88.5	89	89	89	90	90	89.7	90.3	90.7	89	90
Estland	65.2	65.2	65.2	70.2	66.1	65.5	64.7	64.1	61.3	66.5	64.3	67.8	66	65
Finland	83	83	83	83	83	83	83	83	83	83	83	83	83	83
France	81.5	81.5	82.4	82.4	82.4	82.4	82.4	82.4	82.4	82.4	82.4	81.5	82	82
Germany	94.2	94.5	94.9	95.2	95.3	96	96	97.1	97	97	97.3	98	95	97
Gerorgia	16	17	18	20	25	29	31	33	34	36	41	43	22	36
GreatBritain	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97	97.3	97.3	97	97
Greece	80	80.7	81.5	82.2	83	84	84	85	86	87.3	87.3	88.1	82	86
Hungary	51	53.4	56	57.5	59.5	60.6	63.4	66.5	67.7	68.8	71.8	72.5	57	69
Iceland	35.1	35.1	49.8	49.8	49.8	86.4	88.0	89.4	92.0	91.0	66.4	66.7	54	82
Ireland	61.3	62.3	63.3	64.3	65	66	67	65	69	72	71	69	64	69
Isle of Man	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.0	97.3	97.3	97	97
Italy	94	94	94	94	94	94	94	94	94	94	94	94	94	94
Kosovo	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0	0
Lebanon	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	3	3
Lettland	65.2	65.2	65.2	70.2	66.1	65.5	64.7	64.1	61.3	66.5	64.3	67.8	66	65
Lichtenstein	51.4	54.7	55.0	58.7	60.9	62.2	75.1	85.0	84.3	85.6	86.5	87.4	58	84
Litauen	45.7	46.0	46.0	50.5	47.3	47.3	47.0	70.7	70.3	71.5	72.0	73.0	47	67
Luxembourg	100	100	100	100	100	99	99	98	98	98	97.1	99	100	98
Makedonia	5	6	6	6	6	6.5	7	7	7	7	7.5	7.7	6	7
Malta	16	16	18	18	15	15	11	10	17	23	20.7	37	16	20
Moldowa	8.06	8.7	9.35	9.99	11.2	11	12	13	12.9	12.9	13.7	16.7	10	13
Montenegro	8.45	8.45	8.45	8.45	8.5	8.5	8.7	22	26	30	34	33.0	9	26
Netherlands	98.2	98.4	98.5	98.6	98.8	99	99.1	99.1	99.3	99	99.4	99.5	99	99
Norway	79.9	80.8	80.5	81.2	81	83.7	84.1	83.7	83.8	85.2	85	85	81	85
Poland	53.6	55.3	56.7	58.2	59	59.2	59.8	60.3	61	61.5	62	63.5	58	61
Portugal	70.4	71.3	73	74	74	74	76.7	74	78.2	81.3	80	81	73	79
Romania	8.06	8.7	9.35	9.99	11.2	11	12	13	12.9	12.9	13.7	16.7	10	13
Russia	65.2	65.2	65.2	70.2	66.1	65.5	64.7	64.1	61.3	66.5	64.3	67.8	66	65
Serbia	3.7	3.7	3.7	3.7	3.7	3.7	3.9	10.7	13.6	16.4	17.8	20.5	4	14
Slovakia	54.7	55.2	55.3	55.9	56.5	57.1	57	58.2	59.3	59.5	60.4	61.6	56	59
Slovenia	62.6	62.6	62.6	62.6	62.6	62.6	62.6	62.6	62.6	62.6	62.6	62.6	63	63
Spain	79.5	79.7	81.1	82.6	84.1	85.6	88.4	88.7	86.5	91.8	96.0	95.0	83	91
Sweden	86	86	86	86	86	86	86	86	86	86	86	86	86	86
Switzerland	95.4	96	96	96.2	96	96.8	97	97	97	97	97.3	98	96	97
Syria	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	2.5	3	3
Turkey	16	17	18	20	25	29	31	33	34	36	41	43	22	36
Ukraine	8.06	8.7	9.35	9.99	11.2	11	12	13	12.9	12.9	13.7	16.7	10	13

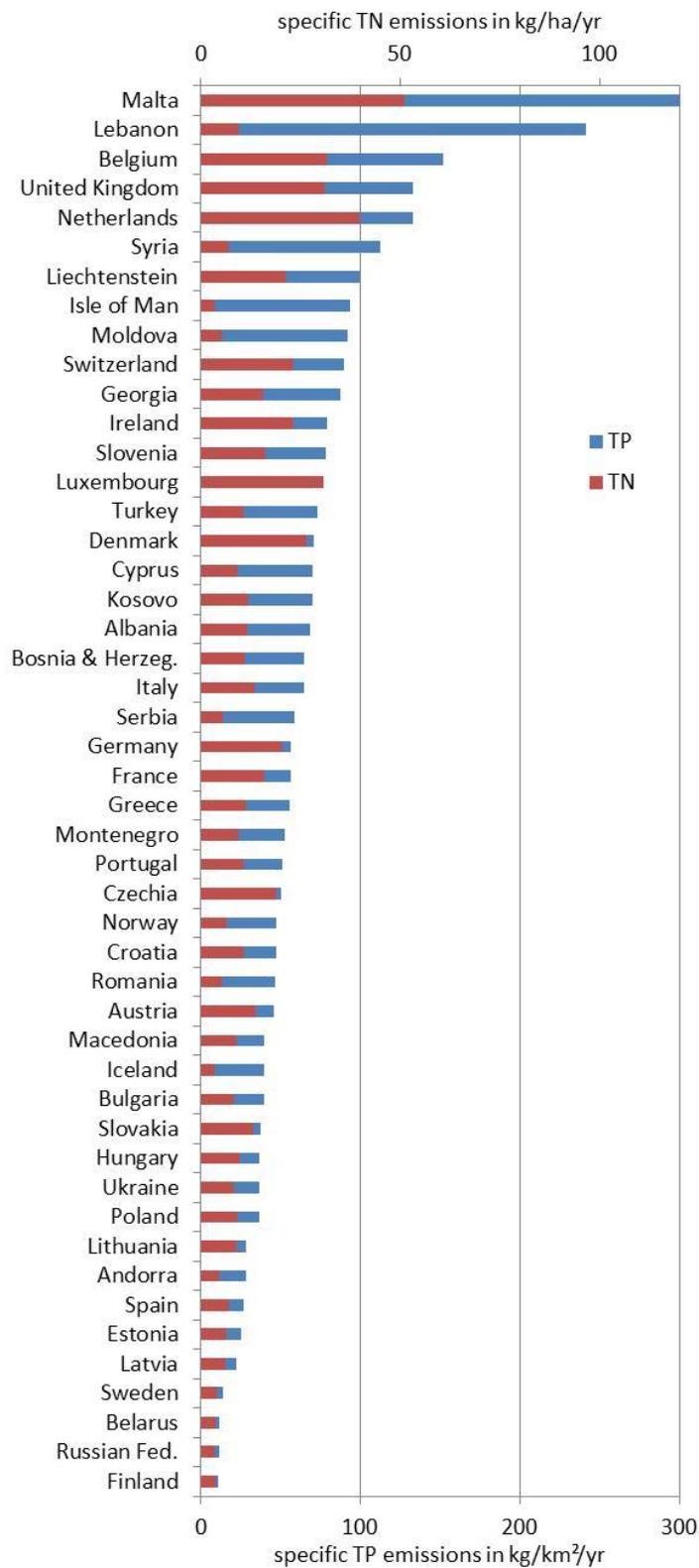


Figure-A 3: Mean specific total nitrogen and total phosphorus emissions for the years 2001-2010 per country in the MARS modelling extent.

Funded by the European Union within the
7th Framework Programme, Grant Agreement 603378.
Duration: February 1st, 2014 – January 31th, 2018



Deliverable 7.2-2: Bayesian Belief Networks: Linking abiotic and biotic data

Authors: Gerben van Geest, Lilith Kramer, Tom Buijse (**Deltares, Netherlands**), Jannicke Moe, Raoul-Marie Couture, Anne Lyche Solheim (**Norwegian Institute for Water Research NIVA**), Eugenio Molina-Navarro, Hans Estrup Andersen, Dennis Trolle (**Aarhus University, Denmark**), Katri Rankinen (**Finnish Environment Institute SYKE**), Pedro Segurado (**University of Lisbon, Portugal**).

Editor: Gerben van Geest & Lilith Kramer (**Deltares, Netherlands**)

Due date of deliverable: **Month 45, October 2017**

Actual submission date: **Month 47, December 2017**

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

Contents	2
Preface.....	5
Abstract.....	6
1 Introduction	9
1.1 Background	9
1.2 Bayesian Belief Networks	12
1.3 Case studies	14
2 Case studies	17
2.1 Norway: Vansjø.....	17
2.1.1 Introduction.....	17
2.1.2 Purpose.....	18
2.1.3 Model construction	19
2.1.4 Class boundaries	20
2.1.5 Conditional probability tables.....	21
2.1.6 Scenarios	23
2.1.7 Results.....	24
2.1.8 Validation.....	30
2.1.9 Obstacles, pros and cons	31
2.2 Finland: Lepsämäenjoki	33
2.2.1 Introduction.....	33
2.2.2 Purpose.....	35
2.2.3 Model construction	35
2.2.4 Conditional probability tables.....	37
2.2.5 Scenarios	37
2.2.6 Results.....	38
2.2.7 Pros and cons	39
2.3 Denmark: Odense	40
2.3.1 Introduction and purpose	40
2.3.2 Model construction	42

2.3.3	Class boundaries	43
2.3.4	Conditional probability tables.....	44
2.3.5	Scenarios	45
2.3.6	Validation.....	46
2.3.7	BBN modelling	46
2.3.8	Results and discussion	46
2.4	Netherlands: Regge and Dinkel.....	57
2.4.1	Introduction.....	57
2.4.2	Purpose.....	58
2.4.3	Model construction	59
2.4.4	Class boundaries	61
2.4.5	Conditional probability tables.....	63
2.4.6	Scenarios	64
2.4.7	Results and discussion	66
2.5	Portugal: Sorraia.....	69
2.5.1	Introduction.....	69
2.5.2	Purpose.....	70
2.5.3	Model construction	70
2.5.4	Class boundaries	72
2.5.5	Conditional probability tables.....	72
2.5.6	Scenarios	73
2.5.7	Validation.....	73
2.5.8	Results and discussion	74
3	Synthesis.....	79
3.1	Scenario studies.....	79
3.2	Assessment of the BBN modelling approach.....	82
3.3	Validation	84
3.4	Differences between BBNs for diagnostic and prognostic purposes	85
4	Acknowledgements	87
5	References	88

Appendix 1: BBN versions of case study Odense (Denmark).....	94
Appendix 2: Conditional probability tables (CPTs) of case study Odense (Denmark)	97
Appendix 3: Stressors probability distributions of case study Odense (Denmark)	101
Appendix 4: Results of scenarios of case study Regge and Dinkel (The Netherlands).....	103

Preface

This document is a synthesis of the work done in five regional catchment case studies of work package 7.3: Combining abiotic and biotic models for river basin management planning of the FP7 MARS (Managing Aquatic ecosystems and water Resources under multiple Stress) project. The five regional case studies were: Vansjø (Norway; Northern region), Lepsämäenjoki (Finland; Northern region), Odense (Denmark; Central region), Regge and Dinkel (Netherlands; Central region), and Sorraia (Portugal; Southern region). The work described in this report was started in March 2016. In May 2016 Christian Feld (UDE), Jannicke Moe (NIVA), Harm Duel and Ellis Penning (Deltares) organised a workshop on Bayesian Belief Networks hosted by NIVA in Oslo, which was followed up by several web conference calls. The project finished in December 2017.

Abstract

Aquatic ecosystems in Europe have been heavily degraded over the past century, as a result of stressors including eutrophication, hydromorphological alterations and overfishing. Accordingly, many measures have been carried out or are planned to improve the ecological status of water bodies. For this, models are required to forecast the effects of the measures planned. Over the past decade, Bayesian Belief Networks (BBNs) models are increasingly applied to aquatic ecosystems. BBNs have a number of advantages, such as explicit incorporation of uncertainty in the outcome, the ability to handle incomplete datasets, expert opinions and model simulations, and a relative simple graphical representation of complex ecosystems interactions. Accordingly, there is an increasing interest in the construction and application of BBNs in water management.

Aim of MARS work package 7.3

The aim of MARS work package 7.3 is to combine abiotic and biotic models for river basin management planning. In this work package, BBNs have been used for the coupling of these models.

Case studies

In this report we have developed predictive BBN models for five case studies catchments across Europe to explore the effects of future scenarios on biological responses and ecological status of water bodies. The case studies cover many dimensions of the MARS project, such as:

- Three regions of Europe (North, Central, South), with case studies from Finland (Lepsamänjoki), Denmark (Odense), The Netherlands (Regge and Dinkel), Portugal (Sorraia), and Norway (Vansjø);
- The two water categories: rivers and lakes;
- The three story lines: Techno, Fragmented and Consensus world that have been used in MARS work package 4.2;
- Various stressor types: Total P, Total N, hydrology, hydromorphological alterations, temperature, etcetera;.
- Biological indicators: chlorophyll a in rivers and lakes, cyanobacteria in lakes, macrophytes, macroinvertebrates, fish, and total ecological status of the water body.

Results

For all case studies, the BBN method enabled the coupling of abiotic and biotic models, and facilitated predictions of biological responses under the different future storylines. Therefore, BBNs had a clear additional value compared to the abiotic process-based catchment models (MARS work package 4). Below, the main results are presented for the case studies.

- Norway: The process-based models (INCA-P and MyLake) predicted temperature, TP and Chl-a in the lake, while the BBN added predictions on cyanobacteria and their

- response to increased temperature, and included this in the assessment of ecological status. The cyanobacteria node resulted in a stricter assessment of ecological status than the chl-a node alone. Due to the cyanobacteria node, this BBN also indicated a slight negative effect of increased temperature in the Techno and Fragmented world scenarios on ecological status, although effects of land-use changes and nutrients were dominating.
- Finland: The process-based model (INCA-P) predicted TP and Chl-a in the river for one catchments (Lepsämäenjok), while the BBN added predictions on EQR (ecological quality ratio, based on macrophytes, macroinvertebrates and fish) and total ecological status. The ecological status node gave stricter assessments than chl-a alone. As for Norway, increased temperature had a negative impact ecological status, but the effects of land use were stronger.
 - Denmark: The process-based model (SWAT) predicted flow and nutrients (TP, TN) in the river, while the BBN added three biological quality elements: macrophytes, macroinvertebrates and fish. Some of the predicted changes in ecological status by the BBN contrasted the initial expectations. For example, the probability of High-Good status of macrophytes was higher for Techno and Fragmented world than for Consensus world, although the latter storyline is more sustainable. A plausible explanation is that the differences in ecological status were driven by the hydrological parameters, which depend mostly on climate change rather than land-use.
 - Portugal: The process-based models (SWAT) predicted hydrological and nutrient variables in the river, while the BBN added four BQEs (phytobenthos, macrophytes, macroinvertebrates and fish) and total ecological status. The BBN predicted that different BQEs respond differently to scenarios and to mitigation options. Phytobenthos and macroinvertebrates responded most strongly to land use. Hence, other mitigation measures implemented in Techno world (e.g. dam removal) had no effect in the BBN. Macrophytes and fish, in contrast, responded to both nutrients and hydrological stressors, and therefore also to scenarios of measures such as a more efficient irrigation and an optimization of fertilizers.
 - The Netherlands: This case study demonstrates how stakeholder engagement and expert judgement can be utilized to develop a BBN, and even run it for future scenarios. We found that presenting a BBN in a group of stakeholders helped them in constructively discussing their water systems. The BBN helped the waterboard in discussion among colleagues, to obtain common understanding, and with communication towards the public. The BBN showed that the impact of human alterations in the streams caused a larger impact on the macrophyte abundance than climate change. In fact, in the current defined scenarios the impact of the dams nullified the impact of climate change. Only changes in riparian zone maintenance would sort an effect.

Advantages of using BBNs

One of the most obvious advantages of BBN is that it can integrate different sources of information, such as expert judgement, empirical modelling and process-based modelling into a single framework. Another advantage of the BBN is the simplicity, allowing a faster scenario analysis than corresponding process-based models. Furthermore, the BBN approach provides an opportunity to include biological elements, as demonstrated by our studies, which is not the case in many existing process-based models. Moreover, our case studies showed that the BBN methodology can facilitate the determination of potential impacts of climate change on ecological status of water bodies.

Disadvantages of using BBNs

There are several limitations associated with the BBN methodology in the context of environmental management. A drawback of the high simplicity of BBNs is the necessity to constrain the information in the BBN, which may diminish the credibility of the results. In addition, the design of the BBN structure and cause-effects links involves many exploratory analyses, and decision making may not be straightforward. The fact that the network cannot contain loops puts also constraints on the ecological processes that can be modelled. Furthermore, the definition of probabilities involves a big amount of effort, while the necessity of defining subjective discretization of variables into interval classes is sometimes problematic. At last, the accumulation of uncertainty with the length of the network implies that it can be difficult to draw conclusions from the final output.

Validation

Developed BBNs should be validated to assess the ability of the model in representing the ecosystem. The procedure for validation of models strongly depends on the purpose of the model, and – hence - model validation is highly case-specific and it is difficult to generalise statements. Based on our case studies, the following criteria have been identified, viz. (1) the BBNs should capture the most important causal relationships of the ecosystem modelled, (2) the quantification of each of these relationships should be validated separately, and (3) the results of the BBNs should be able to fit observed data fairly well.

Differences between BBNs for diagnostic and prognostic purposes

In task 7.2 of MARS, BBNs have been used for diagnostic purposes, while in this report (task 7.3) the focus is on prognosis. This led to the question whether different adaptations in the BBN design are required when its main purpose is to perform diagnostic versus prognostic analyses. The overall general causal structure of a BBN could be used for both diagnostic and prognostic purposes. However, a BBN needs to be simplified to become useful for the purpose it is designed for. When the focus is on diagnosis, the choices for this simplification may be made in a different way than when the model has a prognostic purpose. Accordingly, the design of the final BBN model may differ substantially between BBNs for either diagnostic or prognostic purposes.

1 Introduction

1.1 Background

Aquatic ecosystems in Europe have been heavily degraded over the past century, as a result of stressors including eutrophication, hydromorphological alterations and overfishing. The Water Framework Directive (WFD) commits all European Union member states to achieve ‘good’ ecological status of all water bodies by 2027. Accordingly, many measures have been carried out or are planned to improve the ecological status of water bodies.

For the formulation of effective measures, insight should be gained about the causes of the ecological deterioration (‘diagnosis’). Additionally, predictive models are required to forecast the effects of the measures planned. In MARS Workpackage (WP) 4 (Multiple stressors at the river basin scale), predictive process-based models have been applied to forecast effects of climate and land use scenarios for a large number of case studies across Europe (Ferreira et al. 2016) (<http://fis.freshwatertools.eu/index.php/casestudies.html>). The construction of such models however is complicated, because aquatic ecosystems are characterized by complex and unknown interactions of abiotic and biotic processes. Additionally, available datasets are incomplete or inaccurate with regard to the relationships of interest, and generally there is a lack of knowledge about response of the ecosystems to multiple stressors. Accordingly, in many cases the only available information to make decisions may be expert knowledge and opinion, along with limited empirical data. Consequently, water managers are required to make important decisions for measures in the face of a large amount of uncertainty.

For assessment of ecological status and risk of not achieving management targets, water managers need tools which incorporate the existing knowledge, as well as their uncertainty. Modelling tools have been proven as useful for such an assessment (Devia et al., 2015; Trolle et al., 2012). Process-based catchment scale models are well known as powerful tools to address water quantity and quality issues and to simulate different kind of scenarios (Arnold et al., 1998). Such models however require some specific features for successful application for management. First, the existing (expert) knowledge and data should be easy to integrate, and it should be able to update them rapidly with new data. Secondly, the models should explicitly incorporate uncertainties in their structure and in their predicted outcomes, as these uncertainties may play an important role for the final selection of measures. Third, such models should be readily updated to incorporate new scientific knowledge or evolving policy needs. Fourth, these models should be meaningful to the broad range of persons involved in the decision making process and therefore a clear presentation of the model structure and the inference process is required. Finally, process-based catchment models typically provide predictions of physico-chemical conditions such as temperature and nutrient

concentrations (Ferreira et al. 2016), but not on biological indicators that are needed for ecological status assessment according to the WFD.

Over the past decades, several methods have been used that take into account the requirements stated above. Of these methods, Bayesian Belief Networks (BBNs) are increasingly being applied to aquatic ecosystems to conduct quantitative ecological risk assessments and integrate science and management interactions through a decision making framework. BBNs have a number of features, such as explicit incorporation of uncertainty in the outcome, the ability to handle incomplete datasets, expert opinions and model simulations, and a relative simple graphical representation of complex ecosystems interactions. Accordingly, there is an increasing interest in the construction of BBNs, and from 2002-2013, 74 peer reviewed papers containing BBNs have been published for freshwater and estuarine ecosystems (reviewed by McDonald et al., 2015).

MARS workpackage 7.3: Combining abiotic and biotic models for river basin management planning

In the EU project MARS, the effects of multiple stressors on ecosystems of rivers, lakes and estuaries are investigated. Additionally, MARS provides an overview of tools for river basin management under conditions of multiple stress. For this purpose, a number of practical tools are presented (<http://mars-project.eu/index.php/tools.html>) along with a model selection tool (<http://fis.freshwatertools.eu/index.php/mst.html>), because there is no single tool that can cover the variety of conditions across Europe.

In MARS workpackage 7.2 the diagnostic use of BBNs is presented (Feld et al., 2017). In addition, a ‘cook book’ has been made for the construction of such BBN-models (Feld, 2016).

In this workpackage (MARS 7.3), the predictive use of BBNs is presented for several case studies for rivers and lakes in Europe. The construction of these BBNs is based on the results of MARS WP 4, in which causal relationships are constructed according to the DPSIR-approach, ensuring causal relationships between causes for deterioration, pressures, state variables and biota. In WP4, these relationships were subsequently statistically tested with large datasets (Ferreira et al., 2016).

Workpackage 7.3 has built further upon the results of WP4. For this, NIVA organised in May 2016 a BBN-workshop in Oslo. In this workshop, guidance was given for the application of BBN by two specialists, viz prof. S. Mäntyniemi (University of Helsinki, Finland) and dr. D. Barton (NIVA, Oslo). Subsequently, BBNs have been developed for case studies by the individual partners. The progress and results of the different case studies were regularly discussed during video conferences, as well as (potential solutions for) problems that arose during construction of the BBNs.

The aim of WP7.3 is to combine abiotic and biotic models for river basin management planning. For this purpose, we have applied the BBN approach.

The individual models are mapped to the DPSIR chain of the MARS conceptual framework (Figure 1) to structure the discussion between modellers and stakeholders. This framework also links the assessment of risk, ecological status and ecosystem services within the framework of River Basin Management Plans (RBMPs), therefore providing a close integration of stressors, ecosystems and services.

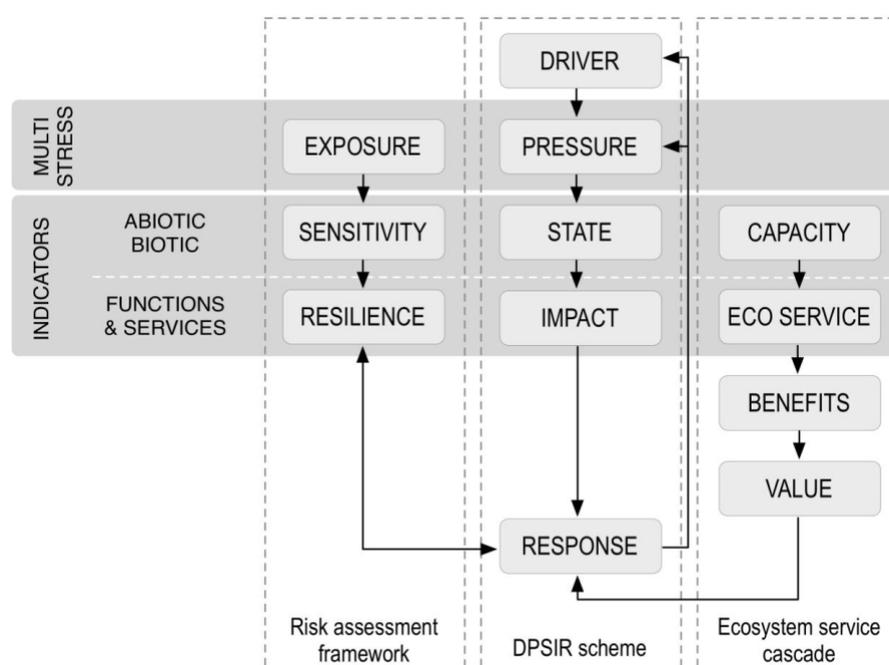


Figure 1. The MARS conceptual model for an integrated assessment framework

MARS storylines: climatic and socio-economic scenarios

Future climatic and socio-economic scenarios have been developed within MARS (deliverable 2.1-4) (Faneca Sanchez, 2015). These scenarios provide both a qualitative framework and, where possible, quantitative data for modellers to run simulations. A selection of scenarios has been used to define the three MARS storylines (<http://fis.freshwatertools.eu/index.php/infolib/scenarios.html>): "Techno world", "Consensus world" and "Fragmented world". *Techno world* sketches a future in which the world will be driven by economy. Policies are focussed on enhancing trade and not on the environment. It is based on a combination of the Representative Concentration Pathway (RCP) 8.5 and the

Shared Socioeconomic Pathway (SSP) 5. *Consensus world* is a world in which the economy and population keep on growing, but they work on protecting the environment. It is based on a combination of the RCP 4.5 and the SSP2. *Fragmented world* is based upon inequality. Each country needs to fight for its own survival and the environment is only protected locally by rich countries. This world is based on a combination of the RCP 8.5 and the SSP3.

In some case studies we have also tested only the social (land use) changes but with observed climate, but for simplicity purposes we have given it the name of the storyline, even if the storyline, as defined in MARS, also includes climate change.

In this study we have made use of the future climate data provided by the MARS scenarios, to predict changes in ecological status under different scenarios for two time horizons (2030 and 2060).

The storylines are described in fact sheet "MARS scenarios and storylines" (http://mars-project.eu/files/download/fact_sheets/MARS_fact_sheet03_storylines.pdf).

Goal of this report

The following goals have been identified for this report:

- A description of MARS case studies that have developed and applied BBNs in freshwater ecosystems (rivers, lakes) in Europe for predictive purposes;
- A discussion of methodological issues (and solutions) that have arisen during the development and testing of BBNs;
- A discussion about different methods of validation for BBNs;
- A discussion about storylines of the MARS scenarios.

1.2 Bayesian Belief Networks

The BBN approach begins by conceptualising a model of interest as a graph or network of nodes and linkages (Figure 2). A network node represents an important system variable and a link from one node to another (depicted as an arrow) represents a dependency relationship between these variables. These relationships may indicate direct causal dependencies or the combined effect of more complex associations. A node has a number of possible discrete states (e.g. differentiation between high/low, or 10-20% cover, 30-40% cover), each of which has an associated probability of occurring. The likelihood of all categories sums to unity ('1' or 100%) within a single node (see Table 1).

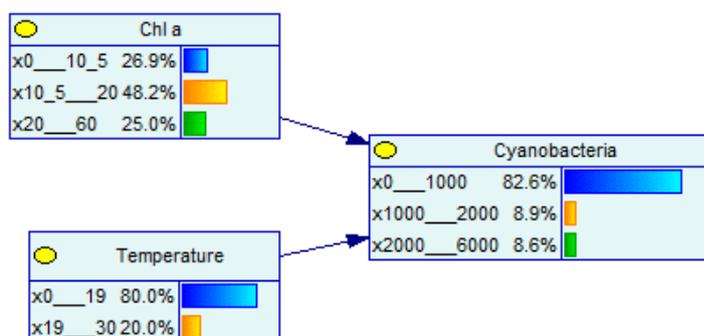


Figure 2 Example simple BBN network. (For more details, see Chapter 2.1).

When no link exists between nodes they are said to be conditionally independent. The concept of conditional independence helps in simplifying a complex system by deconstructing it into subsets. Input parameters are those nodes that can be measured in the field, having no other nodes entering them. Each node can have a series of prior (or unconditional) probabilities of being found in a particular state.

Intermediate nodes depend on input nodes or other intermediate nodes, and are useful for linking variables measured on different scales and for linking subsets of the network. These nodes are represented by conditional probabilities, which represent the likelihood of the state of the node given the states of input parameters affecting it. The probability distribution for each given child node is determined by the probability of each state of its parent nodes. For example, the probability of the node ‘probability for cyanobacteria blooms’ is conditional upon the states of input nodes ‘chlorophyll’ and ‘temperature’ and can be represented using a conditional probability table (CPT, Table 1). For further details about BBNs, see https://en.wikipedia.org/wiki/Bayesian_network.

Table 1 Examples of a conditional probability tables (CPT): CPT for Cyanobacteria conditional on Chl and water temperature. Each column contains the probability distribution of the child node (cyanobacteria) for a given combination of states of the parent nodes (Chl and temperature). For more information, see Chapter 2.1.

Chl (µg/L)	0-10.5		10.5-20		20-60	
	0-19	19-25	0-19	19-25	0-19	19-25
Cyanobacteria (µg/L)						
0-1000	1.000	1.000	1.000	0.923	0.333	0.323
1000-2000	0.000	0.000	0.000	0.077	0.333	0.290
2000-6000	0.000	0.000	0.000	0.000	0.333	0.387

BBNs have a number of features that make them useful for ecological risk assessment and prediction (McDonald et al, 2015). First, they can incorporate and compile qualitative and quantitative data from incomplete datasets, model simulations and expert opinions. Secondly, they provide predictive analyses of uncertain, complex and multi-state ecosystems.

Furthermore, BBNs can be easily created, updated modified and extended. Additionally, uncertainties associated with both the model and data can be quantified, and the effects of uncertainty are an explicit part of the outcome of the results. At last, BBNs provide a graphical representation of complex ecosystems interactions that can be useful in management and science integration (McDonald et al, 2015).

Marcot et al. (2006) have given an number of guidelines for constructing BBNs. In our case studies, these guidelines have been adhered mostly, viz:

1. As far as possible, the number of parent nodes to any give node has been kept to three or less, and the number of states for each node is five or less. This keeps the associated conditional probability tables (CPT's) small enough to be tractable and understandable.
2. Parentless (input) nodes – typically representing predictor habitat and environmental variables – are those items that can be pre-processed or empirically evaluated from existing data;
3. Intermediate nodes are used to summarize the major themes denoted in the ecological causal web.
4. Preferably, all nodes are observable, quantifiable or testable.
5. The fewest discrete states necessary within any given node are used to represent their effects, thereby ensuring that enough states are distinguished to ensure the desired precision of the estimates and the range of input values in the model.
6. The number of layers of nodes (viz. the depth of the model) is kept to four or less, if possible. This is desirable for at least three reasons:
 - (1) deep models with many intermediate nodes (latent variables) may contain unnecessary uncertainty propagated from input to output nodes;
 - (2) the sensitivity of the output node to input nodes may be swamped and dampened by intermediate nodes;
 - (3) output nodes in models with asymmetric structures may be far less sensitive to more distant input nodes with many intervening intermediate nodes than the modeller intended.

1.3 Case studies

Five regional case studies contributed to this report. These case studies were: Lepsamänjoki (Finland), Odense (Denmark), Regge and Dinkel (The Netherlands), Sorraia (Portugal), and Vansjø (Norway). These case studies cover a swathe across Europe from the Lepsämanjoki catchment in the North East, to the Sorraia catchment in the South West. Figure 3 shows the locations of these catchments.

Although the case studies share some common features they also show many distinctive characteristics. A broad overview of these characteristics is shown in Table 2. More in depth information on each case study can be found in the subsequent chapters and in the MARS Case study synthesis final report (D4.1, Ferreira et al., 2016).



Figure 3. Map showing the locations of the case study basins.



Table 2. General characteristics of each case study.

Name	Country	European region	Water category	Main stressors	Main response variables	MARS scenarios	CPT methods	Validation methods
Vansjø	Norway	Northern	Lake	TP, temperature, wind	Chl-a, Cyanobacteria; ecological status	Techno, Fragmented, Consensus	Based on data (observed and modelled)	External datasets
Lepsämäenjoki	Finland	Northern	River	TP, temperature, discharge, land-use, hydro-morphology	Chl-a, ecological status	Fragmented, Consensus	Based on data (observed and modelled)	
Odense	Denmark	Central	River	Hydrology, TP	Ecological status (fish, macrophytes, macroinvertebrates)	Techno, Consensus, Fragmented	Data (observed and modelled)-based	Independent, observed ecological status data
Regge and Dinkel	Netherlands	Central	River	Hydrology, hydrological structures, maintenance	Macrophyte abundance	Techno, Fragmented, Consensus	Expert judgement and data-based.	
Sorraia	Portugal	Southern	River	Diffuse pressures, abstraction, barriers, hydrological	Phytobenthos, macrophytes, macroinvertebrates, fish	Techno, Fragmented, Consensus	Based on data (observed and modelled)	

2 Case studies

2.1 Norway: Vansjø

2.1.1 Introduction

The Norwegian case study represents the Morsa catchment, consisting of the river Hobøl and the lake Vansjø (Figure 4). Land cover of the Vansjø-Hobøl catchment is dominated by forestry (78%), agriculture (15%) and lakes (7%). The agricultural land-use is mainly cereal production (89%), with small areas under grass (9.8%), vegetables (0.6%) and potatoes (< 0.1%). The local catchment for Vanemfjorden has a much higher proportion of agriculture, especially vegetable crops. Agricultural activities contribute about 48% of the total P input to the river basin, followed by natural runoff (39%) and waste water treatment plants (WWTP) (5%) and scattered dwellings (8%) (Skarbøvik and Bechmann, 2010). The lake has a long history of eutrophication from at least the 1970s when systematic monitoring of the lake began. Total P concentrations in the western basin Vanemfjorden lie between 20-40 $\mu\text{g/L}$ P, above the threshold of 20 $\mu\text{g/L}$ for the good ecological status as required by the Water Framework Directive (Skarbøvik et al., 2016). The lake, and in particular the basin Vanemfjorden, suffers from toxin-producing cyanobacterial blooms.

The Morsa catchment is also a case study in MARS WP4 (catchment modelling; <http://fis.freshwatertools.eu/index.php/vansjo-hobol.html>). A detailed description of the catchment can be found in deliverable D 4.1 (Ferreira et al. 2016), part3 (Northern Basins region). The key physical and chemical processes of the catchment, river and lake process have been modelled by a chain of process-based models: PERSiST (catchment hydrology), INCA-P (river) and MyLake (lake). This chapter describes the effects of the MARS storylines (future land use and climate scenarios) on the river and lake water quality. An earlier application of this model chain to Lake Vansjø for a different set of future scenarios (in the EU project REFRESH) was reported by Couture et al. (2014).

The process-based lake model used for Lake Vansjø can predict the MARS benchmark indicators (BInd) Total P (BInd02) and Chl a (BInd08), which are also relevant for classification of ecological status (BInd01). However, ecological status assessment of lakes in the Northern GIG should also be based on more details on the phytoplankton community composition, including the seasonal maximum cyanobacteria concentration (BInd10). Cyanobacteria are not yet predicted by MyLake or other suitable lake models for Lake Vansjø. Instead, we use the empirical relationship between observed cyanobacteria values and physical and chemical variables (temperature and Chl a) in the BBN for Vansjø. The BBN model is set up for the western basin of Lake Vansjø, named Vanemfjorden (all the following references to Lake Vansjø will refer to this basin).



Figure 4 Geographical location and land-use distribution of the Vansjø-Hobøl catchment. The MyLake model is evaluated at the Vanemfjorden station (open circle). A black arrow indicates the outlet of the lakes to the Oslo fjord.

2.1.2 Purpose

The purpose of the BBN model (Figure 5) was to link the outcome of the process-based hydrologic, catchment, and lake models under different climate and land-use scenarios to a biological response, i.e. the abundance of cyanobacteria. The ecological status of water bodies should be determined primarily by biology and secondarily by supporting physico-chemical element such as nutrients. However, process-based models typically do not predict biological responses other than the Chlorophyll a concentration. By including the cyanobacteria abundance in this model, we can obtain a more correct assessment of the ecological status in Lake Vansjø.

2.1.3 Model construction

The BBN model took the output of lake temperature, TP concentration and Chl concentration from the MyLake simulations. The development and application of the BBN model for Lake Vansjø under different scenarios are described by Moe et al. (2016). Here we use a slightly modified version of this BBN model. The current version has been further developed by including a climatic variable: wind speed. The reason is that cyanobacteria blooms are more likely to occur during periods of low wind speed (Ibelings et al. 2003)

The conceptual model has five modules (Figure 5): (1) Climate and land use scenarios (yellow nodes); (2) output from the process-based lake model MyLake (blue nodes); (3) climatic data (red nodes), (4) monitoring data from Lake Vansjø (green nodes); and (5) the national classification system for ecological status of lakes (grey nodes).

Data sources

The future climate data (obtained from MARS WP4) contain daily values of air temperature, precipitation, wind and other variables on a 0.5 ° x 0.5 ° grid, for the period 2006-2095. These data are obtained from two different climate scenarios, the Representative Concentration Pathways (RCP) 4.5 and 8.5. RCP4.5 is used in the "Consensus world" storyline, while RCP8.5 is used in both the "Techno World" and "Fragmented world" storylines. The predicted lake temperature, total P and Chl-a for the different MARS scenarios were obtained from MyLake and processed for use in the BBN model in the same way as described by Moe et al. (2016). The corresponding monitoring data from Lake Vansjø as well as the cyanobacteria (Haande et al. 2016) and the processing of these data have also been described by Moe et al. (2016).

The full set of stressor variables considered for the BBN were those predicted by the climate model (air temperature, precipitation and wind speed) and by MyLake (secchi depth, TP, Chl-a and lake temperature). The previous BBN version (Moe et al. 2016) included only the four output variables from MyLake; we now wished to investigate whether the BBN would be improved by including climatic variables as well. Other potential stressor variables (e.g. total N or water colour) were not included because it was not possible to predict these for the different scenarios. Secchi depth was omitted from the revised BBN because the most recent version of the national classification guidance states that current class boundaries are not applicable for lakes with high turbidity such as Lake Vansjø.

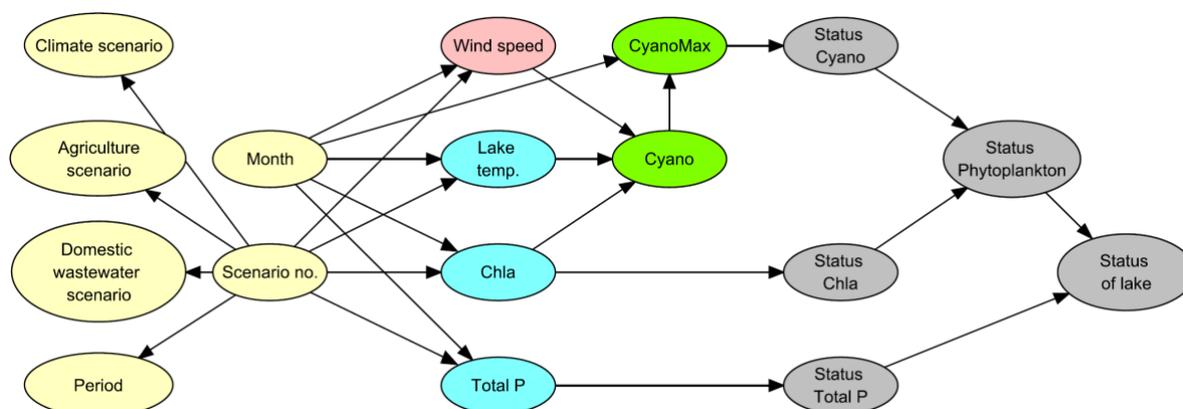


Figure 5 Conceptual version of BBN model for Lake Vansjø. The conceptual model has five modules, which can be mapped to the DPSIR framework (Figure 1): (1) Climate and land use scenarios (yellow nodes; Drivers); (2) output from the process-based lake model MyLake (blue nodes; Pressures and abiotic States); (3) climatic data (red nodes; Pressures or abiotic States), (4) monitoring data from Lake Vansjø (green nodes; biotic States); and (5) the national classification system for ecological status of lakes (grey nodes; Impacts). The set of arrows pointing to one node represents the conditional probability table for this node. The possible states and examples of probability distributions for each node are shown in Figure 4.

Model structure

The relationships between meteorological data and lake monitoring data from the same period were explored by GAM, linear models and regression trees. Precipitation did not show any relationship with the lake data, and was therefore not included in the BBN. Air temperature showed a very strong correlation with lake temperature ($R^2 = 0.99$) and was therefore also omitted, assuming that lake temperature would have a more direct effect on the biology than air temperature. Wind speed had a significant on the cyanobacteria concentration in regression trees: lower wind speed (≤ 3.3 m/s) was associated with higher probability of high concentrations. This outcome is consistent with preliminary experimental results from MARS task 3.1 (Stechlin See): higher wind speed provides better mixing of the water column, which reduces the competitive advantage of cyanobacteria relative to other phytoplankton groups. The final set of predictor variables for cyanobacteria were therefore Chl-a, lake temperature and wind speed.

2.1.4 Class boundaries

Variables for assessment of ecological status were TP, Chl-a and CyanoMax (seasonal maximum of cyanobacteria concentration). Each of these nodes had three intervals that corresponded to concentrations of High-Good, Moderate and Poor-Bad status classes, respectively (see Table 3). The node Lake temperature had two intervals: below and above 19 °C. This boundary was determined by a regression tree analysis of the monitoring data: the probability of cyanobacteria abundance exceeding $1000 \mu\text{g L}^{-1}$ was significantly higher

when the lake temperature was above 19 °C (Moe et al. 2016). The node Wind speed also had two intervals: below and above 3.3 m/s. This threshold matches well with the thresholds of 3.3 m/s used to defined calm vs. windy days in MARS task 3.3.1 (Moe et al. 2017, Chapter 2.3).

Table 3 Boundaries of status classes for biological and chemical elements included in the BBN model, according to the Norwegian classification system*) for lakes of type L-N8a (large, lowland, moderately calcareous, humic).

Variable	Good/Moderate	Moderate/Poor
Chl ($\mu\text{g L}^{-1}$)	10.5	20
CyanoMax ($\mu\text{g L}^{-1}$)	1000	2000
Total P ($\mu\text{g L}^{-1}$)	20	39

*) <http://www.miljodirektoratet.no/Documents/publikasjoner/M587/M587.pdf>

2.1.5 Conditional probability tables

The discrete probability distributions in the CPTs are obtained by different approaches in the different BBN modules (see Moe et al. 2016 for more details). In Module 2 (Process-based model output), the conditional probability distribution of each child node was calculated as the frequency distribution of this variable across each of its parent nodes in the reference scenario (extended baseline), for all 60 runs of MyLake pooled together. In Modules 3 (Climate) and 4 (Monitoring data), likewise, the links from the predicted MyLake outcome to the observed data (wind, laketemperature, Total P and Chl-a-a) were based on the joint frequency distributions of the two variables. The observed data were paired with the corresponding predicted data for the same week, and the concentration intervals were compared.

The CPT for the Cyanobacteria node was also based on counts of observations in the previous version (Moe et al. 2016). However, with the additional parent node (wind) the number of cells in the CPT was doubled from 18 to 36, and the number of observations (90) would not be sufficient for generating probability distributions merely based on counts. Instead, we used an ordinal regression method to estimate the probability of the three states of cyanobacteria status with the three parent nodes as predictor variables. Ordinal regression is similar to logistic regression (where the response variable is 0/1), but allows for more than two ranked categories as response variable (such as HG, M, PB), and estimates the cumulative probability of each state (see **Figure 6**). We used the function "clm" (cumulative logit model) in the R package "ordinal" (Christensen, 2015). Temperature and wind were categorical predictor variables with two states (low/high), while Chl-a was used as a continuous predictor variable (it was not possible to have all three predictor variables

categorical). To obtain the probability distribution for the different combinations of the parent nodes, we used the predict function (predict.clm). To obtain a single input value for each interval of Chl-a (as required for the predict function), we used the median Chl-a concentration of each Chl-a interval. An extract of the resulting CPT for cyanobacteria is shown in Table 4.

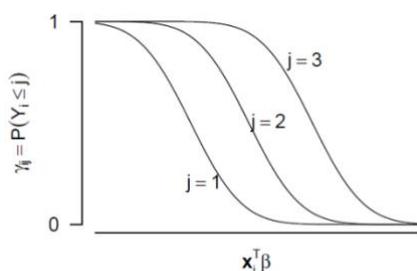


Figure 6 Illustration of a cumulative link model with four response categories (after Christensen, 2015, Fig. 1).

Table 4 Extract of the conditional probability table (CPT) for the cyanobacteria node. Note that the table contains only 3 of the 6 intervals of Chl-a.

Chl-a (ug/L)	5 - 10.5				15 - 20				>25			
Temperature (°C)	<19		>19		<19		>19		<19		>19	
Wind (m/s)	<3.3	>3.3	<3.3	>3.3	<3.3	>3.3	<3.3	>3.3	<3.3	>3.3	<3.3	>3.3
Cyanobacteria (ug/L)												
0 - 1000	1	1	0.990	1	1	1	0.829	1	1	1	0.065	1
1000 - 2000	0	0	0.009	0	0	0	0.150	0	0	0	0.337	0
>2000	0	0	0.001	0	0	0	0.021	0	0	0	0.598	0

The CPT for CyanoMax (the maximum of Cyano for each year; Table 5) was obtained by counting the number of observed Cyano in each concentration interval and each season, and calculating the frequency distribution across the corresponding CyanoMax intervals for all of these observations. For example, out of the 34 observations of Cyano concentration below 1000 $\mu\text{g L}^{-1}$ in the months May-June, 10 observations (29%) came from a year where the CyanoMax in the same year exceeded 2000 $\mu\text{g L}^{-1}$. Thus, even if the predicted cyanobacteria concentration for a single date in June is below 1000 $\mu\text{g L}^{-1}$, there is still a 29% probability that the CyanoMax value will be >2000 $\mu\text{g L}^{-1}$ later that year.

The BBN assigns the probability of ecological status of the lake given estimates total P, Chl-a and cyanobacteria abundance, according to the status class boundaries for each variable (Table 3). The ecological status is determined by two biological indicators (Chl-a and cyanobacteria) and one physico-chemical indicator (total P), using the following two combination rules. (1) If the cyanobacteria status is lower than Chl-a status, then the combined phytoplankton status

is set to the average of the Chl-a and cyanobacteria. (If the cyanobacteria status is equal to or higher than the Chl-a status, then the cyanobacteria status is not included). (2) If the phytoplankton status is High or Good, and Total P status is lower than the phytoplankton status, then the combined lake status is reduced by one status class. (The status assessment of Chl-a and total P based on individual months is not strictly correct, since the assessment should be based on the average of observations from the whole period May-October).

Table 5 The conditional probability table for the CyanoMax node. "Experience" is the count of observations for each combination of the parent node states.

Cyano ($\mu\text{g L}^{-1}$)	0 - 1000			1000 - 2000			>2000		
	May - Jun	Jul - Aug	Sep - Oct	May - Jun	Jul - Aug	Sep - Oct	May - Jun	Jul - Aug	Sep - Oct
CyanoMax ($\mu\text{g L}^{-1}$)									
0 - 1000	0.62	0.72	0.67	0	0	0	0	0	0
1000 - 2000	0.09	0.14	0.11	0.17	0.17	0	0	0	0
>2000	0.29	0.14	0.22	0.83	0.83	1	1	1	1
Experience	34	29	27	6	6	2	1	12	2

Table 6 Mean values (and standard deviations) of variables predicted by climate model IPSL (wind) or lake model MyLake (lake temperature, Total P and Chl-a) for different scenarios and time horizons. Scenarios: BL = extended baseline, 4.5 = climate scenario RCP 4.5, 8.5 = climate scenario RCP 8.5, CW = Consensus World, FW = Fragmented World, TW = Techno World.

Scenario	Wind		Lake temperature		Total P		Chl-a	
	2030	2060	2030	2060	2030	2060	2030	2060
BL	2.16 (0.48)	2.17 (0.55)	14.8 (4.8)	14.9 (4.7)	16.7 (6.0)	14.1 (5.6)	8.3 (4.5)	6.3 (3.7)
4.5	2.13 (0.54)	2.11 (0.53)	17.4 (4.7)	18.4 (4.5)	17.6 (5.3)	15.9 (5.2)	8.7 (4.0)	7.2 (3.5)
8.5	2.08 (0.46)	2.05 (0.51)	17.6 (4.9)	19.2 (4.4)	17.1 (5.3)	16.0 (5.0)	8.6 (4.1)	7.3 (3.5)
CW	2.13 (0.54)	2.11 (0.53)	17.4 (4.7)	18.4 (4.5)	14.1 (4.1)	12.4 (3.5)	7.6 (3.7)	6.2 (3.0)
FW	2.08 (0.46)	2.05 (0.51)	17.6 (4.9)	19.2 (4.4)	19.1 (6.2)	18.3 (6.1)	9.2 (4.3)	8.0 (3.9)
TW	2.08 (0.46)	2.05 (0.51)	17.6 (4.9)	19.2 (4.4)	21.4 (7.3)	21.1 (7.5)	9.9 (4.7)	9.0 (4.3)

2.1.6 Scenarios

The BBN has been run for two types of scenarios:

1. Explorative scenarios for the cyanobacteria node. Four environmental scenarios were defined by the different states of lake temperature (above/below 19 °C) and wind (above/below 3.3 m/s): cold and windy (CoWi), cold and calm (CoCa), warm and windy (WaWi), and warm and calm (WaCa). These four scenarios have been run for the three states of eutrophication (Chl-a status). The purpose was to investigate the

response of the cyanobacteria to these stressor combinations, and to identify stressor interactions

2. MARS storylines. The BBN has been run for the 3 MARS storylines (Consensus World, Fragmented World and Techno World), as well as for scenarios with climate change only. The selected time horizon was 2030 (i.e. years 2020-2040) and 2060 (i.e. years 2050-2080). A summary of the output from the climate model and the MyLake model (Couture et al. 2017) used in this BBN is given in Table 6.

2.1.7 Results

The Bayesian Belief Network was used to project the changes in the ecological status of Lake Vansjø given the alternative future climate scenarios and storylines. The BBN provides the necessary quantitative links from the projections of Chl-a provided by MyLake to estimates of the probability of cyanobacteria abundance that exceed the threshold for good ecological status, and to integrate this result with other results from the process-based models using the combination rules of the national classification system.

The combined effect of Chl-a, lake temperature and wind speed on the cyanobacteria status and on the phytoplankton status was explored for all possible combinations of the parent nodes, for the month July. Two examples of combinations are shown in **Figure 7**: high (a) vs. low (b) wind speed in combination with high lake temperature and intermediate Chl-a concentrations. The phytoplankton status is determined by the Chl-a status in combination with the cyanobacteria status: the cyanobacteria status can only contribute to reducing the phytoplankton status, but not improving it. In all cases, the phytoplankton status was worse than the Chl-a status, because there is always a probability of a cyanobacteria bloom, which reduces the probability of high-good phytoplankton status.

The results (**Figure 8**) show that there was a strong interaction between the effects of the three stressors on the cyanobacteria status (left panel). When the Chl-a status was high or good (i.e. low concentration of Chl-a), then the two other stressors had no effect on the cyanobacteria (**Figure 8a**). Therefore, the combined phytoplankton status was also equal under these four scenarios (**Figure 8b**). When Chl-a status was moderate, then the combination of warm water and calm days (scenario WaCa) resulted in higher probability of cyanobacteria blooms: probability of High-Good status was reduced from ca 70% to 60% (**Figure 8c**).) This effect on cyanobacteria was transferred to the phytoplankton status: probability of poor-bad status increased from 10 to 20% (**Figure 8d**). When Chl-a status was poor or bad, then the warm and calm conditions had an even stronger effect on cyanobacteria: the probability of High-Good status was further reduced to <10% (**Figure 8e**). However, since the Chl-a status was already poor-bad, the change in cyanobacteria status had no further effect on the combined phytoplankton status.

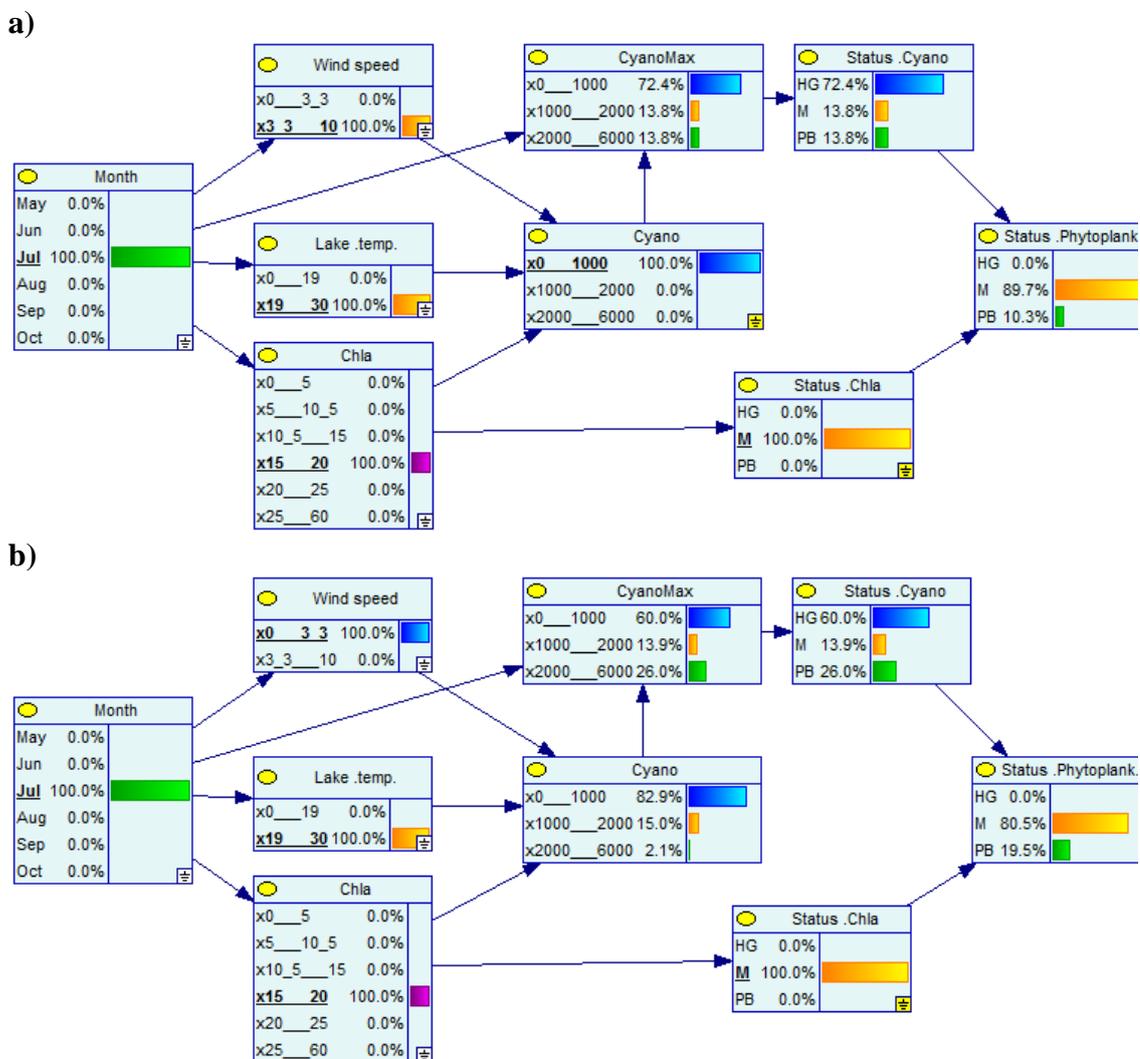


Figure 7 Examples of results from the BBN model run for two exploratory scenarios: (a) Warm and windy and (b) Warm and calm, both with intermediate Chl-a concentrations. The figure illustrates that the lower wind speed (b) results in higher probability of CyanoMax above 2000 µg/L and thereby increases the probability of Poor-Bad phytoplankton status (from 10% to 20%).

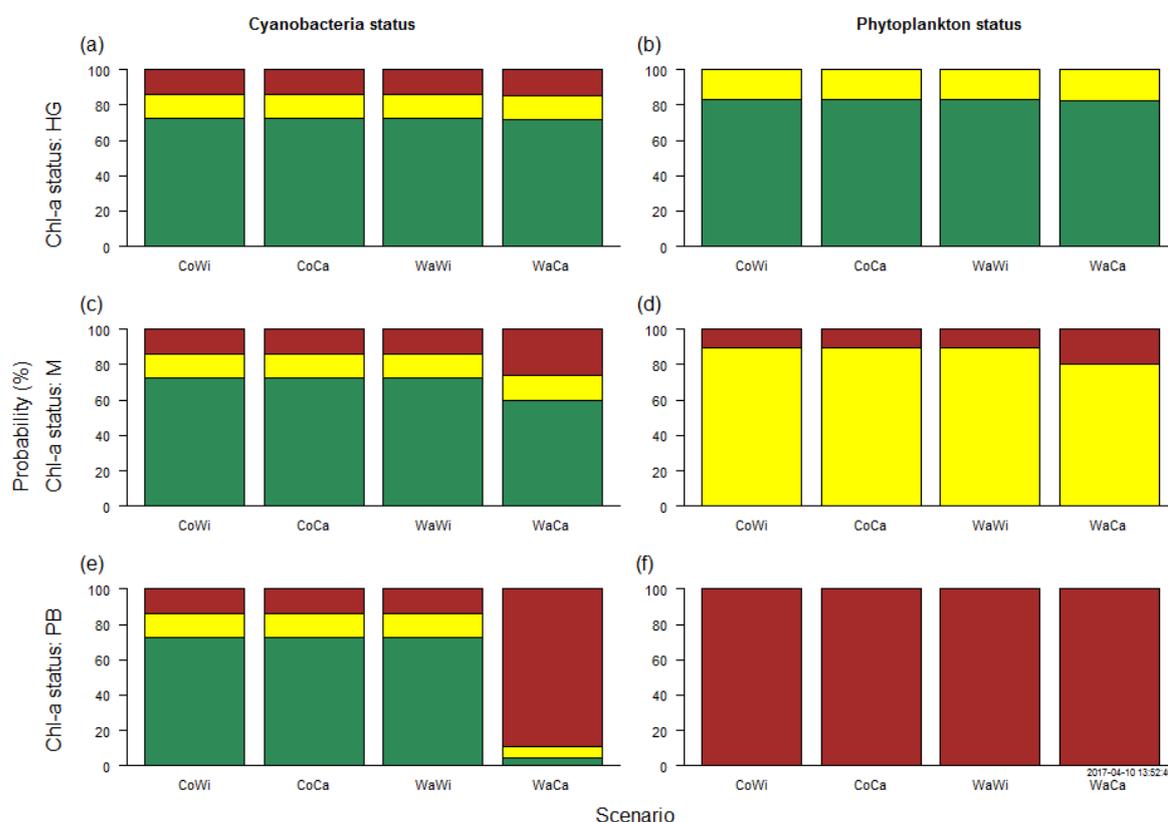


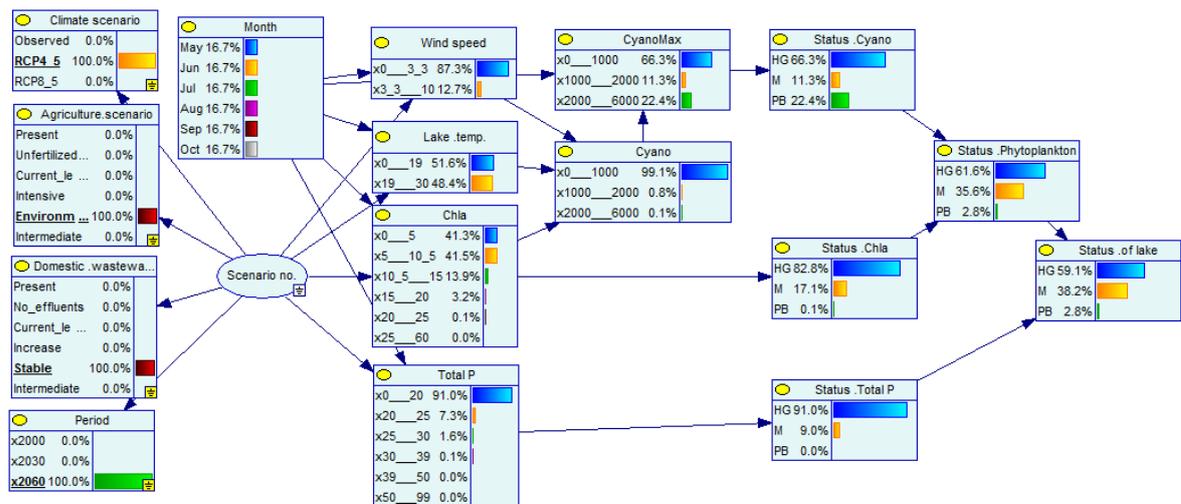
Figure 8 Results from exploratory scenarios of environmental conditions: Cold (Co), Warm (Wa), Windy (Wi) and Calm (Ca). The four environmental scenarios are combined with three states of Chl-a status: High-Good (upper panel), Moderate (middle panel) and Poor-Bad (lower panel). Barplots show the probability distribution across classes of cyanobacteria status (left panel) and phytoplankton status (right panel), which combine Chl-a and cyanobacteria status.

To summarise the exploration of environmental scenarios, (1) the combination of high temperature and calm wind will promote high concentrations of cyanobacteria, (2) this stressor interaction has the strongest effect on cyanobacteria when Chl-a concentration is high, (3) but the effect of this stressor combination on the combined phytoplankton status is highest when Chl-a concentration is intermediate.

Two examples of the full BBN's model predictions for MARS scenarios are shown in **Figure 9**: the posterior probability distributions for all nodes under the two storylines Consensus World (a) and Techno World (b), both using the climate model IPSL and the time horizon 2060. In the Consensus World scenario (**Figure 9a**), the probability High-Good (HG) status for Chl-a as predicted by the MyLake model is 84%. The probability of High-Good status for cyanobacteria, however, is only 67%, and the combined phytoplankton status has even lower probability (63%) of being acceptable (High-Good). The Total P predicted by MyLake in this scenario has a high probability (93%) of High-Good status, and therefore contributes only little to the overall lake status (59% probability of High-Good).

In the Techno World scenario (Figure 9b), the probability of High-Good Chl-a status (63%) is much lower than in the Consensus World. The probability of High-Good cyanobacteria status is practically unaltered (66%), but the probability of combined phytoplankton status is reduced to 46%. In this scenario the Total P status has only 58% probability of being High-Good, resulting in an overall lake status with only 30% probability of High-Good.

(a)



(b)

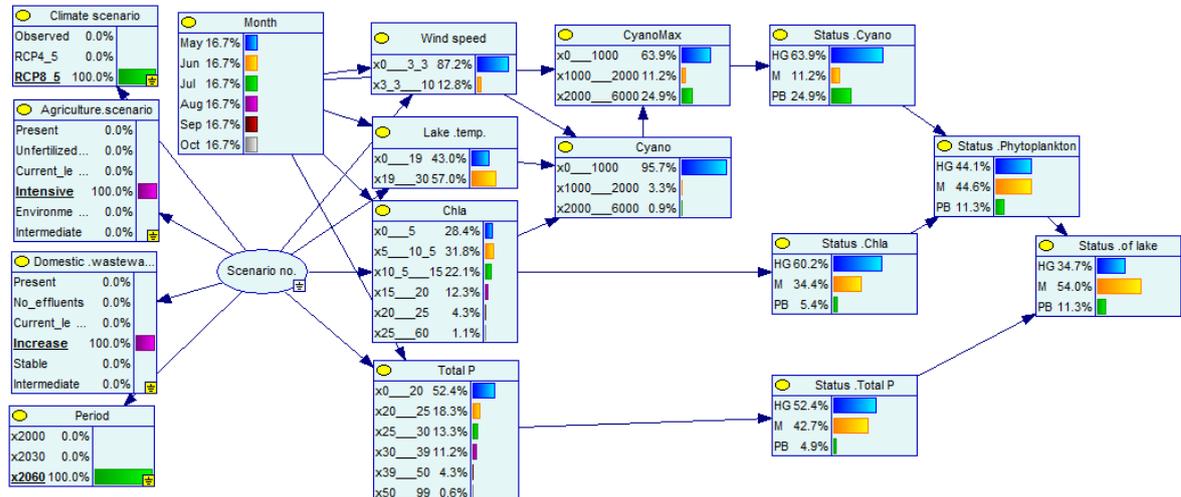


Figure 9 Examples of results from the BBN model run for two storylines: (a) Consensus World, (b) Techno World. Both use the climate model IPSL and data from the time horizon 2060. The figure illustrates e.g. that these two storylines result in different probabilities of High-Good status of Chl-a (83% and 60%, respectively). The probability of High-Good status for Cyanobacteria are quite similar for the two storylines (66% and 64%, respectively). Nevertheless, the "Status Cyano" node has an important impact on the combined "Status Phytoplankton" node, reducing the probability of High-Good (compared to the status of Chl-a) to 62% and 44%, respectively. The Total P node further reduces the probability of High-Good status of the lake to 59% and 35%, respectively, for these two storylines.

The posterior probability distributions (i.e. predictions) for selected nodes has been compared across all climate scenarios and MARS storylines (**Figure 10**). Relative to the extended baseline ("BL"), the Consensus World storyline would improve the ecological status, the Fragmented World storyline would produce a deterioration of the status, and the Techno World storyline would give a large deterioration of the status. The results indicated that climate change alone, as represented by the two RCPs 4.5 versus 8.5, will not produce great differences in the probability of obtaining good ecological status for Lake Vansjø.

The land-use changes, on the other hand, gave significant changes in the probability of obtaining good ecological status in the future. For Consensus World, which includes climate scenario RCP4.5, the sustainable land use (environmental agriculture scenario and stable domestic wastewater scenario) would result in lower concentrations of total P and Chl-a, compared to the climate scenario with no change in land use (columns "4.5"). For Fragmented World and Techno World, the intermediate or more intensive land use would result in a lower ecological status of total P and Chl-a, as well as slightly lower ecological status of cyanobacteria. The effects of these two land use scenarios were most significant for total P: the probability of High-Good status decreased by more than 10 percent points (pp) from RCP8.5 (75.2% High-Good) to FW (63.8%), and likewise by more than 10 pp to TW (52.4%). The effects were less prominent for phytoplankton, where the probability of High-Good status dropped by only 4-5 pp from RCP8.5 (53.5%) to FW (49.1%) and likewise to TW (44.1%). The total lake status, which was determined primarily by phytoplankton and secondarily by total P, reflected the changes in phytoplankton status more closely than the changes in total P status. This result implies that even large reductions in total P may not result in corresponding improvements in ecological status, which is in accordance with experiences from Lake Vansjø (Couture et al. 2016).

The BBN results indicated that cyanobacteria concentrations shows very little response to change in scenarios and storylines. The difference in Chl-a status for the three storylines is not reflected in the cyanobacteria status. One reason is that the probabilities of concentrations $>1000 \mu\text{g L}^{-1}$ are very low (typically $<1\%$; not shown in plots). The probability of Cyano status Moderate or Poor-Bad (i.e. seasonal CyanoMax $> 1000 \mu\text{g L}^{-1}$) is higher (around 30%), but also shows very little response to the scenarios. Another reason for the lack of response is that the CPT for cyanobacteria contains high uncertainty. The ordinal regression model is based on relatively few observations (103 samples), of which only 13 samples have cyanobacteria concentration exceeding $1000 \mu\text{g L}^{-1}$. As a result of this uncertainty, the probabilities of different outcomes do not vary much between different scenarios. In future work with this BBN we will continue to investigate more options for constructing and updating this CPT, e.g. including data from other relevant lakes and expert judgement. Moreover, the two climate scenarios had little difference in wind speed (**Table 6**), therefore the probability distribution of the Wind speed node remained almost unaltered (**Figure 10a**). The wind speeds predicted by the climate model were usually calm (ca. 90%),

while the observed meteorological data for Lake Vansjø had only 56% calm days. This mismatch was probably due to insufficient downscaling of the climate model predictions to the local conditions of Lake Vansjø. Therefore, under the given climate scenarios, the stressor interaction between wind and lake temperature seen in exploratory scenarios (Figure 8e) did not emerge in different BBN runs.

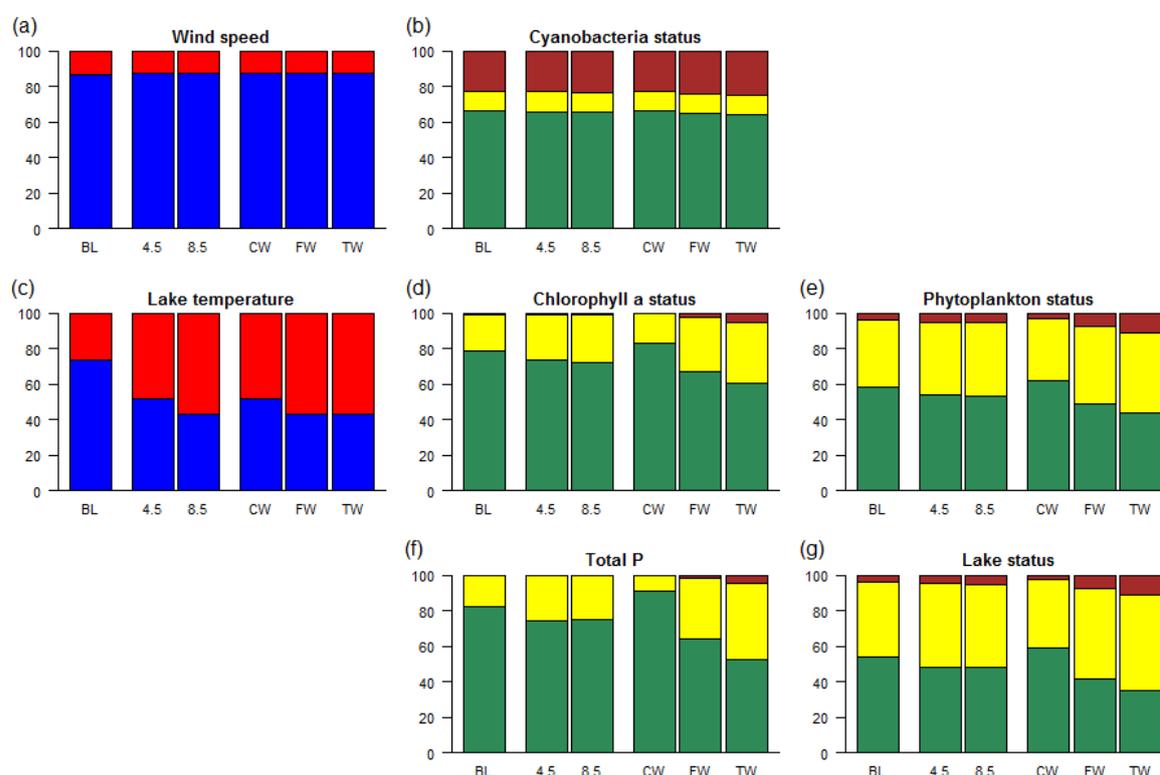


Figure 10 Results from selected nodes of the BBN model of Lake Vansjø, using outputs from MyLake. Shown are the projections for the various climate scenarios and storylines using the IPSL climate model, for the 20-year period 2050-2070. Colour codes: Temperature: blue = ≤ 19 °C , red = >19 °C. Wind speed: blue = ≤ 3.3 m/s, red > 3.3 m/s. Status nodes: green = High-Good, yellow = Moderate, brown = Poor-Bad. Cyanobacteria concentration: green = <1000, yellow = 1000-2000, brown > 2000 (µg L⁻¹). Scenarios: BL = extended baseline, 4.5 = climate scenario RCP 4.5, 8.5 = climate scenario RCP 8.5, CW = Consensus World, FW = Fragmented World, TW = Techno World.

Although the cyanobacteria did not show a clear response to the scenarios in the study, this variable still has an important role in the overall status classification of the lake. For all scenarios, cyanobacteria contributes to worsening the status assessed by phytoplankton, compared to Chl-a alone. In this way, the status of cyanobacteria, as predicted by the BBN, contributes to the overall response of phytoplankton status and of total lake status to the scenarios.

2.1.8 Validation

A previous version of the this model was evaluated by three critical aspects (Moe et al. 2016): (1) the link from process-based model predictions to observed values, (2) The CPT for cyanobacteria, and (3) the effects of water temperature.

(1) The accuracy of the MyLake model predictions varied highly among the different indicator variables. The model performance was discussed in detail by Couture et al. (2014); here we only considered the accuracy at the level of node states (intervals) and focus on the implications for the BBN model. Our decision not to include the mismatch between MyLake predictions and observations in the final BBN version can be justified by the fact that this uncertainty should already have been accounted for in the calibration of MyLake. The resulting 60 parameter sets were instead included as a source of uncertainty in the BBN. Incorporating the prediction – observation mismatch as an additional source of uncertainty would not only make the BBN model non-responsive to the scenarios, but also introduce a systematic error for TP.

(2) The CPT for cyanobacteria was a key aspect of the BBN, because this CPT provided the link from the abiotic to the biological components. Due to the limited number of cyanobacteria observations, to reserve a subset of the cyanobacteria data for evaluation purposes would not be meaningful. Instead, we used an independent dataset ("EUREGI") to construct an alternative CPT for cyanobacteria and compared the outcome of this version with that the original version. The EUREGI dataset gave similar probability distributions in the CPT for cyanobacteria to those from Lake Vansjø. Consequently, the model version with EUREGI data predicted effects of climate and management scenarios on ecological status of cyanobacteria that were very similar to the default model version. The fact that an independent, large-scale dataset gave similar CPTs and consequently very similar model predictions as the original data from Lake Vansjø strengthened our confidence in the cyanobacteria component of the model.

(3) A critical component of this BBN is the effect of water temperature on cyanobacteria. Moreover, since the conditional probabilities used for calculating posterior probabilities for cyanobacteria are based on very few observations for some of the parent state combinations, it is important to check that these CPTs do not provide spurious results. We therefore inspected more closely relationship between temperature, Chl-a and cyanobacteria by setting evidence (fixating probabilities) for the nodes Temperature and Chl-a. The results (Moe et al. 2016) showed that the model behaved as expected regarding seasonal variation in temperature and in indicator variables, and that the BBN generated reasonable predictions.

For the revised version of the model (Figure 17), one of the main changes is the inclusion of the node Wind speed as a parent node for cyanobacteria. The effect of wind speed in combination with temperature and Chl-a was explored separately (Figure) and provided

meaningful results. However, corresponding wind speed data are not available for the larger EUREGI dataset, therefore it is not possible to validate this revised part of the model using the larger dataset.

A BBN model can in principle be used in a diagnostic way. For this case study, this could be done by setting evidence (100% likelihood) for a given status (e.g. High-Good) for one or more status node and inspect the environmental conditions and scenarios leading to the selected state. However, the BBN model for Lake Vansjø was not designed for this purpose, and there is therefore a risk of obtaining results that are not informative. For example, setting evidence for different states of Cyanobacteria will alter the probability distributions for Period (time horizon) and for Months, unless these nodes are also fixed. It is possible to develop a BBN model also for diagnostic use for this case study, but it might not provide much insight beyond what is already known about this system.

2.1.9 Obstacles, pros and cons

There are several limitations associated with the BBN methodology in the context of environmental management. The fact that the non-dynamic network cannot contain loops puts constraints on the ecological processes that can be modelled; phosphorus and phytoplankton dynamics in lakes are typically dominated by feedback processes (Saloranta and Andersen, 2007). For example, high phytoplankton biomass can reduce the Secchi depth; on the other hand, lower Secchi depth can limit further phytoplankton growth due to light limitation. In our study, such feedback loops were handled by dynamic models (INCA-P and MyLake), while the BBN summarised the outcome of the catchment and lake process. Moreover, the accumulation of uncertainty with the length of the network implies that it can be difficult to draw conclusions from the final output nodes (Borsuk et al., 2004). Other challenges associated with the use of BBNs have been discussed previously (Landuyt et al., 2013; Uusitalo, 2007; Varis and Kuikka, 1999).

Compared to existing process-based models for ecological status of rivers and lakes, the BBN approach provides an opportunity to include biological elements, as demonstrated by our study. Even when data are sparse, theory or expert knowledge on selected biological indicators can be used as a first step to construct causal links (CPTs) between abiotic and biotic responses. Since the WFD requires that assessments are based primarily on biology (EC, 2000), this is clearly an added value for use of models in water management in Europe. Moreover, the WFD requires that potential impacts of climate change are considered in the next set of river basin management plans (EC, 2000). Although much knowledge is available on effects of climate change on ecosystems, including specific effects on biological quality elements in lakes (Moe et al., 2016), incorporating such information in predictive models is a challenge. The BBN methodology can facilitate the use of such knowledge, manifested as

expert judgement of probabilities under given climatic scenarios. Furthermore, a BBN model may be relatively easy to understand for end users who do not have any modelling background (Marcot et al., 2006). Therefore, BBNs are promising tools for supporting informed decision making and thus the work of water managers.

2.2 Finland: Lepsämänjoki

2.2.1 Introduction

The Finnish case study represents the river basins in Southern Finland where main soil types are till and clay (Figure 10). These soil types have been formed from acidic bedrocks due to melting processes during ice age, forming clay fields in the lowland areas in the middle of more coarse upland areas. These clays are unstable and thus vulnerable to erosion due to low calcium content. On these soil types there is productive agriculture, mainly spring cereal cultivation. Around larger towns there are also pressures from intensified urban land use.

The relatively small study catchment (Lepsämänjoki) in the river basin Vantaanjoki alone did not contain enough data for statistical analysis performed in WP4 and thus a larger set of representative river basins in southern Finland (Figure 11) were included in the analysis (Couture et al. 2016). This BNN modelling is based on all river basins of Rankinen et al. (2016) which has river formations dominantly on clay soil types. These river basins cover a wide range of different land uses (Table 7).

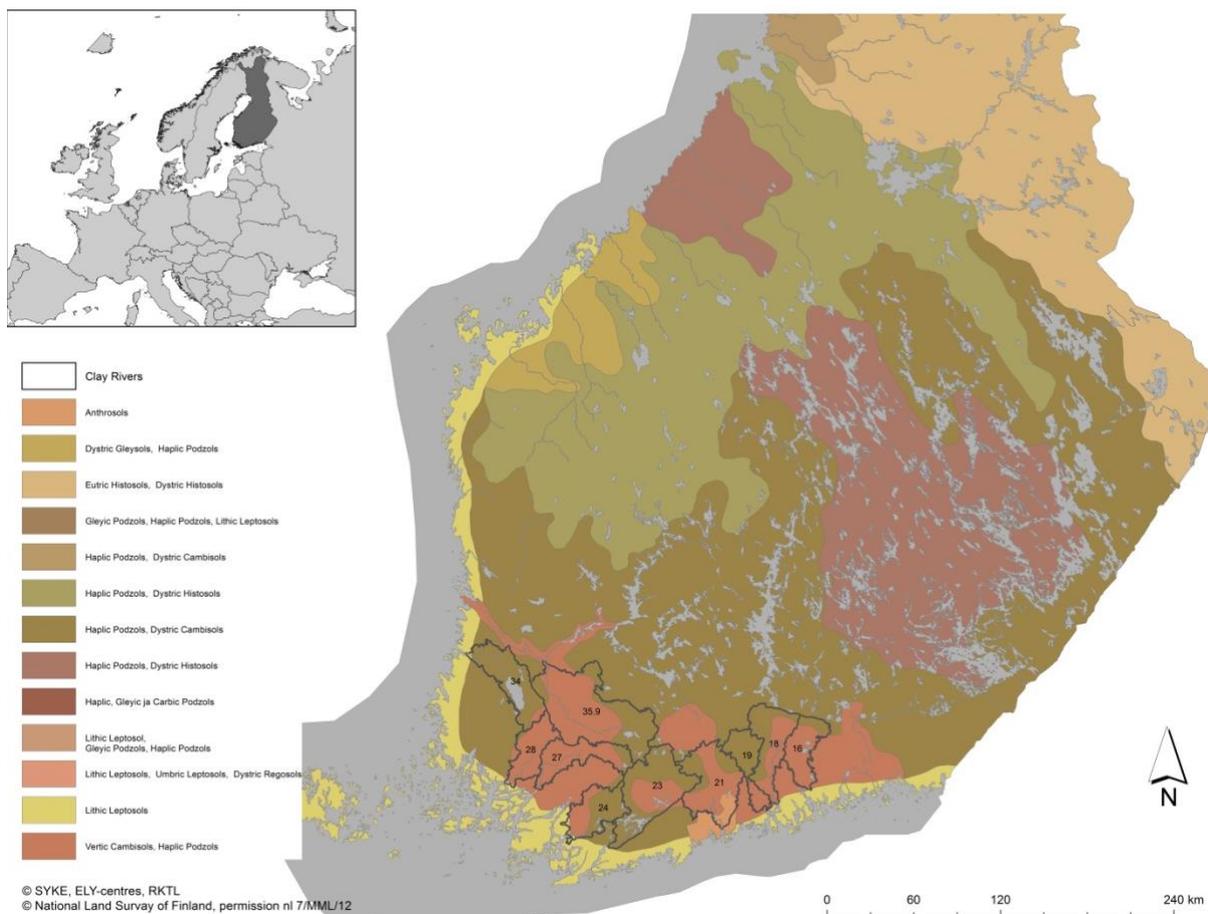


Figure 11. Location of the clay type river basins

Table 7 Land use in the river basins (Corine 2006: for details, see Haakana et al. (2008); Sucksdorff et al, 2001). Artificial = areas paved by humans (e.g. towns, roads, etc)

Code	River	Catchment area (km ²)	Lakes (%)	Fields (%)	Artificial (%)
16	Koskenkylänjoki	895	4.25	30.92	6.30
18	Porvoonjoki	1273	1.42	31.77	10.77
19	Mustijoki	783	1.61	31.04	9.15
21	Vantaanjoki	1686	2.19	24.11	20.69
23	Karjaanjoki	2046	11.66	18.44	9.54
24	Kiskonjoki	629	5.76	23.48	7.41
27	Paimionjoki	1088	1.85	42.92	8.34
28	Aurajoki	874	0.44	37.15	12.17
34	Eurajoki	1336	12.89	23.44	7.16

Ecological status of the clay type rivers is typically moderate or poor. In this type only 10% of the rivers has good ecological status (Figure 12). The classification is based on deviation of biological quality elements and supporting elements (Total Phosphorus (TP) concentration and hydro-morphological alteration) from their type-specific reference conditions (Aroviita et al. 2012).

The stressors of interest are runoff and TP concentration as well as summer temperature of river water. Changes are driven by climate and land use. In previous studies temperature and TP concentration seem to have a synergistic influence on Chl-a concentration (WP4 statistical analysis). Runoff and water temperature also interact because slowly flowing water warms up more rapidly than fast flowing water.

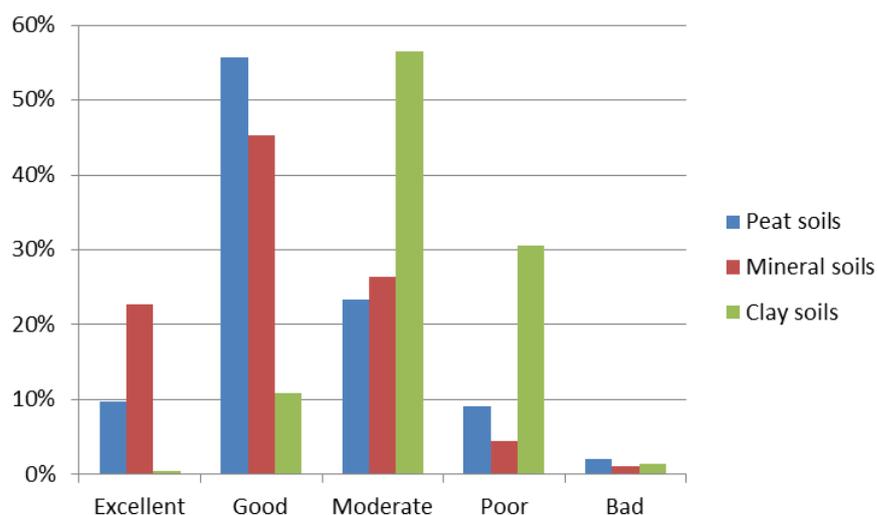


Figure 12 Ecological status classification of peatland, mineral land and clay soil river types in Finland (Peat N = 972, Mineral N = 640, Clay N = 159).

2.2.2 Purpose

The purpose of the BBN model was to link the outcome of the process-based eco-hydrological model to ecological status under different climate and land-use scenarios. Process-based models typically do not predict biological responses other than the Chl-a concentration. Chl-a concentration is not an official classification indicator in Finland, but it is included in the BBN, because it is a MARS benchmark indicator (BInd08). The boundary limits for Chl-a used in this study are those determined for lakes. The BBN model is based on observed river temperature, TP concentration and Chl-a concentration, land use and hydro-morphological modification of the rivers and river basins. Class boundaries are based on Jenks natural breaks method that seeks to reduce the variance within classes and maximize the variance between classes

2.2.3 Model construction

The conceptual model of the Finnish BBN is constructed based on expert judgement, previous studies and statistical analysis of indices and stressors from WP4. It follows the simplified structure of MARS conceptual model to reduce the amount of nodes and thus uncertainty.

The BBN model is based on those variables that were found influential in previous statistical analysis of WP4.

Present day situation is based on long time series of land use, climate and nutrient concentrations in Finnish rivers (Rankinen et al. 2016) and observed ecological status of the rivers. Conditional probabilities between stressors and scenarios were derived using Persist and INCA simulation results. *A posteriori* data is a sample (N = 40) representing rivers on clay soil types. Rivers in this sample flow to the Baltic Sea and have continuous discharge measurement, which allows accurate runoff calculations. In the validation the rest of river water bodies on clay soil types were used (total N = 119) (Table 8 and Table 9).

Statistical analyses and hydro-ecological modelling is described in detail by Couture et al. (2016) in Deliverable 4.1-3 Case study synthesis: Report on case studies from Northern river basins.

Table 8 Hydromorphological modification

	Not highly modified	Highly modified
Clay rivers	67%	33%
Sample	75%	25%

Table 9 Agricultural field percentage

	<10%	10-25%	>25%
Clay rivers	4%	32%	64%
Sample	4%	44%	52%

The BNN is presented in **Figure 13**.

According to BNN the ecological status is worse than good (WTG) in 68% of the cases. Ecological Quality Ratio (EQR index = Observed value / Reference value) is calculated from data on macrophytes, macroinvertebrates and fish). The EQR index shows better status than what is estimated by TP concentration. (EQR = Observed value / Reference value). The class boundary of TP concentration between good and worse than good is $60 \mu\text{g l}^{-1}$.

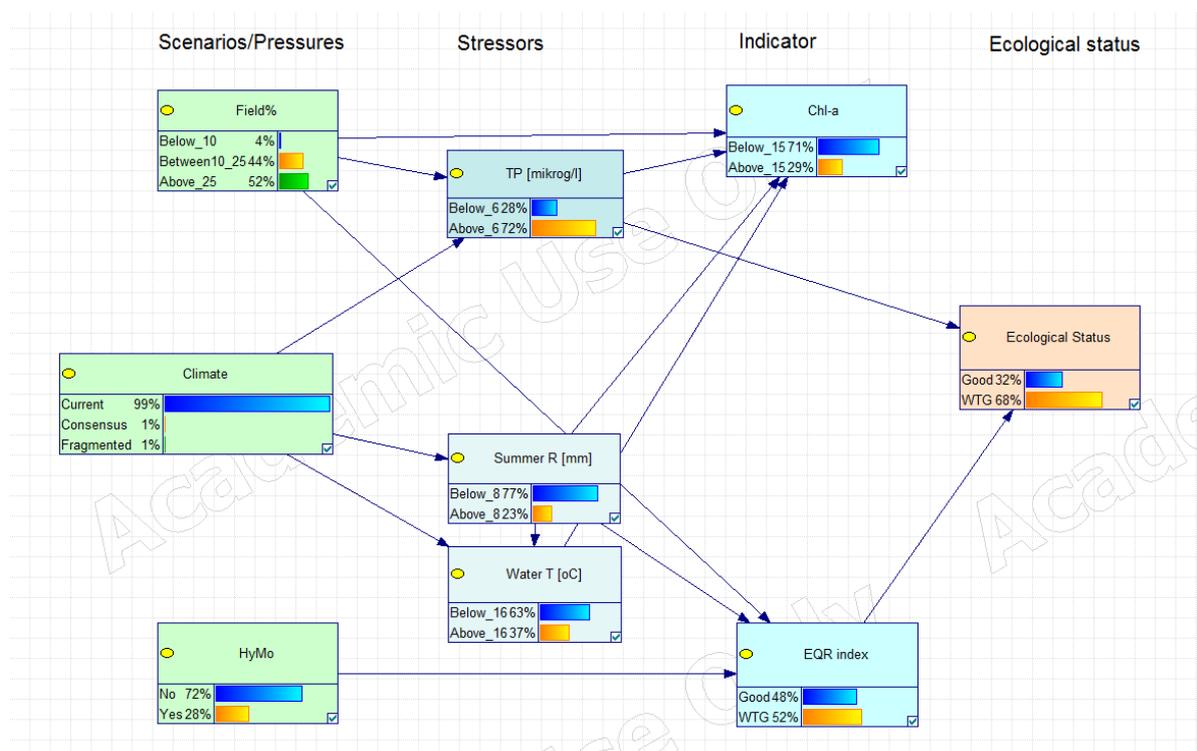


Figure 13. BNN model for the Finnish clay type rivers (WTG= worse than good)

Model results were validated against independent data that was not used in construction of the model (Table 9). In validation data 43% of river formations did not have biological classification, and 8% did not have physical-chemical classification. Hydromorphological modification is missing from 8% of the validation data. Model results showed slightly lower percentage of river basins not to achieve good ecological status than validation data based on physical-chemical classification. On the contrary, based on biological classification model

results were 20% higher than in validation data. Taking into account the high number of missing data in river basins used in validation, this difference may be acceptable.

Table 10 Percentage of river formations that do not achieve good ecological status

Indicator	Model result	Validation data
Physical-chemical class	72%	77%
Biological class	52%	31%

2.2.4 Conditional probability tables

The discrete probability distributions in the CPTs are obtained by different approaches in the different BBN modules. The linkage between field percentage and TP concentration is based on empirical data (Rankinen et al. 2016), as well as that of ELS index (Rääpysjärvi et al. 2016) and Chl-a concentration (unpublished data). Linkage between hydro-morphological modification and ELS index is based on empirical data of Turunen et al. (2016). Linkages between summer runoff and water temperature in different climate scenarios are based on INCA and Persist modelling in WP4. Finally, linkage between ecological status, biological index (ELS) and physicochemical index (TP) is based on data the second planning period. Conditional probability tables for Ecological Status are presented in Table 11.

Table 11 Conditional probability table of ecological status

ELS	Good	Worse than good	Worse than good	
TP	Good	Worse than good	Good	Worse than good
Good	0.83	0.3	0.21	0.2
Worse than good	0.17	0.7	0.79	0.8

2.2.5 Scenarios

According to climate change scenarios both temperature and precipitation is assumed to increase favouring agricultural production in Southern Finland. Nevertheless, in BNN climate change and land use change are separated to make the model more general. Increase of agricultural field area may not be an opportunity in some of these relatively small river basins, due to less suitable soil types for agriculture, or competition with urbanization. MARS storylines are Consensus and Fragmented World, which used climate change scenarios GFDL, RCP4 for Consensus and RCP8 for Fragmented for 2025-2034.

- Consensus world: the main objective of the government and citizens is to stimulate economic activity but also to promote sustainable and efficient use of resources. The current guidelines and policies are continued. As future climate is assumed to favour agricultural production by increasing yields in Finland, field percentage is assumed to increase, limited only by soil types and field slopes (>10%) which are not suitable for cultivation.

- **Fragmented world:** The focus of this storyline is to survive as a country instead of as part of Europe. National institutions focus on economic development and no attention is paid to the preservation of the ecosystems. In this storyline field area is assumed to increase up to 90% of the sub-catchment area as future climate favour agricultural production. As current environmental guidelines are not valid, the main production type will be monoculture of cereals with increased fertilization level. As the catchment is located relatively close to Helsinki, also increase in human settlements is assumed.

2.2.6 Results

Temperature increase changes ecological status less than land use change (Figure 14). When climate change is combined by radical increase in field area, ecological status will decrease (Figure 15a). When high temperature increase is combined with increase in field area and increase in hydro-morphological modifications, up to 77% of the sites will show worse than good status (Scenario Fragmented). Chl-a concentration changes in the same direction as ecological status, but the message is in general more optimistic (Figure 15b). It is worth noticing that in the BNN model there is no connection between hydro-morphological status and Chl-a concentration due to a lack of data.

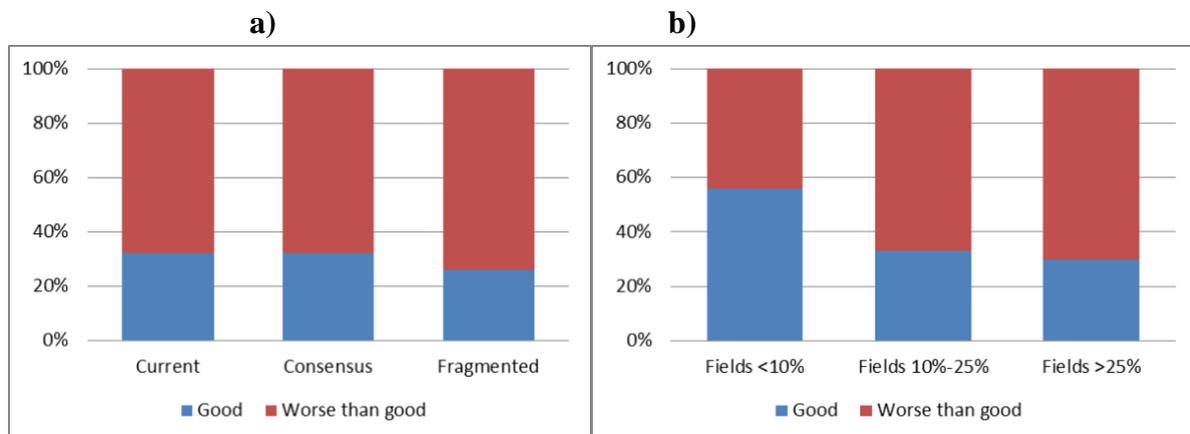


Figure 14 Effect of pressure on ecological status a) climate change only, b) land use change only

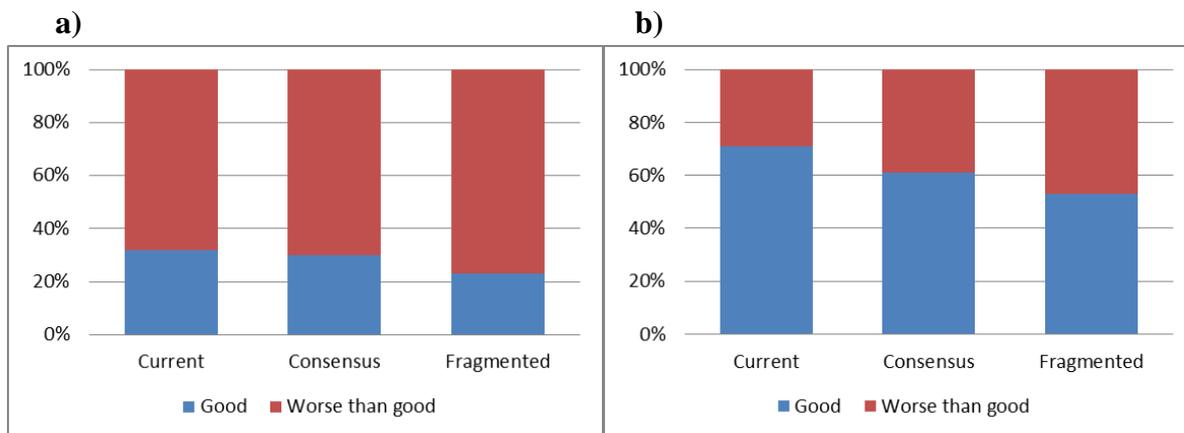


Figure 15 Deviation of a) ecological status, and b) Chl a concentration in different scenarios

2.2.7 Pros and cons

In BBN the option to use expert opinion makes it a useful tool in cases when the quantity or quality of measured data does not allow specific statistical analysis. The method allows combining data from different sources and levels, expert opinions and literature values. The results are not absolutely accurate, as they inclusively contain the uncertainty of input data. On the other hand, the better data is used the more reliable are the model results.

In this specific case study the BNN was used as an upscaling tool. Effect of climate change on water amount and quality was modelled in well studied relatively small catchment, where data enough for detailed dynamic eco-hydrological modelling was available. These results were upscaled to larger set of river basins with similar climate, geology and agricultural production by BNN by combining them with Chl-a and water quality data from this larger area.

BBN needs simplification of the conceptual model of the problem. Simplification makes it more visible for researchers and stake holders. For stake holders also thinking by probabilities may be a familiar pattern. On the other hand, BBN is a static model which does not say anything of the time needed for change, a question that stake holders are often interested in.

2.3 Denmark: Odense

2.3.1 Introduction and purpose

The Odense Fjord basin is one of those under multiple pressures. Located in the island of Funen (Denmark), it has an area of approximately 1,100 km² (Figure 16). Aquatic ecosystems in the basin include lakes, rivers and transitional waters (the Odense Fjord is an estuary). In spite of several action plans, many of these waters do not meet the Water Framework Directive criteria of good ecological status (Miljø- og Fødevarerministeriet, 2017). This status has been conditioned due to urbanization, hydro-morphological modifications (channelization and tile draining), summer droughts, groundwater abstraction and fertilizers and pesticides from the agricultural sectors. More information about the study area can be found in Ferreira et al. (2016) and Molina-Navarro et al. (2018).

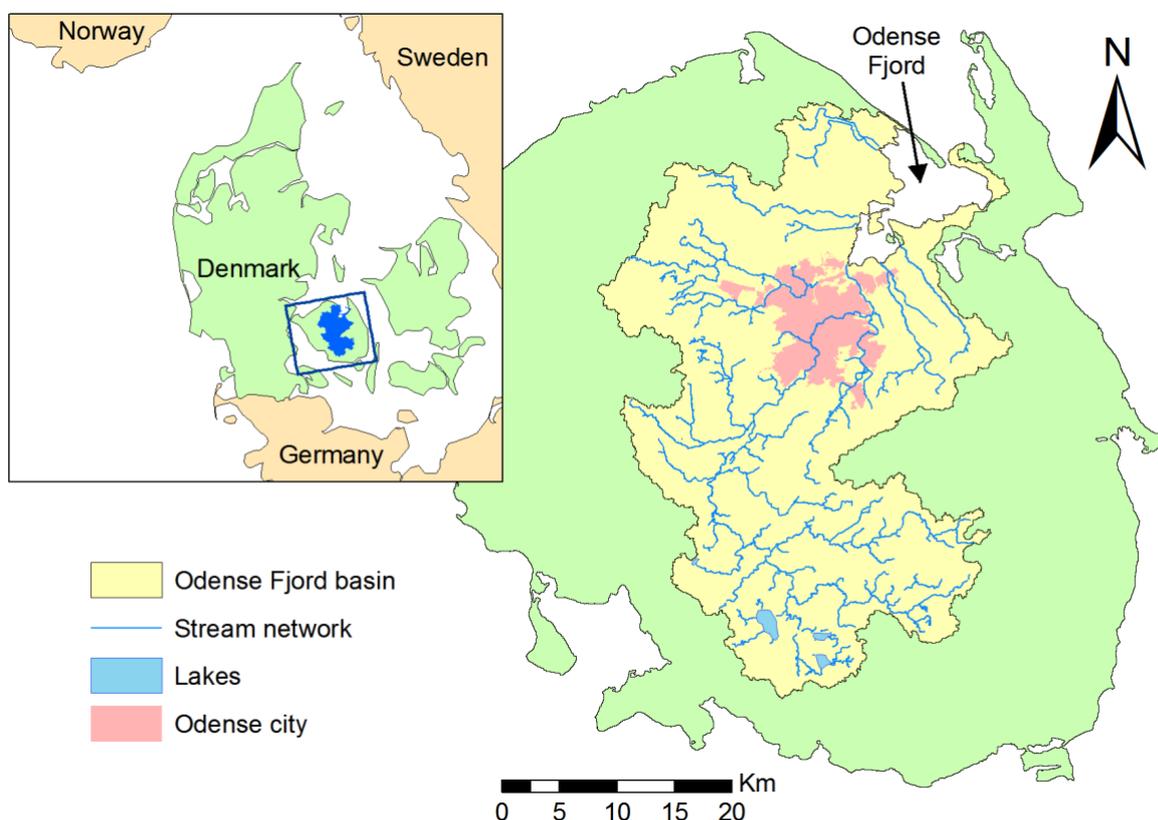


Figure 16 Location of the Odense Fjord basin and its main aquatic ecosystems. The extent of Odense, the main city in the study area, is also depicted.

Previous work in the Odense Fjord basin within the MARS WP4 (<http://fis.freshwatertools.eu/index.php/odense.html>) involved the set-up of a process-based hydro-ecological catchment model with the Soil and Water Assessment Tool (SWAT) (Molina-Navarro et al., 2017, 2018). Besides, MARS storylines (Faneca Sanchez et al., 2015) were used to design three future scenarios to be simulated with SWAT. Since the study area

is eminently agricultural, the MARS storylines were downscaled to the Odense catchment focusing on the farming context:

- **Storyline 1 - Techno world (TW):** Agricultural area remains similar, with some conversion to permanent grass and willow. Slight increase in livestock density and slight decrease in artificial fertilizer application rates. Climate change scenario RCP 8.5 (Representative Concentration Pathway 8.5), a rising scenario with very high greenhouse gas emissions, is assigned to this storyline.

- **Storyline 2 - Consensus world (CW):** Agricultural area decreases and changes towards forest and less intensive farming types. Artificial fertilizer application decrease slightly as well. RCP 4.5 climate change scenario, a stabilization scenario, was assigned to this storyline.

- **Storyline 3 - Fragmented world (FW):** Agricultural area increases and changes towards intensive pig and dairy farm types. Livestock density and fertilizer application increase. Again, climate change scenario RCP 8.5 rules in this storyline.

The time horizons for scenarios simulation, explained in Faneca Sanchez et al. (2015), are 2030 (interval 2025-2034) and 2060 (2055-2064). More detail on the scenarios implementation and simulation can be found in the “Odense” chapter from the MARS Case study synthesis final report (D4.1, Ferreira et al., 2016), where scenarios were given agriculture-related names since the downscaling is focused on the farming context. Thus, the downscaled scenarios for techno, consensus and fragmented world MARS storylines were named “High tech agriculture” (HT), “Agriculture for nature” (AN) and “Market driven agriculture” (MD).

The SWAT model is able to provide simulated values for a number of variables (flow, nutrient concentrations) that might act as water ecosystems stressors, since they can affect ecosystem quality. Additionally, empirical models were developed to establish mathematical relationships between physico-chemical stressors and biological indicators (fish, macrophytes and macroinvertebrate indices) of stream ecological status in Denmark (“Odense” chapter in Ferreira et al., 2016).

A BBN modelling approach can be useful to combine the process-based modelling results with the existent data used to develop the empirical modes. That is, it can help to establish a relationship between those stressors derived from SWAT (and that were seen as relevant for predicting biological indices in the empirical modelling process) and the biological indicators, so the probability of a certain ecological status can be derived. Thus, the ultimate purpose of the BBN modelling in the Odense Fjord catchment is to provide a graphical, simple but effective tool to assess the impact of future scenarios in the ecological status of the streams in the catchment.

The BBN modelling in the Odense Fjord catchment comprises a multiple stress framework, evaluating the impact of land use, agricultural management and climate changes over both physical and chemical variables, which in turn act as stressors for the biological variables that indicate the ecological status of the streams in the catchment. Such a modelling framework might be used as a decision-support tool when designing water management policies or applying mitigation actions for those multiple stressors and their interactions.

2.3.2 Model construction

The BBN model for the Odense Fjord catchment was designed with the software GeNIe, created by BayesFusion and freely available for academic and scientific use (www.bayesfusion.com). The network designed aims to connect a number of physico-chemical variables (“stressors”) with biological variables (“indicators”) that act as ecological status indices, following the EU Water Framework Directive criteria (European Parliament and Council, 2000).

Among the stressors seen as relevant for predicting the different biological indicators in the empirical modelling work (Ferreira et al., 2016, Andersen et al., forthcoming), only those that can be derived from SWAT at a reach level were selected to be included in the BBN. As a result, four variables dependent on the hydrological regime and two related with nutrients were selected (Table 12). Despite not being identified as a significant stressor in the development of the empirical models, total nitrogen concentration was also added to the BBN. It might serve as a chemical water quality proxy (nitrogen loads to aquatic ecosystems is one of the main environmental problems in Denmark (Kaspersen et al., 2016), and its inclusion in the BBN allowed evaluating how the different scenarios might affect it.

Table 12 Variables selected as stressors for BBN modelling in the Odense Fjord catchment.

Stressor	Description
BFI	Baseflow index, defined as baseflow volume divided by total volume
Q90	Flow below the 90 th percentile* of the flow-duration curve divided by median flow (Q ₅₀)
FRE25	Annual frequency of flow events above the 25 th percentile of the flow-duration curve**
DUR3	Annual duration of extreme flow events three times above the flow at Q ₅₀ (days)
TP	Annual mean concentration of total phosphorus (mg/L)
TN	Annual mean concentration of total nitrogen (mg/L)

* 90th percentile = the flow value that is exceeded 90 % of the time, i.e. a low flow indicator.

** 25th percentile = the flow value that is exceeded 25 % of the time, i.e. a high flow indicator.

Among these stressors, TP and TN are also MARS Benchmark Biological Indicators (BInd2 and BInd3, Birk et al., 2015). Besides, DUR3 is related to MARS BInd4 (mean duration of

high pulses within each year), since both represent duration of high pulses, although defined different.

Regarding the indicators, those covered in the empirical modelling work were included in the BBN. Namely, three biological indices of ecological status (for fish: DFFV, macrophytes: DVPI and macroinvertebrates: DVFI) designed specifically for Denmark by their Danish acronym) and one of the MARS Benchmark indicators, the Average Score Per Taxon (ASPT, BInd12). More information about the Danish indices can be found in Kristensen et al. (2014, fish), Larsen et al. (2014, macroinvertebrates) and Larsen and Baattrup (2015, macrophytes). The Ecological Quality Ratio (EQR, observed index divided by its reference value) of each index was used.

Finally, indices values were translated into ecological status classes (bad, poor, moderate, good or high) in the network, following the EU Water Framework Directive guidelines (European Parliament and Council, 2000). A final node combining the ecological status classes (worst case classification) given by the three Danish indicators closed the network, which actually is the MARS Benchmark indicator BInd1. ASPT was excluded in the status class definition because there are no boundaries defined for Danish streams, but it was kept to maintain a MARS Benchmark indicator in the indicator level of the network and to compare it with the Danish macroinvertebrate index.

The BBN created can be seen as a simplified version of the MARS conceptual model for the catchment. It also included as drivers the agriculture and climate changes derived from SWAT scenarios, as pressures some SWAT outputs related with the hydrological alteration besides TP and TN concentrations, and as biotic state indicators the ASPT index and the ecological status Danish indices, besides the final ecological status. However, the MARS conceptual model contains other elements not included in the BBN, such as outputs from lake modelling or ecosystem services and responses. For further details, see the MARS conceptual model for the Odense Fjord catchment in Ferreira et al. (2016).

2.3.3 Class boundaries

Once the structure of the network was designed, the variables represented in network nodes needed to be categorized. Three levels, “Low”, “Medium” (or “Med”) and “High”, were defined for every stressor (Table 13). The boundaries were determined taking into account the whole dataset for Danish streams that was also used for empirical modelling (Ferreira et al., 2016, Andersen et al., forthcoming). This way, the boundaries represent a whole range of field conditions that can determine the subsequent indicator values (conditional probabilities between stressors and indicators are also obtained from the national dataset, due to lack of observed data in the Odense catchment, as explained later). The discretization was done using the “Uniform counts” tool in GeNIe, which creates three classes with the same number of data points each.

Table 13 Threshold values used for nodes (stressors and indicators) categorization.

	Stressors: Levels		
	Low	Medium	High
BFI	< 0.34	0.34 – 0.62	> 0.62
Q90	< 0.65	0.65 – 0.79	> 0.79
FRE25	< 7.4	7.4 – 10.6	> 10.6
DUR3	< 3.4	3.4 – 6.2	> 6.2
TP	< 0.10	0.10 – 0.14	> 0.14
TN	< 3.5	3.5 – 5.1	> 5.1
	Indicators: Ecological status classes (EQR)		
	Poor/Bad	Moderate	High/Good
Danish Fish Index	< 0.40	0.40 – 0.72	≥ 0.72
Danish Macrophyte Index	< 0.35	0.35 – 0.50	≥ 0.50
Danish Macroinvertebrate Index	< 0.32	0.32 – 0.52	≥ 0.52
ASPT	< 0.77	0.77 – 0.89	≥ 0.89

Regarding indicators, three classes were established too (Table 13), considering the upper and lower threshold values for the “Moderate” ecological status class of each index (its ecological quality ratio, EQR) according to the literature (Kristensen et al. 2014, Larsen et al. 2014, Larsen and Baattrup, 2015). Thus, the classes’ boundaries for each indicator correspond to the “Poor/Bad” (PB), “Moderate” (M) and “High/Good” (HG) ecological status classes, and so they were translated in the third level of the network. For ASPT, however, no boundaries were found for Danish streams. Thus, ASPT node was not connected to the third level of the network, but United Kingdom thresholds (Birk and Hering, 2006) were used to include it in the “indicators” level for comparison purposes.

2.3.4 Conditional probability tables

Conditional probabilities between stressors and scenarios were derived using SWAT simulation results. For each stressor, the average value for each reach modelled in the basin (31 reaches) in a 10-years model run was considered. The stressor’s inclusion in the network was based on those that were found relevant in the empirical modelling work, and this work used time-series averages (Ferreira et al., 2016, Andersen et al., forthcoming), so the same caveat had to be applied here.

Real data from the Odense Fjord catchment to calculate probabilities between stressors and indicators is insufficient. Thus, CPTs in this case were derived from the aforementioned Danish streams dataset, which contains real data from 131 streams. Denmark as a whole is a lowland country (Windolf et al., 2011), so the use of a national dataset might represent no concern in this sense. For the macrophyte index, two stressor-indicator classes combinations had no data available (low DUR3 and low Q90; high DUR3 and high Q90). Probabilities of the closest neighbour combination were assigned in these cases, following Moe et al. (2016).

The ecological status classes have a 1:1 correspondence with the indicator classes. The final ecological status class probability (last network node) was derived taking into account the worst class obtained within the three indicators, following the “one out, all out” principle established in the EU Water Framework Directive (Van de Bund and Solimini, 2007).

2.3.5 Scenarios

In order to facilitate results extractions and comparison, and following the outline of the Odense catchment report elaborated for MARS WP4 (Ferreira et al., 2016), three versions of the BBN for the Odense Fjord catchment were created:

- **Version 1:** Considers only land use changes (LUC) in MARS storylines, and not climate changes, to account for the isolated effects of land use changes within scenarios. Thus, only LUC scenarios were run in SWAT with observed climate (2001-2010) (TW_OBS, CW_OBS and FW_OBS), including a fourth baseline scenario run with present land use (PLU_OBS).
- **Version 2:** Includes the baseline scenarios for the MARS storylines, i.e. present land use and projected climate data from 2011-2020 for both RCPs (PLU_4.5, baseline of CW; PLU_8.5, baseline of TW and FW).
- **Version 3:** Models the MARS storylines downscaled to the Odense Fjord catchment for time horizons 2030 (TW_30, CW_30, FW_30) and 2060 (TW_60, CW_60, FW_60), as described above.

It might be highlighted that in this report the original MARS storyline nomenclature, i.e. “Techno world” (TW), “Consensus world” (CW) and “Fragmented world” (FW), is kept to facilitate comparison with other BBN study cases. However, readers must take into account that downscaled scenarios in Odense for techno, consensus and fragmented world MARS storylines were named “High tech agriculture” (HT), “Agriculture for nature” (AN) and “Market driven agriculture” (MD) in other Odense catchment studies derived from this project (e.g. “Odense” chapter in Ferreira et al., 2016; Andersen et al., forthcoming; Molina-Navarro et al., 2018).

Although the initial Odense catchment report for MARS WP4 considered two climate models, GFDL-ESM2M and IPSL-CM5A-LR; for BBN modelling purposes only the second climate model was used to simulate climate change in the catchment, because it yielded the best median output both regarding cumulative precipitation and cumulative runoff relative to observations (MARS internal document).

2.3.6 Validation

Probability predictions of stressors from scenarios can be considered as verified since CPTs were created with a process-based SWAT model that was already calibrated and validated (Ferreira et al., 2016, Molina-Navarro et al., 2017). Similarly, stressors selection influencing indicators was based on abiotic-biotic statistically significant models (Ferreira et al., 2016, Andersen et al., forthcoming) that were created with the same dataset used for CPTs calculation, which somewhat guarantees the reliability of the BBN. However, it must be acknowledged that not all the stressors accounted in those empirical models were included in the BBN, so an additional validation of BBN probability results would be desirable. Such a validation was addressed comparing the probability distribution for ecological status classes calculated by the BBN for the PLU_OBS scenario with the real data published in the latest Odense Fjord basin management plans (Miljøministeriet, 2011; Miljø- og Fødevarerministeriet, 2016, 2017).

2.3.7 BBN modelling

Probability distributions for the ecological status classes in each scenario were obtained setting the corresponding evidence in the BBN and updating the beliefs. First, the effects of isolated land use changes were evaluated (BBN version 1, see Appendix **Error! Reference source not found. Error! Reference source not found.**) comparing TW_OBS, CW_OBS and FW_OBS with PLU_OBS. Then, the effects of future storylines were analysed comparing the probabilities obtained in the scenarios with their respective baselines (BBN versions 3 and 2 respectively, see Appendix **Error! Reference source not found. Error! Reference source not found.**).

Posterior probabilities were explored setting evidence for a time horizon and for a certain ecological status in the final node of the BBN, updating the beliefs afterwards. This allowed analysing which scenario might be more probable given a certain ecological status.

2.3.8 Results and discussion

BBN design

All the considerations in the previous section resulted in the network design shown in Figure 17. Additionally, in Appendix **Error! Reference source not found. Error! Reference source not found.**, the three versions created for this network are depicted and the CPTs for each node are provided. It must be acknowledged that an initial draft of the BBN included an additional linkage, because the empirical modelling showed that Q90 was also found relevant for the Danish fish index prediction (Ferreira et al., 2016, Andersen et al., forthcoming).

However, this caused that many combinations in the CPT for the index had no observed data, so the calculation of probabilities based these data was not possible. Thus, in order to get a more consistent CTP, considering that both Q90 and BFI are related with baseflow, and that Q90 showed a weaker influence than BFI for the fish index prediction (Ferreira et al., 2016, Andersen et al., forthcoming), the Q90-Fish index linkage was deleted.

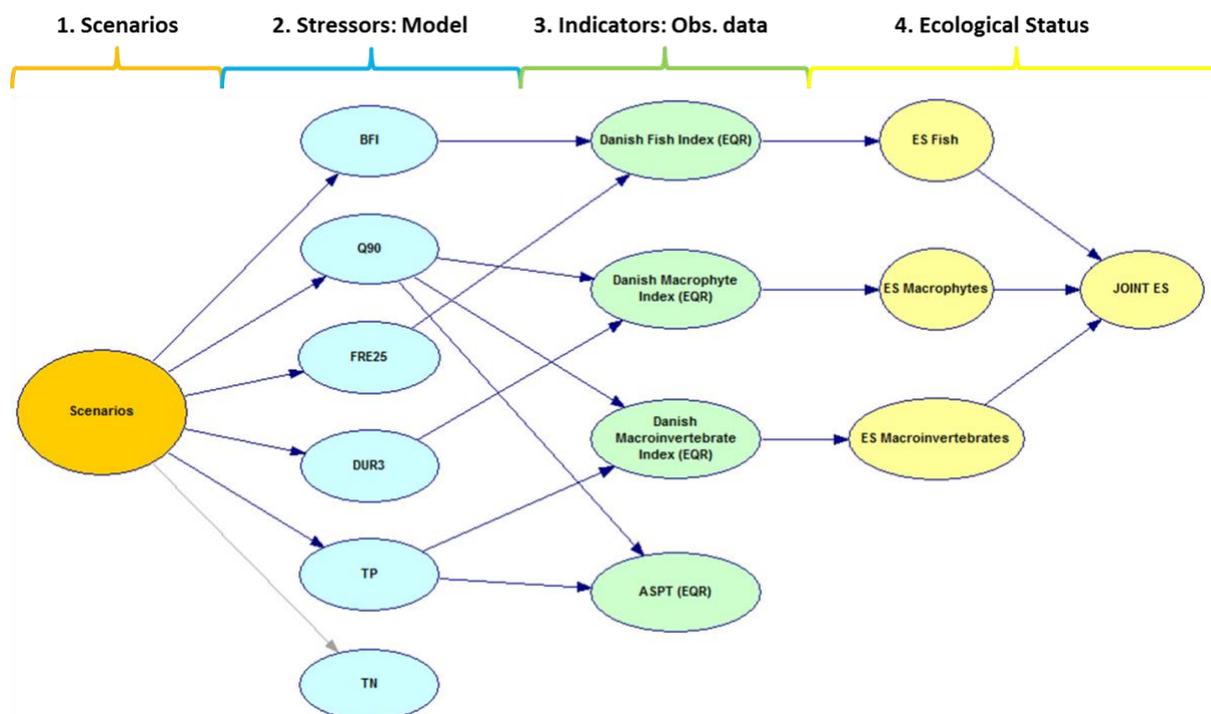
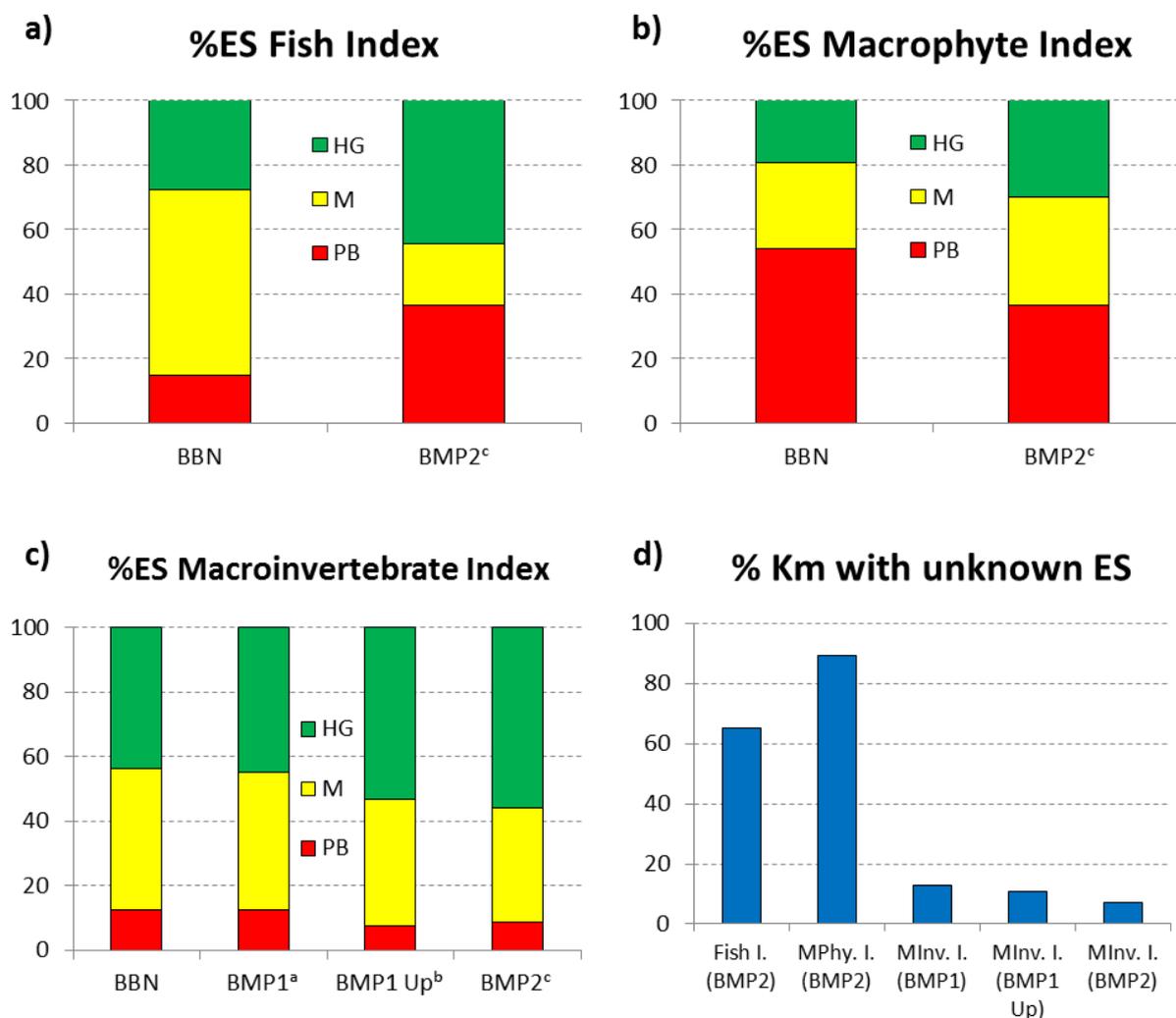


Figure 17 Bayesian Belief Network (BBN) designed for the Odense Fjord catchment. It contains four levels of nodes: scenarios (orange), stressors (blue), indicators (green) and ecological status (yellow), as well as their respective dependences (arrows).

Validation

Figure 18 shows the probability distribution of ecological status classes calculated by the BBN for the PLU-observed climate data (2001-2010) scenario and the real probabilities derived from real data published in the latest Odense Fjord basin management plans (Miljøministeriet, 2011; Miljø- og Fødevarerministeriet, 2016, 2017). Unfortunately, the availability of Danish fish and macrophyte indices data in the Odense Fjord catchment is scarce, and only for the second management plan (unknown ecological status reported for 65% and 89% of the total length of the stream network, Figure 18d). Thus, a reliable comparison of BBN results with observed data cannot be done for these two indices. Regarding the Danish macroinvertebrate index, results show that the BBN modelled probability distribution is very similar the one derived from observed data, validating the modelling approach. It has to be acknowledged, however, that the observed data refers to a percentage of river length

with a certain ecological status, while the BBN was elaborated with river segment data (sub-basins), regardless its length.



^a Data from the first Odense Basin Management Plan (2010-15, includes data from 2003-10, Miljøministeriet, 2011).

^b Data from the first Odense Basin Management Plan updated in 2013 (median from 01-2008, GIS corrected for the SWAT delineated catchment; Miljø- og Fødevarerministeriet, 2017).

^c Data from the second Odense Basin Management Plan (2015-2021, includes latest data up to 2012 -2013 for fish-, GIS corrected for the SWAT delineated catchment; Miljø- og Fødevarerministeriet, 2016, 2017).

Figure 18 Modelled (BBN) and observed (BMP, Basin Management Plan) probability distribution of ecological status classes for fish (a), macrophytes (b) and macroinvertebrates (c), and percentage of river length without information for the observed data (d, I.=index, MPhy.=Macrophyte, MInv.=Macroinvertebrate).

BBN modelling: Scenario´s simulation results

Isolated Land Use Change (LUC) scenarios

First, the isolated effects of land use change (LUC) scenarios were explored. The probability distributions of different ecological status across LUC scenarios were very similar, showing just discrete variations (Figure 19). The fish index showed slight differences: the ecological status slightly moved towards the extreme classes in TW_OBS (mainly towards HG) and in CW_OBS (mainly towards PB), while it moved to the moderate status in FW_OBS. This matches with slight variations seen in BFI and FRE25 across scenarios (Appendix **Error! Reference source not found. Error! Reference source not found.**). Regarding the macrophyte index, the probability of PB status decreased in the TW_OBS scenario, increasing the probabilities of M and HG status (Figure 19). This might be a response to a DUR3 decrease. Changes in other scenarios, however, were minor. For the macroinvertebrate index, probability distributions across scenarios were virtually the same. As a result of the combination of the three indices under the “one out, all out” principle (Van de Bund and Solimini, 2007), probability distributions for the final ecological status classes remained very similar across scenarios (Figure 19). Thus, results observed suggested that isolated LUC scenarios might not affect much the ecological status of streams in the Odense catchment.

These results are in agreement with those obtained in MARS WP4 for the Odense catchment. Ecological status nodes in the BBN are dependent on four stressors associated with hydrological regime and TP. Process-based modelling results (Ferreira et al., 2016, Molina-Navarro et al., 2018) showed that changes in land use alone affected catchment runoff only to a small degree. TP concentration is to a large extent driven by hydrology, so it was also unaffected by the land use scenarios. The empirical abiotic-biotic modelling also showed very small effects of land use change alone for the EQR value of the four indices, which only varied at the second decimal level for the 31 sub-basins averages (see table 5.38, Ferreira et al., 2016). Thus, regarding LUC scenarios, BBN results were consistent with previous findings in the catchment. This gives credibility to the network, and allows addressing the next step, modelling combined LU and climate change scenarios.

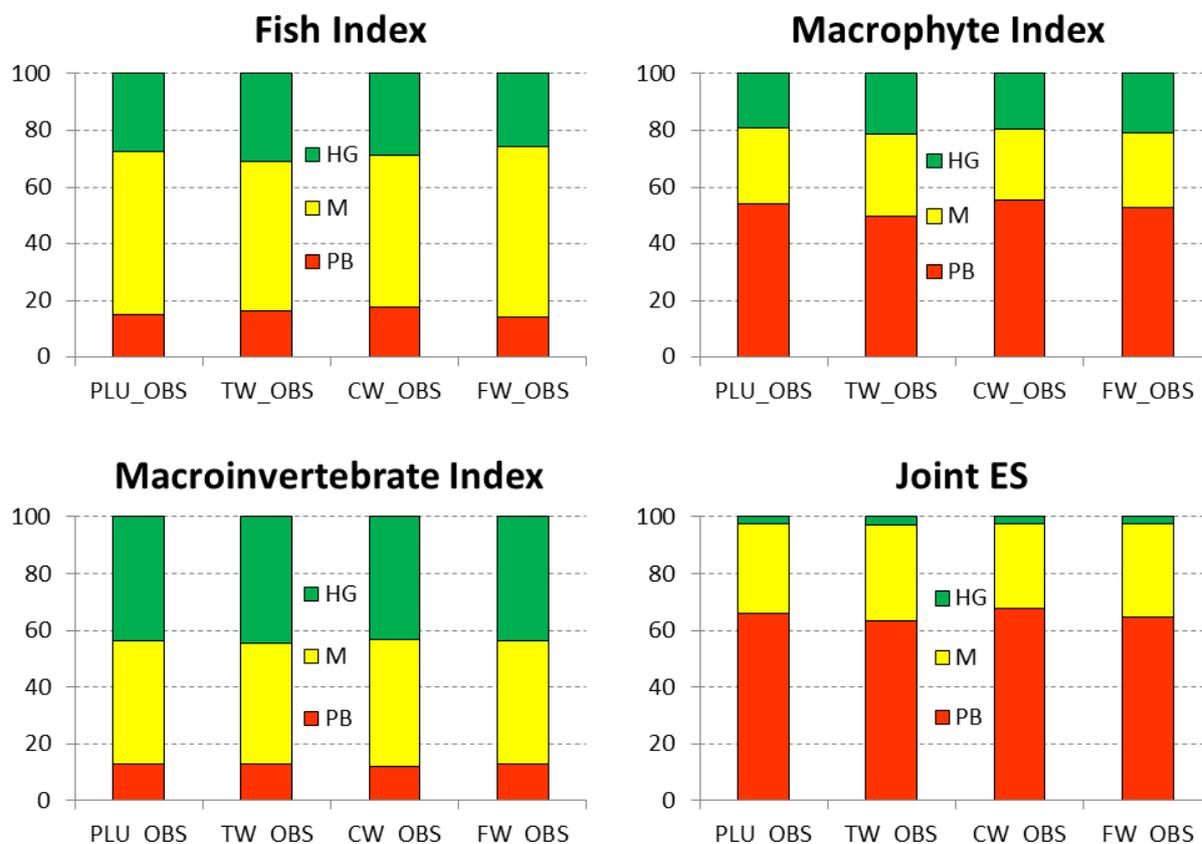


Figure 19 Probability distributions (%) for ecological status classes (PB: Poor/bad, M: Moderate, HG: High/Good) for the different indices and LUC scenarios.

Despite not being at the indicator level, nitrogen load is one of the main concerns for Danish aquatic ecosystems (Kaspersen et al., 2016; Thodsen et al., 2016), so it is worth to discuss the scenario impacts observed for TN concentration after BBN modelling. Contrary to the other stressors, isolated LUC scenarios showed a very noticeable impact on its probability distribution. Probabilities of low and, especially, medium TN levels increased in TW_OBS and CW_OBS, decreasing drastically the probability of high TN (Appendix **Error! Reference source not found. Error! Reference source not found.**). The opposite trend was observed in FW_OBS, in which high TN probability became 100%. Despite not included TN as a predictor of biological indices when developing abiotic-biotic empirical models, Ferreira et al. (2016) and Andersen et al. (forthcoming) also assessed the scenario impacts on its concentration. They obtained similar results (which was expected, since the same data were used in these studies): TN values 33% higher in FW_OBS than in the baseline, but 28% and 20% lower in CW_OBS and TW_OBS, respectively. These results evidence that not only nitrate loads (Molina-Navarro et al., 2018), but also TN concentrations, will be impacted by differential fertilization across isolated LUC scenarios.

MARS storylines

Similar to isolated LUC scenarios, the probability distributions of ecological status classes for the different scenarios remained quite stable across MARS storylines. Major differences were again observed for fish and macrophytes indices (Figure 20). The general trend for the fish index in the future storylines was an increase of probability of HG status, decreasing M status. When simulating isolated LUC, however, the trend was different within scenarios (Figure 19). This revealed a stronger effect of climate over land use changes, and might be explained mostly by higher probabilities of high BFI in all the scenarios (Appendix **Error! Reference source not found. Error! Reference source not found.**). The process-based catchment modelling results (Molina-Navarro et al., 2018) also revealed larger groundwater contribution in all the scenarios except from FW_30 (MD30 in the Odense-specific nomenclature). Anyhow, a slight increase of the fish index PB status was also predicted by the network, balancing the probability distributions so that the average ecological status was not expected to change much.

Probability variations in the macrophyte index were especially relevant for the 2060 horizon (Figure 20). In TW and FW scenarios, HG status probability increased, while probabilities of M and PB status probabilities decreased, and vice-versa for CW (M remained stable). This trend was not observed in isolated LUC modelling, so it was a consequence of the different climate inputs in these scenarios (RCP 4.5 in CW, RCP 8.5 in TW and FW), actually a response to the variations in DUR3 (Appendix 3).

The effects of scenarios observed for the Danish macroinvertebrate index probabilities were negligible (Figure 20). The ASPT showed different probability distribution than the Danish index in the “indicators” level, which was expected since there were no EQR thresholds criteria for Danish streams (and United Kingdom thresholds were used instead; Birk and Hering, 2006). However, the behaviour of both indicators was similar and the probabilities of its levels barely changed when simulating scenarios (data not shown for ASPT). This is not surprising, since both indices are influenced by the same stressors and they also depended on the same predictors in the abiotic-biotic empirical modelling (Ferreira et al., 2016, Andersen et al., forthcoming). Thus, despite no MARS benchmark indicator was used at the “Ecological Status” level of the BBN, the Danish macroinvertebrate index seemed to be a good substitute of ASPT (MARS BInd 12, Birk et al., 2015).

As a result of applying the “one out, all out” principle (Van de Bund and Solimini, 2007), the probability distributions of the final ecological status classes became more uniform (Figure 20). Probabilities barely changed for the short term (2030). For the 2060 horizon, however, PB status probability decreased for TW and FW (-6.1% and -4.3%, respectively), increasing slightly both M and HG status probabilities. Conversely, for the CW scenario, PB status probability increases (+8.2%), while M probability decreases (-7.4%). Despite these slight variations, the BBN predicted that the ecological status of rivers in the Odense Fjord

catchment would remain mostly PB in the future (% of PB probability varies between 56.7% and 69.2%, Figure 20). Comparing these values with the probabilities obtained with observed climate and present land use (Figure 19), results suggested that both climate and land use changes might not exert a large effect on the ecological status of the rivers in the catchment during the coming years.

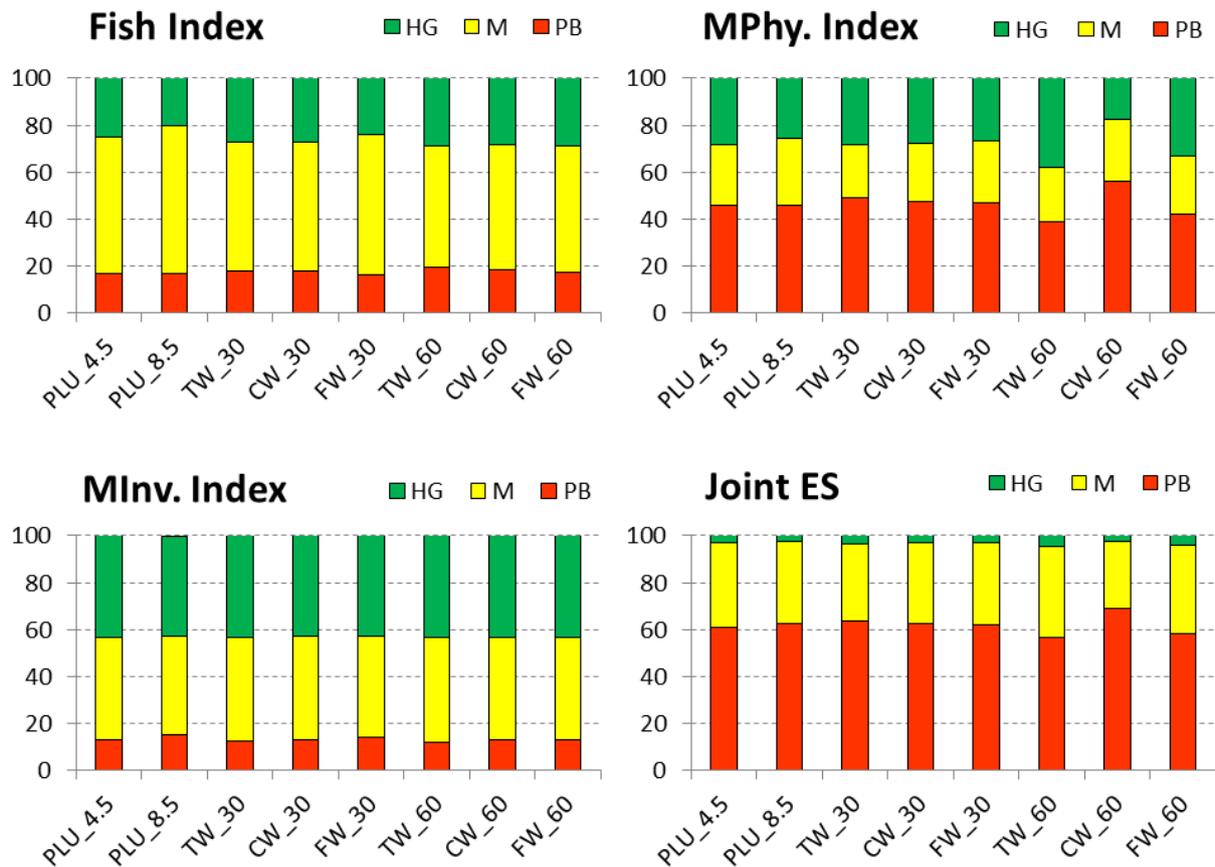


Figure 20 Probability distributions (%) for ecological status classes (PB: Poor/bad, M: Moderate, HG: High/Good) for the different indices in baseline (PLU_4.5, PLU_8.5) and combined MARS storylines scenarios (MPhy.=Macrophyte, MInv.=Macroinvertebrate).

Nevertheless, the slight variations observed in the fish and the macrophyte indices were in line with the results obtained in the work done for MARS WP4. Process-based modelling suggested that climate changes and different climate inputs within scenarios (and not LUC) were the main drivers of hydrology and phosphorus transport (Ferreira et al., 2016, Molina-Navarro et al., 2018). Since the stressors in the BBN are hydrology and phosphorus related, it is coherent that the probability variations observed in these indices are related with climate inputs. The process-based model revealed important effects on phosphorus transport in TW and FW scenarios (Ferreira et al., 2016, Molina-Navarro et al., 2018). However, P seemed to

have barely any impact in the simulation of future ecological status classes through BBN modelling. This is explained because phosphorus load increase across scenarios was associated to a runoff increase, so TP concentration, which is the magnitude included in the BBN, was not expected to vary much.

The empirical abiotic-biotic modelling done for MARS WP4 revealed that the effects of combined LUC and climate change scenarios were minor for the fish and the macroinvertebrate indices when evaluating averages for the entire catchment (Ferreira et al., 2016, Andersen et al., forthcoming). The strongest response was observed for the macrophyte index, finding the lowest values for the CW scenario in the 2060 horizon, while TW showed the highest positive impact. These results match up with the BBN modelling findings, which was somewhat expected since the empirical models were used to select the stressors to be included in the network. Nevertheless, not all the stressors found relevant to predict biological indices were selected, only those that SWAT can predict at a sub-basin level and providing enough data to fill the CPT; i.e. only two stressors per ecological status index. Finding the same trends when predicting ecological status endorses the robustness of the stressors selection in the BBN from the empirical modelling work.

TN concentration, despite not linked to the biological indices, might serve as a chemical water quality proxy. TN was the process-modelling derived stressor showing greater variations in its probability distribution after BBN modelling (**Appendix Error! Reference source not found. Error! Reference source not found.**). The trend observed was similar to isolated LUC scenarios, as already seen by Ferreira et al. (2016) and Andersen et al. (forthcoming). However, probability of high TN in TW in 2030 did not decrease as much as when analysing isolated LUC, which means that climate change somewhat favoured higher TN concentrations in this scenario, attenuating the effect of LUC. It might be explained because, despite the lower fertilization, nitrate loads did not decrease due to a counteraction of a simulated higher discharge, and organic N load increased too (Ferreira et al., 2016, Molina-Navarro et al., 2018). In the 2060 horizon, however, lower TN concentration was modelled by the BBN, probably due to a higher flow that favours dilution (Molina-Navarro et al., 2018). In summary, the variations observed when modelling only LUC remained in the MARS storylines, but modulated by climate change inputs. These results are of great interest since Denmark is one of the world's most intensively farmed countries. Despite the efforts already done reducing nutrient loading of the aquatic environment (total concentration of nitrogen and phosphorus in the streams have decreased by approximately 43% and 40% respectively since 1989; Thodsen et al., 2016), further nitrogen reductions are required in Denmark for the successful application of the WFD (Kaspersen et al., 2016). However, TN is not a limiting nutrient in the streams of the Odense Fjord basin, so its variability might not exert a noticeable impact on their ecological status, which estimation is the ultimate purpose of the BBN.

Posterior Probabilities

If no evidence is set up in the network, 33% probability for each MARS storyline was expected. Figure 21 shows the posterior probabilities for each storyline after setting evidence for a certain ecological status in each time horizon and for each index (plus the joint ecological status). Results revealed that the percentages barely changed for the 2030 horizon, which would mean that in a short term none of the scenarios would favour any ecological status class. For the 2060 horizon, only the macrophyte index showed variability, which was reflected in the joint ecological status after applying the “one out, all out” principle (Van de Bund and Solimini, 2007). The largest differences were observed in the HG status in 2060 (Figure 21), which might be more probable in TW and less probable in CW, in accordance with the results in Figure 20.

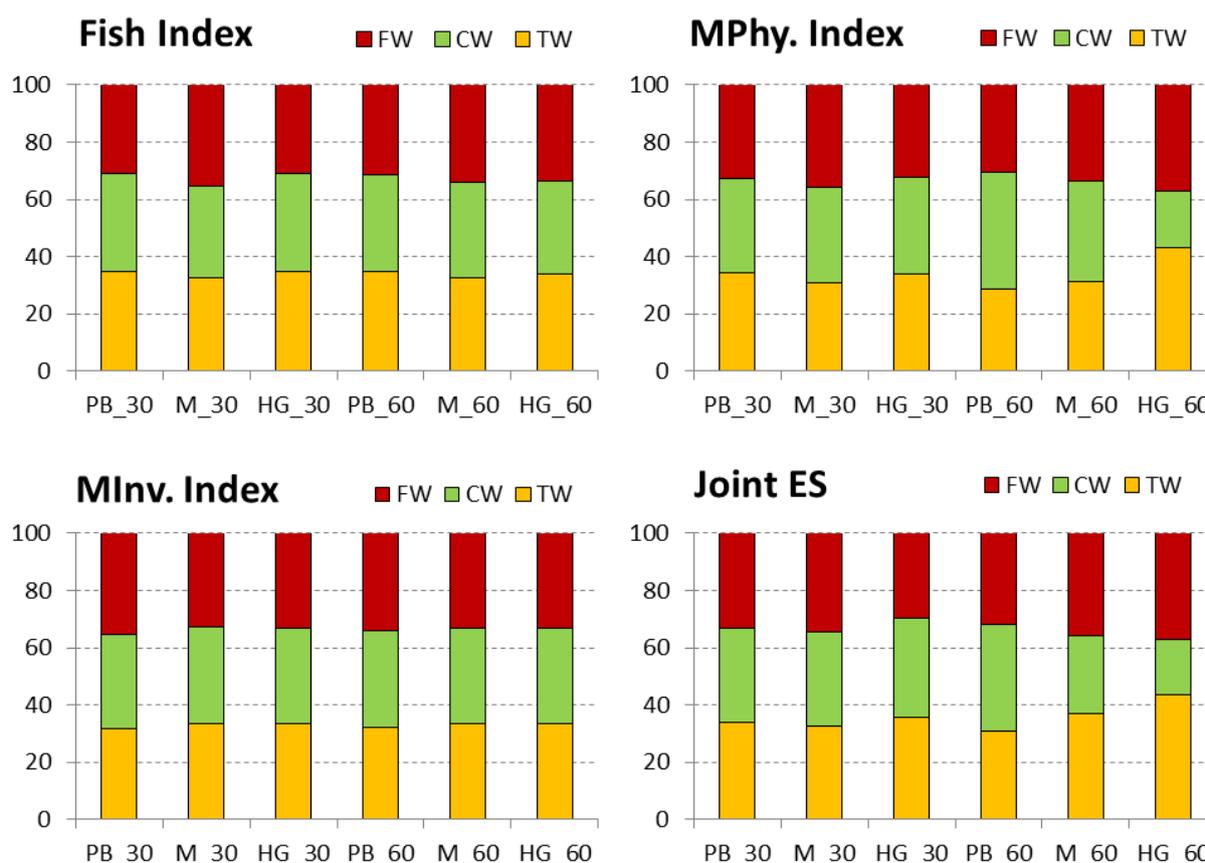


Figure 21 Posterior probabilities of each MARS storyline after setting evidence for a certain ecological status in each time horizon and for each index (MPhy.=Macrophyte, MInv.=Macroinvertebrate).

These results contrast with the initial expectations of the scenarios, since CW (“Agriculture for Nature” or AN in the Odense-specific nomenclature) was conceived as a scenario to favour the good ecological status of waters. Good practices in this scenario included an

increase in the forest surface and a decrease in fertilization (Ferreira et al., 2016, Molina-Navarro et al., 2018). Nevertheless, BBN modelling results described above suggested that the differences in ecological status probabilities were a consequence of variation of hydrological parameters of scenarios, and those depend mostly on climate changes. CW scenario was developed under a lower emissions (RCP 4.5) than TW and FW (8.5), yielding different climate inputs that ultimately affected the Bayesian modelling of ecological status probabilities. These results are in agreement with the results obtained in the work done under the MARS WP4 (Ferreira et al., 2016, Andersen et al., forthcoming, Molina-Navarro et al., 2018). However, these changes were minor and the main conclusion derived from the BBN modelling framework implemented in this work is that no large changes in ecological status are expected in the different storylines modelled.

Obstacles, pros and cons constructing and using the BBN

Using a BBN to model the impact of future scenarios on the ecological status of the rivers of the Odense Fjord catchment has both pros and cons. Compared to the other approach used by this research team, i.e. linking a process based model with abiotic-biotic empirical modelling, the most obvious advantage of the BBN is a larger simplicity. Once the BBN has been created, the procedure to simulate scenarios is simpler and faster, allowing the user to have a quick overview of the scenario impacts on both the stressor values and the ecological status indices at the same time. Besides, everything is done in the same interface (Appendix **Error! Reference source not found. Error! Reference source not found.**), facilitating the visualization of modelling results. This simplicity makes the BBN an approach that would be suitable to be used as a decision support tool by water managers or other stakeholders without high expertise in eco-hydrological modelling.

On the other hand, a handicap of this larger simplicity is the necessity to limit the input of information in the BBN. In order to have a simple, visual model, that might facilitate its use, the number of stressors has to be limited, which might diminish the credibility of the results. Besides, data availability might be not enough to create CPTs for all the desired nodes and combinations, which forces the user to simplify the network. In our particular case, one of the caveats that conditioned the data availability was that the BBN was created based on previous modelling efforts (process-based and empirical, Ferreira et al., 2016), with given and limited data inputs and outputs. Another disadvantage of the BBN is that it does not allow assessing the intra-catchment variability, which is possible to evaluate with the linking of process-based and empirical modelling (Ferreira et al., 2016, Andersen et al., forthcoming, Molina-Navarro et al., 2018).

2.4 Netherlands: Regge and Dinkel

2.4.1 Introduction

This case study deals with (sections of) two medium sized streams (Regge and Dinkel) in the catchment of the river Rhine (Figure 22 left). The catchments of these streams have a temperate marine climate with annual precipitation of 800 to 850 mm per year, mean evaporation of 560 mm per year and a mean annual air temperature around 9.9 °C (KNMI, 2016).

The catchments of the Regge and Dinkel are situated within the authority of Waterboard Vechtstromen. The Dinkel catchment (Figure 22 right) includes nine surface water bodies according to the Water Framework Directive (WFD) and the Regge catchment covers 22 surface water bodies (Ministry of Infrastructure and the Environment, 2015).

An overview of the chemical and ecological status of the surface water bodies of the Regge and Dinkel according to the WFD is presented in Figure 23 and Figure 24. Most of the surface water bodies of the Dinkel catchment do not meet the objectives for chemical status (Figure 24) and only one water body has a good ecological status (Figure 24).

Land use in the Regge and Dinkel catchments consists primarily of agriculture (60-70%), for which an extensive system of artificial watercourses and drainage is present. Apart from agriculture, water chemistry is influenced by waste water treatment plants. Urban areas are scattered throughout the catchments. Ground and surface water are intensively used for irrigation, industry, drinking water and recreation. Parts of the streams and rivers have been straightened or deepened for flood prevention; the same applies to its tributaries. As such, the main drivers¹ of the catchments are: urban development, flood protection, agriculture and climate change. Accordingly, the most important reasons for not reaching a good ecological status are combinations of insufficient water flow, stream regulation, insufficient connectivity for fish and moderate to poor chemical status (nutrients and other pollutants).

For more information on the Regge and Dinkel catchment we refer to the MARS deliverable D4.1 part 2 (Ferreira et al, 2016), as well as <http://fis.freshwatertools.eu/index.php/regge-dinkel.html>.

¹ The definition of a driver within the MARS project is: an anthropogenic activity or climate change phenomenon that may have an environmental effect (Birk et al., 2015).

2.4.2 Purpose

The purpose of the BBN was to predict the probability for excessive growth of submerged aquatic vegetation, based on the combined effects of hydrology, mowing, and nutrients, at the spatial scale of a river reach. This variable is related to the MARS benchmark indicator “abundance of aquatic vegetation” (Bind11).



Figure 22 Left: River basins according to the RBMP in the Netherlands. Right: Map of the study area. In red the Dinkel catchment, in blue the Regge catchment.

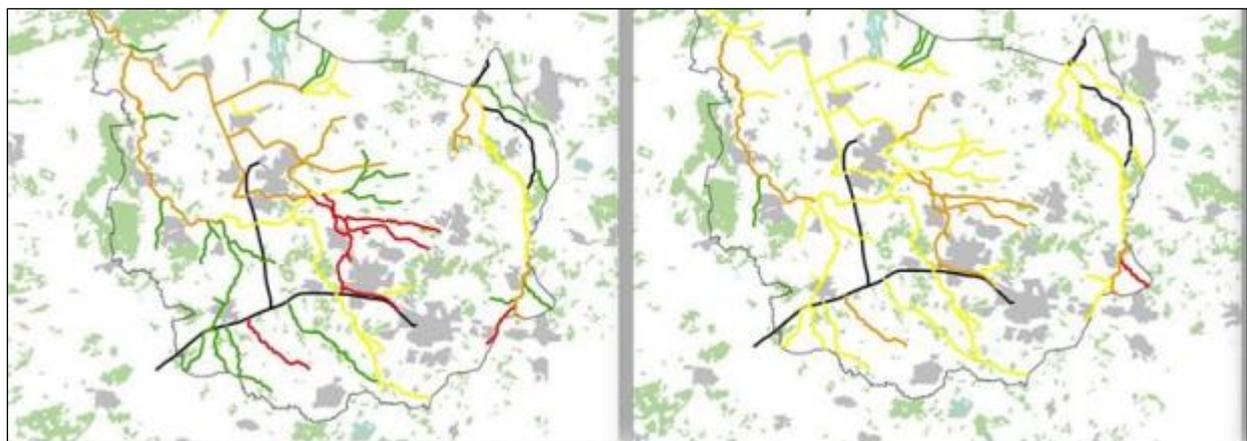


Figure 23 Chemical status of surface water bodies in the Regge and Dinkel catchments (Source: Waterboard Vechtstromen, 2015). Left: Total Phosphate, Right: Total Nitrogen. Green: good, yellow: moderate, orange: poor and red: bad status. Black: no WFD designated water body.

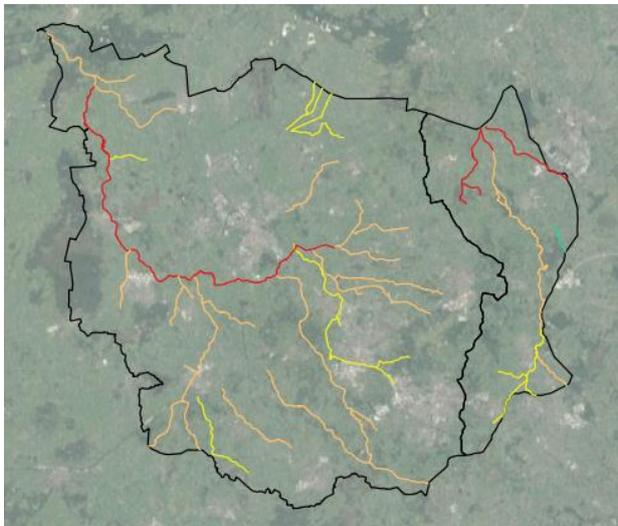


Figure 24 Ecological status of surface water bodies in the Regge (left) and Dinkel (right) catchments (Source: Waterboard Vechtstromen, 2015). Green: good, yellow: moderate, orange: poor and red: bad status.

2.4.3 Model construction

The causal network of the Bayesian Belief Network (BBN) of the Regge and Dinkel was set up according to the guidelines for developing and updating a BBN by Marcot et al. (2006). This meant that an initial influence diagram has been drawn based on expert knowledge and literature, after which this initial diagram has been discussed and adjusted with the expert knowledge of the stakeholder Waterboard Vechtstromen. The resulting diagram is shown in Figure 25.

The software packages that were used for calculations and display are: GeNIe (available from www.bayesfusion.com) and SamIam (available from <http://reasoning.cs.ucla.edu/samiam/>).

In the next step, we simplified and converted the influence diagram into a BBN (Figure 26). Nodes that were relatively difficult to measure and deemed unnecessary were removed, according to the guidelines of Marcot et al. (2006).

The simplifications that were applied were:

- Removal of the nodes influencing discharge. Discharge is measured in parts of the catchment, while surface runoff and groundwater seepage are not.
- Removal of the slope of the river bed. There were four parent nodes influencing water velocity. The slope of the river beds in the entire area is nearly the same. Therefore, it was deemed as one of the lesser interesting nodes influencing the water velocity in the Regge and Dinkel area.

- The sediment section is converted into one node: phosphorus (P) in the sediment. This section of the diagram was deemed too complex for the current application. The amount of phosphorus in the sediment has been shown to increase the growth of macrophytes (Smolders et al., 2015). Due to the fact that the surroundings of the Regge and Dinkel are for more than two thirds farm lands (Ferreira et al., 2016), we assume that the phosphorus concentration in the sediment (under suitable conditions, e.g. low enough flow velocity) will always sufficient for abundant macrophyte growth.
- The four nodes influencing the macrophyte abundance were kept in the model. All variables associated with these nodes play an important role for macrophyte abundance in streams.
- Maintenance was split into two nodes. To take different types of maintenance into account. This simplification was actually applied during the filling in of the CPT's, as we found out that it would be difficult to interpret the combined effect of maintenance of the riparian zone vs maintenance of the channel.

The initial influence diagram was developed in addition to the conceptual MARS model of WP 4.1-3. Parts of the conceptual MARS model of the Regge and Dinkel case study can still be found in the BBN, however there has been adjustments to fit the BBN.

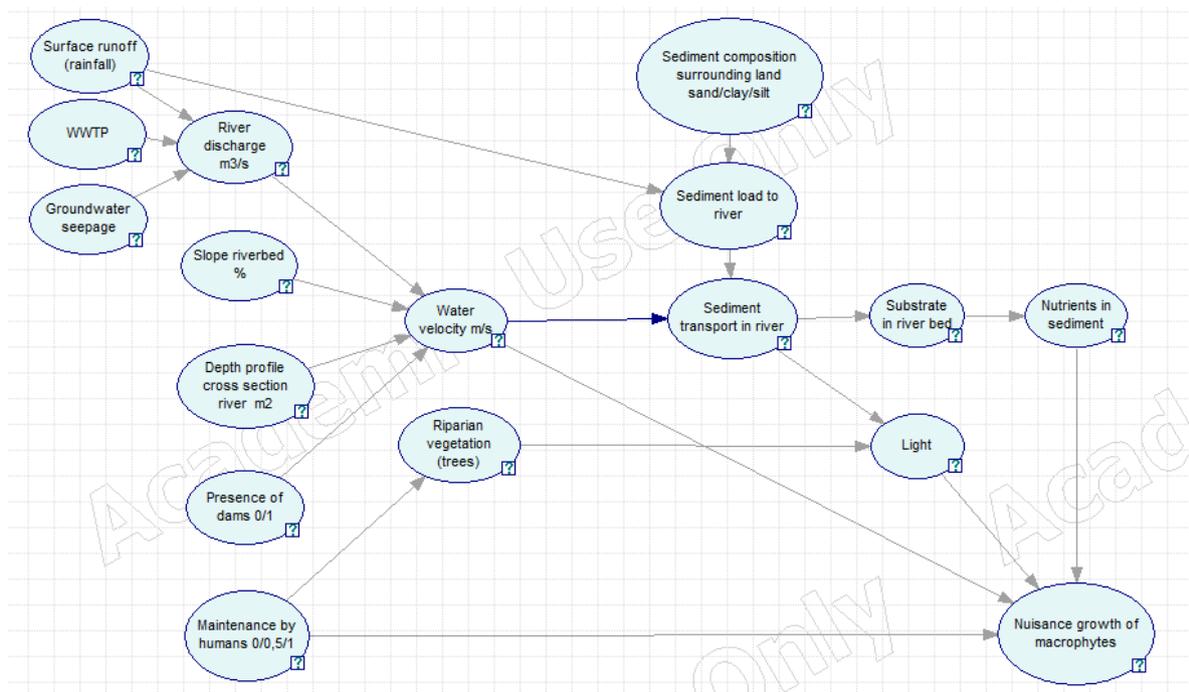


Figure 25 The influence diagram of the Regge and Dinkel case study. For phosphorus, it is assumed that are adsorped on sediment particles (e.g. clay, silt)

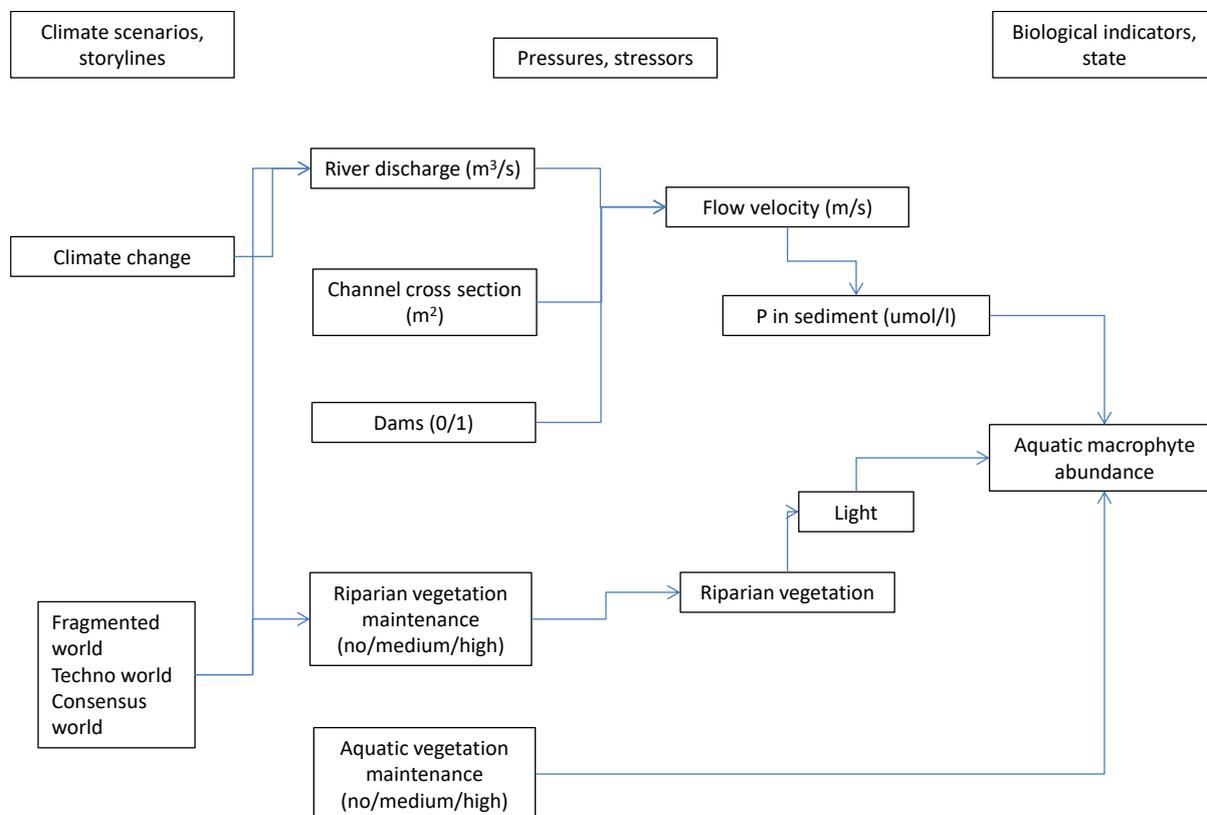


Figure 26 The BBN of the Regge and Dinkel case study. For phosphorus, it is assumed that are adsorped on sediment particles (e.g. clay, silt). For a more detailed explanation: see text.

2.4.4 Class boundaries

After the construction of the causal web, the classes were derived by using expert knowledge and available literature (Smolders et al., 2017; Janauer et al. 2010; Nichols, 2010). Key to the definition of the classes was to keep the BBN as simple as possible.

If clear numerical thresholds were available, these threshold were used. In some cases however, we made use of ordinal states, such as “low”, “medium”, and “high”. These ordinal states were chosen due to the lack of location specific data. In this way general knowledge of the working of rivers can still be implemented into the model and the possibility to refine the BBN remains open.

To identify the necessary classes for all parent nodes, we discovered that it is best to let these classes be (at least partially) steered by what is needed in the child nodes. Therefore, we started by looking at the necessary classes for the “Macrophyte abundance” node and worked our way up from there.

The “Macrophyte abundance” was divided into three classes “low”, “medium” and “high”. The state of this node is determined by four parent nodes, namely: Light, P in sediment, Maintenance and Water velocity.

For macrophytes to grow a minimum amount of light is needed. Therefore, the light node was split into two classes: “yes” (meaning that the minimum amount for macrophyte growth is reached) and “no” (meaning that the minimum amount for macrophyte growth is not reached).

Besides light, maintenance will have a direct effect on the abundance of the macrophytes. Many streams in the catchment experience maintenance for flood control. This means that the water managers sometimes remove macrophytes from the entire water bed to keep the water flowing. However, the maintenance can be classified in different intensities of mowing. Therefore, this node was split into the classes “not” (no maintenance), “medium”, and “strong”.

The amount of phosphorus in pore water of sediment needs to reach a certain threshold before it starts to influence the abundance of macrophytes (Smolders et al., 2017). This threshold (30 μmol total-P per liter pore water (in sediment)) was translated into two classes: “above 30 $\mu\text{mol/l}$ ” (associated with high submerged macrophyte cover) and “below 30 $\mu\text{mol/l}$ ” (low cover).

The classes of water velocity were determined partially according to the study of Janauer (2010). In this study the effect of water velocity on the abundance of several submerged aquatic macrophytes was measured. The classes that we used were: “0-0.5cm/s”, “0.5-4 cm/s”, “5-30 cm/s”, “31-69 cm/s” and “69 cm/s or higher”. The classes 0-0.5cm/s and 0.5-4 cm/s were made in accordance with the water velocity needed for sedimentation of very small particles (see the section on phosphorus).

The amount of light in the system is determined by the amount of riparian vegetation, which can be either “high”, “medium”, or “low”. The amount of riparian vegetation is mainly determined by its maintenance, which has the same classes as the maintenance on the submerged vegetation.

Phosphorus is mainly transported into the water system via overland flow. It enters the system mostly bound to fine sediments. Therefore, we assumed that for the amount of P in the sediment to reach the threshold for a high macrophyte abundance the water velocity needs to be low enough for small silt particles to settle. This implies that according to the Hjölstrom diagram the water velocity needs to be almost zero (max. 0.5 cm/s) (Nichols, 2010).

Dams or weirs can either be present or not, therefore this node was given two classes, either “yes” or “no”.

Finally, the classes for river discharge and depth profile cross section were determined by what was present in the catchment. For the average discharge in the growing season (May-September) the min-max range of these data was roughly 0-11 m³/s. To get to the ranges of surface area the range of depths and widths present in the Regge and Dinkel catchment have been checked. These were 0 – 33 m and 0 - 2.70 m respectively. This led us to a range of surfaces between 0-89.1 m², which was then rounded off to 0-100 m². We then divided the discharges and surface area into classes meaningful for the flow velocity.

2.4.5 Conditional probability tables

In the next step, the conditional probability tables (CPTs) for the BBN have been determined. These probability tables are the heart of any BBN. Parentless nodes have unconditional probability tables that represent prior knowledge on frequencies of each state. Child nodes have CPTs that represent combinations of all states of its parent nodes.

The probabilities in the CPTs were filled in based on general knowledge, internal consistency and calculations. If the nodes were filled in based on general knowledge the probabilities were translated via Table 14. To keep internal consistency, sometimes slight adjustments to the probabilities were made (e.g. to allow for slight increases in probability, 0.05 was added or subtracted from initial chances, therefore resulting in fractions such as 0.55).

For the dam node we assumed that during the growing season the discharge is usually low. This means that the presence of dams will usually result in (more or less) stagnant water. Therefore, we assume that in the presence of dams the water velocity will always be below 0.5 cm/s.

The probability on a total-phosphate concentration of 30 µmol/l in porewater in sediments of streams with a water flow < 0.5 cm/s was assumed to be high, since most sediment particles to which the phosphorus is bound should settle at this velocity. If this velocity was exceeded the probability on a high phosphate concentration was assumed to be low.

The probability table for flow velocities was calculated based to the following equation:

$$\text{m/s} = (\text{m}^3/\text{s})/\text{m}^2$$

The relations between maintenance, riparian vegetation and light were estimated.

Subsequently, the CPT's as well as the structure of the BBN were discussed and validated with experts of the water board.

As a last step the parts of the CPT's for which measurements were available, were updated using these data. The MARS storylines have been adjusted to fit the BBN model of the Regge and Dinkel catchment regarding changes in discharge, phosphate and the riparian zone. In the

Results section, a more detailed explanation is given about the structure of the BBN, the variables that are (not) included, and the boundaries of the classes.

Table 14 Translation of ordinal chances into fractions.

Ordinal chance	Fraction
Extremely high	0.95
Very high	0.9
High	0.8
Moderately high	0.6-0.7
Not high, not low	0.5
Moderately low	0.3-0.4
Low	0.2
Very low	0.1
Extremely low	0.05

2.4.6 Scenarios

Explorative scenarios

The BBN was run for several explorative scenarios to determine how each input node (river discharge, channel cross section, presence of dams and the two types of maintenance) influenced the macrophyte abundance. All input nodes were set once to their maximum and once to their minimum class after which the BBN was run. This led to $2^5 = 36$ runs.

Storylines

The suitable elements of the WP4 Regge and Dinkel process model storylines (Table 5.42 in Ferreira et al., 2016) were translated to the BBN. These elements applied to a change in discharge, phosphate and the riparian zone.

The baseline was chosen based on generic characteristics of a small stream in the Regge and Dinkel catchment (Table 15).

Table 15. Baseline settings

Node	Class
River discharge	0.01-0.05 m ³ /s
Channel cross section	1-2 m ²
Dam	Present
Maintenance in the riparian zone	Medium
Maintenance in the submerged zone	Medium

The MARS storylines are shaped by a combination of land use and climate scenarios.

Running the BBN for each storyline and each separate scenario plus a baseline lead to *nine BBN runs* (Table 16).

The amount of change in river discharge for each storyline and each separate scenario was calculated based on Sobek-model results for the time horizon 2060 as presented in the WP4 synthesis report (Ferreira et al., 2016). The amount of change in river discharge of each storyline or separate scenario was translated into fractions, which were then applied to the median of the BBN baseline river discharge class (Table 16).

Table 16. Overview of discharge levels for each storyline and each scenario. Discharges are based on a percentage of the median of the BBN baseline discharge class based on Sobek-model data (Table 5.49 - Ferreira, 2016).

Storyline	Land use scenario	Climate scenario	Decrease (%)	Discharge (m ³ /s)
Baseline	-	-	0	0.030
-	Fragmented world	-	-28	0,022
-	Techno world	-	-44	0,017
-	Consensus world	-	-9	0,027
-	-	IPSL-CMA-LR RCP8.5	-28	0,022
-	-	IPSL-CMA-LR RCP4.5	-19	0,024
Fragmented world	Fragmented world	IPSL-CMA-LR RCP8.5	-50	0,015
Techno world	Techno world	IPSL-CMA-LR RCP8.5	-66	0,010
Consensus world	Consensus world	IPSL-CMA-LR RCP4.5	-25	0,023

The phosphate concentrations in sediments increase for every storyline in comparison to the baseline. In the current BBN it is assumed that the phosphate concentration is always high enough for a high abundance of macrophytes to occur. This means that the inclusion of this element would not cause an effect in the current BBN and it was therefore left out.

The intensity of maintenance in the riparian zone differed per land use scenario. The maintenance of the riparian zone in the Consensus world is expected to become minimal due to a pro-environment attitude. In the Techno world the intensity increases, but not enough for it to become ‘high maintenance’, because people prefer to keep their land use optimal, but also understand the need for bufferzones. Finally, the maintenance intensity of the riparian zone in the Fragmented world will become high, because people in this scenario do not care about the environment, but prefer to have as large lands as possible for farming.

The generic characteristics of the stream include the presence of a dam, which affects the discharge patterns of streams. The calculated scenario discharges however all still remain within the same discharge class, despite a trend for decreasing discharge. Therefore, to make the effect of a decrease in discharge discernible, *three additional scenarios* were run without the presence of a dam and with a decreasing discharge.

2.4.7 Results and discussion

Setting up a model

It must be noted that the final structure of the BBN needed to change because of new insights about interactions between different nodes. This was the case with our BBN, when maintenance turned out to mix two types of maintenance which could affect the macrophyte abundance in different ways.

Explorative scenarios

The results of the storylines and separate scenarios are shown in Figure 27. For facilitation of the interpretation of the results, multiple graphs were made, each based on one input node. For the other graphs see: Appendix **Error! Reference source not found. Error! Reference source not found.**

The results showed that dams only have an effect at high river discharge or with the combination of low discharge and small cross section profile. Overall, the probability for extensive macrophyte vegetation increases with the presence of dams. This effect is especially strong with a high intensity maintenance on riparian vegetation.

The cross section profile only affects macrophyte abundance if there are no dams present. A small profile decreases the possibility of a high macrophyte abundance to zero as a result of increased flow velocity. Similarly, river discharge only has an effect if there are no dams present. Overall, a high river discharge decreases the chances on a high macrophyte abundance. A high maintenance on submerged macrophytes means a lower probability on a high abundance, while a high maintenance of the riparian zone leads to higher probability.

Overall, it can be concluded that these results are in line with expert knowledge.

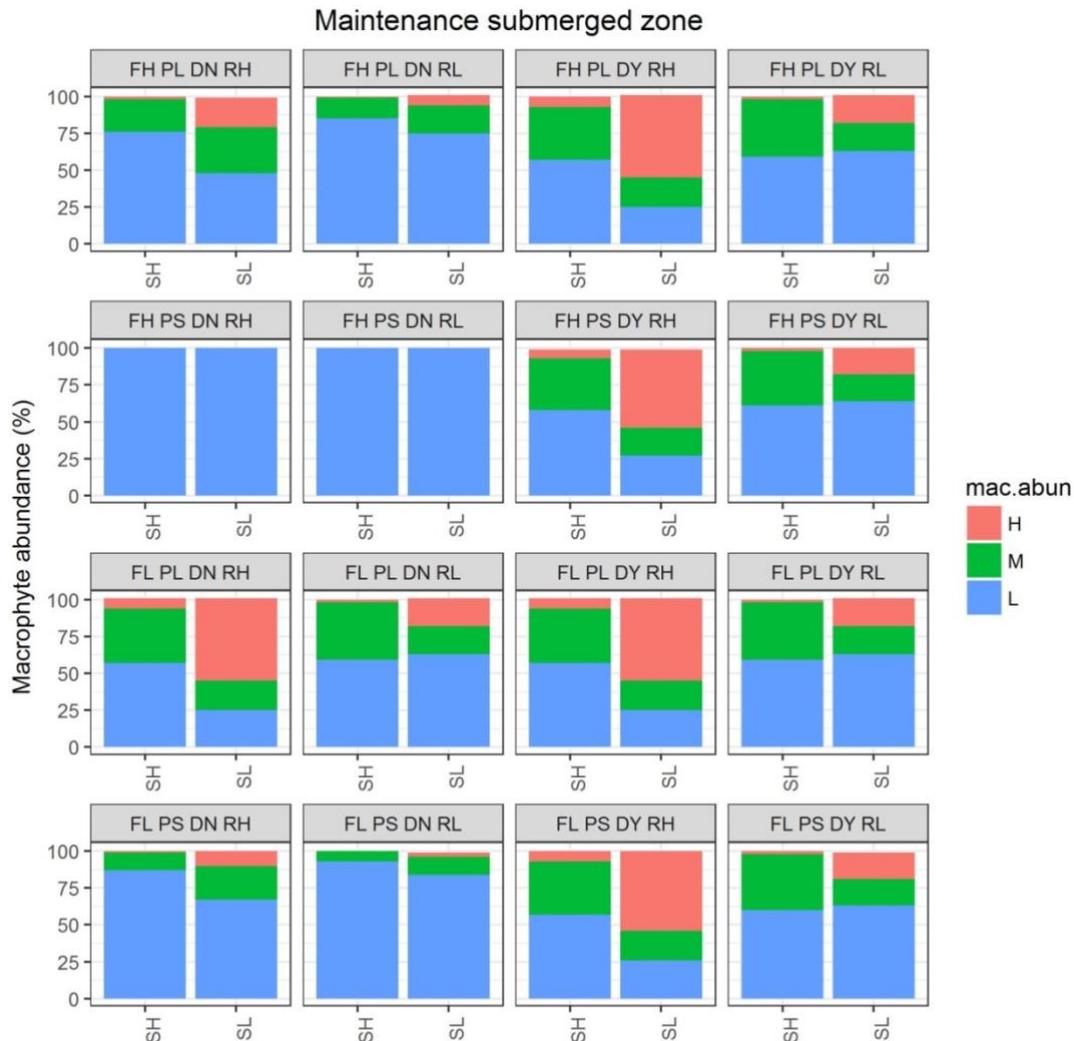


Figure 27 Results of the maximum and minimum scenario for the node “vegetation maintenance in the submerged zone”. FH: Discharge high, FL: Discharge low, PL: cross section large, PS: cross section small, DN: no dams present, DY: dams present, RL: maintenance in the riparian zone low, RH: maintenance in the riparian zone high, SH: maintenance in the submerged zone high, SL: maintenance in the submerged zone low.

Storylines

There was no discernible effect of the climate change on the outcome of the BBN runs in comparison to the baseline. The reasons for this were twofold, due to the fact that climate only affected the discharge:

1. There was a dam present in the baseline scenario and this would have excluded any effect of differences in discharge.
2. The amount of change in discharge due to climate was still within the boundaries of one discharge class.

The storylines had little effect on the probability for excessive macrophyte growth. Only the changes in riparian zone maintenance sorted an effect. The influence of the storylines on discharge was not visible due to the reasons already mentioned for climate.

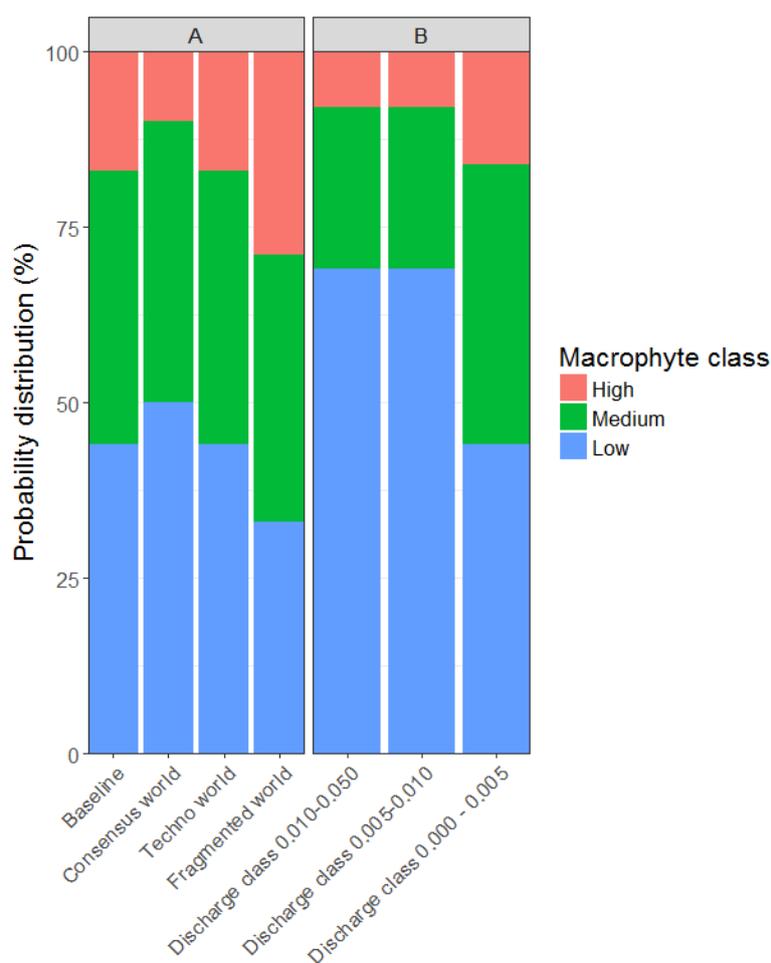


Figure 28. Results of storyline runs (A). In these runs dams are present. Results of the extra discharge scenario runs (B). In these runs no dams are present.

The scenarios that differed from the baseline are shown in Figure 28 (A). There is a decrease in the probability for excessive abundance of macrophytes in the Consensus world scenario compared to the baseline, while there is no change in the Techno world and in the Fragmented world scenario.

The effect of a decrease in flow velocity when no dams would be present becomes apparent in Figure 28(B). In the absence of a dam the water can flow faster, which causes a lower chance on a high (nuisance) abundance of macrophytes. Only when the stream almost dries out, which is the case in the lowest discharge class, the chance on a high abundance of macrophytes increases and becomes equal to the baseline scenario including dams.

2.5 Portugal: Sorraia

2.5.1 Introduction

The Portuguese case study is the Sorraia basin which occupies an area of 7730 km² and flows along a length of 155 km (Figure 29). It merges with the river Tagus at the estuary and it is also the Tagus tributary with the largest basin area. Roughly half of the Sorraia basin is covered by cork-oak forest, mainly in the headwaters, while the remaining is covered by one of the largest area of irrigated crops in Portugal, with a total of circa 15500 ha. Overall, approximately 41% of the area is forest, 28% range-grasses, 17% agriculture, 9% pine, 2% orchard, 2% urban and industrial and 1% pasture (Mateus et al. 2009; <http://fis.freshwatertools.eu/index.php/sorraia.html>).

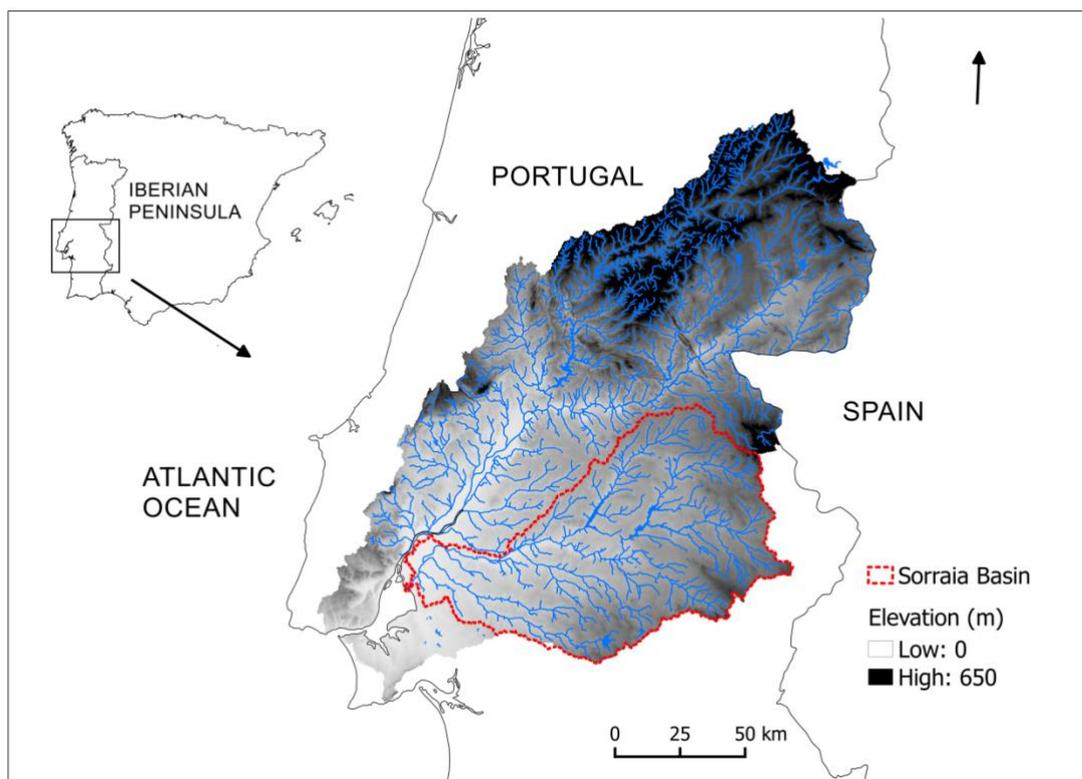


Figure 29 Tagus river basin with the location of the Sorraia case-study basin

Sorraia watershed is characterized by a Mediterranean climate, with high temperature and dry summers, and low temperature wet winters. The average annual precipitation is around 600 mm, from 400 mm in the dry years to up to 900 mm in the wet years. The average monthly precipitation is 50 mm, ranging from 25 mm in hot months (April - September) to 70 mm in cold months (October - March). The presence of two large reservoirs in the basin affects flow

patterns and runoff downstream. Additionally, the natural flow is substantially reduced by water abstraction for irrigation. The Sorraia watershed has a total of 153099 habitants with a density of 20 hab/km², mainly concentrated in three core areas: Ponte de Sôr (16722 hab), Samora Correia (17123 hab) and Coruche (19944 hab) (INE, 2011). It has only minor issues regarding urban pollution and urban wastes.

According to the River Basin Management Plan (RBMP), the main pressures on the basin are: (1) Hydromorphological changes, (2) Diffuse pollution, (3) Municipal discharges, (4) Flow regulation and (5) Water abstraction. Key ecosystem services identified by the RBMP are water for irrigation, recreation services and waste water treatment. The Water Framework Directive (WFD) status of 122 water bodies is: 54 good (44%), 15 moderate (12%), 12 poor (10%), 2 bad (2%) and 39 (32%) unclassified. The main causes of poor or failing status in the basin are mainly related with the water demand for agricultural purposes, which in the Sorraia basin is the highest within the Tagus river basin region (26% of total need). Nutrient loads from agriculture, livestock and urban origin, mainly in the alluvial valley, are also important potential causes of poor status in the basin. The Sorraia catchment was also used as a case study in MARS WP4 and a more detailed description of the catchment can be found in MARS deliverable D 4.1-2 (Ferreira et al., 2016).

A basin wide conceptual model based on the DPSIR framework was developed for the Sorraia Basin (see Ferreira et al., 2016), which defined the main drivers and pressures acting upon the Sorraia Basin that were considered in this case study. This conceptual model was used to frame the process-based and empirical-based modelling. This link between both modelling approaches and the basin-model allow abiotic and biotic state predictions under climate and land-use change scenarios, as well as under different response (measures) scenarios, defined according to the MARS storylines. Two drivers, agriculture and climate change, giving rise to four pressures, (1) diffuse pressure, (2) abstraction/flow diversion, (3) dams, barriers and locks and (4) hydrological alteration, were considered. The model focused on the main stressors identified for the basin: hydrological regime and nutrients acting on several biotic quality indicators.

2.5.2 Purpose

The construction of a Bayesian Belief Network (BBN) to the Sorraia case study is mainly to provide a prognostic tool that will allow end users to assess the general evolution of ecological status under future climate and socio-economic scenarios. The BBN also includes the implementation of alternative measures to mitigate global changes.

2.5.3 Model construction

The BBN model was based on a subset of the MARS conceptual model for the Sorraia basin (see Ferreira et al., 2016), essentially focusing on the effects of low flow and nutrients on the

biotic quality status of rivers (Figure 30). We considered four main components within the BBN, from parent to child nodes: (1) climate scenarios, socio-economic storylines and measure scenarios, (2) pressures and stressors, (3) natural environmental background and (4) biotic state. We followed the climate scenarios and socio-economic storylines adopted by the MARS Project (Faneca-Sanchez et al. 2015). Pressures consisted on three land use cover in the upstream catchment: % of agriculture, % of irrigated crops and % of urban areas. The stressors included three hydrological variables related with low flow – flow alteration, mean number and duration of low flow events – and one nutrient stressor – total N. Natural environmental background included a variable expressing longitudinal gradient (river slope) and annual mean air temperature. Biotic state was given by EQR values of four biotic quality elements (BQE): phytobenthos, macrophytes, macroinvertebrates and fish (Almeida et al., 2014; Aguiar et al., 2014; Feio et al., 2014; Segurado et al., 2014).

The structure of the BBN models was essentially based on outputs from both process-based models (SWAT - Soil and Water Assessment Tool), that simulated hydrological and nutrient stressors from baseline and future climate and socio-economic scenarios, and empirical models (GLMM - Generalized Linear Mixed Models) (see Ferreira et al., 2006 for further details on the use of SWAT and GLMM), that quantified relationships between hydrological and nutrient stressors with biotic indicators. Land use and natural environmental background (river slope and mean annual temperature) were also included as predictor variables in the empirical modelling. Because biomonitoring data from Sorraia was limited, and in order to encompass a wider environmental gradient, we also used data from the remaining Portuguese Tagus basin to run empirical models. The database comprised 240 sites from the Water Frame Directive biomonitoring program (Portuguese Environmental Agency, APA), with two sampling occasions (2010-11) (see Ferreira et al., 2016 for further details). The cause-effect links between stressors/land use/environmental background nodes and biotic indicator nodes were defined essentially according to the variables selected in GLMM (Segurado et al., submitted). The BBN was implemented using GeNIe v2.1 (freeware for academic use; available at <https://download.bayesfusion.com/>).

CLIMATE SCENARIOS,
STORYLINES, MEASURES

PRESSURES /
STRESSORS

ENVIRONMENTAL
BACKGROUND

BIOTIC INDICATORS / STATE

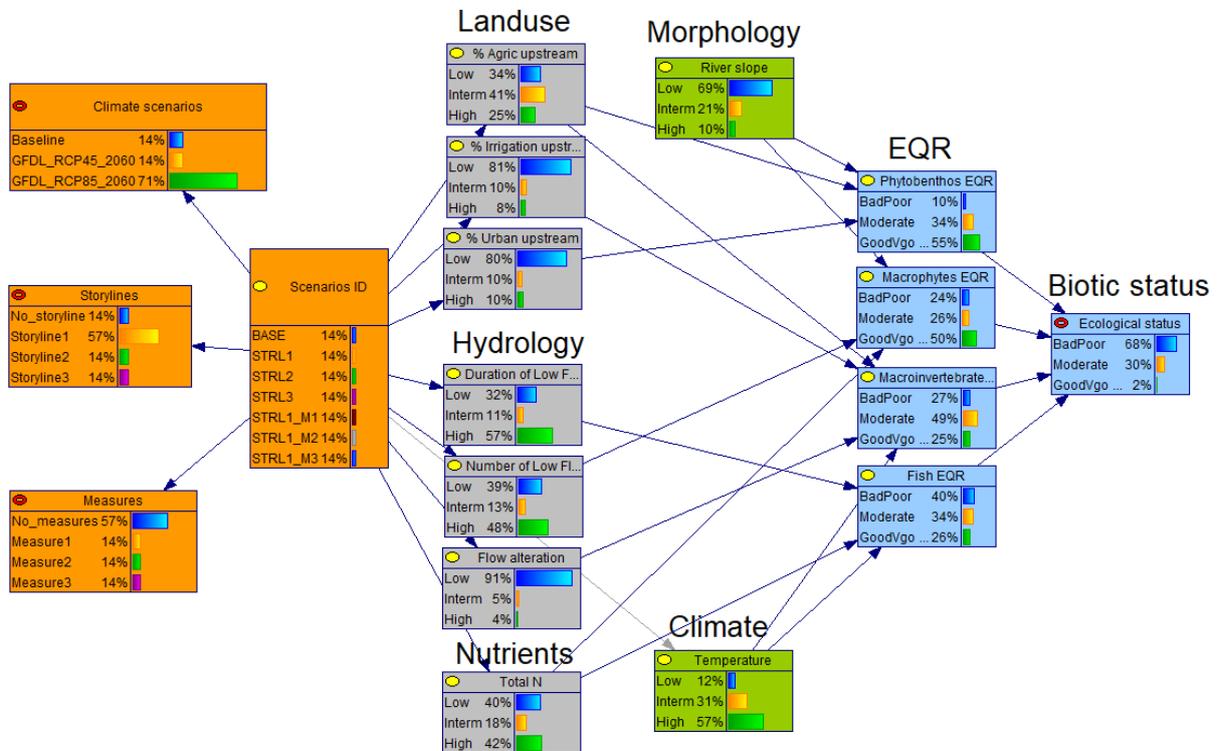


Figure 30 Bayesian Belief Network model developed for the Sorraia Basin

2.5.4 Class boundaries

The definition of class boundaries for pressures, stressors and background environmental variables was supported by partial response curves of biotic indicators derived from empirical models (Boosted Regression Trees – BRT). We considered three classes: low, intermediate and high. The low/intermediate boundary was located near the initial inflection of the curve and the Intermediate/high boundary was located near the final inflection of the curve.

The class boundaries of biotic indicators were based on the quality boundaries of the biotic quality indices. We considered three classes: bad/poor, moderate and good/very good. The links between biotic indicators and biotic quality were deterministic, based on the “one out all out” rule, as followed by the Water Framework Directive (European Commission, 2000).

2.5.5 Conditional probability tables

The construction of the conditional probability tables (CPT) linking stressors, land use and environmental background to biotic indicators were based on an expert judgement partially informed by the effect sizes and partial responses given by empirical models (GLMM). The

parameters were then learned with biomonitoring data from the year 2010. The conditional probability tables linking scenarios to stressors were based on outputs from SWAT (Soil and Water Assessment Tool) simulations.

2.5.6 Scenarios

In the design of the BBN the scenarios were based on climate models, emission scenarios and storylines adopted in the MARS Project. To simplify the BBN structure we only considered one climatic model (GFDL) and one future projection (year 2060) (Table 17). We used the three storylines adopted in MARS - the Techno world (STRL 1), the Consensus world (STRL 2) and the Fragmented world (STRL 3) – using a downscaling adapted to the environmental and socio-economic context of the Sorraia Basin (see Table 16 and further details in Ferreira et al., 2016). We also considered the implementation of three simple alternative measures within STRL 1: (1) removal of the two large dams (Measure 1); (2) optimization of irrigation and fertilization practices (Measure 2); both dam removal and optimization of irrigation and fertilization practices (Measure 3). We used SWAT (ArcSWAT 2012.10.19 interface for ArcGIS) to simulate daily values of the hydrological and nutrient variables for the six scenarios from 2055 to 2064.

Table 17 Mean values for each scenario considered in the BBN design

Variables/scenarios	Baseline	Techno world	Consensus world	Fragmented world	Measure 1	Measure 2	Measure 3
Mean annual temperature	15.58	17.15	16.78	17.15	17.15	17.15	17.15
% Agriculture	32.75	39.30	26.20	39.30	39.30	39.30	39.30
% Irrigated crops	1.60	2.07	1.28	2.07	2.07	2.07	2.07
% Urban areas	0.67	0.73	0.70	0.73	0.73	0.73	0.73
Duration of low flow events	13.72	62.58	63.21	46.79	62.29	62.48	62.20
Number of low flow events	2.89	7.11	6.75	7.02	7.14	7.12	7.14
Flow alteration	1.29	1.29	1.29	1.29	0.00	1.29	0.00
Total Nitrogen	0.43	1.03	0.83	3.28	1.02	0.34	0.32

2.5.7 Validation

To validate scenarios, we used the dataset of the year 2011 to construct new CPT for the Stressors/pressures nodes and predict class probabilities of the EQR for each BQE and the

final biotic quality. These predictions were then compared with real EQR classes and resulting biotic quality from the 2011 biomonitoring data.

2.5.8 Results and discussion

Projections of biotic status under scenarios

The projections under scenarios for the 2060 horizon performed with the BBN resulted in consistent trends - but with differing intensities - of the biotic state among each BQE (Figure 31). Globally, the “fragmented world” (STRL 3) resulted in the highest increase of Bad/poor status. The status class probability distributions between “Techno world” (STRL 1) and “fragmented world” (STRL 3) resulted in very similar for phytobenthos and macroinvertebrates. This is because the status of these BQE were essentially affected by land use, which were considered to be the same in both storylines. This is a major caveat of the BBN designed for the Sorraia basin. According to empirical modelling we only found significant responses of phytobenthos and macroinvertebrates EQR to land use, which is in fact a proxy for many possible individual or interacting stressors. These BQE might have been affected by individual or interacting stressors that were not considered in our analyses. Because a single sampling occasion was available in each year, it is also possible that these BQE do not respond strongly to a coarse temporal resolution. Due to the stronger link of these two BQE to land use, mitigation measures implemented in the “techno world” (dam removal, optimization of fertilizer use and irrigation), which do not involve land use changes, had barely any effect on the probability distribution of phytobenthos and macroinvertebrates status classes.

Macrophytes and fish showed more marked variations among scenarios, which is due to a stronger effect of nutrient and hydrological stressors. For macrophytes the bad/poor class probability varied roughly from 15% in the baseline to 38% in the “fragmented world” (STRL 3) and the good/very good class probability varied from 57% in the baseline to 37% in the “fragmented world”. For fish the bad/poor class probability varied roughly from 28% in the baseline to 58% in the “fragmented world” (STRL 3) and the good/very good class probability varied from 35% in the baseline to 19% in the “fragmented world” (STRL 3). Mitigation measures implemented in the “techno world” (STRL 1) showed to have different efficiencies in improving biotic status for different BQE’s. Dam removal did not show any noticeable effect on any BQE and global biotic status. An increased efficiency of irrigation and optimization in the use of fertilizers had a marked effect on macrophytes and fish, approaching the status class probability to the baseline situation. Hence, these results show that an extra effort on the programme of measures might be needed to attain the WFD targets in the future.

The probability distribution of global biotic status projected under the different storylines showed similar trends to those shown for each BQE (Figure 32). The trends were more marked, which is due to the “one out all out” rule used to compute the global biotic status. The bad/poor class probability was 56% in the baseline and predicted to increase to 72% in

the “techno world” (STRL 1), 65% in the “consensus world” (STRL 2) and 80% in the “fragmented world” (STRL 3). The good/very goods class probability was 3,8% in the baseline and predicted to decrease to 1,8% in the “techno world” (STRL 1), 2,8% in the “consensus world” (STRL 2) and 1,1% in the “fragmented world” (STRL 3). Mitigation measures involving irrigation and fertilizer use optimization in the “techno world” (STRL 1) were proven to be effective to improve the global biotic status classes to levels near to those found in the baseline situation.

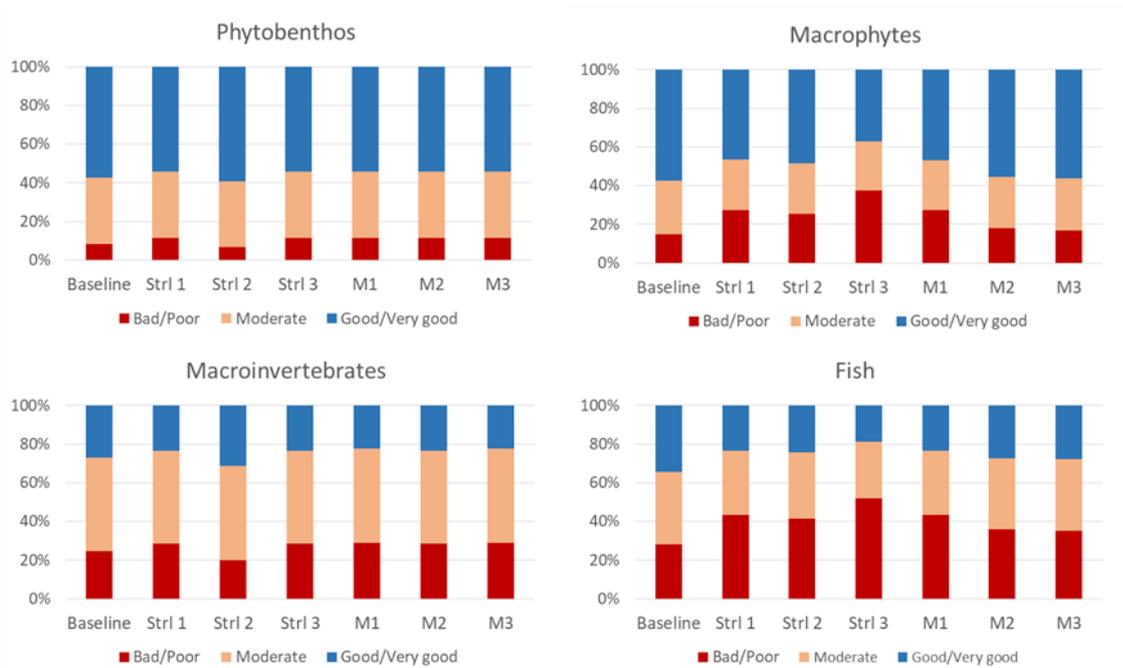


Figure 31 Probability distributions of biotic status for each BQE, as predicted by the BBN under the considered scenarios for the 2060 horizon

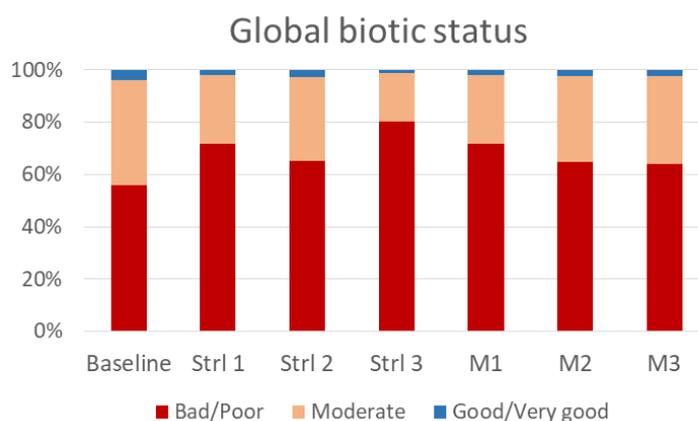


Figure 32 Probability distributions of global biotic status, as predicted by the BBN under the considered scenarios for the 2060 horizon

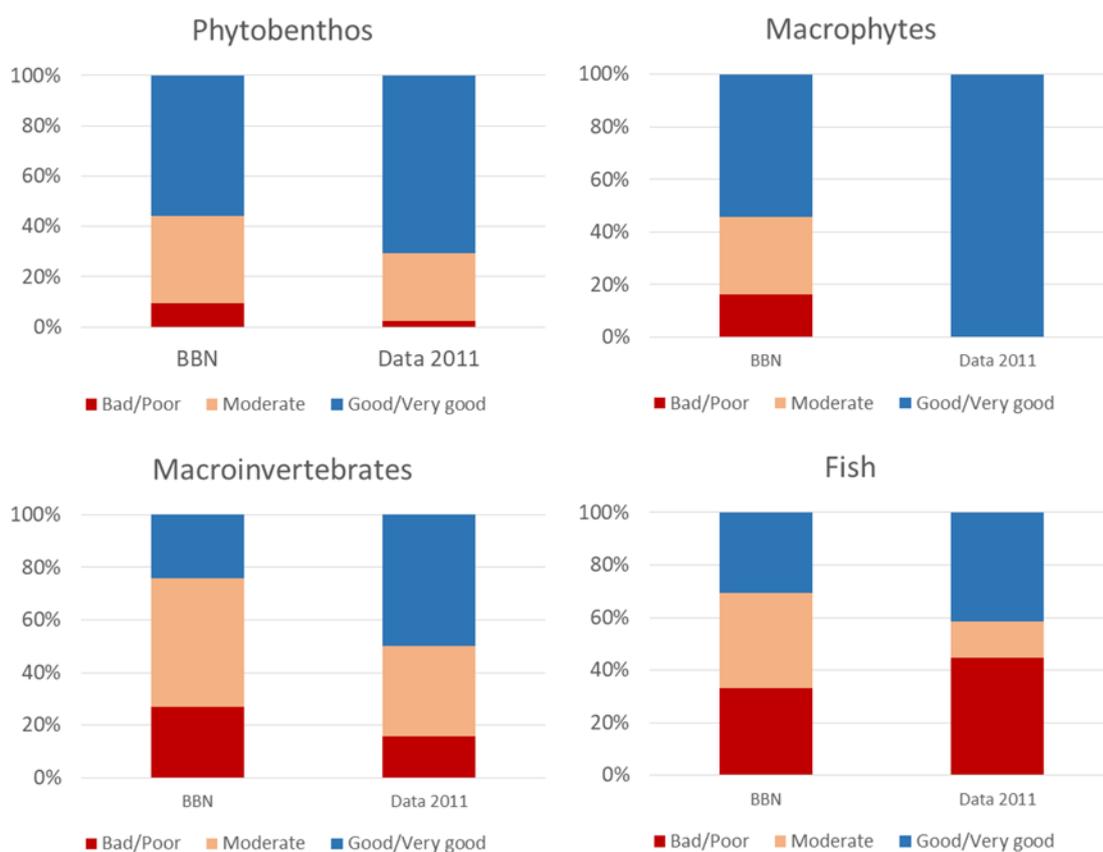


Figure 33 Probability distribution of biotic status classes estimated by the BBN using probability distributions of stressors derived from the 2011 biomonitoring data vs. observed biotic status classes in 2011 for each Biotic Quality Element

Validation of BBN

The probability distribution of biotic status classes estimated by the BBN using probability distributions of stressors derived from the 2011 biomonitoring data and the observed biotic status classes in 2011 are shown in Figure 33 for each BQE and Figure 34 for the global biotic status. There are obvious discrepancies between probabilities derived from the BBN and the observed probabilities. Namely, the BBN predictions tend to underestimate the good/very good class probabilities consistently among BQE's. Unfortunately, this validation has a strong limitation because not all sites in the database have complete data regarding EQR values for each BQE index. As a consequence, this validation is comparing BBN probabilities for the whole 2011 biomonitoring dataset with observed probabilities that represent only a subset of the data.

In the case of the global biotic status, the probability distribution derived from the 2011 biomonitoring data tend to exaggerate the number of bad/poor sites (Figure 34) because, due to the “one out all out” rule, when there is a missing classification for one or more BQE the only way to classify a water body is as bad/poor, i.e., when it is classified as bad/poor according to at least one BQE. This is caused by missing data for some BQEs. Sites with missing data can only be classified in case there is a bad/poor classification for at least one BQE. Consequently, sites that did not have a bad/poor classification for at least one BQE could not be classified when there were missing data for at least one BQE.

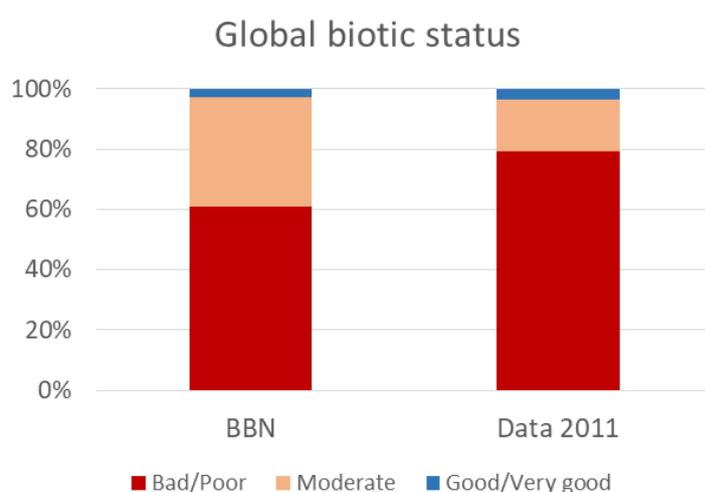


Figure 34 Probability distribution of global biotic status classes estimated by the BBN using probability distributions of stressors derived from the 2011 biomonitoring data vs. observed biotic status classes in 2011

Obstacles, pro´s and con´s constructing and using the BBN

Several obstacles were encountered during the implementation of the Bayesian Belief Network for the Sorraia Basin case study. First, the design of the BBN node structure and cause-effects links involved a lot of exploratory analyses and decision making that was not very straightforward and most often achieved by trial and error. The definition of probability conditional tables involves a big amount of effort, especially for nodes that have more than two parental nodes. Also, the process of variable discretization required much exploratory effort. A main caveat of BBN is that it is necessarily an over-simplification of reality. Not only because it is difficult to handle very complex cause-effect chains but also because of the necessity to perform a subjective discretization of variables into interval classes. In our case, data limitations did not allow us to validate in the most appropriate way the resulting BBN, because we had many missing data on EQR values for several BQE. Despite these obstacles and limitations there are clearly advantages of using BBN over other modelling approaches such as empirical modelling. A first advantage of BBN is that it can mix different sources of information, such as expert judgement, empirical modelling and process-based modelling into

a single framework. A second advantage is the simplicity of the final product allowing to easily visualize cause-effect links between ecological patterns. A third advantage is that it provides a mean to perform both prognostic analysis – e.g., to make outcome predictions under given management decisions or global change scenarios – and diagnostic analysis – e.g., to prioritize management targets to achieve given outcomes. A third and maybe the main advantage of BBN is that it allows to compute almost instantaneously several projections under alternative decisions in a very user-friendly environment. Therefore, BBN is potentially a very powerful tool for decision-making.

3 Synthesis

3.1 Scenario studies

In this report we have developed BBN models for five case studies catchments across Europe (Table 1) to explore the effects of future scenarios on biological responses and ecological status of water bodies. The case studies cover many dimensions of the MARS project:

- The three European regions: North, Central and South
- The two water categories: rivers and lakes
- The three story lines: Techno, Fragmented and Consensus world
- Various stressor types: Total P (MARS Benchmark Indicator BIn02), Total N (Bind03), hydrology (Bind04-06), hydromorphological alterations, temperature, etc.
- Biological indicators: chlorophyll a (Bind09) in rivers and lakes, cyanobacteria in lakes (Bind10); macrophytes (BInd11), macroinvertebrates (Bind12-13) and fish (Bind15)
- total ecological status of the water body (BInd01)

Four of the case studies (Norway, Denmark, Finland and Portugal) build directly upon the process-based catchment models used in MARS WP4. The Dutch case study, where a process-based model was not available, relied on a different approach involving data and expert knowledge of stakeholders and scientists. All case studies have used the three MARS storylines, except Finland which did not include Techno world. However, since Techno and Fragmented world are based on the same climate scenario (RCP 8.5), both climate scenarios are represented in all case studies. Future climate data were simulated by two different climate models, GLFD and ISPL, for both climate scenarios. For each case study, the climate model most suitable for the given region was selected (IPSL for Norway, Denmark and Netherlands; GLFD for Finland and Portugal). For all case studies, results are shown for the time horizon 2060 (i.e. 2050-2070), except for Finland showing results for the time horizon 2030 (i.e. 2020-2040) and for Denmark, which shows both time horizon periods.

In this synthesis, we highlight how the BBN methodology was used to link biological elements and ecological status to stressors and other indicators. Because of the many differences among the case studies - both inherently and in the implementation of the MARS scenarios, we focus on the main commonalities. We assess the added value of this modelling effort compared to the starting point for each case, i.e. the process-based catchment modelling from MARS WP4.

The linking of biological variables to physico-chemical variables in the BBN models was built upon the extensive empirical data analysis in WP4, which involved first the design of conceptual models and then the quantification of relationships between biological response

variables and various stressors. The CPTs for biological variables were mostly based directly on observations; i.e. the count of observations of a response variable in its different states co-occurring with the stressor variables in their different states. In many cases the count of observations was not sufficient to fill in a CPT, especially if there was more than one parent node (stressor). The count-based method was then supplemented with other methods, such as statistical modelling (e.g. cyanobacteria in Norway), or expert judgement (e.g. macrophytes in the Netherlands).

For all case studies, the BBN method enabled the coupling of abiotic and biotic models, and facilitated predictions of biological responses under the different future storylines. Therefore, BBNs had a clear additional value compared to the abiotic process-based catchment models (MARS work package 4). Below, the main results are presented for the case studies.

- Norway: The process-based models (INCA-P and MyLake) predicted temperature, TP and Chl-a in the lake, while the BBN added predictions on cyanobacteria and their response to increased temperature, and included this in the assessment of ecological status. The cyanobacteria node resulted in a stricter assessment of ecological status than the chl-a node alone. Due to the cyanobacteria node, this BBN also indicated a slight negative effect of increased temperature in the Techno and Fragmented world scenarios on ecological status, although the effects of land-use changes and nutrients were dominating.
- Finland: The process-based model (INCA-P) predicted TP and Chl-a in the river for one catchments (Lepsämäenjok), while the BBN added predictions on EQR (ecological quality ratio, based on macrophytes, macroinvertebrates and fish) and total ecological status. The ecological status node gave stricter assessments than chl-a alone. As for Norway, increased temperature had a negative impact ecological status, but the effects of land use were stronger.
- Denmark: The process-based model (SWAT) predicted flow and nutrients (TP, TN) in the river, while the BBN added three biological quality elements: macrophytes, macroinvertebrates and fish. Some of the predicted changes in ecological status by the BBN contrasted the initial expectations. For example, the probability of High-Good status of macrophytes was higher for Techno and Fragmented world than for Consensus world, although the latter storyline is more sustainable. A plausible explanation is that the differences in ecological status were driven by the hydrological parameters, which depend mostly on climate change rather than land-use. The total ecological status showed less response to scenarios than the individual BQEs, because the combination of the different responses made all probability distributions more uniform.
- Portugal: The process-based models (SWAT) predicted hydrological and nutrient variables in the river, while the BBN added four BQEs (phytobenthos, macrophytes, macroinvertebrates and fish) and total ecological status. The BBN predicted that different

BQEs respond differently to scenarios and to mitigation options. Phytobenthos and macroinvertebrates responded most strongly to land use. Hence, other mitigation measures implemented in Techno world (e.g. dam removal) had no effect in the BBN. Macrophytes and fish, in contrast, responded to both nutrients and hydrological stressors, and therefore also to scenarios of measures such as a more efficient irrigation and an optimization of fertilizers. In this case, the total ecological status showed a more marked response than the individual BQEs, because the one-out-all-out (OOAO) rule resulted in stricter assessment. This result illustrates an inherent problem of the OOAO, rather than the BBN method: a higher number of BQEs tends to result in a stricter assessment, and uncertainty in the assessments of individual BQEs tend to increase the risk of underestimating the total ecological status (Moe et al. 2016).

- The Netherlands: This case study was not based directly on the process-based model of work package 4.2, because it was difficult to retrieve reliable statistical relationships. Instead, it demonstrates how stakeholder engagement and expert judgement can be utilized to develop a BBN, and even run it for future scenarios. We found that presenting a BBN in a group of stakeholders helped them in constructively discussing their water systems. The BBN helped the waterboard in discussion among colleagues, to obtain common understanding, and with communication towards the public. The BBN showed that the impact of human alterations in the streams caused a larger impact on the macrophyte abundance than climate change. In fact, in the current defined scenarios the impact of the dams nullified the impact of climate change. Only changes in riparian zone maintenance would sort an effect.

A common strength of these five BBN models was the large amount of data available, both observed and simulated by the process-based catchment models. The statistical modelling of relationships between biological and physico-chemical variables, which had been carried out by WP4, also facilitated the development of the BBNs by providing causal and quantified relationships. This background provided a unique opportunity to link future climate and land-use scenarios through the whole DPSIR chain from changes in drivers to ecological status until the time horizon 2060. However, building the BBN models directly on these conceptual models also constrained the selection of variables to those predicted by the process-based models, and constrained the causal relationships to those identified as significant by the statistical modelling. For example, the BBN for Norway did not contain TN or water colour (humic content), since these variables are not predicted by the lake model (MyLake), although they may affect cyanobacteria. In the BBN for Portugal, the responses of phytobenthos and macroinvertebrates were dominated by the strong effect of land use estimated by WP4, and therefore showed no response to other types of mitigation measures. Therefore, when interpreting the outcome of future scenarios in the different case studies, it is important to be aware of such constraints on the BBN development. Nevertheless, one result was common across the whole set of case studies: the inclusion of one or more biological

components by BBN modelling resulted in stricter assessment of ecological status, compared to the process-based models alone. This outcome demonstrates the importance of applying modelling methods such as BBN to supplement more traditional models.

3.2 Assessment of the BBN modelling approach

The approach of constructing BBNs for forecasting the effects of scenarios enabled us to combine information from different sources (models, data, expert opinion) into one methodology. This is a clear advantage over other approaches, e.g. a statistical modelling approach, where variables cannot be included if no data are available. Overall, the case studies indicate that BBNs can be a successful approach, considering that suitable information can be retrieved from different sources. Below, the pros and cons of the use of BBNs are treated in more detail.

Advantages

One of the most obvious advantages of BBN is that it can integrate different sources of information, such as expert judgement, empirical modelling and process-based modelling into a single framework. This makes BBN an useful tool in cases when the quantity or quality of measured data do not allow specific statistical analysis.

Another advantage of the BBN is the simplicity. Once the BBN has been created, the procedure to make predictions based on scenarios is simpler and faster than corresponding for process-based models, allowing the user to have a quick overview of the scenario impacts on both the stressor values and the ecological status nodes at the same time. Consequently, the inputs and outputs of the BBN model may be relatively easy to understand for end-users without any modelling background. This simplicity makes the BBN very suitable as a decision support tool by water managers or other stakeholders without expertise in eco-hydrological modelling. For our case studies, the BBNs enabled us to compute instantaneously projections under alternative decisions in a user-friendly environment. BBN is therefore a potential powerful tool for decision-making.

Furthermore, the BBN approach provides an opportunity to include biological elements, as demonstrated by our studies, which is not the case in many existing process-based models for ecological status of rivers and lakes. Even when data are sparse, theory or expert knowledge on selected biological indicators can be used as a first step to construct causal links (CPTs) between abiotic and biotic responses. Since the Water Framework Directive (WFD) requires that assessments are based primarily on biology (EC, 2000), this is clearly an added value for use of models in water management in Europe.

Moreover, the WFD requires that potential impacts of climate change are considered in the next set of river basin management plans. Although much knowledge is available on effects

on climate change on ecosystems, including specific effects on biological quality elements in lakes (Moe et al., 2016), incorporating such information in predictive models is a challenge. The BBN methodology can facilitate the use of such knowledge, manifested as expert judgement of probabilities under given climatic scenarios.

Overall, it can be concluded that BBN is a promising approach to develop tools for supporting informed decision making and thus to facilitate the work of water managers.

Disadvantages

There are several limitations associated with the BBN methodology in the context of environmental management. Challenges that are associated with the use of BNs have also been discussed previously by Landuyt et al. (2013), Uusitalo (2007) and Varis and Kuikka (1999).

A drawback of the high simplicity of BBNs is the necessity to constrain the information in the BBN. In order to have a simple, visual model, which may facilitate its use, the number of stressors has to be limited, which may diminish the credibility of the results. In addition, the design of the BBN structure and cause-effects links involves many exploratory analyses, and decision making may not be straightforward and often achieved by trial and error. Furthermore, the necessity of defining subjective discretization of variables into interval classes is sometimes problematic.

The fact that the network cannot contain loops puts constraints on the ecological processes that can be modelled; phosphorus and phytoplankton dynamics in lakes are typically dominated by feedback processes (Saloranta and Andersen, 2007). For example, in the Lake Vansjø case study, high phytoplankton biomass can reduce the Secchi depth; on the other hand, lower Secchi depth can limit further phytoplankton growth due to light limitation. In the BBNs, such feedback loops were handled by dynamic process-based models, while the BBN summarised the outcome of these processes. In addition, the accumulation of uncertainty with the length of the network implies that it can be difficult to draw conclusions from the final output nodes (Marcot et al., 2006).

The definition of probability conditional tables involves a big amount of effort, especially for nodes that have more than 2 parental nodes. Also, the process of variable discretization required much exploratory effort. Besides, data availability might be not enough to create CPTs for all the desired nodes and combinations, which forces the user to simplify the network.

3.3 Validation

Developed BBNs should be validated to assess the ability of the model in representing the ecosystem. According to a review of McDonald et al. (2015), only 61% freshwater and estuarine BBNs published between 2002-2014 were validated. The procedure for validation of models strongly depends on the purpose of the model, and – hence - model validation is highly case-specific and it is difficult to generalise statements. Based on our case studies, the following criteria have been identified, viz.

1. the BBNs should capture the most important causal relationships of the ecosystem modelled,
2. the quantification of each of these relationships should be validated separately,
3. the results of the BBNs should be able to fit observed data fairly well.

Inclusion of most important causal relationships

The inclusion of the most important causal relationships can be checked by carefully reviewing these relationships and the assumptions made in the model formulation and comparing them to the most up-to-date knowledge as possible. This is probably the most important step in the construction of BBNs, and can be done by peer review by experts or managers before adoption of the BBN into management decision-making frameworks (McDonald et al., 2015). However, it may be not possible to achieve this in any quantitative terms. Another approach is to play with the model and check if it matches with available expert knowledge of the ecosystem under study or with behaviour of similar ecosystems.

Quantification of causal relationships

The relationships in the BBN can also be validated separately. This may be especially relevant when few data are available. An example of this approach is the BBN of Lake Vansjø. Because there was a limited number of cyanobacteria observations, an independent dataset ("EUREGI") was used to construct an alternative CPT for cyanobacteria and compared the outcome of this version with that the original version. The EUREGI dataset gave similar probability distributions in the CPT for cyanobacteria to those from Lake Vansjø, which strengthened the confidence in these probability distributions (Chapter 2.1).

Another approach to deal with few data is to fix probabilities in the CPTs, and to check if the model behaves according to expert knowledge. In the BBN of Lake Vansjø the relationship between temperature, Chl-a and cyanobacteria was examined more closely by setting evidence (fixating probabilities) for the nodes temperature and Chl-a. The results showed that the model behaved as expected regarding seasonal variation in temperature and cyanobacteria, and that the BBN generated reasonable predictions (Moe et al. 2016).

Fit between actual observations and modelled results

The results of the BBN-model should be able to fit observed data fairly well. This can be quantified in different ways, based on the comparison of model predictions and actual observations. Depending on the type of data, residuals, misclassification rates or something similar can be calculated to evaluate how well the BBN matches independent datasets of field observations. Many metrics are available to denote model prediction or classification accuracy and error rates (Marcot, 2012). From these, the metric that best fits the purpose of the model can be selected.

When data are unavailable or cannot be gathered, it may be possible to query other experts and compile a case file of the experience and known or predicted outcomes, and use that to validate the model. However, such results would then be more appropriately interpreted as verification of the model against outcomes provided by a set of independent experts' knowledge (Marcot, 2017).

For our case studies, the outcome of the BBNs of the Odense Fjord (Denmark) and Soraia (Portugal) performed reasonable to good when compared to field observations (if data available). For Finland and The Netherlands, the validation was problematic due to a shortage of field data.

3.4 Differences between BBNs for diagnostic and prognostic purposes

In task 7.2 of MARS, BBNs have been used for diagnostic purposes (Feld et al., 2017), while in this report (task 7.3) the focus is on prognosis. This led to the question whether different adaptations in the BBN design are required when its main purpose is to perform diagnostic versus prognostic analyses.

In general, the final structure of the BBN should be determined by the purpose of the model. Unclear modelling objectives typically results in constantly changing or expanding models, resulting in poorly-designed models with low to no scientific credibility. Therefore, Marcot (2017) strongly advocates to choose just one purpose for BBN, viz. in this case either diagnosis or prognosis.

In principle, a BBN could be applied for both diagnostic and prognostic purposes. For example, the BBN for Denmark was used in both in a diagnostic and prognostic way, to calculate the posterior probabilities of each MARS storyline after setting evidence for a certain ecological status in each time horizon and for each index (Figure 21). However, a BBN needs to be simplified to become useful for the purpose it is designed for. When the focus is on diagnosis, the choices for this simplification may be made in a different way than when the model has a prognostic purpose. Accordingly, the design of the final BBN model may differ substantially between BBNs for either diagnostic or prognostic purposes.

Another caveat on the use of prognostic BBNs is the fact that CPTs are often filled with data or expert knowledge from sites ranging from ‘bad’ to ‘(very) good’ ecological quality. Accordingly, by using BBNs that are filled with these data, it is assumed that species composition will respond ‘directly’ to improved (abiotic) conditions. In many ecological rehabilitation projects however, it is shown that ecological recovery can be strongly delayed (years to decades), because many species need to recolonize the rehabilitated areas, even when abiotic conditions are optimal for these ‘target’ species. This recolonisation process strongly depends on the distance to source population of these species, as well as their dispersal characteristics. Hence, the use of datasets differing in ecological quality gives an indication of the potential increase in ecological quality, rather than the realized ecological quality on the short term. If the purpose of the BBN is to forecast the realized ecological recovery, then datasets should be used which incorporate sites *before* and *after* ecological rehabilitation.

4 Acknowledgements

The work described in this report draws on and is reliant on a large body of work that existed prior to the start of the MARS Project (FP7 grant number 603378), including the development of catchment models and data collection for each case study, as reported in MARS Deliverable 4.1. In particular, special thanks are given to Prof. Samu Mäntyniemi (University of Helsinki) and Dr. David Barton (Norwegian Institute for Nature Research) for teaching the workshop on BBN modelling.

Gerben van Geest, WP7.3 Task leader

Deltares, Utrecht. December 2017.

5 References

Aguiar, F.C., Segurado, P., Urbanič, G., Cambra, J., Chauvin, C., Ciadamidaro, S., Dörflinger, G., Ferreira, J., Germ, M., Manolaki, P., Minciardi, M.R., Munné, A., Papastergiadou, E., Ferreira M.T., 2014. Comparability of river quality assessment using macrophytes: a multi-step procedure to overcome biogeographical differences. *Science of the Total Environment* 476–477, 757–767.

Almeida, S.F.P., Elias, C., Ferreira, J., Tornes, E., Puccinelli, C., Delmas, F., Dorflinger, G., Urbanic, G., Marcheggiani, S., Rosebery, J., Mancini, L., Sabater, S., 2014. Water quality assessment of rivers using diatom metrics across Mediterranean Europe: A methods intercalibration exercise. *Science of the Total Environment* 476-477, 768-776.

Andersen, H.E., Baattrup-Pedersen, A., Trolle, D., Lu, S., Molina-Navarro, E., Larsen, S.E., forthcoming. Impact assessment of multiple stressors on macrophytes, benthic macroinvertebrates and fish in Danish streams.

Arnold, J.G., Srinivasan, R., Muttiah, R.S., Williams, J.R., 1998. Large area hydrologic modeling and assessment Part I: Model development. *J. Am. Water Resour. Assoc.* 34, 73-89.

Aroviita, J. Hellsten, S., Jyväsjärvi, J., Järvenpää, L., Järvinen, M., Karjalainen, S.M., Kauppila, P., Keto, A., Kuoppala, M., Manni, K., Mannio, J., Mitikka, S., Olin, M., Perus, J., Pilke, A., Rask, M., Riihimäki, J., Ruuskanen, A., Siimes, K., Sutela, T., Vehanen, T. & Vuori, K.-M., 2012: Guidelines for the ecological and chemical status classification of surface waters for 2012–2013 – updated assessment criteria and their application). Environmental Administration Guidelines 7/2012: 1–144. (in Finnish)
<http://hdl.handle.net/10138/41788>

Birk, S., Constantinescu, A., Gieswein, A., Cardoso, A.C., Borja, A., Solheim, A.L. et al., 2015. Framework to select indicators of multi-stressor effects for European river basin management, in: Deliverable 2.1. Four manuscripts on the multiple stressor framework, 184-214. Available at: <http://www.mars-project.eu/index.php/deliverables.html>.

Birk, S., Hering, D., 2006. Direct comparison of assessment methods using benthic macroinvertebrates: a contribution to the EU Water Framework Directive intercalibration exercise. *Hydrobiologia* 566, 401-415.

Christensen, R. H. B. (2015). ordinal - Regression Models for Ordinal Data. R package version 2015.6-28. <http://www.cran.r-project.org/package=ordinal/>

Couture, R.-M., Tominaga, K. , Starrfelt, J., Moe, S.J. , Kaste, Ø., Wright, R.F., 2014. Modelling phosphorus loading and algal blooms in a Nordic agricultural catchment-lake system under changing land-use and climate. *Environmental Science: Processes & Impacts*. DOI: 10.1039/C3EM00630A.

Couture, R.-M., Cremona, F., Rankinen, K., Gutiérrez-Cánovas, C., 2016. Deliverable 4.1-3 Case study synthesis: Report on case studies from Northern river basins.

Raoul-Marie Couture, S. Jannicke Moe, Yan Lin, Øyvind Kaste, Sigrid Haande, Anne Lyche Solheim. 2017. Simulating water quality and ecological status of Lake Vansjø, Norway, under land-use and climate change by linking process-oriented models with a Bayesian network, *Science of The Total Environment*, 621:713-724

Devia, G.K., Ganasri, B.P., Dwarakish, G.S., 2015. A Review on Hydrological Models. *Aquat. Procedia* 4, 1001-1007.

European Parliament and Council, 2000. Directive 2000/60/EC of 23 October 2000 establishing a framework for the Community action in the field of water policy. *Official Journal of the European Union*.

Faneca Sanchez, M., Duel, H., Alejos Sampedro, A., Rankinen, K., Holmberg, M., Prudhomme, C., et al., 2015. Report on the MARS scenarios of future changes in drivers and pressures with respect to Europe's water resources, in: Deliverable 2.1. Four manuscripts on the multiple stressor framework, 215-284. Available at: <http://www.mars-project.eu/index.php/deliverables.html>.

Feio, M.J., Ferreira, J., Buffagni, A., Erba, S., Dörflinger, G., Ferréol, M., Munné, A., Prat, N., Tziortzis, I., Urbanič, G., 2014. Comparability of ecological quality boundaries in the Mediterranean basin using freshwater benthic invertebrates. *Statistical options and implications. Science of the Total Environment* 476-477, 777-784.

Feld C.K., Saeedghalati M, Lemm J., Birk S., Hering D., Strackbein J., Melcher A., Humer N., Mischke U., Venohr M., Bremerich V., Mahnkopf J., Loos S., Duel H. (2017) MARS Diagnostic Tools to diagnose the causes of deterioration of water bodies. MARS Deliverable D7.1-Task 7.2.

Feld, C.K. (2016). Protocol of the BN analysis to develop the MARS T7.2 Diagnostic Tool. Protocol of the preliminary analysis using the AQEM/STAR sand-bottom lowland river dataset (D01/D03); Analysis of stressor-response relationships, stressor hierarchy and variable dependences using generalised linear and boosted regression modelling (GLM, BRT), regression trees (CART) and cluster analysis (Cluster)

Ferreira, T., Panagopoulos, Y., Bloomfield, J., Couture, R.-M., Ormerod, S., et al. (MARS Project WP4 Team), 2016. Case study synthesis - Final Report. MARS Project Deliverable 4.1. Available at: <http://mars-project.eu/index.php/deliverables.html>.

Haande, S., Moe, S.J., R.-M. Couture, 2016. Phytoplankton and other monitoring data from Lake Vansjø. *Freshwater Metadata Journal* 15: 1-8. <http://dx.doi.org/10.15504/fmj.2016.15>

Haande, S., Solheim, A., Moe, J., Brænden, R., Lyche Solheim, A., Moe, J., Brænden, R., 2011. Klassifisering av økologisk tilstand i elver og innsjøer i Vannområde Morsa iht. Vanndirektivet. (Classification of ecological status in rivers and lakes in river basin Morsa according to the Water Framework Directive). NIVA report no. 6166. NIVA report no. 6166.

Haakana, M., Hatunen, S., Harma, P. Kallio, M., Katila, M., Kiiski, T., Makisara, K., Perasaari, J. Piepponen, H., Repo, R. Teiniranta, R. Tomppo, E. Torma, M. 2008. Finnish CORINE 2006 project: determining changes in land cover in Finland between 2000 and 2006, Proc. SPIE 7110, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology VIII, 71100V. <http://dx.doi.org/10.1117/12.800148>.

Handboek Hydrobiologie Appendix 7. KRW Watertypen.

Ibelings, B. W.; Vonk, M.; Los, H. F. J.; Molen, D. T. van der; Mooij, W. M. 2003. Fuzzy modeling of cyanobacterial surface waterblooms: validation with NOAA-AVHRR satellite images. *Ecological Applications* 13:1456-1472.

INE, 2012. Censos 2011 - XV Recenseamento Geral da população e V Recenseamento Geral da Habitação. Resultados definitivos. INE, I.P., Lisboa.

Janauer, G.A., Schmidt-Mumm, U., Schmidt, B., 2010. Aquatic macrophytes and water current velocity in the Danube River. *Ecological Engineering* 36:1138-1145.

Kaspersen, B.S., Jacobsen, T.V., Butts, M.B., Boegh, E., Müller, H.G., Stutter, M., et al., 2016. Integrating climate change mitigation into river basin management planning for the Water Framework Directive – A Danish case. *Environ. Sci. Policy* 55(1), 141-150.

Kragt, M.E., 2009. A Beginners Guide to Bayesian Network Modelling for Integrated Catchment Management. Landscape Logic. Technical Report n. 9.

Kristensen, E.A., Jepsen, N., Nielsen, J., Pedersen, S., Koed, A., 2014. Dansk fiskeindeks for vandløb (DFFV). Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi.

KNMI, 2016. Monthly values over the last 30 years (1986-2016). <https://www.knmi.nl/nederland-nu/klimatologie/maandgegevens>. Royal Netherlands Meteorological Institute

- Landuyt, D., Broekx, S., D'hondt, R., Engelen, G., Aertsens, J. & Goethals P.L.M. (2013) A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling & Software* 46: 1-11.
- Larsen, S.E., Baattrup-Pedersen, A., 2015. Matematisk beskrivelse af Dansk Vandløbsplante Indeks. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi.
- Larsen, S.E., Friberg, N., Skriver, J., Kjellerup Larsen, L., Wiberg-Larsen, P., 2014. Konvertering af DVFI faunaklasser til EQR-værdier (Økologisk Kvalitets Ratio). *Vand & Jord* 21, 12-16.
- Marcot, B.G. (2017) Common quandaries and their practical solutions in Bayesian network modeling. *Ecological Modelling* 358: 1-9.
- Marcot, B.G. (2012) Metrics for evaluating performance and uncertainty of Bayesian network models. *Ecological modelling* 230: 50-62.
- Marcot, B.G., Steventon, J.D. , Sutherland, G.D., McCann, R.K., 2006. Guidelines for developing and updating Bayesian belief networks applied to ecological modelling and conservation. *Canadian Journal of Forest Research*. 36: 3036-3074.
- MARS project, 2015. MARS fact sheet #03. MARS scenarios and storylines.
- Mateus, V., Brito, D., Chambel-Leitão, P., Caetano, M., 2009. Produção e utilização de cartografia multi-escala derivada através dos sensores LISSIII, AWiFS e MERIS para modelação da qualidade da água para a Bacia Hidrográfica do Rio Tejo. Conference Paper. <https://www.researchgate.net/publication/272885901> (accessed 30 July 2017).
- McDonald, K.S., D.S. Ryder, M. Tighe (2015) Developing best-practise Bayesian Belief Networks in ecological risk assessments for freshwater and estuarine ecosystems: A quantitative review. *Journal of Environmental Management* 154: 190-200.
- Miljø- og Fødevareministeriet, 2016. Vandområdeplan 2015-2021 for Vandområdedistrikt Jylland og Fyn. Styrelsen for Vand- og Naturforvaltning, Copenhagen.
- Miljø- og Fødevareministeriet, 2017. MiljøGIS. <http://miljoegis.mim.dk/> (accessed 26.04.17).
- Miljøministeriet, 2011. Vandplan 2010-2015. Odense Fjord. Hovedvandopland 1.13. Vanddistrikt: Jylland og Fyn. Miljøministeriet, Naturstyrelsen, Odense.
- Ministry of Infrastructure and the Environment, 2015. Stroomgebiedbeheerplan Rijn 2016-2021. <https://www.waterkwaliteitsportaal.nl/>
- Moe, J. Alahuhta, J. Aroviita, M. Beklioglu, T. Bucak, L. Carvalho, R.-M. Couture, S. Ergodan, L. Globevnik, H. Gundersen, S. Hellsten, E. Jeppesen, M. Järvinen, M. Koprivšek,

N. Kotamäki, E. E. Levi, S. Maberly, O. Malve, M. Mjelde, P. Nõges, T. Nõges, M. O'Hare, J. A. Richardson, J.E. Sample, T. Shatwell, B. Skjelbred, A. L. Solheim, P. Taylor, S. Thackeray, S. Vilbaste, H. Woods. 2017. Effects of multiple stressors on ecosystem structure and services of phytoplankton and macrophytes in European lakes. MARS deliverable 5.1-4.

Moe, S.J., Haande, S., Couture, R.M., 2016. Climate change, cyanobacteria and ecological status of lakes: a Bayesian network approach. *Ecological Modelling* 337:330-347.

<http://dx.doi.org/10.1016/j.ecolmodel.2016.07.004>

Molina-Navarro, E., Andersen, H.E., Nielsen, A., Thodsen, H., Trolle, D., 2017. The impact of the objective function in multi-site and multi-variable calibration of the SWAT model. *Environ. Model. Softw.* 93, 255-67.

Molina-Navarro, E., Andersen, H.E., Nielsen, A., Thodsen, H., Trolle, D., 2018. Quantifying the combined effects of land use and climate changes on stream flow and nutrient loads: A modelling approach in the Odense Fjord catchment (Denmark). *Science of the Total Environment* 621, 253-264.

Nichols, G. 2010. *Sedimentology and stratigraphy*, second edition. Wiley Blackwell.

Rääpysjärvia, J., Hämäläinenb, H., Aroviita, J., 2016. Macrophytes in boreal streams: Characterizing and predicting native occurrence and abundance to assess human impact. *Ecological Indicators* 64: 309–318.

Rankinen, K., Keinänen, H, Cano Bernal, J. (2016). "Influence of climate and land use changes on nutrient fluxes from Finnish rivers to the Baltic Sea." *Agriculture, Ecosystems & Environment* 216: 100-115.

Rankinen, K., Keinänen, H., Cano Bernal J., 2016. Influence of climate and land use changes on nutrient fluxes from Finnish rivers to the Baltic Sea. *Agriculture, Ecosystems & Environment* 216: 100-115.

Saloranta, T.M., Andersen, T. (2007) MyLake—A multi-year lake simulation model code suitable for uncertainty and sensitivity analysis simulations. *Ecological Modelling* 207: 45-60.

Segurado, P., Caiola, N., Pont, D., Oliveira, J, Delaigue, O., Ferreira, M.T., 2014. Comparability of fish-based ecological quality assessments for geographically distinct Iberian regions. *Science of the Total Environment* 476–477, 785–794.

Skarbøvik E, Bechmann M.E., 2010. Some Characteristics of the Vansjø-Hobøl (Morsa) Catchment. 5. Bioforsk Soil and Environment, Ås, pp. 44.

Skarbøvik E., Haande S., Bechmann M., Skjelbred B., 2016. Monitoring Morsa 2014-2015 (in Norwegian). NIBOS report 2-42. NIBOS, Ås, Norway, pp. 98.

Smolders, A., Lucassen, E., Roelofs, J. Kramer-Hoenderboom, A., Lenssen, J., 2017. Woekering van waterplanten in beken tot op de bodem uitgezocht. H2O-Online. 16 februari 2017.

Smolders, A., Roelofs, J., Lucassen, E. 2015. De onderwaterbodem in relatie tot de ontwikkeling van waterplanten in watergangen van het beheergebied van Waterschap Rijn en IJssel. B-ware. Report number: 2015.48.

Sucksdorff, Y. and R. Teiniranta (2001). Land Cover and Land Use Mapping in Finland. Strategic landscape monitoring for the Nordic countries. G. Groom and T. Reed, TemaNord 2001:523. 523: 89-108.

Thodsen, H., Windolf, J., Rasmussen, J., Bøgestrand, J., Larsen, S. E., Tornbjerg, H., et al., 2016. Vandløb 2015. NOVANA. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi.

Trolle, D., Hamilton, D., Hipsey, M., Bolding, K., Bruggeman, J., Mooij, W., et al., 2012. A community-based framework for aquatic ecosystem models. *Hydrobiologia* 683, 25-34.

Turunen, J., Muotka, T., Vuori d, K-M., Karjalainen, S.M., Rääpysjärvi a, J. Sutela, T., Aroviita, J. ,2016. Disentangling the responses of boreal streamassemblages to lowstressor levels of diffuse pollution and altered channel morphology. *Science of the Total Environment* 544: 954–962.

Uusitalo, L. (2007) Advantages and challenges of Bayesian networks in environmental modelling. *Ecological Modelling* 203: 312-318.

Van de Bund, W., Solimini, A., 2007. Ecological Quality Ratios for ecological quality assessment in inland and marine waters. European Commission, Directorate-General Joint Research Centre, Institute for Environment and Sustainability.

Varis, O. & Kuikka, S. (1999) Learning Bayesian decision analysis by doing: lessons from environmental and natural resources management. *Ecological Modelling* 119: 177-195.

Windolf, J., Thodsen, H., Troldborg, L., Larsen, S.E., Bøgestrand, J., Ovesen, N.B., Kronvang, B., 2011. A distributed modelling system for simulation of monthly runoff and nitrogen sources, loads and sinks for ungauged catchments in Denmark. *J. Environ. Monit.* 13, 2645-2658.

Appendix 1: BBN versions of case study Odense (Denmark)

Figure A1.1. BBN Version 1: Isolated land use change scenarios with observed climate (2001-2010). The BBN shows the conditional probability for each node level considering equal probabilities in scenarios.

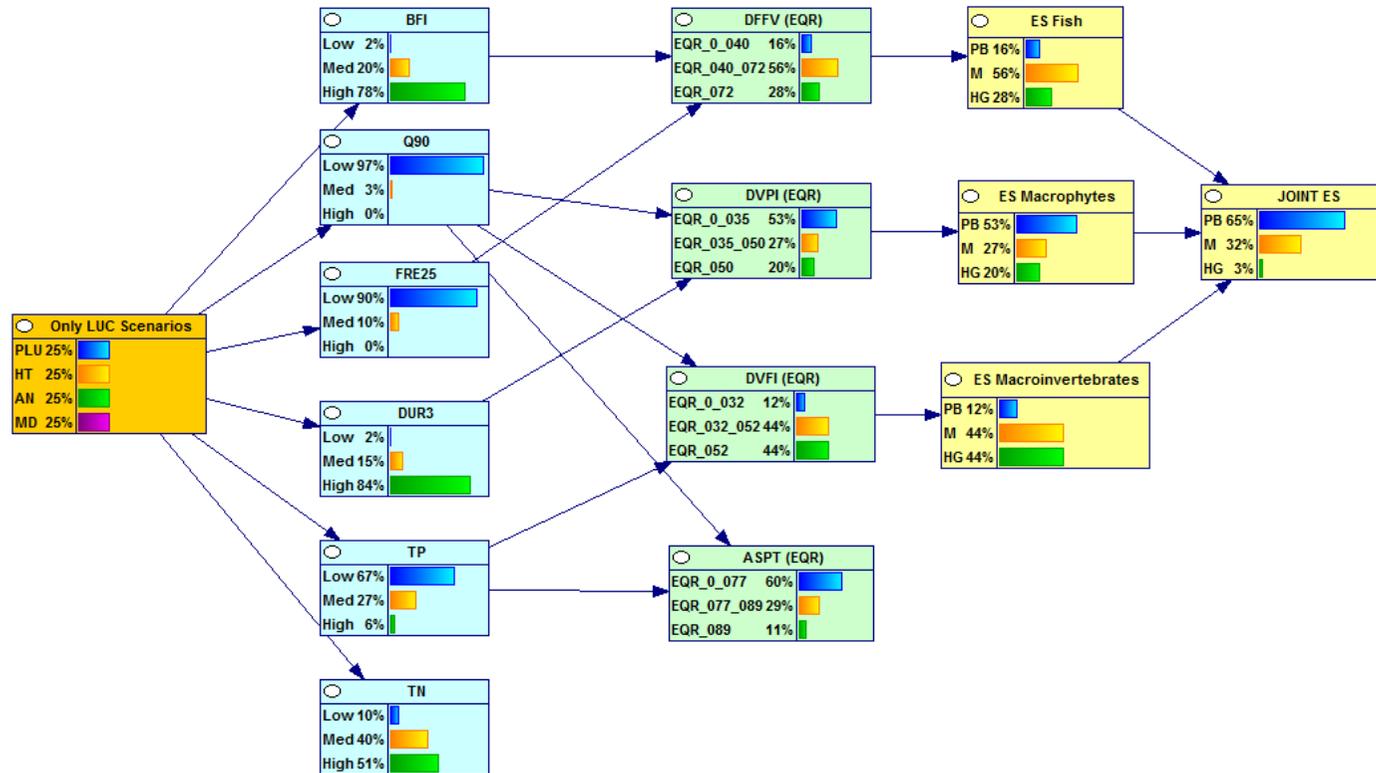


Figure A1.2. BBN Version 2: Baseline scenarios for the MARS Storylines with modelled climate (2011-2020). The BBN shows the conditional probability for each node level considering equal probabilities in scenarios.

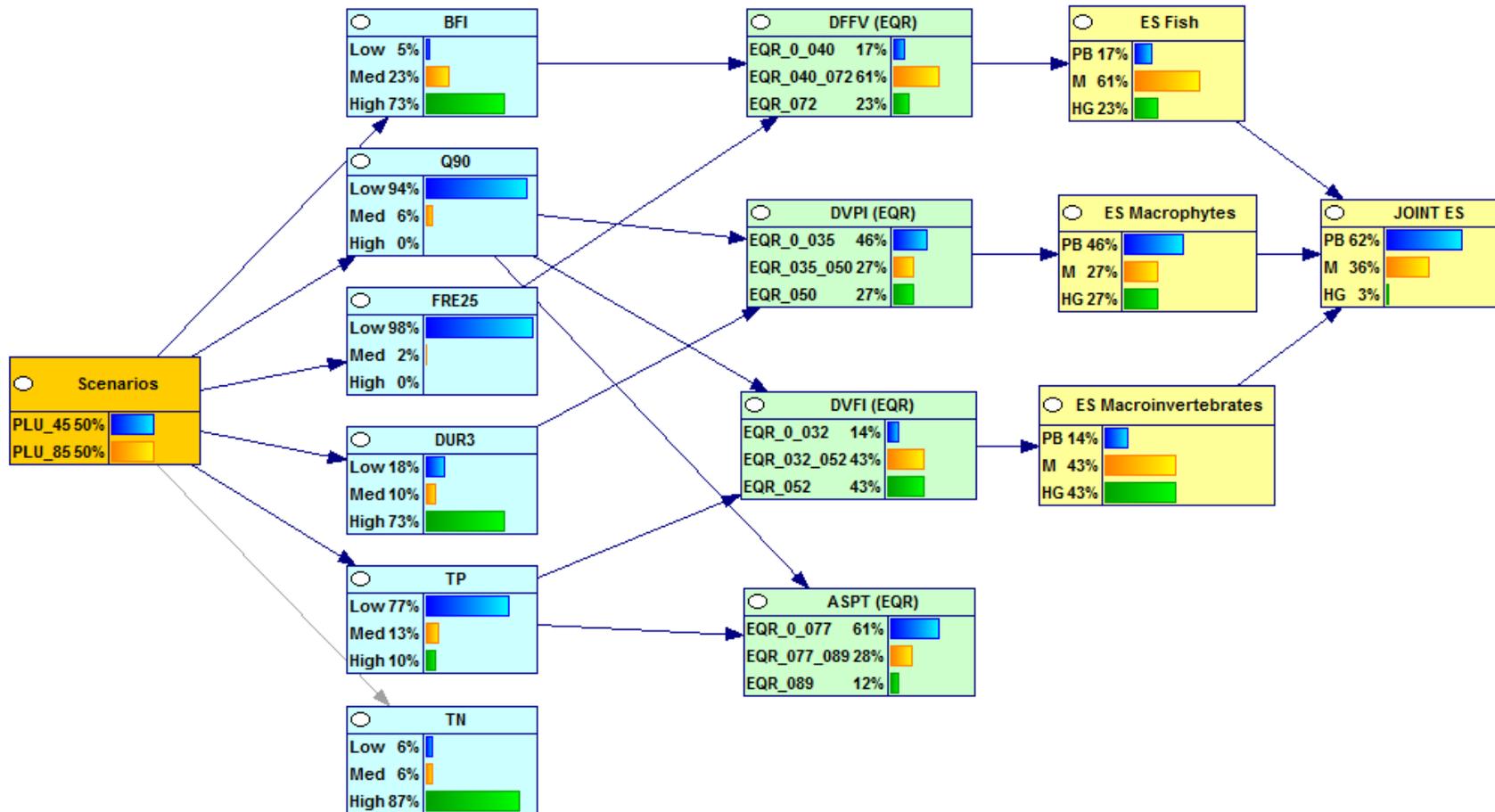
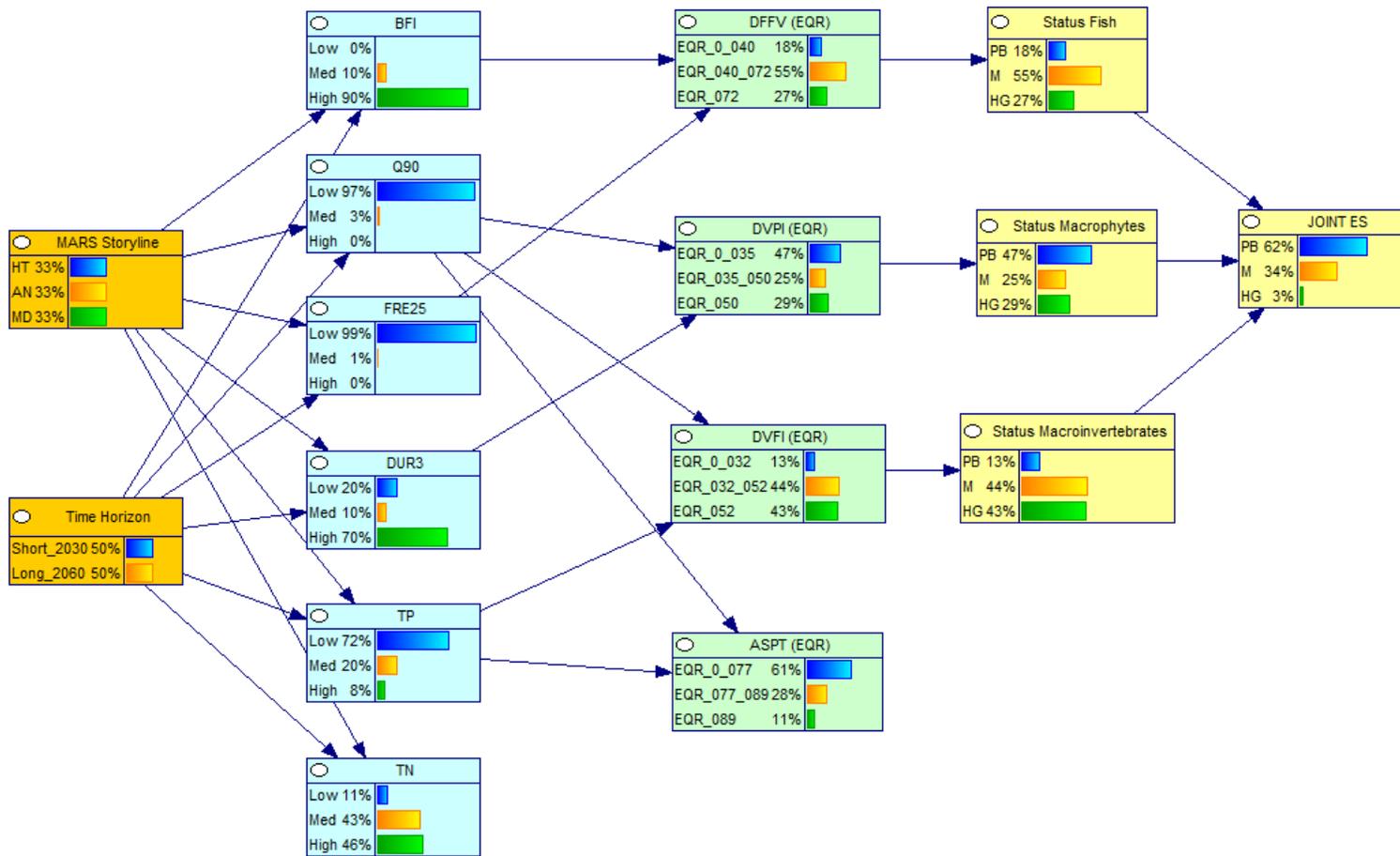


Figure A1.3. BBN Version 3: MARS Storylines combined scenarios. The BBN shows the conditional probability for each node level considering equal probabilities in scenarios.



Appendix 2: Conditional probability tables (CPTs) of case study Odense (Denmark)

Table A2.1. Conditional probabilities between isolated land use change scenarios and baseline scenarios and stressors.

	PLU	HT	AN	MD	PLU_4.5	PLU_8.5
	BFI				BFI	
Low	0.032258	0	0	0.032258	0.032258	0.064516
Med	0.258065	0.129032	0.096774	0.322581	0.193548	0.258065
High	0.709677	0.870968	0.903226	0.645161	0.774194	0.677419
	Q90				Q90	
Low	0.967742	0.935484	1	0.967742	0.967742	0.903226
Med	0.032258	0.064516	0	0.032258	0.032258	0.096774
High	0	0	0	0	0	0
	FRE25				FRE25	
Low	0.870968	0.903226	0.967742	0.870968	0.967742	1
Med	0.129032	0.096774	0.032258	0.129032	0.032258	0
High	0	0	0	0	0	0
	DUR3				DUR3	
Low	0	0.032258	0	0.032258	0.16129	0.193548
Med	0.129032	0.193548	0.129032	0.129032	0.193548	0
High	0.870968	0.774194	0.870968	0.83871	0.645161	0.806452
	TP				TP	
Low	0.677419	0.645161	0.677419	0.677419	0.709677	0.83871
Med	0.258065	0.322581	0.258065	0.258065	0.193548	0.064516
High	0.064516	0.032258	0.064516	0.064516	0.096774	0.096774
	TN				TN	
Low	0.064516	0.096774	0.225806	0	0.064516	0.064516
Med	0.096774	0.774194	0.709677	0	0.064516	0.064516
High	0.83871	0.129032	0.064516	1	0.870968	0.870968

Table A2.2. Conditional probabilities between MARS storylines (combined land use and climate changes scenarios) and stressors.

Time Horizon	HT		AN		MD	
	2030	2060	2030	2060	2030	2060
BFI						
Low	0	0	0	0	0	0
Med	0.096774	0.032258	0.096774	0.064516	0.193548	0.096774
High	0.903226	0.967742	0.903226	0.935484	0.806452	0.903226
Q90						
Low	1	1	0.967742	0.967742	0.935484	0.967742
Med	0	0	0.032258	0.032258	0.064516	0.032258
High	0	0	0	0	0	0
FRE25						
Low	1	1	1	1	1	0.967742
Med	0	0	0	0	0	0.032258
High	0	0	0	0	0	0
DUR3						
Low	0.193548	0.354839	0.193548	0	0.193548	0.258065
Med	0.064516	0.193548	0.064516	0.032258	0.032258	0.193548
High	0.741935	0.451613	0.741935	0.967742	0.774194	0.548387
TP						
Low	0.741935	0.709677	0.677419	0.709677	0.741935	0.709677
Med	0.193548	0.258065	0.193548	0.225806	0.129032	0.225806
High	0.064516	0.032258	0.129032	0.064516	0.129032	0.064516
TN						
Low	0.096774	0.096774	0.193548	0.290323	0	0
Med	0.483871	0.83871	0.580645	0.580645	0	0.096774
High	0.419355	0.064516	0.225806	0.129032	1	0.903226

Table A2.3. Conditional probabilities between stressors and indicators. The same CPT is used in the three versions of the BBN.

DFFV									
BFI	Low			Med			High		
FRE25	Low	Med	High	Low	Med	High	Low	Med	High
EQR Low	0.5000	0.2000	0.5000	0.0000	0.3333	0.1667	0.2000	0.0000	0.1429
EQR Med	0.5000	0.6000	0.1667	1.0000	0.1667	0.1667	0.5000	0.2000	0.0000
EQR High	0.0000	0.2000	0.3333	0.0000	0.5000	0.6667	0.3000	0.8000	0.8571
DVPI									
Dur3	Low			Med			High		
Q90	Low	Med	High	Low	Med	High	Low	Med	High
EQR Low	0.1667	0.1667	0.0435	0.3333	0.2727	0.3333	0.5882	0.0000	0.0000
EQR Med	0.1667	0.1667	0.3913	0.3333	0.3182	0.3333	0.2353	1.0000	1.0000
EQR High	0.6667	0.6667	0.5652	0.3333	0.4091	0.3333	0.1765	0.0000	0.0000
DVFI									
Q90	Low			Med			High		
TP	Low	Med	High	Low	Med	High	Low	Med	High
EQR Low	0.1429	0.0667	0.0909	0.3750	0.1538	0.2143	0.0000	0.0833	0.0000
EQR Med	0.4286	0.4667	0.5455	0.1250	0.2308	0.6429	0.0588	0.0833	0.0000
EQR High	0.4286	0.4667	0.3636	0.5000	0.6154	0.1429	0.9412	0.8333	1.0000
ASPT									
Q90	Low			Med			High		
TP	Low	Med	High	Low	Med	High	Low	Med	High
EQR Low	0.5714	0.6111	0.9286	0.5000	0.6923	0.7333	0.0588	0.1667	0.0000
EQR Med	0.2857	0.3333	0.0714	0.3750	0.2308	0.2667	0.4706	0.6667	0.6667
EQR High	0.1429	0.0556	0.0000	0.1250	0.0769	0.0000	0.4706	0.1667	0.3333

Table A2.4. Conditional probabilities between EQR levels for the biological indices and ecological status classes. The same table is applied for every index.

Biological Index			
	EQR Low	EQR Med	EQR High
PB	1	0	0
M	0	1	0
HG	0	0	1



Table A2.5. Conditional probability table to determine the joint ecological status from the three Danish biological indices.

Status Fish	PB									M									HG											
Status Macrophytes	PB			M			HG			PB			M			HG			PB			M			HG					
Status Macroinvert.	PB	M	HG																											
PB	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	0	0	1	1	1	1	0	0	1	0	0	1	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	0	0	0	0	1	1	0	1	0	0	1	0
HG	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Appendix 3: Stressors probability distributions of case study Odense (Denmark)

Figure A3.1. Probability distributions (%) for stressors for the isolated land use change scenarios.

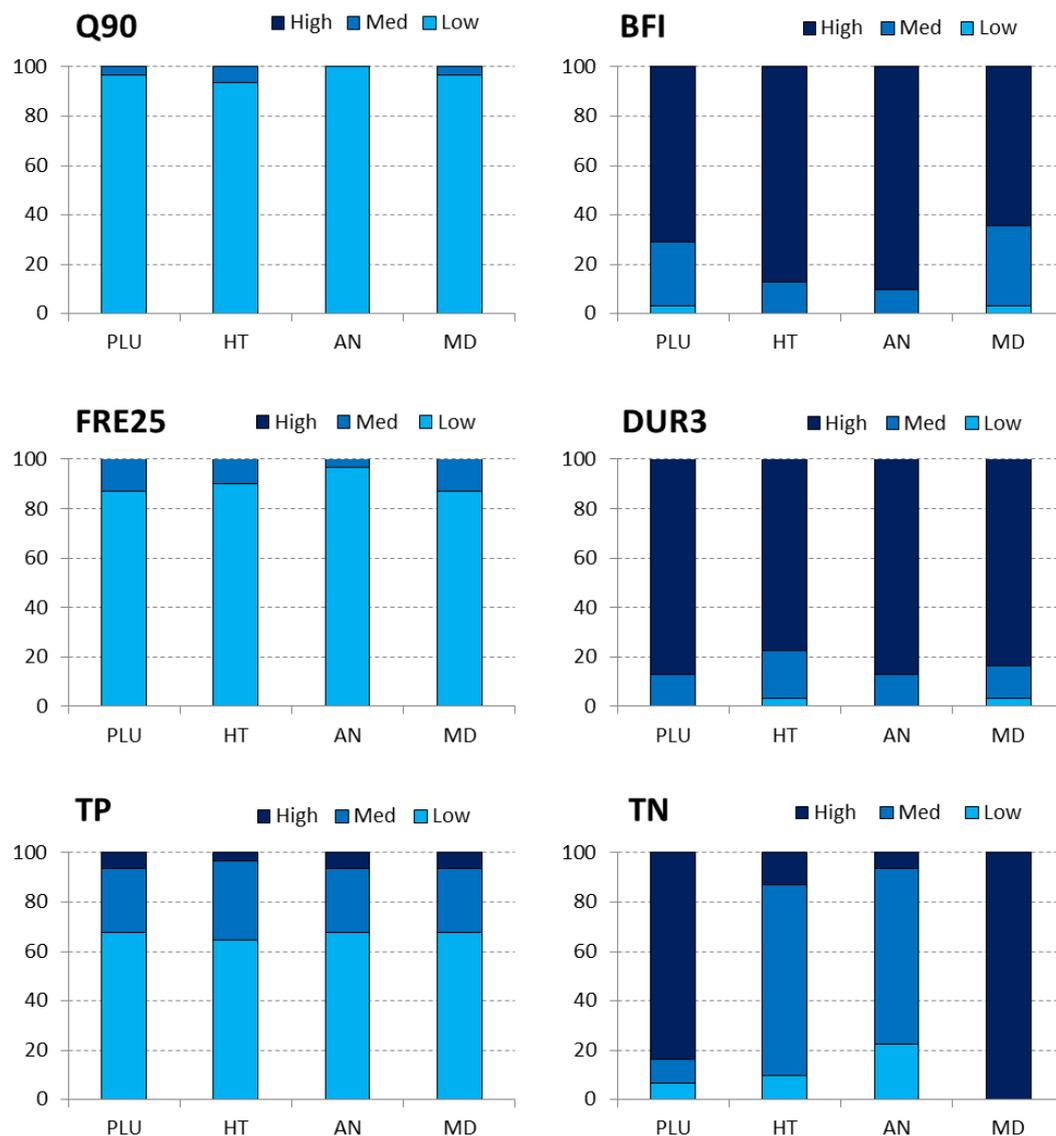
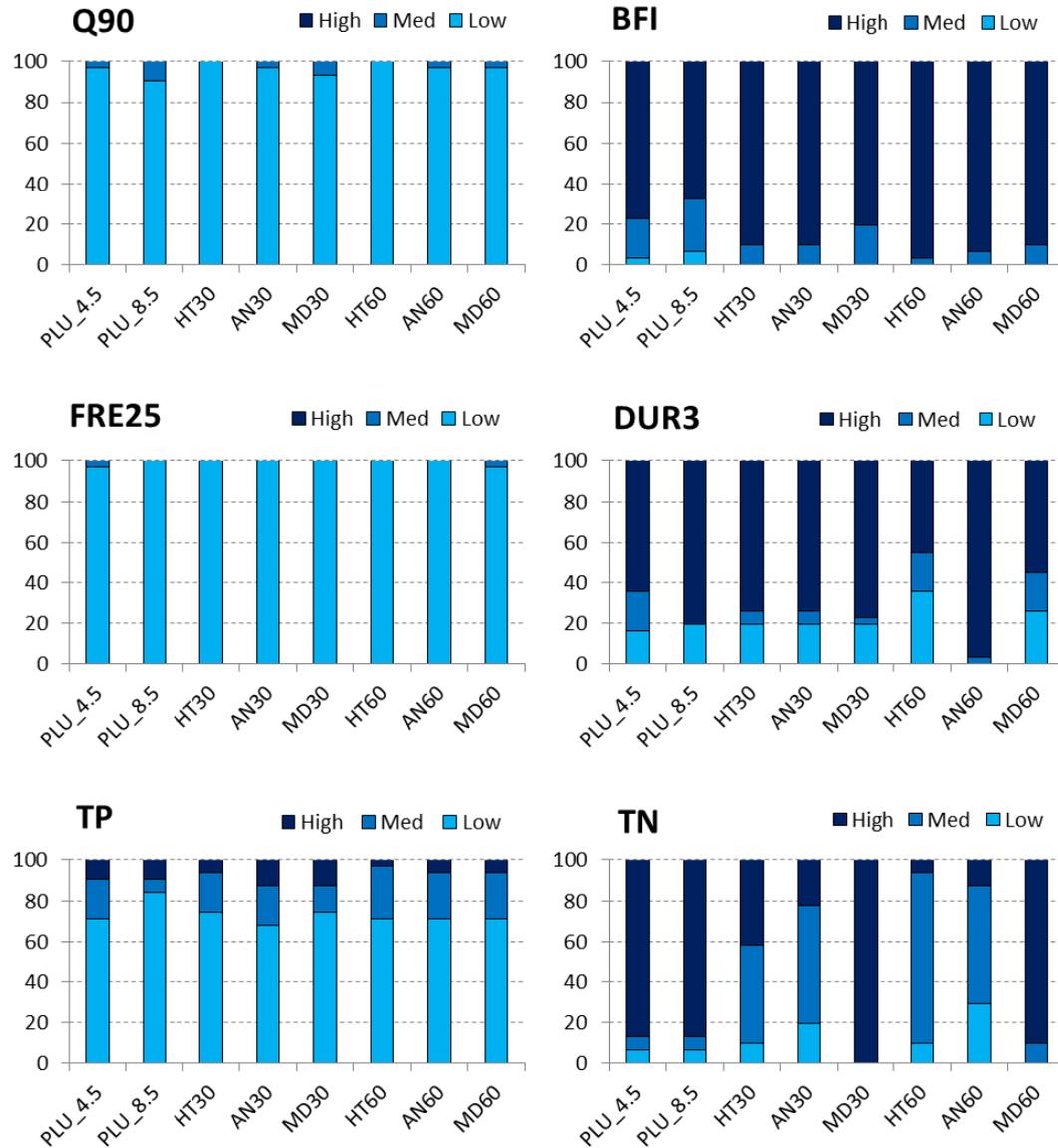


Figure A3.2. Probability distributions (%) for stressors for baseline (PLU_4.5, PLU_8.5) and MARS storylines scenarios.



Appendix 4: Results of scenarios of case study Regge and Dinkel (The Netherlands)

Figure A.4.1. Results of different scenarios for maintenance of submerged vegetation

Abbreviations: FH: Discharge high, FL: Discharge low, PL: Depth profile surface area large, PS: Depth profile surface area small, DN: no dams present, DY: dams present, RL: maintenance in the riparian zone low, RH: maintenance in the riparian zone high, SH: maintenance in the submerged zone high, SL: maintenance in the submerged zone low.

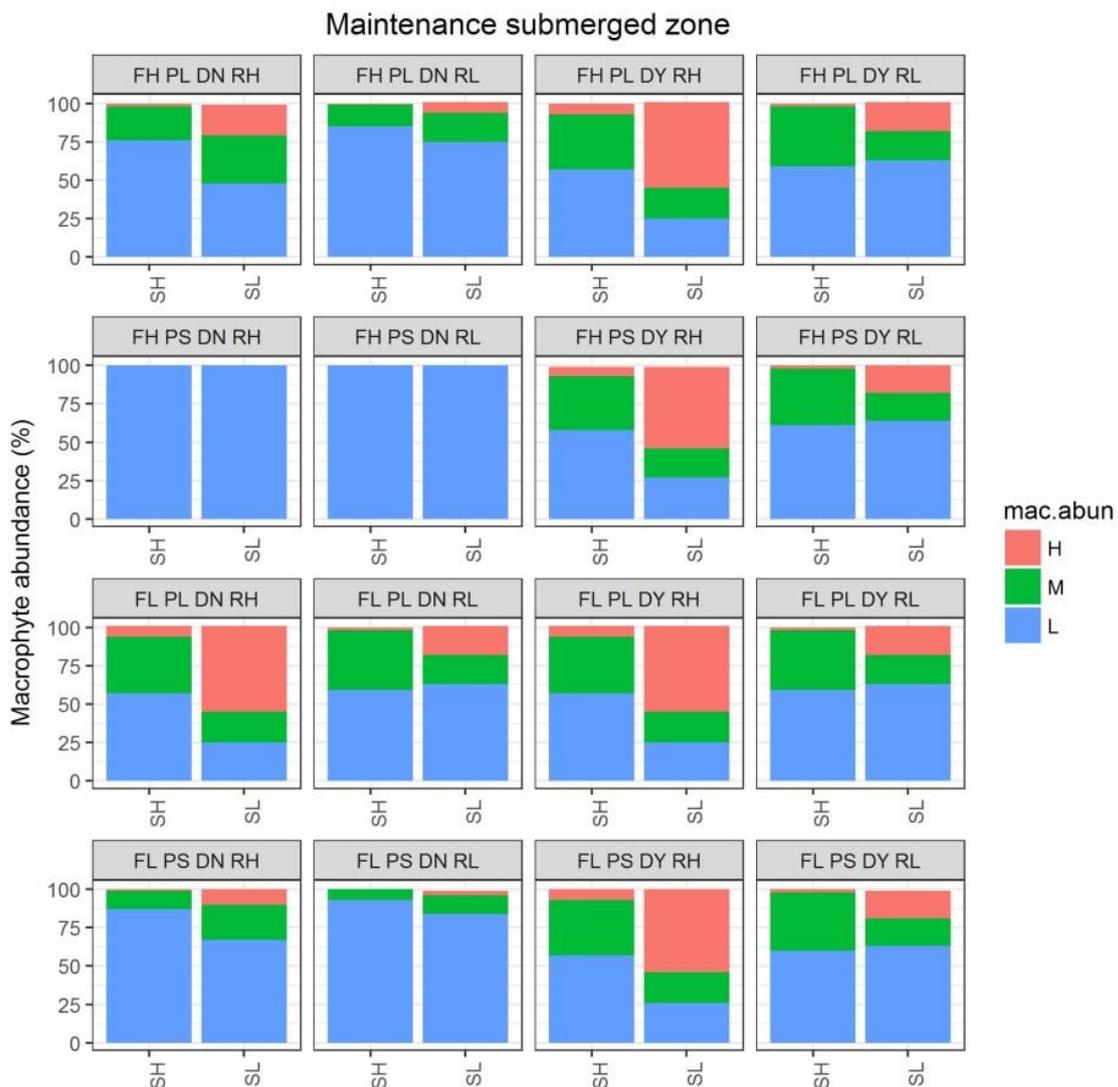


Figure A.4.2. Results of different scenarios for river discharge

Abbreviations: FH: Discharge high, FL: Discharge low, PL: Depth profile surface area large, PS: Depth profile surface area small, DN: no dams present, DY: dams present, RL: maintenance in the riparian zone low, RH: maintenance in the riparian zone high, SH: maintenance in the submerged zone high, SL: maintenance in the submerged zone low.

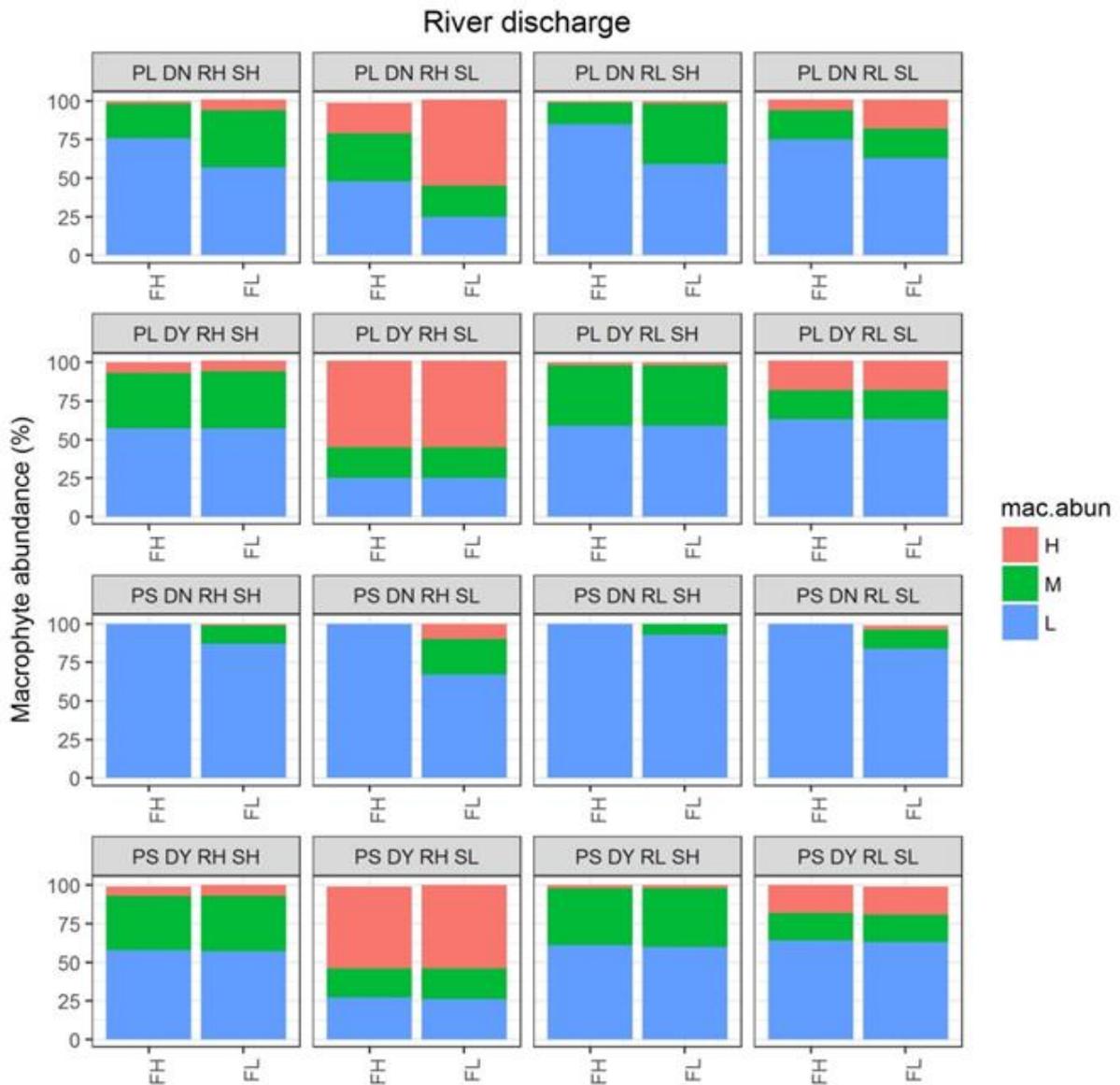


Figure A.4.3. Results of different scenarios for maintenance of riparian vegetation

Abbreviations: FH: Discharge high, FL: Discharge low, PL: Depth profile surface area large, PS: Depth profile surface area small, DN: no dams present, DY: dams present, RL: maintenance in the riparian zone low, RH: maintenance in the riparian zone high, SH: maintenance in the submerged zone high, SL: maintenance in the submerged zone low.

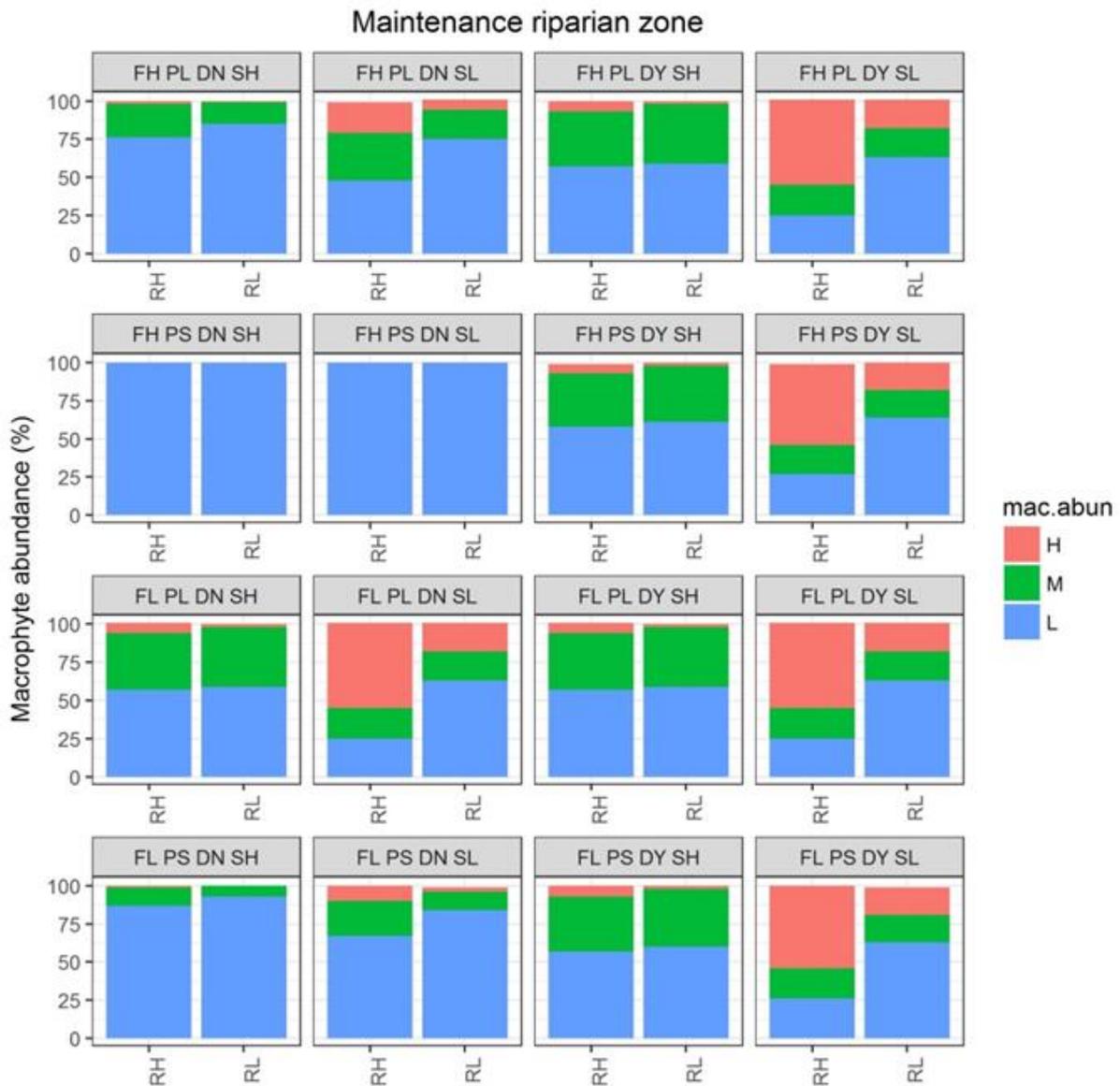


Figure A.4.4. Results of different scenarios for depth profile

Abbreviations: FH: Discharge high, FL: Discharge low, PL: Depth profile surface area large, PS: Depth profile surface area small, DN: no dams present, DY: dams present, RL: maintenance in the riparian zone low, RH: maintenance in the riparian zone high, SH: maintenance in the submerged zone high, SL: maintenance in the submerged zone low.

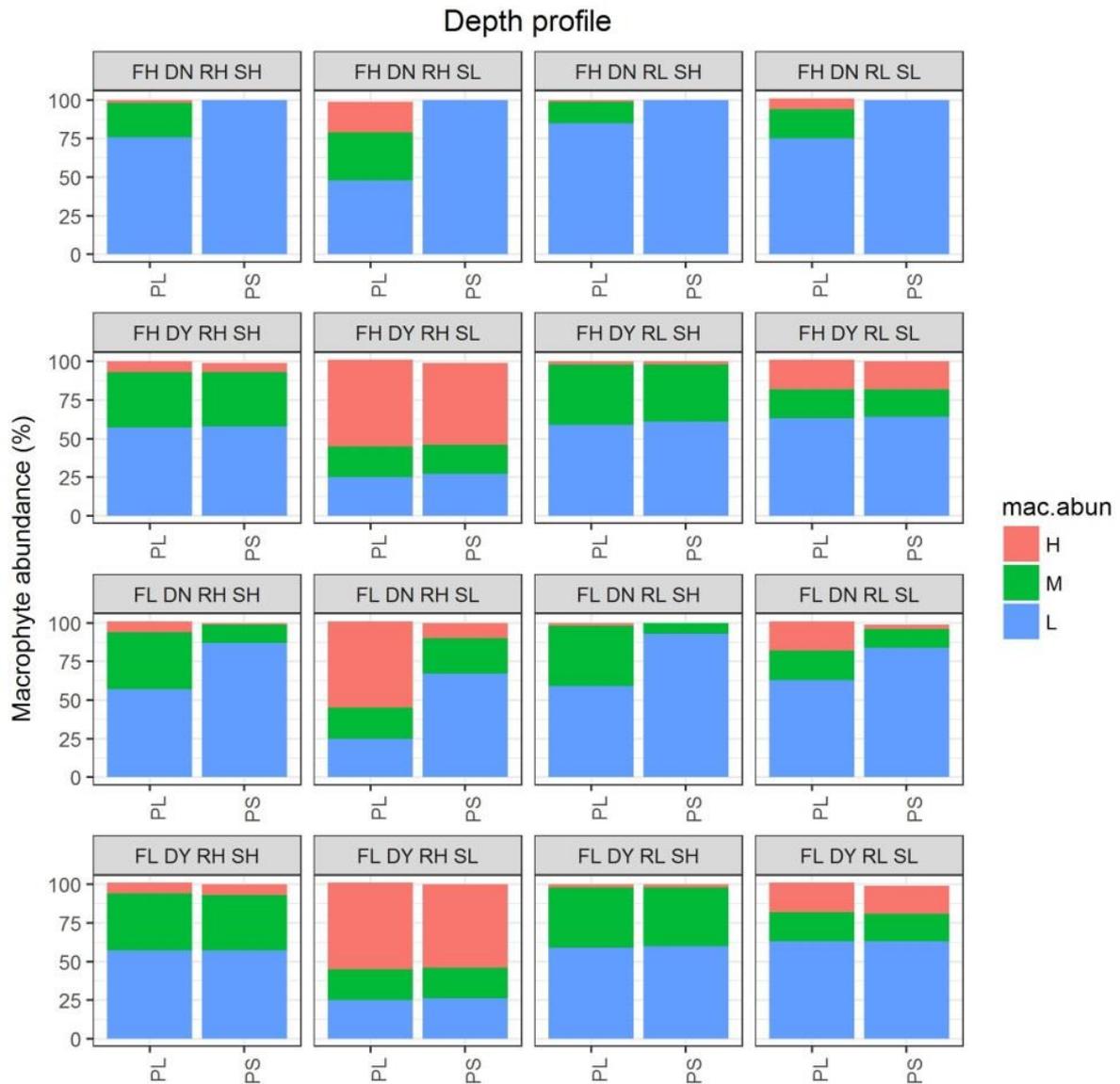


Figure A.4.5. Results of different scenarios for presence of dams

Abbreviations: FH: Discharge high, FL: Discharge low, PL: Depth profile surface area large, PS: Depth profile surface area small, DN: no dams present, DY: dams present, RL: maintenance in the riparian zone low, RH: maintenance in the riparian zone high, SH: maintenance in the submerged zone high, SL: maintenance in the submerged zone low.

