The Rotte River Fish Migration Project

A Bayesian Network approach for fish habitat suitability

Msc thesis graduation report

Gerard J.L. Vos

August 15, 2023

The River Rotte Fish Migration Project

A Bayesian Network approach for fish habitat suitability

by

G.J.L. Vos

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on $August 18^{th} 2023$.

Student number:	9906213	
Thesis committee:	Prof. dr. M. McClain,	IHE, supervisor
	Dr. E. Mostert,	TU Delft
	Dr. ir. G. Schoups,	TU Delft

Preface

And with this report of my Msc-research comes an end to my student life. In this period I have made new friends, gathered incredible knowledge and developed technical and analytical skills. I am looking forward to apply all these achievements in the working field of ecological restoration and nature preservation, which I have become acquainted with during my years of study. This underlying research once more illustrated the enthusiasm and joy of the people involved in this field of work.

For me, ecological research is also an entwining helix of political reality, the (almost) boundless ecological perspectives and sensible technical and mathematical applications. This mixture is represented by my highly appreciated graduation committee, as well. I would like to thank Prof. dr. Michael McClain for his always positive and progressive approach which motivated me to push boundaries, but to move along determinations and observations as well. I would also like to thank Dr. Erik Mostert for his realism in terms of political needs and in terms of the requirements for my graduation. And last but not least, I would like to thank Dr. ir. Gerrit Schoups for his support to explore the probabilistic approach in Bayesian Networks.

Furthermore, I would like to thank my family and friends for the excitement and interest that they shared with regard to my study. And of course my thanks to my beloved Tara and Loïc, who give me so much love and laughter.

G.J.L. Vos
Delft, August 2023

Summary

The political interest in fish habitat suitability is increasing. An important action in ecological restoration of freshwater bodies is (re)establishing longitudinal connections between upstream and downstream waters, by removing dams and creating fish passages. The presence of favourable habitat conditions is essential for the effectiveness of these actions. Also, the fish population partly determines the quality status of water bodies within the Water Framework Directive (WFD) in the Netherlands. Moderating the fish population, for example by changing fish habitat conditions, can contribute to this ecological quality rating.

It is clear that habitat conditions are a key factor for the presence of different fish species. But the relationship between these conditions and the presence of fish species is complex, not fully understood and certainly not always applied directly in political decision making. In this respect, ecological models can help to simplify the analysis of these complex systems and to analyse the main factors for habitat suitability.

The aim of the thesis is developing an ecological model for fish habitat suitability in the river Rotte basin, usable in policy development for ecological restoration. The model functions as a decision support system to explain the differences in fish population, as a result of the local conditions in different parts of the Rotte. Next to that, the model is a tool to evaluate the impact of management actions to modify fish habitat suitability specifically for plant-loving and benthivorous fish species.

The river Rotte is a heavily modified water body in the Province of South-Holland, the Netherlands. Recently, the connectivity between the Rotte and the Nieuwe Maas has been restored by implementing a fish passage and making the main pumping station (Mr. U.G. Schilthuis) ”fish passable”. The connections have been restored in order to facilitate fish migration, like the European Eel, but also to facilitate fish movement in general, to expand living conditions and open up new habitat.

Furthermore, the Rotte is a WFD designated water body, which means that there´s a political goal to realise a ”good ecological potential”. However, the current status varies between ”poor” and ”moderate” and this is partly caused by the unbalance between plant-loving and benthivorous fish species. The WFD-status also means that regular monitoring of the physio-chemical, hydromorphological and biological conditions is mandatory. The physio-chemical conditions include, among others, water quality and nutrient loads. Hydromorphological conditions describe the shape of the water bodies, water depth and riparian zone. Biological conditions focus on the presence of aquatic plants, invertebrates and fish species.

The Rotte makes for an important case study, because there are many water bodies with similar characteristics, especially in lowland river areas. In the Netherlands alone there are at least 36 similar water bodies, including implementation fish passages and monitoring of physio-chemical, hydromorphological and biological data.

The model developed in this thesis is a Bayesian Belief Network model (BBN) that predicts habitat factors for food preference and preference for habitat structure. This model is built using the NETICA software package by Norsys Systems. The relatively simple structure consists of two target variables, which are the habitat suitability for the plant-loving and benthivorous fish species, and ten explanatory variables for physio-chemical, hydromorphological and biological conditions. The model predicts the probability distribution for the fish habitat suitability for plant-loving and benthivorous fish species and is based on machine learning with a set of cases from the monitoring data. These cases, 2,000 in total, consist of a combination of variable states at a specific time and location as observed in the Rotte basin.

The BBN model has been applied to assess the impact of local conditions on fish habitat differentiation. This application is based on *probabilistic inherence*, for which observations from the monitoring data will be entered in the model as evidence. By adding this evidence to the explanatory variables, the target variables will converge towards one of the states ”Good” or ”Moderate”. The direction and the magnitude of this convergence implies the impact of the explanatory variable on the fish habitat suitability.

Next to that, a methodology has been developed to evaluate the impact of management actions on fish habitat suitability for plant-loving and benthivorous fish species. For this application, the prior probability distribution of the physio-chemical and hydromorphological variables in the model have been adjusted, according to intended effect of the management action. For example: to apply a scenario for the reduction of nutrient loads, the probability distribution for the model variable *Trophic state* has been adapted, to reduce the frequency of the higher trophic states in favour of the lower trophic states. The adjustment of the prior probability distributions finally affect the probability distribution of the target variables. From the changes in the probability distribution of the target nodes, the impact of the management actions on the habitat suitability and, eventually, fish population composition can be derived.

The research has resulted in a decision support system for ecological policy making and a methodology to apply this model for the assessment of local conditions or to simulate scenarios for management actions. However, the current model version shows to be inadequate in terms of model prediction accuracy. Model evaluation shows that the model predictions depend largely on sampling data and not on model structure, which indicates that the data set for learning is not sufficiently balanced. This data limitation can be solved by expanding the scope of the model to other water bodies in the Netherlands, by adding learning cases from the monitoring data. Another option to improve the model is to use metamodels for (ecological) system processes for specific model variables.

The model prediction accuracy can also be improved by expanding the model structure with additional habitat factors, beside the current preferences for food and habitat structure. A major advantage of BBN models is the flexibility to expand the model structure, without a complete revision.

To conclude this research, the BBN model is functional and usable for policy making. Also, the BBN approach has proven to be a powerful communicative tool to engage a variety of stakeholders in ecological restoration. Nevertheless, the current version of the model needs improvement to enhance model prediction accuracy. The main improvement is tuning the model by expanding the scope to similar water bodies in the Netherlands. After that, the model has to be evaluated for its performance, sensitivity of the model variables and the quality of the set of learning cases via cross-validation. As a third step, expansion of the model structure should be discussed with ecologists and other stakeholders, to include more habitat factors to determine fish habitat suitability.

Contents

Chapter 1

Introduction

Approximately 70% of the rivers world wide are non-free flowing due to human interventions and in two-third of these cases the longitudinal connection is obstructed with dams, weirs, culverts or sluices [19]. In the European rivers there are over 1 million barriers, resulting in a barrier density of one barrier per 1.5 kilometre river length. textcolorredIn this respect, the Netherlands is the front runner with a barrier density of 20 barriers per kilometre river length [3]. These figures illustrate that fragmentation and disruption of longitudinal connectivity is a major human pressure on freshwater systems; a pressure with huge consequences for ecosystems and biological systems, such as the living conditions of ichthyofauna (or fish).

The barriers impede migration and dispersal of fish species and negatively impact fish distribution, abundance and productivity ([6], [33], [39], [53]). The disruption of longitudinal connectivity is a threat for ecosystem functioning, due to the ecosystem services that fishes provide and for biodiversity ([3], [18], [19], [51]). For good reason, the restoration of longitudinal connectivity is a key pillar for many ecology and water system management strategies, such as the United Nations Decade on Ecosystem Restoration [5], the European Biodiversity Strategy 2030 [11] and the Wild and Scenic Rivers Act in the U.S. [39]. A frequent intervention, in this respect, is dam removal or the implementation of fish passages.

The availability of suitable habitat conditions is an essential factor for the efficacy of restoring longitudinal connectivity, next to the ecological targets and the physical design of the fish passage ([40], [51]). But the presence of suitable conditions is not certain. And this applies the more so for heavily modified water systems, based on the high level of poor and moderate ecological potential in these specific water bodies [13]. For the case in which a lentic (stagnant) water system, such as a canal system or polder system, is connected to a lotic (fluvial) water system, the distance between required and preferred habitat conditions is most likely even greater. Thus, one of the main challenges for effective restoration of longitudinal connectivity is the realisation of favourable upstream habitat conditions. This challenge is becoming more complicated in heavily modified and regulated water systems, which have, in general, more restrictive constraints for water management options due to other user requirements (e.g. flow dynamics, water level variations, extension to flood plains). And in some cases these challenges are directly related to the (re)connection between lentic and lotic water systems. The complexity of ecological systems and processes hamper the analysis of (causal) relationships and approaches for ecological restoration. But habitat modelling techniques can be deployed to unravel these complexities. The models can be used to assess the main drivers for habitat suitability and to determine whether improvements of the habitat conditions are possible.

The focus in this study is on the particular challenge for the restoration of heavily modified water systems and the connection between lotic and lentic water systems. This specific challenge is common in lowland rivers, where polder and canal systems are being reconnected to rivers and transitional waters. In the Netherlands, for example, there are at least 36 projects in designated EU Water Framework Directive (WFD) water bodies which connect stagnant systems to fluvial systems [42]. The river Rotte, a storage canal in the Netherlands, will be used as study area to investigate a modelling approach for the fish habitat suitability (FHS) to support political decission making. This former peat drainage river is under authority of the Hoogheemraadschap Schieland en Krimpenerwaard (HHSK) and is designated as a heavily modified water body for the WFD. The WFD imposes requirements on ecological quality in the Rotte, but achieving the quality levels is a challenge; the physio-chemical status is poor and the ecological status varies between insufficient and moderate. The poor ecological conditions is partly caused by the current fish population distribution, because of the high percentage of benthivorous fish species in relation to plant loving species ([26], see also paragraph 2.2). Furthermore, a fish passage has been realised between the river Rotte and the Nieuwe Maas in 2022. Thus, there is an interest to look into FHS for the Rotte, both from the perspective of the fish passage and the political objectives with respect to the WFD.

The main purpose of the research is to develop a decission support system (DSS) for the ecological policy with a focus on fish habitat conditions. In this respect, the target groups in this research are the plant loving fish species, with a plant-based diet and a habitat in or above vegetation, and the benthivorous fish species, with a diet consisting mainly of benthic life (macrofauna) and a habitat preference for soft bottoms. The DSS has to be suitable for the assessment of current fish habitat conditions as well as improvements of these conditions due to (management) interventions. Obviously, these interventions will be limited by other functions and priorities in the Rotte basin. This model will be based, primarily, on available data resources of the water authority.

The research questions in this study focus on model development of the DSS and the application of the model for policy making. The research questions with respect to model development are:

1. What modelling technique suits best given the context and data availability of the Rotte basin?
2. How does the choice for the model variables determine the accuracy of the model predictions?

The research questions with respect to model application are:

3. How can the model determine the impact of local conditions on fish habitat differentiation among the target groups?
4. How does the model respond to different management options and what does this imply for the FHS and population composition?

The aim of this study is to contribute to the research and actions for ecological restoration of freshwater bodies within the WFD and beyond. The general topic in this research is the evaluation and improvement of habitat conditions for heavily modified water systems. In the European Union $12-13\%$ of all surface water bodies is designated as "heavily modified" [14]. Management options for these water systems to improve ecological conditions are equally impaired, just like the Rotte. The habitat modelling approach, assessed in this case study, serves as an example to explore management options and their impact on habitat conditions.

The modelling approach developed in this project will be applicable for a wider spectrum of water systems within the European Union (EU). The starting point of the habitat modelling technique is the use of available data resources of the water authority (HHSK). Since the Rotte is a designated WFD water body with specific monitoring requirements, it is very likely that comparable data are available for other water systems in the EU. Therefore the developed modelling technique can be applied with limited adaptations to other designated WFD water systems.

Finally, HHSK has restored the connectivity between a lentic and a lotic water system. Fish passages between those water types facilitate eurytope fish species and specific diadromous fish species such as the European Eel (*Anguila anguila*) and the Three-thorn stickleback (*Gasterosteus aculeatus*). Eurytope fish species are, unlike rheophilic species, less vulnerable for habitat loss, because they are less critical [53]. This research contributes to the discussion in the Netherlands for the usefulness and necessity to implement fish passages between lentic and lotic water systems and for the definition of ecological targets defined in these projects.

Reading guide

The next section (chapter 2) presents the study area. This chapter includes the characteristics of the Rotte basin, the policy framework and the available data resources. Chapter 3 discusses the methodology for model development and the application of the model. Chapter 4 presents the results, followed by the discussion in chapter 5. This report closes with the conclusions in chapter 6. Furthermore, an explanation of specific terms and terminology with respect to biological issues and water management can be found in the Glossary.

Chapter 2

Study area

Introduction

This chapter gives a brief description of the Rotte basin in paragraph 2.1. Followed by the policy framework (paragraph 2.2), an inventory of the available data (paragraph 2.3) and a description of the fish population in the Rotte basin and connected water bodies.

2.1 Description of the study area

The Rotte is a former brook for the drainage of peat lands into the river Maas. In the 10^{th} century, farmers start to dig channels for dewatering agricultural land into the Rotte. In the 12^{th} and 13^{th} century, the residents of the surrounding villages created dikes and dams for flood protection. The most downstream dam was constructed in 1270 at the mouth of the river Nieuwe Maas. The construction of this dam was the starting point for the development of the city of Rotterdam, today the second city of the Netherlands in terms of inhabitants and the largest port in Europe.

The dewatering and land reclamation around the Rotte results in land subsidence and, eventually, the land has become lower than the river. Windmills had to be built to maintain the water level in these polders and the Rotte transformed from a natural brook into a storage canal. Various projects of canalisation, for both navigation and drainage, have resulted in the current water system for the Rotte basin, including the position and elevation in the landscape, the design of the riparian zone and shore lines and the locations of infrastructure for drainage or water inlets.

The total length of these branches is $22 kilometres$ (km) and the total water surface of $180 hectare$ (ha). The total drainage area of the Rotte basin is $14,250 ha$. The pumping station Mr. U.G. Schilthuis (maximum capacity of $1,350 m^3/minute$) located downstream of the Boezem in Rotterdam, discharges from the basin into the Nieuwe Maas.

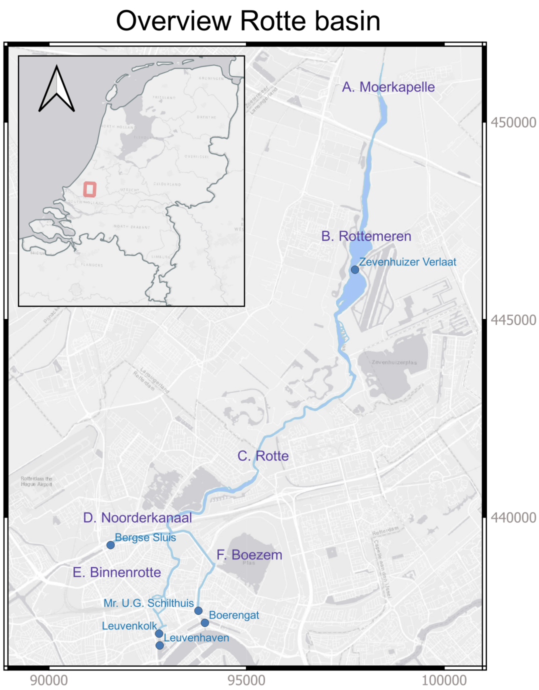


Figure 2.1: Overview Rotte basin.

The Rotte basin consists officially of five sections: Moerkapelle, Rottemeren, Rotte, Noorderkanaal, Binnenrotte and Boezem (see figure 2.1). The sections in the northern part of the basin have lake shaped water bodies (Moerkapelle and Rottemeren), while the line shaped Rotte section connects these lakes to the southern sections. This northern part is characterised by the elevated position of the river above the landscape. The river shores are dominated by clay levees covered with grass, dense reed structures (helophyte) in the littoral zone and a relatively steep riverbed profile following directly after the littoral zone. The dominant land use in the surrounding polders is agriculture and urban land use. Furthermore, the Rottemeren is a popular location for recreation (e.g. fishing and navigation). The southern sections (Noorderkanaal, Binnenrotte and Boezem) are all located in urban area. In this area, vertical embankments are dominant and littoral vegetation is scarce and fragmented. The dominant pressures in the Rotte basin are nutrient loads from agricultural production and canalization. Water management is prioritized by flood protection and water level regulation for agriculture and for the protection of the foundations of constructions ([22]).

The objective for the Rotte within the WFD is a "good ecological potential", which includes, among others, a healthy fish stock and composition [8]. HHSK has recently restored the (ecological) connectivity between the Rotte basin and the Nieuwe Maas. During a renovation in 2015 the pumping station Mr. U.G. Schilthuis has been made "fish passable" in both directions. Due to the renovation, fish can safely move between the Nieuwe Maas into the Rotte basin during operation as well as in inactive mode. In 2022 a new fish passage has been opened between Leuvenhaven and Leuvenkolk. The fish passage will be closed when salinity of the Nieuwe Maas is becoming too high (more than $400ppt$) or when flow velocity in the fish passage is becoming to fast due to high water levels in the Nieuwe Maas (more than $1.80mNAP$). Beside from the facilitation of fish migration, the realisation of the fish passages serve other functions. The inlet of relatively nutrient poor water from the Nieuwe Maas into the Rotte improves the water quality. Furthermore, the fish passage is part of the urban development plan and contributes to the environmental quality of the city centre.

The Rotte basin is, for the most part, disconnected for fish from the surrounding polders and adjacent water systems. Fish can only incidentally move between the Rotte and the polders via various sluices or via pumping stations. An increased connectivity between Rotte and polders is desired by the water authority to increase habitat availability and an expansion of available habitat types. Different projects for the realisation of fish passages are being studied, for example at the Bergse Sluis and Zevenhuizer Verlaat, but realisation of this connectivity is not expected within $3-4$ years from now (M.Meier HHSK, personal communication, $September 26^{th} 2022$).

2.2 Policy framework

The relevant policy frameworks and regulations in relation to fish habitat concern water management regulations, the ecological directives WFD and the European Eel Regulation, the mowing strategy and the regulation of nutrient levels. These frameworks will be discussed briefly in the following section. Focus is on the management options, and the limitations and opportunities these options impose on the FHS.

Flood protection and water level regulation

The main priority for the water authority in the Rotte basin is flood protection and water level regulation for the adjacent polders. The minimum height of the embankment of the Rotte reaches until $-0.30mNAP$. The water authority applies a flexible water level in the Rotte basin, with water levels varying between $-1.20mNAP$ and $-0.90mNAP$. During extreme calamities a (temporally) water level of $-0.65mNAP$ is allowed [22]. Main challenges in water management for the water authority are climate adaptation and limitations of the total drainage capacity of the Rotte basin. Lateral spread out is available, though limited, to the designated floodplains, but this storage capacity is insufficient for extreme weather conditions. HHSK realised in 2013 a water retention area in the Eendragtspolder as part of the climate adaptive measurements. This multi-functional retention area has a storage capacity of 4 million cubic meters and will be used when water levels in the Rotte exceed -0.90 mNAP.

The drainage capacity of the Rotte basin is confined by the capacity of the pumping station Mr. U.G. Schilthuis but even more so by the hydraulic conveyance capacity of the Boezem. At full pumping capacity, the slope of the water line in the Boezem is becoming too steep due to the limited conveyance capacity, which causes inconveniences for the housing boats in the Rotte basin. Thus, adjustments to the design of the Boezem are restricted and further reduction of the conveyance capacity are not allowed ([22] and E. Venneman HHSK, personal communication, $November 23^{rd} 2022$).

European Water Framework Directory and Eel Regulation

The Rotte is classified as M3-type: a buffered regional canal because of the lentic conditions and the predominant linear shape [15]. As mentioned earlier, the Rotte is a heavily modified water body and the objective within the WFD is a "good ecological potential" according to the ecological reference values for this specific class. Evaluation of the ecological potential is based on biological status (including fish stock), physio-chemical status and hydromorphological status. The adoption of fish as a quality element in the WFD has ensured that fish abundance and composition has become an ecological task for the water authorities, while this task used to be delegated to the regional fishing associations before the ratification of the WFD.

Table 2.1 presents the quality elements for the evaluation of M3-type water bodies:

Biological status	Physio-chemical status	Hydromorphological status
Composition and abundancy phytoplankton	Nutrient levels	Hydrological regime
Composition and abundancy other aquatic flora	Transparency conditions	Morphology
Composition and abundancy macro-fauna	Thermal conditions	
Composition and abundancy fish	Acidification status	
	Oxygenation conditions	
	Salinity conditions	

Table 2.1: WFD quality elements for M3-type water bodies.

HHSK performs a three-annual monitoring of the physio-chemical and biological status of the Rotte basin, including a feasibility assessment for the WFD objectives in 2027. Monitoring results point out that the physio-chemical status is poor, while the biological status varies between insufficient and moderate for all quality elements. For the biological quality element fish, the moderate status in the Rotte is mainly caused by the high abundance of Bream (*Abramis brama*) compared to plant loving fish species and other eurytope species ([46][12], see also paragraph 2.4). The system of Ecological Key Factor (EKF) has been developed, within the Dutch approach for the WFD, to structure the complex ecological relationships within water systems. Evaluation of these key factors give an indication for the deep-rooted causes for the biological and physio-chemical status and determine management options to realise ecological objectives [41]. Following the evaluation of the key factors, the deep-rooted causes for the poor ecological conditions in the Rotte are the high nutrient loads in water and sediment, the turbidity caused by chlorophyl-a and insufficient habitat suitability due to the design of the embankment and the limited availability of shallow waters ([26], [56]).

The water authority has implemented various measures to improve the ecological status in the Rotte basin, for example the design of the riparian zone, the reduction of nutrient pressures and improving connectivity for fish migration. The water authority has developed shallow zones with natural foreshores to stimulate the development of littoral vegetation. For the reduction of nutrient pressures from agriculture, the water authority has launched the programmes Zero-Emission Greenhouse and Delta Plan for Agricultural Water Management. Nevertheless, feasibility of the objectives in 2027 remains uncertain given the monitoring results and the impact of the ecological measures by the HHSK ([26], [27]).

In 2007 the EU put in place the European Eel Regulation (Council Regulation (EC) No 1100/2007) for the long-term management of the eel stock in Europe. The European Eel (*Anguila anguila*) is a diadromous species, or more specific a catadromous species, and migrates from freshwater systems to the Sargasso Sea for spawning. Since 2014, the European Eel is on the IUCN-red list for threatened species and this endangered status is partly caused by the barriers for upstream and downstream migration ([55]).

In line with the EU regulation, the Dutch government developed a National Eel Management Plan. Focal point of the national plan is the removal of all main barriers for Eel migration before 2027 and strict limitations for catching of Eel and guidelines for restocking. HHSK facilitates the National Eel Management Plan with the development of sustainable fish migration measures and the enforcement of the guidelines for fishery and restocking of Eel. The measures and policies with regard to the WFD and the National Eel Management Plan are integrated in the Fish Management Plans of HHSK ([23], [25]).

Eco-colour Course: mowing strategy

Since 2009 HHSK has adopted the Eco-colour Course (ECC) as an ecological mowing strategy. The goal of the ECC is to improve ecological quality, while ensuring other water management functions, such as flood protection, water level regulation and (recreational) navigation. The ECC provides guidelines for the abundance of aquatic and littoral vegetation as well as guidelines for the mowing frequency. Within the ECC, HHSK has determined three categories for the main waterways ([21]).

Table 2.2 presents the guidelines for the ECC and gives an illustration for the ecological classifications:

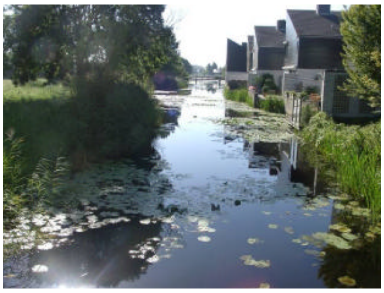
	Blue	Yellow	Green
Principle	Transit has priority	Attention for ecology	Ecology has priority
Abundance	• Max 30% submerged vegetation • Max 20% width of littoral vegetation • Max 50% floating vegetation	• 30%-50% submerged vegetation • 10%-50% width of littoral vegetation • Max 50% floating vegetation	• 30%-75% submerged vegetation • 20%-50% width of littoral vegetation • Max 50% floating vegetation
Mowing frequency	Min 2 times/year	2 times/year (staged)	1 time/year (staged)
Illustration			

Table 2.2: Guidelines for the Eco-colour Course of the HHSK ([21])

The total shore length of the Rotte basin is approximately 56.6 kilometers. Eco-colours Green and Blue are mostly applied in the Rotte basin with, respectively, 85% and 11% of total shore length. The eco-colour Blue can be found in the Southern section of the Boezem because of the primary function of drainage (based on data provided by HHSK February 28^{th}, 2023).

Table 2.3 presents the total shore length for all eco-colours and the relative length:

Eco Colour	Shore length (km)	Rel. length (%)
Green	48.0	85.0
Blue	6.3	11.0
Yellow	2.4	4.0

Table 2.3: Application of eco-colours in the Rotte basin. Data provided by HHSK (February 2023)

2.3 Data availability

This section gives an overview of the relevant data for the Rotte basin with regard to the assessment of fish habitat. The available data will provide the building blocks for the ecological assessment and modelling in this research. Three data categories will be discussed, namely physio-chemical data, biological data and hydromorphological data. Appendix A presents the maps with the monitoring locations for the physio-chemical data, hydromorphological data and the biological data, including the fish monitoring.

Physio-chemical data

The water authority issues 12 locations for the monitoring of physio-chemical data in the Rotte basin. Two additional monitoring locations are located in the basin of the Nieuwe Maas, at Leuvenhaven and Boerengat. Monitoring of physio-chemical data supports the water authority to control water quality, to detect toxic hazards and for the management of water allocation between the Rotte basin, the adjacent polders and the Nieuwe Maas. In the Rotte basin, monitoring takes place for approximately 320 variables, consisting of chemical conditions (e.g. nutrient levels, toxic substances and pesticides), water quality (e.g. pH, salinity, turbidity and chlorofyl-a) and biological data (e.g. vegetation coverage and presence of birds). In recent years, there have been no severe exceeding of toxic levels in the Rotte basin which could have had a serious impact in fish populations [28]. With regard to nutrients and water quality, the parameters total phosphor, total nitrate, turbidity, temperature and chlorophyl-a are most important to determine the trophic state. Time series for these data are available from 1990 until 2023 and measurements take place monthly. Chloride concentrations, pH and dissolved oxygen do impact FHS, but for pH and chloride concentrations the levels remain within the harmful boundaries for fish species, and therefor have limited impact on FHS. Oxygen levels can vary enormously due to external factors (e.g. time of day, weather conditions, or temperature). Monthly measurements of oxygen levels provide limited insight in FHS, therefor this data is regarded as less useful for this research (personal communication J.P. Kalkman, A. Regtien and T. van Rooijen on 22nd of March 2023).

Biological data

Next to the physio-chemical data, the water authority collects biological data in the Rotte basin. Focus of monitoring of biological conditions is on the biodiversity and abundance of different taxa, such as fish species, phytoplankton, diatoms, macrophytes and macrofauna. The monitoring of the biological conditions in the Rotte basin takes place three annually and are part of the water system analysis for the WFD. The methodology for the assessment of the biological monitoring has been recorded in protocols and guidances ([4], [49]). As a rule, the water authority uses fixed measurement locations, but the number of locations, monitoring frequencies and timing of the monitoring vary for the different taxa. The water authority performs the monitoring of phytoplankton and diatoms at 3 locations in the Rotte basin, uses 17 locations for the assessment of macrophytes and 18 locations for macrofauna. Monitoring data of fish is collected from approximately 44 locations using electro fishing and quail fishing. The monitoring locations for fish may vary for different years. A map with all biological monitoring locations is presented in appendix A. The results of the fish monitoring will be discussed extensively in paragraph 2.4.

Phytoplankton and Diatoms

The regular monitoring of phytoplankton takes place in three locations in the sections Rottemeren and Rotte. The monitoring includes 340 different species. In 2010 an incidental assessment of phytoplankton has taken place in the section Boezem as well. The monitoring results have been registered for the period between 1994 and 2020, although for some years monitoring data is not available. On average, there are four monitoring occasions per year and most of the monitoring occasions have been planned in April, July and September. The total number of measurements for phytoplankton (distinguished in location and time of year) is 136.

The monitoring of diatoms takes place on 3 locations in the sections Rottemeren and Rotte. The monitoring includes more than 170 different species. In the the period 1989 - 2020 monitoring data is available for 20 years. On average, there are 4 - 5 monitoring occasions per year and most of the monitoring occasions have been planned between April and September. The total number of measurements for diatoms (distinguished in location and time of year) is 136.

Macrophyte
The monitoring of macrophyte takes place on 17 locations in the sections Rottemeren and Rotte. These locations have been selected on the presence of macrophyte. The monitoring includes almost 80 different species, and includes helophyte, hydrophyte, duckweed and riparian vegetation. Data for macrophytes is published in two different methods. The first method is the vegetation coverage (in%) relative to the maximum possible coverage given the conditions which is limited by the water depth. The second method is in STOWA classes and describes the abundancy in different classes (1 - 9) following Braun - Blanquet [4].

In the the period 1996 - 2020 monitoring data is available for 11 years. Between 2008 and 2020, structural monitoring for the WFD takes place once every three years for, on average, 8 locations in the Rotte basin. The monitoring of macrophytes is planned between June and September. Both monitoring locations and timing correspond to the monitoring of macrofauna. The total number of measurements (distinguished in location and time of year) is 52.

Macrofauna
The monitoring of macrofauna takes place on 18 locations in all sections, except for the Noorderkanaal and the Binnenrotte. The monitoring includes more than 380 different species. In the the period 1993 - 2020 monitoring data is available for 10 years. Between 2008 and 2020, structural monitoring for the WFD takes place once every three years for, on average, 9 locations in the Rotte basin. The monitoring of macrofauna is planned between May and September. The total number of measurements (distinguished in location and time of year) is 52.

Table 2.4 illustrates the extent of the biological data. The table presents for each taxa an indication for the number of species and the specifications of the monitoring data:

Taxa	No. of species	No. of monitoring years	No. of monitoring locations	Frequency	Planning	Total measurements
Phytoplankton	>340	18	3	4 / year	April, September	136
Diatoms	>170	20	3	4 - 5 / year	April, September	136
Macrophyte	<80	11	17	1 / 3 year	June - Sept.	60
Macrofauna	>380	10	18	1 / 3 year	May - Sept	52

Table 2.4: Characteristics of biological data

Hydromorphological data

The hydromorphological data describe the stream flow characteristics and the structure and form of the water bodies. In general, the most relevant hydromorphological characteristics are the shape of the water body, flow velocities, water level fluctuations, resident times, sediment type and the design of the shores and embankments. The available hydromorphological data for the Rotte is available from measurements by the water authority in the support of water management and the WFD combined with the observations from the field visits executed in the context of this research.

First of all, the water authority measures water levels for operational management in the Rotte basin at 12 locations at a frequency of every 5 minutes. For this study, HHSK has made data for the water level measurements available at five locations with continuous measurements for the years 2019 until 2022. In accordance with the water level regulations, this water level varies between -1.2 and -0.9 mNAP (see also paragraph 2.2). The average water levels are between -1.08 and $-1.04 mNAP$ depending on the location, with on average higher water levels in the more upstream locations.

HHSK also uses a hydrodynamic model (SOBEK 1DFLOW Rural) for the simulation of the behaviour of the Rotte basin for different scenarios. The model can be used to provide information on flow velocities, discharges, resident times, water level fluctuations and the slope of the water line under various water supply and discharge conditions. The water authority has examined the river profiles in the Rotte in 2012. Measurements have taken place for the water level ($mNAP$), the water depth between water line and mud layer (m) and the distance between the water line and the solid bottom. The assessment has resulted in 49 profiles for the water level and gives insight in the elevation of the solid bottom ($mNAP$) and the thickness of the mud layer on top of the solid bottom (m). An overview of the locations of the profiles is presented in Appendix A, while appendix B presents typical bed profiles for the sections in the Rotte basin. Also, the profiles give an indication for appearances of steep or faint shores and the presence of shallow waters. In general, the bed profiles in the Rotte basin are symmetric, relatively steep and flood plains are hardly present. Therefor, the impact of water level fluctuations of $0.30m$, between $-1.20and - 0.90mNAP$, on the water surface area and the interaction between the aquatic zone and the riparian zone due to flooding is limited.

Additional data for the morphological conditions have been collected by the HHSK for the sample location characteristics for the monitoring of macrophyte and for the physio-chemical monitoring. The sample location characteristics include, among others, local turbidity (in terms of Secchi disc depth), water depth and the thickness of the muddy layer. The measurements for water depth and turbidity are particularly valuable for this research, because the Water depth-Turbidity ratio (WDT-ratio) gives an indication for the light availability for the productivity (or growth) of hydrophyte.

The location and time period for the measurements corresponds to the monitoring occasions for biological data. The water depth has been measured at all 17 monitoring locations for macrophytes and for 6 different years. Though, the number of measurements vary for each year. For example, if we compare the WFD monitoring in 2011 and 2017, data is available from, respectively, 12 and 5 monitoring locations. Also the assessment of the monitoring varies for different years. For most of the years the water depth is measured at different distances from the shore, at an interval of $0.5m$. But for 2020 the water depth measurement has been done at a single location. Also the results of the measurements in 2020 differentiate from the earlier measurements, because the relatively high turbidity results in a higher WDT-ratio which indicates limited light availability. When we leave out the data for 2020, the measurements for the sample locations characteristics result in 32 valuable cases for the WDT-ratio from the sample location characteristics. For the period from 2008 until 2016, the HHSK has calculated the WDT-ratio as part of the physio-chemical monitoring. Data for the WDT-ratio from this monitoring effort are available for more than 820 cases.

Data for the thickness of the muddy layer is available for 16 out of 17 locations for macrophyte monitoring, but the available data vary per year. For the year 2011, the muddy layer is monitored at 12 locations, while in 2014 this analysis has not been performed. The total number of monitoring occasions for the thickness of the muddy layer is 27.

Finally, an inventory has been made of the structure of the riparian zone by observation during field visits in the period September - October 2022. This inventory gives classification for type of shores and its presence, such as natural banks (51%), vertical embankments (44%) and foreshores (5%). The inventory also gives an overview of the locations of different typologies of riparian vegetation, such as underwater structure and overhanging vegetation. These typologies represent the interaction between the riparian zone and the ecological water system due to the creation of structure and shelter or the contribution to the food web.

2.4 Fish populations

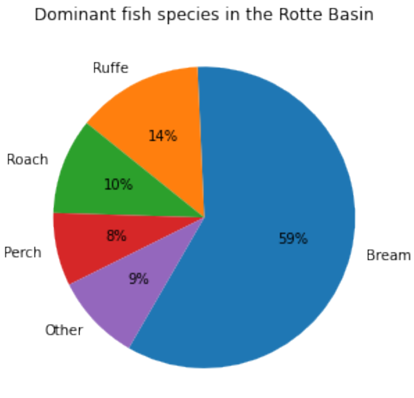
This section presents the fish biodiversity and abundance for the relevant water bodies for migration and dispersal of fish to and from the Rotte basin. These water bodies include the fresh water bodies of the Rotte basin and the directly connected water bodies of the Schie basin, the Bergse Plassen and the Kralingse Plas as well as the transitional water body of the Nieuwe Maas. Fish data of the three most recent monitoring occasions have been used to create a comprehensive picture of the fish population. The data has been provided by HHSK, Hoogheemraadschap Delfland (HHDL) and Rijkswaterstaat (RWS), the executive agency of the Ministry of Infrastructure and Water Management. These different institutions provide data for different years. The water authorities HHSK and HHDL provide three-annual monitoring data for different years; HHSK has published data for the years 2014, 2017 and 2020 ([46], [12]) and an additional fish monitoring in the Binnenrotte section took place in 2013. HHDL presents the data for the years 2015, 2018 and 2021 ([47], [2], [34]). Data on the fish population by RWS for the Nieuwe Maas are available for the consecutive years 2018, 2019 and 2020 [32]. Appendix C presents the average fish species abundance for the Rotte basin and the connected waters and the abundance of species in the different sections of the Rotte basin. The appendix also indicates the classification of the fish species in specific guilds for, among others, food preferences, flow preferences and substrate.

The Rotte basin inhabits 22 species, including *hybrids*. The majority of fish species belongs to the eurytope flowguild ($n = 14$), followed by limnophilic species ($n = 4$). Three observed species in the Rotte basin belong to the rheophilic flowguild, namely Bullhead (*Cottus gobio*), Gudgeon (*Gobio gobio*) and Ide (*Leuciscus idus*). These rheophilic species are regarded as less critical for streaming waters ([53]). The freshwater bodies also inhabit the diadromous species European Eel (*Anguila anguila*) and Three-spined Stickleback (*Gasterosteus aculeatus*). The contribution, in terms of biodiversity, of the connected freshwater bodies to the Rotte basin is limited. The fish biodiversity of the Bergse Plassen and Kralingse Plas overlap completely with the Rotte basin. The Schie boezem inhabits four species, which were not observed in the Rotte basin, namely Bleak (*Alburnus alburnus*), Flounder (*Platichthys flesus*), Gibel Carp (*Carassius auratus gibelio*) and Sunbleak (*Leucaspius delineatus*). In this respect, the Flounder is also regarded a diadromous species.

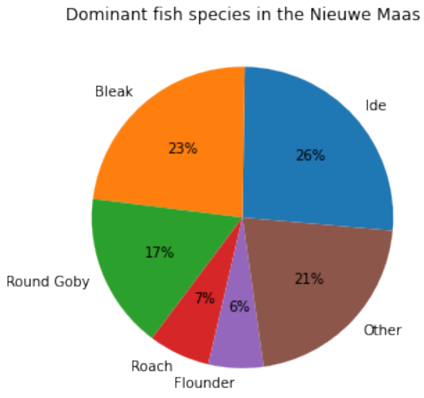
The average fish abundance in the Rotte basin over three monitoring occasions is almost $6,450n/ha$, which is low compared to the Bergse Plassen ($14,000n/ha$) and the Kralingse Plas ($6,900n/ha$), but much higher than the Schie boezem ($2,750n/ha$). Fish abundance in the Rotte basin is dominated by Bream (*Abramis brama*), Ruffe (*Gymnocephalus cernuus*), Roach (*Leuciscus rutilus*) and Perch (*Perca fluviatilis*), bearing in mind temporal and spatial variations of fish population compositions. The top four species represent more than 90% of total fish abundance (see also figure 2.2). The high percentage of Bream is a main cause for the poor assessment of the fish population for the WFD because of the assumed turbidity caused by benthivore activities. This species is mainly found in the Northern branche of the Rotte basin, whereas the Southern branches inhabit more Tubenosed Goby (*Proterorhinus marmoratus*), Rudd (*Scardinius erythrophthalmus*), Tench (*Tinca tinca*) as well as European Eel. In this respect, the Tubenosed Goby is regarded an invasive species. Rudd and Tench are in the guild of plant loving fish species which have a positive valuation for the WFD.

The fish population composition for the connected freshwater bodies is comparable to the Rotte basin, although the density of Perch and Roach is relatively higher at the expense of the abundance of Bream. Given the similarities in fish populations, the habitat conditions for fish species in the Rotte basin and the connected fresh waters are comparable. This means that fish migration and dispersal can easily take place between these water systems, thus providing escape routes in case of hazards and contribute to genetic mixture as well.

The monitoring data for the fish population in the Rotte basin have been combined with the classification of guilds for spawning, substrate preference and food as published by de Leeuw & Backx [10]. Although the classification of fish species per guild by de Leeuw & Backx [10] is not complete, it does give an indication for the habitat preferences of the current fish population. The analysis shows that more than 90% of the fish population in the Rotte basin use vegetation for spawning. This guild includes, among others, Bream, Roach, Perch and Rudd. A minority in the Northern sections of the basin lay their eggs on the river bed (benthic spawning), including various Gobies, Asp (*Aspius aspius*) and Pikeperch (*Sander lucioperca*). The Southern section presents fish species with floating eggs (pelagic spawning), such as European Eel, as well as parent protected eggs (e.g. Three-spined Stickleback), but both guilds have a very low representation of 7% and 4% respectively. In this respect, it is important to notify that European Eel does not spawn in the



(a) Rotte basin



(b) Nieuwe Maas

Figure 2.2: Average fish population composition for Rotte basin and Nieuwe Maas.

Rotte basin, but in the Sargossa Sea in the Northern Atlantic Ocean.

de Leeuw & Backx [10] have determined five different substrate guilds, based on the preference of fish species for habitat structure. The substrate guilds range from a preference for a habitat with vegetation to soft bottoms, sandy bottoms and rocky bottoms. The distinction for the substrate guild is not as clearly defined as the spawning guild and fish species have been classified for more substrate guilds (see also appendix C). The large part of the population prefers locations above or in vegetation or has no preference at all (both $\pm 45\%$). Fish species Bream, Bitterling (*Rhodeus sericeus*), Carp (*Cyprinus carpio*) and Pike (*Esox lucius*), for example, have been classified in both fish guilds. Almost 10% of fish species in the Rotte basin has a preference for soft bottoms of fine sand or mud, for example European Eel, Ruffe and Ten-spined Stickleback (*Pungitius pungitius*) and a small part has a preference for rocks or sandy bottoms, such as different Gobies (2%). The substrate guilds show a distinction between the Northern and Southern sections. In the northern sections, the guilds for vegetation and the fish species with no specific substrate preference are dominant. In the Southern section, the fish species with preference for soft bottoms of mud or fine sand are majority (30%), followed by the species with no specific preference or a preference for vegetation (both 27%). But Also, the abundance of fish species with a preference for sandy bottoms or rocky substrate are also significant (both $\pm 7\%$).

For food preferences, classification for food preferences is not clearly defined as fish species can have a diversified diet and can act opportunistic based on availability. However, from the analysis of the food guilds as defined by de Leeuw & Backx [10], planktivores and insectivores (e.g. Bream, Roach, Perch, European Eel and Rudd) are dominant in the Rotte basin over zoobenthos (e.g. Bleak and Barbel or *Barbus barbus*), herbivores (e.g. Bitterling and Roach) and piscivores (Pike, Pikeperch and Three-spined Stickleback). In the Northern sections planktivores and insectivores are dominant, containing $35 - 40\%$ of total fish abundance, followed by herbivores and zoobenthos with both ($\pm 10\%$). In the Southern sections, planktivores and insectivores are dominant, but to a lesser extend ($30 - 40\%$). The abundance of herbivores in the Southern sections is comparable to the Northern sections, but the piscivores are relatively more present with $\pm 14\%$ of total fish abundance.

The Nieuwe Maas is a transitional, mesohaline, water body and presents a different fish composition compared to the fresh water bodies described above. During monitoring 24 different fish species have been observed and no hybrids have been found. Most species are freshwater fish (n = 11), followed by estuarian residents (n = 6) and seasonal migrants (n = 3). During monitoring two diadromous species have been observed, namely the Flounder (*Platichthys flesus*) and the Three-spined Stickleback. And finally, two fish species from the guild of marine juvenile have been observed, namely the Atlantic herring (*Clupea harengus* and the European Seabrass (*Dicentrarchus labrax*).

Also in terms of abundance, the freshwater species are the majority (70% of the fish population), followed by the estuarian residents (24%) and the diadromous species (10%). The most abundant species in the Nieuwe Maas are Ide, Bleak, Roach, Flounder and Round Goby (*Neogobius melanostomus*). The top 5 species account for almost 80% of the total fish population (see also figure 2.2).

The restored connection between the Rotte basin and the Nieuwe Maas can be used for migration of diadromous species and dispersal of freshwater species and estuarian residents who are likely to migrate to limnetic sections, like some of the Goby types. Dispersal will be most likely for eurytope or limnophilic flowguild and the less critical rheophilic species who prefer to, temporarily, exchange flowing sections for stagnant waters and vice versa. In that respect, the connections between Nieuwe Maas and Rotte could result in the introduction of new species into the Rotte basin, such as Bleak, Round Goby, Flounder, Common Goby (*Pomatoschistus microps*), Sand Goby (*Pomatoschistus minutus*) and Monkey Goby (*Neogobius fluviatilis*).

Chapter 3

Methodology

Introduction

The research approach centers around the development of an fish habitat suitability model, based on available monitoring data for the Rotte, and the methodology to apply the model for policy decisions regarding the assessment of current habitat suitability and the impact of management options. Figure 3.1 illustrates the development process from initial research questions until the application of the model. The results of the data inventory (step 1) has been described in chapter 2. This section describes the methodology for step 2: Model development and step 3: Model application. Paragraph 3.1 presents the choice for the modelling approach and the starting points involved. Paragraph 3.2 explains the consecutive steps for the model development and (intermediate) deliverable. The methodology described in these paragraphs are directed at answering research questions 1 and 2. Paragraph 3.3 describes the application of the model. The methodology described in this paragraph is directed at answering research questions 3 and 4.

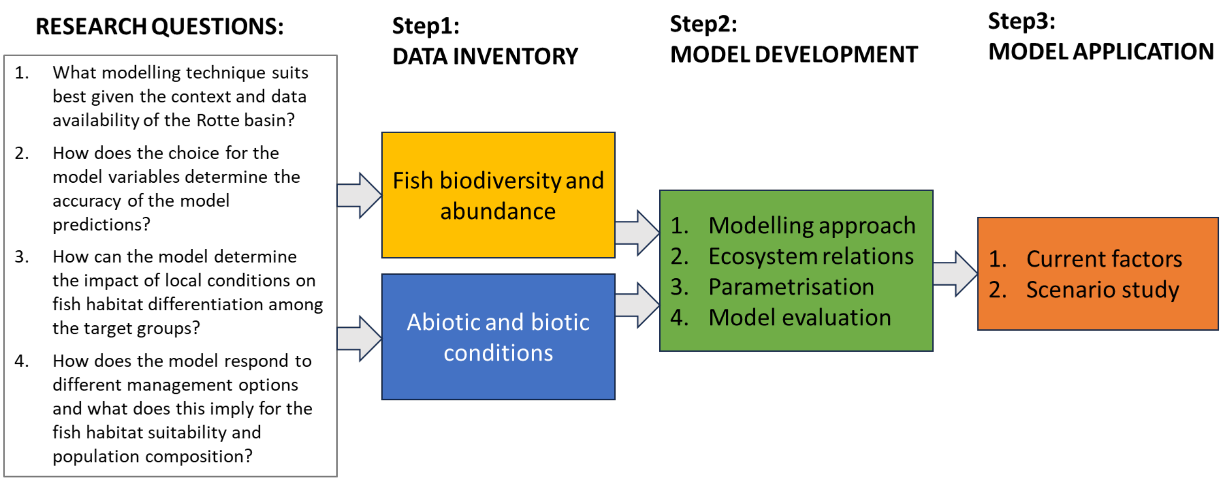


Figure 3.1: Methodology workflow from research questions to model application.

The development process involved the expertise and interests of many different stakeholders in the Rotte basin. HHSK, being the responsible organisation for water management in the Rotte basin, played a key role in defining the political framework and providing valuable data resources for the Rotte basin (see also paragraph 2.3). Ecologists, biologists and ecological modelling experts have been involved in the research process for understanding the ecological system, the interaction between abiotic and biotic factors and the applications of different modelling approaches. The angler association, both at the local and federal level, have been involved in the classification of fish species in target groups and for exploring the limiting factors for FHS. An extensive list of the participants in this research process is provided in appendix D.

3.1 Habitat modelling tool

Model choice

The dominant data-based model types in ecology are the traditional statistical models or regression models, such as Generalized Linear Models (GLM) or Generalized Additive Models (GAM) ([36], [58], [31]), machine learning models based on (artificial) neural networks or fuzzy logics ([35], [58], [1], [44]) and BBN models ([37], [7], [20]). The regression models and machine learning models are classified as deterministic models, which are based on known relationships between explanatory variables and response variables. BBN models are classified as probabilistic models and they express statistical problems in uncertainties in available information and possible outcomes [48]. In practice this means that equal conditions can give different results, due to uncertainties in underlying processes or data.

The starting point for the ecological model in this research is the use of monitoring data provided by HHSK and the possibility to model scenarios for different management options. The selected modelling approach for this research is the BBN model because of the ability to use different forms of information (e.g. categorical and continuous variables and different measurement frequencies and time series), the expected non-linearity of relationships in the ecological system and the high level of uncertainties for these relationships as well as the uncertainties in monitoring data due to temporal and spatial densities and differentiation.

BBN models are also appreciated for the effective use of data. Traditional statistical models and machine learning models normally require large data inputs ([58],[9]), while BBN models reduce data availability requirements by limiting the number of (required) independent variables. Unobserved variables disappear from the computation, but are expressed in the probability distribution of the end variables ([48], [20]). And finally, the relation diagram at the foundation of the BBN model enhances the communication with various stakeholders, for example political decision makers, ecologists and citizens ([37], [7]).

The downside for the application of BBN models in data scarce situations is that limited data requires simplifications of the model structure, which could impose the validity of the model [20]. BBN models have a strict hierarchical structure based on Directed Acyclic Graphs (DAGs) and do not allow feed-back loops nor modelling of dynamic processes or temporal variations ([31], [7]).

Guidelines for BBN models

The BBN model in this research predicts the probability distribution between a ”good” or ”moderate” habitat suitability, given the species specific preferences (or *traits*) and the physical conditions in the Rotte basin, expressed by various parameters. Also, the model should be able to assess the impact of different management options on FHS. The model variables (or *nodes*) are discrete and the states of the variables are determined by probability. Logically, the probabilities for all the possible states of a (single) variable sum up to one. The variables are linked, based on our understanding of the interactions in the ecosystem. The physical conditions, the habitat suitability and all the linkages between these nodes (or *edges*) result in a relation diagram of explanatory nodes, indicated as parent nodes, and response nodes, indicated as child nodes.

Following from the relation diagram, changes in the physical conditions alter predictions for the habitat suitability for different fish guilds and, therefor, will most likely result in changes in the fish population composition. Thus, the relation diagram of parent nodes and child nodes in a hierarchical graph (noted as $\mathcal{G}$) is the starting point for the BN model. The probabilities for the nodes without any parents are unconditional, or prior, probabilities. Whereas the probabilities for the child nodes are conditional, or posterior, and can be calculated with the Bayes' theorem for conditional probability (or *product rule*, [31], [48]):

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B \mid A)}{P(B)} \tag{3.1}$$

In formula 3.1 $P(A \mid B)$ is the conditional probability for A (child node) given B (parent node), $P(A \cap B)$ is the intersection which satisfies both A and B and will be, in most cases, determined from data analysis. $P(B)$ is the unconditional probability for B and will also result from data analysis. The second part of the formula illustrates the possibility for the Bayes rule to change from a causal direction, where the observed cause B has the effect A, to a diagnostic direction, where the observed effect A implicates the cause B.

Furthermore, the hierarchical structure of the BBN results in conditional independence between the variables based on observations of variables (or *evidence*). Every node in graph $\mathcal{G}$ is, by definition, independent from its non-descendants given the parent nodes [31]. This definition is formulated in equation 3.2:

$$\text{For each variable } X_i : (X_i \perp \text{Non} - \text{descendants}_{Xi} | Par^{\mathcal{G}}_{Xi}) \tag{3.2}$$

In this formulation X_i is the model variable, $\perp$ stands for *is independent from* and $|Par^{\mathcal{G}}_{Xi}$ means the conditional probability of variable X_i given the parent nodes of the variable in the graph $\mathcal{G}$.

In practice this means that evidence of the parent nodes creates conditional independence between a variable and its non-descendants. This principle is expressed in the Bayes' Chain Rule for the joint distribution of graph $\mathcal{G}$:

$$P(X_1, X_2, .., X_n) = \prod_{i=a}^{n} P(X_i \mid Par^{\mathcal{G}}_{Xi}) \tag{3.3}$$

Equation 3.3 calculates the joint probability distribution for all model variables $X_1..X_n$ from the product ($\prod$) of the conditional probabilities for each model variable X_i, given its parents.

Every child node in the model has a Conditional Probability Table (CPT), which shows the probability of the different states of this node for each combination of states of the parent nodes. The CPTs can be the result from expert knowledge (e.g. from expert reviews or literature) or can be data-based and calculated with learning algorithms. The data for the data-based models can be derived from field observations, metamodels for (ecological) system processes or from a combination of both. The CPTs for the model in this study are data-based and have been derived via machine learning, using monitoring data in the Rotte basin (see also paragraph 3.2). To apply the model for an assessment of the current situation, evidence from monitoring data can be introduced for each model variable. Scenario study for different management options can be performed by adjusting the CPTs according to the expected output of the policy. The methodologies for the applications of the model will be further explained in paragraph 3.3.

3.2 Model development

This paragraph describes the model development in five consecutive steps. These five steps are:

1. Development of the relational diagram
2. Development of the initial model
3. Preparation of the learning cases
4. Parametrisation of the CPTs
5. Model evaluation

1. The relation diagram

The relation diagram represents the environmental factors and the (presumed) interaction between these factors which will finally determine the habitat conditions. In this research, the target groups for the habitat suitability are the plant-loving fish species with a preference for a plant-based diet and a habitat in or above vegetation (FG1) and the benthivorous fish species with a diet consisting mainly of benthic life (macrofauna) and a habitat preference for soft bottoms (FG2). The model will determine the habitat suitability for the preference of food and habitat structure in terms of substrate, only. The focus on these two habitat aspects leaves out other habitat aspects and,consequently, the environmental factors which determine habitat suitability for this aspect. For example, for predation (e.g. by water fowl or piscivorous fish species) presence of under water structures is an important factor for protection.

Thus, when including the aspect of predation for the habitat suitability of fish, the model should include a model variable for underwater structure as well. Furthermore, the relation diagram in our research only considers the preferences for mature fish species, because this life stage is dominant in the fish monitoring data.

The first draft of the relation diagram is constructed based on literature of the food web for stream biota [17], biogeochemical interactions [29] and feed-back loops for equilibria in shallow lakes [50]. This initial relation diagram includes environmental factors from biological, physical-chemical and hydromorphological backgrounds. This draft version has been discussed with experts with different backgrounds and adjusted accordingly. The final version of the relation diagram is presented in figure 3.2:

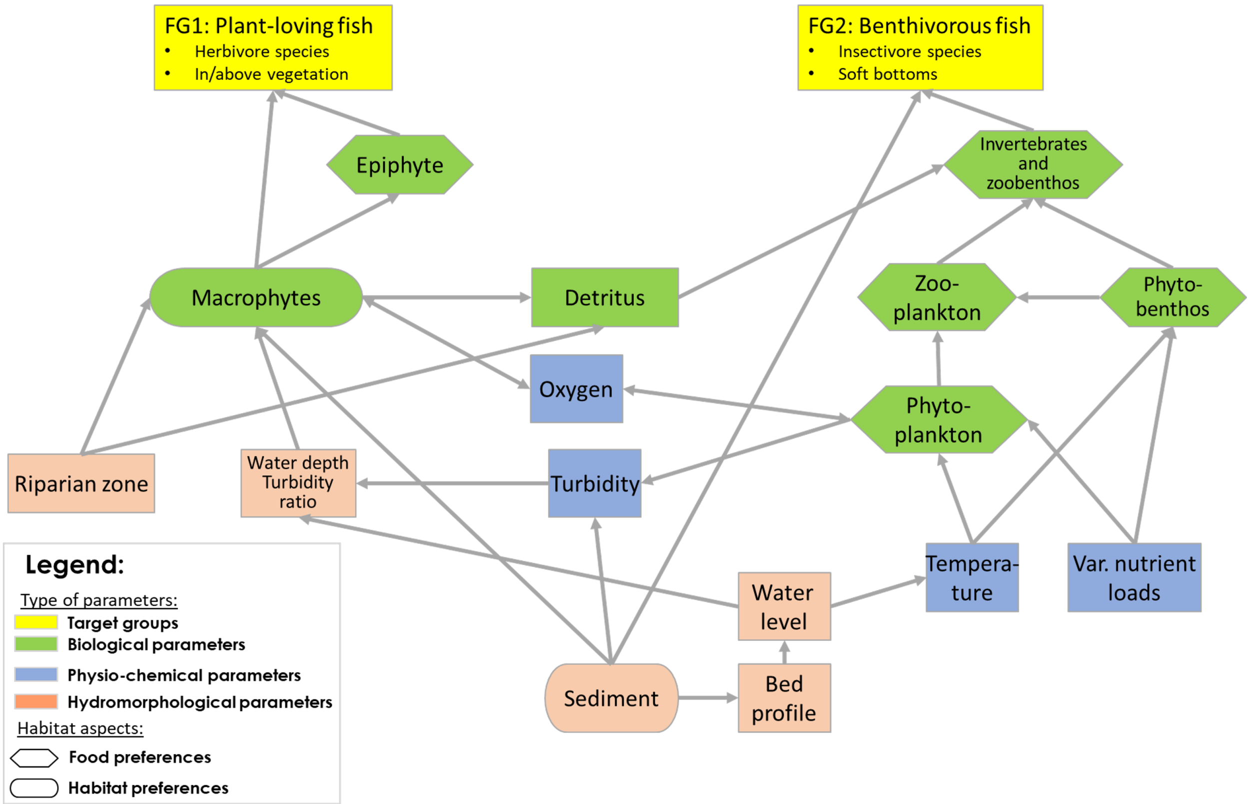


Figure 3.2: Relation diagram for habitat factors food and habitat structure.

The initial model

The initial model, or alpha-model, is created by converting the relation diagram to a BBN model. For the conversion of the relation diagram to the BBN model, I have used the BBN modelling shell NETICA (version 6.09 Limited Mode by Norsys Systems Corp. Vancouver, British Columbia). For further reading, when the text refers to the factors in the ecosystem, they are written in their full name, for example phytopankton and macrofauna. References to the model variables are made by using acronyms written in capitals, for example model variable for phytoplankton (PHYT) and model variable for macrofauna (MAFA).

The conversion involved a simplification of the relation diagram because of the limitations of BBN model approach, concerning data availability, model complexity, the limitations to model dynamic processes and the use of discrete variables. First of all, the variable selection for the initial model is based on data availability and the intended management options for the scenarios. The variable for epiphyte has been left out of the initial model, because there is no monitoring data available for the Rotte basin. Instead, a positive relationship is assumed between abundance of epiphyte the presence of macrophyte, specifically helophyte and hydrophyte, and phytoplankton.

Second, aggregated nodes have been added to reduce the number of variables and model complexity. The aggregated model variable for WDT-ratio (WDT) replaces underlying nodes for water level, bed level and turbidity. The model variable for the trophic state (TS) is defined according to Smith *et al.* [52]. The aggregated node TS combines the variables for the nutrients (nitrate and phosphate) and turbidity but also includes data for chlorophyl-a. Both aggregated nodes can be calculated directly from available data.

Thirdly, BBN models do not facilitate feed-back loops. As a consequence, the interaction between oxygen saturation and macrophytes and phytoplankton, as expressed in the relation diagram, can not be modelled in our BBN model. Neither can we include predator-prey dynamics in our model, such as the interaction between the invasive American crayfish and the state of helophyte and hydrophyte.

And finally, the discrete states for the variables have been determined. For abiotic variables, such as model variable for temperature (TEMP), TS and WDT, threshold values have been determined based on literature and expert views. The threshold value for TEMP is estimated at $14°C$, because monitoring data of the Rotte basin shows that phytoplankton production accelerates above this threshold. The WDT determines light emittance under water for plants to grow. The threshold value for the WDT used in the WFD is 0.6 ([57]). Both threshold values also indicate the non-linearity of the ecosystem. Threshold values for the biotic variables are determined in terms of abundance. The model variables PHYT, MAFA and model variable for hydrophyte (HYDROPH) have a distinction between high abundance/availability and low abundance/availability. The model variable for helophyte (HELOPH) has three states related to functionality for fish, namely 'high density', 'low density' and 'abscence'. In this respect, both 'abscence' and 'high density' have limited functionality. The definition of the threshold values for the biotic variables will be described later in this paragraph under the heading Preparation of the learning cases.

The alpha-model is completed by adding the CPTs. The prior probabilities for the parent-less nodes have been calculated based on frequencies in the monitoring data. The CPTs for the child nodes have been added in two different ways, resulting in two different versions. For the first version, the CPT values for each child node have been added for the most probable and least probable values given the combination of parent nodes. For the remaining combinations, probabilities have been estimated. The second version of the alpha-model has been developed by probabilistic inherence for the most likely state for each variable. The resulting versions of the alpha-model have been discussed with the board of the angler association HSV Groot-Rotterdam. The board members consent to the model structure and endorsed the resulting habitat suitability for the two fish groups based on the initial CPTs. Because of the agreements on the model structure, this review has not been used for an update of the alpha-model structure. In spite of the guidelines for BBN model development presented by Marcot *et al.* [37] and Chen & Pollino [7], no official beta-version of the model has been created.

Figure 3.3 on page 19 presents the resulting model structure including the beliefs that follow from the initial parametrisation of the CPTs.

Preparation of the learning cases

The objective of this step is to create a complete list of variable observations from the monitoring occasions based on unique date and location. This list shows the observed combinations of the variable states, which will be used for the parametrisation of the CPTs through learning. Starting points are the variables from alpha-model and the monitoring data as described in section 2.3. First, the individual cases for each model variable have been abstracted from the monitoring data, based on date and location. The states for the individual cases have been determined based on the adopted threshold values. Second, the individual learning cases have been aggregated in space and time to create learning cases with a combination of variable states. This section briefly describes both processing steps.

Individual learning cases for fish monitoring data:
Fish species from the monitoring data have been classified in target group for plant-loving fish species (FG1), target group for benthivorous fish species (FG2) and the category 'Other'. The classification of fish species is based on the guilds for plant loving species and the group of Bream/Carp according to the WFD [43] and supplemented with the guilds for habitat, substrate and food according to de Leeuw & Backx [10]. This subdivision has been compared with the results from a survey amongst the board members of the angler association HSV-Rotterdam and the Federation for Fishing in the Netherlands. The classification of the fish species into the fish groups is presented in appendix C.

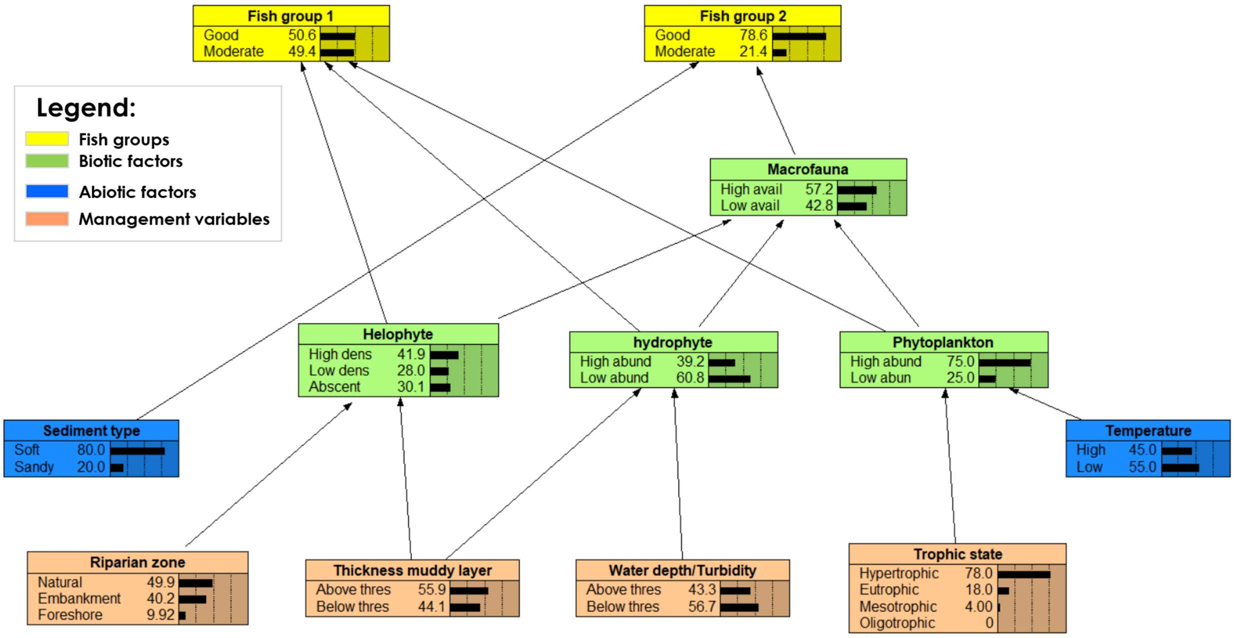


Figure 3.3: Model structure of the alpha-model including beliefs based on initial CPTs.

FG2 is most abundant in the Rotte basin, representing, on average, 82% of the fish population. This group consists for 99% of Bream and a small part of Carp (< 1%). FG1 represents, on average, 13% of the fish counts. The most important fish species in this group are Rudd (64%), Ruffe (17%) and White Bream (*Blicca bjoerkna*, 13%). Finally, the fish group 'Other' represents 3% of the fish abundance in the Rotte basin. Most common fish species in this group are Perch and Pikeperch.

With regard to the migrating fish species, the Three-spinded stickleback and the Flounder belong to FG1 and the European Eel belongs to the group 'Other'. Most of the fish species of the Goby-family belong to the group 'Other' as well.

The distribution of the fish groups shows a strong differentiation between the sections in the Rotte basin. The Northern sections are dominated by FG2, with relative abundance of more than 80%. The Southern sections are dominated by the group 'Other', with abundance varying between $70 - 90\%$. Fish group FG1 is more evenly distributed, with relative abundance varying between $10 - 30\%$.

The state for the habitat suitability for the target groups is determined based on the relative abundancy of the total fish population (in %). The threshold value for a 'Good' habitat suitability is > 40%, which indicates that the fish group is dominant and benefits more from the habitat conditions than the other groups.

In total, there are 80 individual cases based on fish monitoring data (N = 80) with specific date and location. The state for both FG1 and FG2 is 'Good' in 21% of the cases (n = 17). In the end, 80% of all fish monitoring cases will be used for learning (N = 64). The remaining 20% will be used for testing of the model (N = 16) and will be explained later in this paragraph under the heading Model evaluation.

Individual learning cases for other biological monitoring data:
Data for both phytoplankton and macrofauna is collected for the number of species in the monitoring sample. The states for PHYT in the model are classified in terms of abundancy (High and Low abundancy), whereas for MAFA the states are classified in terms of availability (High and Low availability) because access for fish is essential. Nevertheless, the state for both variables is calculated with the sum of all the species in the monitoring samples. The samples from different monitoring occasions can be easily compared, because of the equal sample sizes. A high abundancy is associated with the upper-quartile for all the observations for the variables. For PHYT (N = 136) 74% of the cases has a moderate abundancy (n = 101) and 26% of the cases has a high abundancy (n = 35). For MAFA (N = 52) 73% of the cases have a moderate/low availability (n = 38) and 27% of the cases have a high availability (n = 14).

Data for macrophyte is collected for the STOWA classes and for the relative coverage (in %). The STOWA classes have been converted to the relative coverage (in %) by taking the upper boundary of the coverage as published in [4]. In a second step, vegetation species have been integrated by the vegetation type with a distinction in helophyte, hydrophyte, riparian vegetation and duckweed. This classification follows from the description of the layer and growth form for vegetation as presented in [4] and supplemented information from SynBioSys (version 3.6.3), using the Ellenberg scale for moisture. In line with the model structure, only data for the abundancy of helophytes and hydrophytes will be included in the learning cases. The total abundancy for each case is calculated by summing up the relative coverage for these vegetation types, grouped by monitoring date and location. The sum for the relative coverage per vegetation type could exceed the total area due to the combination of layers. Therefor, the total relative coverage is maximised to 100%.

Finally, the variable has been discretized by applying a threshold values for the relative coverage. For the HYDROPH the threshold value between high and low abundancy is 20%. This threshold value is in accordance with the target value for aquatic plants of the HHSK ([24]). For HELOPH the threshold values are 75% coverage for high density and 2% coverage for absence. The threshold value of 75% for coverage is also the threshold for the classification between riparian vegetation and littoral vegetation in the WFD ([4]). The number of cases for hydrophytes is very limited (N = 16) and in 15 cases (94%) the abundancy is low. For helophytes (N = 51) 22% of the cases have a high density (n = 11), 75% has a low density (n = 38) and for 4% helophyte is abscent (n = 2).

Individual cases for physical-chemical monitoring data:
The learning cases for the variables TEMP and TS have been derived from the physical-chemical monitoring data. Data is available for 12 locations and monitoring takes place monthly. For TEMP (N = 3,648), 49% of the cases has a high temperature (n = 1,796) and 51% has a low temperature (n = 1,852). The trophic state has been calculated from the nutrient levels for nitrate and phosphate, turbidity and chlorofyl-a ([52]). The trophic state for the learning cases is determined by the limiting factor and is equal to the lowest state given the parameters for this aggregated node. For all the cases for the TS (N = 3,361), 68% is hypertrophic (n = 2,283), 30% is eutrophic (n = 1,000) and 2% is mesotrophic (n = 77). And in only one case the trophic state is oligotrophic.

Data for the WDT-ratio is available from physical-chemical monitoring and from the sample location characteristics for biological monitoring. Measurements have been performed for the years 2008 until 2022. For the states of the WDT-ratio a threshold value of 0.6 has been applied. As explained in section 2.3, a value above this threshold indicates there is not sufficient light emittance at the bottom for hydrophytes to grow. Furthermore, a minimum water depth is applied of $0.30m$ to secure access for fish species. The total number of cases for the WDT is 891 (N = 891) for 26 different locations. For 98% of these cases, the WDT is above the threshold (n = 877) and the remaining cases (n = 14) are below the threshold.

Individual cases for hydromorphological data:
Data for the thickness of the muddy layer is collected for sample location characteristics in the period 2008 until 2020. This results in 27 cases for 16 locations (N = 27). For 93% of these cases (n = 25), the model variable for the thickness of the muddy layer (MUDL) is below the threshold of $0.30m$ and for 7% (n=2) the value is above the threshold value. The thickness of the muddy layer is assumed to be static and does not change significantly over time. This assumption is supported by the observation form the cases, where the state of the thickness of the mud layer does not change over time.

It is important to notice that the measurements took place at the monitoring locations for macrophyte, because this parameter is regarded as an explanatory variable for the presence of macrophyte in the WFD. Therefor, the measurements for MUDL are not representative for the total Rotte basin. To illustrate this, we have performed an analysis of the thickness of the muddy layer measured for the 49 profiles in the Rotte. These profiles have, for a large part, a mudlayer which exceeds the threshold value of $0.30m$. For example, for 20 profiles (40%) the thickness of the mud layer exceeds the threshold for 80% of the total profile length. For 12 profiles (24%) halve of the total length of the profile is below the threshold.

The model variable for sediment type (SEDI) is an explanatory variable for the habitat structure for fish species, but this data is not available for the Rotte basin. Therefor, the state of the sediment type (soft or sandy) has been determined based on the thickness of the muddy layer and a quick scan analysis for the geological layers for the subsections in the Rotte basin with DINOLoket (https://www.dinoloket.nl/). The analysis shows that for the sections Rottemeren and Boezem, there is a probability for sandy sediment at the near shore locations, but only when the thickness of the mud layer is below the threshold value of $0.30m$. For the learning cases, a distinction has been made between the cases based on the profiles and the cases based on the sample location characteristics. For the former cases, the state of SEDI is soft for all cases. For the latter, the state of SEDI is sandy only for the locations in the subsections Rottemeren Noord and Boezem and only when the thickness of the muddy layer at these points is below the threshold value. In total, there are 65 cases for SEDI (N = 65). The state of SEDI is sandy for 8% of the cases (n = 5) and for 60 cases the SEDI is soft. The sediment type is assumed to be static over time, which is also the case for the sample location data for the muddy layer.

Finally, the classification for the model variable for the design of the riparian zone (RIPZONE) is based on the observations from the field visits in the period September - October 2022. For every monitoring location in the Rotte basin, the state is determined by the type of the riparian zone of the nearest shore. For all cases for the riparian zone (N = 106), 59% is natural (n = 63), 31% is embankment (n = 33) and 9% is foreshore (n = 10).

Aggregation of the cases:
The spatial and temporal aggregation of the individual cases is required to create the combinations for the observations of the model variables. The objective of the spatial and temporal aggregation of the individual cases is to create combinations of observations for parent nodes and child node. Thus, for the the aggregated cases, the observed state of the child nodes can be (partly) explained by the observations of the parent nodes. Spatial aggregation is particularly required for the child nodes for the fish groups and macrofauna. The required actions for the construction of the aggregated cases included merging of the individual cases based on the location and resampling of data over time. Variables with missing data will be recorded in the aggregated cases because, due to Bayesian reasoning, probabilities for missing data can be estimated during parametrisation. The learning cases have been constructed by appending the (aggregated) cases of parent nodes and child nodes and removing duplicate measurements. Appendix A presents the scale for the spatial aggregation for the fish groups and macrofauna.The Pandas package for Python (Jupyterlab version 3.2.1.) has been used for processing of the monitoring data.

The aggregation of the individual cases has resulted in a set of 3,689 aggregated cases, including the 80 cases for the fish monitoring data. A short explanation for each child node is presented in Box 1. Table 3.1 gives an overview of the number of cases for each child node and the corresponding parent nodes and ancestors:

Child	Parents	Total cases	Cases with data child
FG1 / FG2	PHYT, HELOPH and HYDROPH MAFA and SEDI	80	80
Macrofauna	PHYT, HELOPH and HYDROPH	1,803	47
Phytoplankton	TS and TEMP	3,300	136
Hydrophyte	WDT and MUDL	114	16
Helophyte	MUDL and RIPZONE	51	51

Table 3.1: Overview aggregated cases for the model variables.

Textbox 1

Fish groups FG1 and FG2:
The fish monitoring data for FG1 and FG2 have been combined with the data for the parent nodes, respectively HELOPH, HYDROPH and PHYT for FG1 and MAFA and SEDI for FG2. For the 80 monitoring cases for FG1 and FG2 (N = 80), the data has been aggregated for the parent nodes MAFA and SEDI, because these two variables included all other parent nodes for the fish groups. 80% of the cases (n = 64) for FG1 and FG2 have been added to the Learning Cases and 20% (n = 16) will be used for testing the model.

Macrofauna:
MAFA is a child node from the variables PHYT, HYDROPH, and HELOPH but also a descendant from the physical-chemical variables TEMP and TS. The monitoring locations for macrofauna correspond with its parents, therefor aggregation is possible by merging the monitoring data for macrofauna with the aggregated cases for the parent nodes. The monitoring locations of temperature and trophic state do not correspond to the locations for macrofauna. Therefor, data for temperature and trophic state has added to the aggregated cases for macrofauna based on the adjacent measurement locations. The aggregation considers the minimum distance between the locations as well as comparable local characteristics of the water body, for example for shape, bed profile and the local impact of pumping stations with regard to nutrient loadings. Furthermore, timing for the monitoring of macrophyte and macrofauna do not correspond. To aggregate the data on a temporal scale, all biological monitoring data has been pinpointed at the last day of July. The corresponding physio-chemical and biological conditions (e.g. phytoplankton, temperature and trophic state) have been retrieved from of the nearest monitoring day for each year. The aggregation of the data results in 1,803 cases of which 47 cases include data for MAFA.

Helophyte:
The variable HELOPH is the child node of the parent nodes for MUDL and RIPZONE. Both parent nodes are regarded is time independent, therefor, only aggregation of the spatial data is required. The aggregation for HELOPH has resulted in 51 cases in total. And for every case, all model variables are available. This is due to the fact that the data for the thickness of the muddy layer and helophytes have been measured at the same location and for the same time period.

Hydrophyte:
The variable HYDROPH is the child node of the parentless nodes WDT and MUDL. Monitoring data for the three variables are spatially overlapping and the frequency of the monitoring is annually, but the timing of the monitoring is different. In this respect it is also important to mention that the thickness of the muddy layer is regarded as static over time. For aggregation, data has been merged based on the monitoring location and resampled for each year. The aggregated data frame for hydrophyte consists of the 114 cases in total; 16 cases include data for HYDROPH and 9 cases include data for all three model variables. Eventually, the variable HYDROPH is present in more learning cases than there are actual cases where HYDROPH data is available. The reason is that, due to aggregation, the monitoring data for hydrophyte for one location explains the states for macrofauna and FHS for several locations.

Phytoplankton:
The child node PHYT has the parent nodes TS and TEMP in the model structure. Three monitoring locations in the sections Rottemeren and Rotte observe all three variables, but the frequency and the timing of these observations vary (see also section 2.3). For temporal aggregation of the data, the time stamp of the measurements for these three locations has been resampled to the last day of each month. As part of the physio-chemical monitoring, TEMP and TS is also measured at nine other locations in the Rotte basin. These measurements have been added to the aggregated cases for phytoplankton. The result is a set of more than 3,300 cases for 12 locations. For the three locations where monitoring of phytoplankton takes place, 136 cases with data for PHYT and 128 cases for which all three model variables are available.

2

Parametrisation of the CPTs

The objective of parametrisation of the fish habitat suitability model is to complete the Bayesian Network with the CPTs by machine learning using data in the learning cases. The selection of the learning cases has a high percentage of missing values (72%), mainly because of the three-annual frequency for the biological monitoring. For the situations with high percentages of missing values, the NETICA software package provides two algorithms for machine learning, namely Expectation Maximization (EM) and Gradient Descent (GD). Both algorithms try to find the most likely Bayesian network, that fits with the learning cases. In an iterative process, the learning algorithms adjust the model CPTs in order to find the Bayesian network with the highest likelihood given the available data. The maximisation of the likelihood for the Bayesian network is expressed in equation 3.4, where N is the Bayes network, including the CPTs, and D is the data in the learning cases:

$$P(N \mid D) = \frac{P(D \mid N)P(N)}{P(D)} \tag{3.4}$$

The learning algorithm used for this fish habitat suitability model is EM, because this algorithm functions better than GD for when there are many missing values and when the complexity of the model structure is limited [45]. The limited version of NETICA used for this research only reads the first 1,000 cases. For practical reasons, the parametrisation will be performed in two batches of 1,000 cases. These 2,000 cases have been selected bases on a minimum of missing data for the model variables. The selection includes the aggregated cases for the child nodes FG1/FG2 and MAFA, which contain the least missing values and display most combinations of the model variables. Most cases left out of the selection are the cases for the child node PHYT, which contain observations for TEMP and TS. Furthermore, the CPTs that result from learning of the first batch will be saved and provide the starting points for learning from the second batch.

The distribution of the parentless nodes, such as TS, TEMP, RIPZONE and WDT from the learning cases are not representative for the total basin. Therefor, the parentless nodes will be excluded from learning and a Prior Probability Distribution (PPD). The PPD for the variables TEMP and TS follow directly from the monitoring data and the prior distribution for RIPZONE is based on the total length of the shoreline for each class.

For SEDI, there is only the soft sediment type present in the Rotte basin, according to the profile data. But for nearshore locations in the Rottemeren Noord and the Boezem, there is sandy sediment available as well. Therefor, the PPD for the sandy sediment type, in terms of area, is estimated at 2% of the Rotte basin.

The PPD for the states of MUDL for the total Rotte basin is calculated based on the profile data and the representative area for each profile. And finally, the PPD for WDT has been estimated based on the probabilities for the water level (Low, Medium and High), the bottom level according to the profile data and the distribution of the turbidity according to the physio-chemical monitoring. The water level and the bottom level determine a boundary value for the turbidity, for which the WDT-ratio exceeds the threshold value of 0.6. Thus, the PPD for WDT does not necessarily express the frequency of the state, but above all the risk for exceeding the threshold given the distribution for the turbidity.

Table 3.2 presents the resulting PPDs for the variable states of the parentless nodes:

Variable	States	PPD
TEMP	High / Low	51% / 49%
TS	Hypertrophic / Eutrophic / Mesotrophic / Oligotrophic	68% / 30% / 2% / 0%
WDT	Above threshold / Below threshold	92% / 8%
MUDL	Above threshold / Below threshold	93% / 7%
RIPZONE	Natural / Embankment / Foreshore	51% / 44% / 5%
SEDI type	Soft / Sandy	95% / 5%

Table 3.2: PPD for the parentless nodes.

Model evaluation

The BBN model will be evaluated for model prediction performance, the model sensitivity for the input variables and the quality of the set of learning cases. The set of test cases is at the foundation of the model evaluation, as the observed states of the target nodes in the test cases will be compared with the model predictions based on the observations for the explanatory nodes from this same set. The following paragraph will give a brief explanation of these evaluations. The results of the model evaluation is presented in section 4.1.

The model prediction performance of the model has been evaluated with Scoring Rule Results (SRR) and the error rate for the habitat conditions for both fish groups. For the evaluation, FG1 and FG2 are selected as 'unobserved' and the set of test cases will be used to predict the values for these target nodes given the observed variables by Bayesian belief updating using Bayes theorem (see equation 3.1). The predicted states for the habitat suitability for FG1 and FG2 ('Good' or 'Moderate') will be compared with the actual observed state for the fish groups in the test cases. The SRR expresses the deviation between the predicted state and the actual observed state in a numerical value.

NETICA software offers three methodologies for the calculation of the SRR, namely Logarithmic loss, Quadratic loss and the Spherical payoff. The Logarithmic loss gives a value between 0 and infinity. The values closer to 0 indicates a higher certainty for the prediction. The Quadratic loss gives a value between -1 and 1 and a value closer to 1 indicates a higher certainty for the prediction. Finally, the Spherical payoff gives a value between 0 and 1, and the certainty of the model prediction increases for a value closer to 1.

The NETICA software also calculates the error rate based on the model prediction for the most likely state and the actual state in the confusion matrix. A higher error rate means that the model predictions are less certain. The main difference between the SRR and the error rate is that the SRR takes into account not only the most likely predicted state, but the actual belief levels (probabilities) for each state, whereas the error rate only includes the most likely state and compares this with the actual state ([38], [45]).

The model sensitivity has been evaluated by using the entropy reduction algorithm offered by the NETICA software. The entropy is a measure for the surprise of an outcome, given the available information. If more information is added to the model, for example the probability distribution of a variable, this will reduce the surprise (or entropy) of the outcome of a connected variable. The stronger the entropy reduction by adding the information of a variable, the stronger the impact of the variable on the response variable. Thus, the sensitivity analysis points out which parent nodes have the strongest impact on the child node [38]. The entropy reduction in NETICA software is expressed as the percentage of change of the target node, when the findings of the state for the explanatory node changes. The larger the percentage of this proportional change, the stronger the impact of the explanatory node on the target node [45].

The quality of the set of learning cases has been tested by cross-validation, for which an alternative set of learning cases and test cases has been used. The alternative sample for the cross-validation has been compiled by a random selection of the set of test cases from the aggregated cases and therefore include different combinations of data for the explanatory variables.

Additionally, both models will be evaluated for the Receiver Operating Characteristics (ROC) and Area Under Curve (AUC) derived from the ROC, which express the ratio between *True positive* and *False positive* outcomes from the confusion matrix. If the set for the test cases and learning cases are qualitatively okay and balanced for the presence of variable states, the values for AUC for the samples in the cross-validation have similar outcomes. A large differentiation of the AUC indicates that the quality of the data for learning and testing is insufficient and that model prediction results depend to a large extend on the sampling of the data, thus creating large uncertainties of the model predictions.

3.3 Model application

The aim of the research is to develop a DSS to support policy making with respect to FHS in the context of the WFD and for effective expansion of the living environment for fish species in the context of implementation of fish passages. This objective involves both model development as well as a methodology for the application of the model. The model has been used to answer research questions 3 and 4 (see also figure 3.1 on page 14) in order to develop a methodology for model application and interpretation of the results for the assessment of FHS in the Rotte basin. This paragraph discusses the application of the BBN model to address these questions. The results of the model application are presented in sections 4.2 and 4.3.

Question 3: How can the model determine the impact of local conditions on fish habitat differentiation among the target groups?

To answer this question, the BBN model has been used for probabilistic inference. Monitoring data for a specific location and time has been added to the nodes as findings. This probabilistic inference updates the beliefs for the local situation. The analysis will focus on the beliefs after and before inference. The direction of these changes will be compared with the actual state of the fish habitat suitability FG1 and FG2 according to the monitoring data. If the model functions well, the beliefs of the target nodes will converge with the actual state. The speed of the convergence for each variable indicates the impact of the variable to one of the fish groups. The probabilistic inference will be applied for a total of three monitoring occasions in the branches Moerkapelle, Rottemeren and Noorderkanaal.

Question 4: How does the model respond to different management options and what does this imply for FHS and population composition?

Under current conditions, the poor ecological condition in the River Rotte is caused by high nutrient levels, the design of the riparian zone and the limited availability of shallow waters (see also paragraph 2.2). Management options can be applied to change physical-chemical or hydromorphological factors in a desirable direction, for example by reducing nutrient pressures to the Rotte, the realisation of foreshores or changes in the water level. In turn, the changes of these abiotic factors alter the biological conditions, for example for the state of helophyte, hydrophyte, macrofauna and FHS. The management options will be applied to the BBN model by adjusting the prior probability distribution of the specific (parentless) nodes. Changing the prior probability distributions will result in a differentiation of the probability distributions for FHS. The direction of this differentiation indicates the impact of the management options on the habitat conditions for both fish groups. And as a result, changing the habitat conditions eventually changes the composition of the fish populations in the Rotte.

The management options that have been assessed in this research are:

1. Reduction of nutrient pressures on the Rotte basin. The structural reduction of nutrient loads will increase the frequency of the lower trophic states at the expense of the higher trophic states.
2. Temporarily lowering of water levels. This management option results in a higher frequency of low water levels at the expense of high and medium water and a decrease of the mean water levels in general. Changes of the water level will alter the prior probability for the WDT, which will impact the abundance of hydrophyte.
3. Increase of the area of foreshores in the riparian zone at the expense of the areas with natural banks and embankment. The objective of this management option will stimulate the production of littoral vegetation in terms of helophyte and hydrophyte.

Chapter 4

Results

Introduction

From the inventory of suitable modelling approaches in paragraph 3.1 follows that the preffered approach is the BBN model, given the purpose of the model as a DSS for policy making and data availability for the Rotte basin. The following sections in chapter 3 described the steps for model development including the intermediary deliverables and the methodology for model application.

This chapter discusses the final results of the model development and application. Paragraph 4.1 presents the results for the model development, including model evaluation. Paragraph 4.2 presents the results of the application of the habitat suitability model for the assessment of the current FHS and paragraph 4.3 focus on the impact of management options to alter FHS.

4.1 Fish habitat suitability model

Figure 4.1 presents the final model structure including the beliefs following from the model variable CPTs. The BBN model consists of 12 nodes (N = 12) divided in four hierarchical layers. The 12 nodes are interlinked with 14 edges. The majority of the nodes is binary and has only two states (n = 9). Another two nodes have three states (n = 2) and one node has four states. The total network has 113 conditional probabilities, which is the sum of the sizes of all the CPTs in the network. Appendix E presents the CPTs for all model variables.

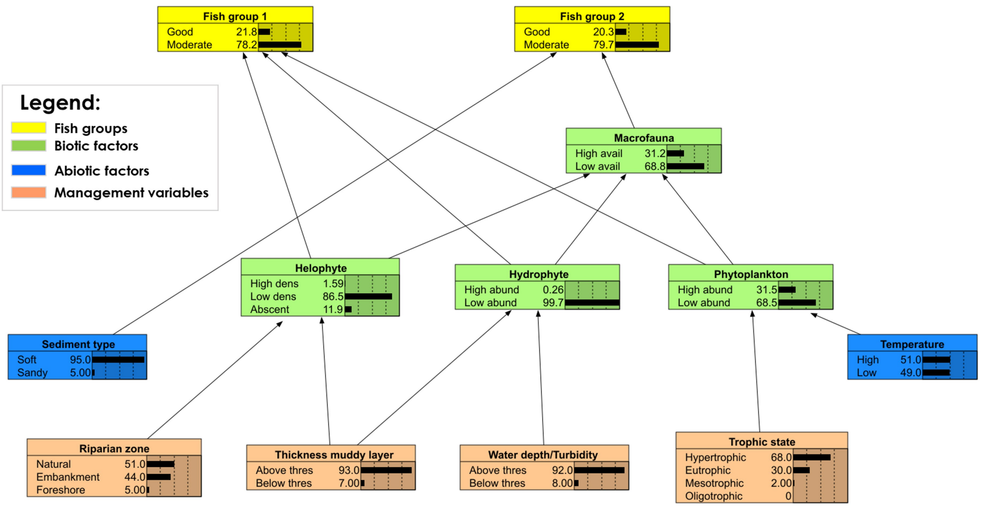


Figure 4.1: Model structure including beliefs.

Model evaluation included the analysis of the model prediction accuracy using the SRR and error rates, a sensitivity analysis to evaluate the impact of the model variables on the target nodes and a cross-validation to test the quality of the set of learning cases.The results from the analysis of the model prediction accuracy for the target nodes FG1 (fish species with a plant-based diet and a habitat preference in or above vegetation) and FG2 (fish species with a diet consisting mainly of benthic life and a habitat preference for soft bottoms) are presented in table 4.1:

Target variable	Logarithmic loss	Quadratic loss	Spherical payoff	Error rate
FG1	1.05	0.33	0.82	18.75%
FG2	0.62	0.43	0.75	31.25%

Table 4.1: Scoring rule results and Error rate

The Spherical payoff approaches 1 for both FG1 and FG2, which indicates a relatively high certainty for the model predictions. Likewise, the values for the Quadratic loss are in the upper part of the range $[-1, 1]$, which also indicates a high certainty. The model prediction for the most likely state for both FG1 and FG2 was, in all cases, Moderate.For all cases, both FG1 and FG2, the model predictions for the most likely state was Moderate. These model predictions resulted in an error rate for FG1 and FG2 of, respectively 18.75% and 31.25%. This error rate indicates that the model predictions for the most likely state corresponds in $70 - 80\%$ of the cases with the actual state.

The sensitivity analysis has been performed for all the child nodes and their parent nodes. Figure 4.2 present the results:

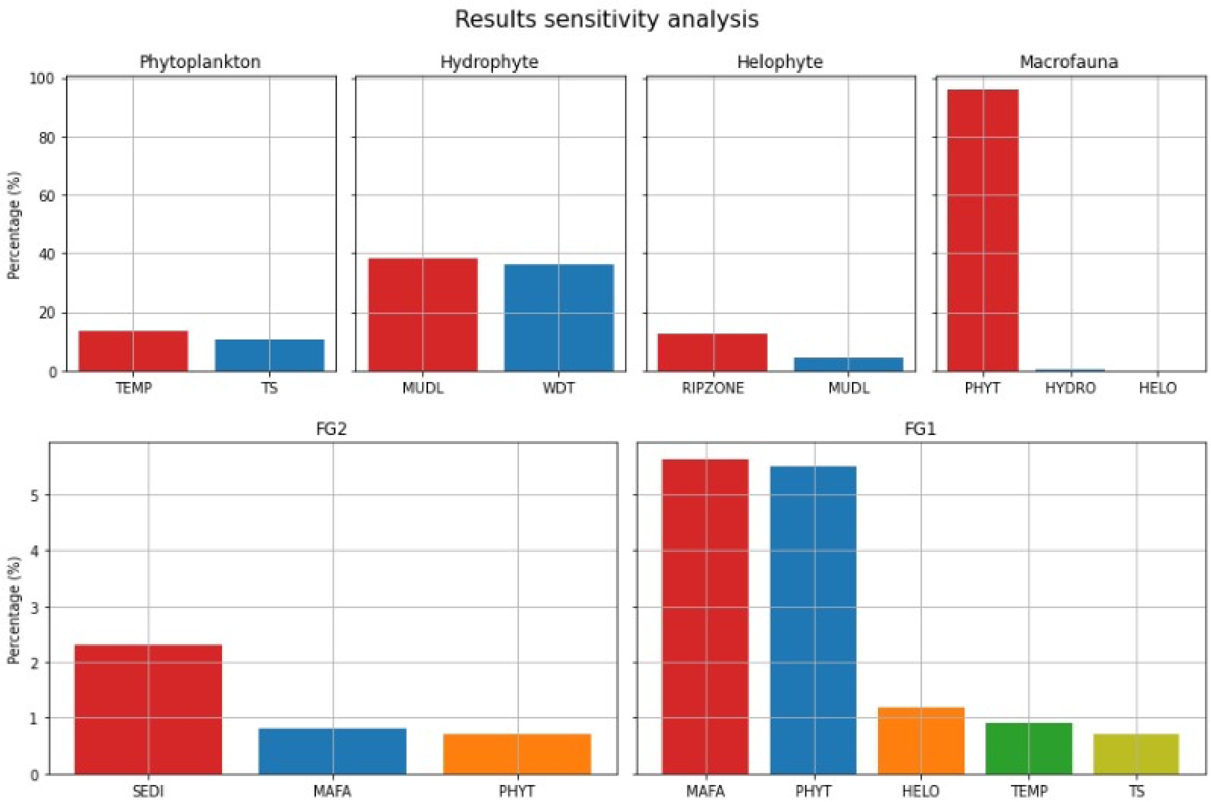


Figure 4.2: Results of the sensitivity analysis for all child nodes

The model variables PHYT, HYDROPH and HELOPH are relatively well explained by the state of their parent nodes, taking into account the complexity of the ecological relationships between the biological and abiotic factors and the simplifications of the model due to the limited the number of parent nodes. The sensitivity analysis points out that the availability of MAFA depends largely on the abundance of PHYT, while the impact of the density of HELOPH (0.2%) and the abundance of HYDROPH (0.1%) are neglectable.

The sensitivity analysis for FHS has been performed for all ancestral nodes. The analysis shows that FG2 depends mostly on its parent node SEDI (2.3%), while the impact of MAFA is neglectable (0.8%). The impact of PHYT on FG2 is equally strong compared to MAFA, which is in accordance with the strong impact of PHYT on MAFA (95%). The parent nodes with the strongest impact on FG1 are PHYT (5.5%) and, to a lesser extent, the HELOPH (1.2%). MAFA also has a strong (indirect) impact on FG1, because, due to the impact of PHYT on MAFA, beliefs for MAFA imply the beliefs of PHYT.

Finally, the model variables TEMP (1.0%) and TS (0.8%) have a considerable impact on the state of FG1, compared to HELOPH. The impact of these variables on FG1 is exercised via their child node PHYT.

Overall, the sensitivity analysis shows that the biotic model variables are to large extend impacted by the parent nodes. The explanatory value of MUDL and WDT for the state of hydrophyte are significant. Likewise, the impact of RIPZONE and MUDL for the state of helophyte and TS and TEMP for the state of PHYT are obviously present, although to a lesser extend. The state of MAFA is predominantly explained by PHYT, while impact of HELO and HYDRO are almost neglectable.

The sensitivity analysis also illustrates that the explanatory value of the biotic model variables for the target nodes is less strong. The limited impact can be explained by the focus of the model on habitat conditions for food and habitat structure, while leaving out other habitat conditions. The reduced explanatory power of the parent nodes for FG1 and FG2 express the complexity of an ecosystem and the interactions between fish species with the surrounding conditions. In this respect, ecologists as well as members of the anglers association have mentioned the flexibility of fish species to adapt to habitat conditions, which suggests that changes in habitat conditions do not directly result in a change in fish population composition (personal communication Erwin Winter (WUR) on December 6th 2022 and angler association HSV Groot-Rotterdam on April 24th 2023).

Table 4.2 presents the results of the AUC for the cross-validation. The table shows the differentiation for the AUC between the original model and the alternative sample set for the cross-validation. The AUC for the cross-validation sample, with values below 0.5, even indicate that the model performance is worse than a random model output [38], which was not the case for the original model.

Target variable	Original model	Resample model
FG1	0.55	0.22
FG2	0.77	0.30

Table 4.2: Results AUC from cross-validation.

Cross-validation and the analysis of AUC illustrates the large impact of data samples on model predictions. This large impact of the sample data indicates that data in the learning cases is insufficiently balanced for parametrisation of the model. Thus, data availability for the current model variables is the basis for model prediction uncertainties and limited accuracy.

The limitations regarding the parametrisation of the model is also reflected in the CPTs of the definitive model (see also appendix E). From a more detailed analysis of the CPTs it follows that FG1 is, according to the results from the sensitivity analysis, mainly influenced by the parent nodes PHYT and HELO, while the model variables TEMP and TS influence FG1 via their child node PHYT. From the CPT for PHYT follows that a high abundance of phytoplankton results, in most cases, in an increase of the probability that FG1 has the state 'Good' (in formal notation $P(FG1_Good)$), while a low abundance of phytoplankton results in a decrease $P(FG1_Good)$. These findings are in line with the expected results, because of the availability of food.

When looking into the CPT for PHYT, the model has some counter intuitive features from a biological perspective. The combinations of TEMP and TS gives a conditional probability $P(PHYT_{High\_abun}|TEMP_{High}, TS_{Hypertrop}$
0.25, while the combination with a lower trophic state results in a higher probability for high abundance of phytoplankton ($P(PHYT_{High\_abun}|TEMP_{High}, TS_{Eutrophic}) \approx 1.0$). The ambiguities in the CPTs are most likely caused by the limited frequencies for the combinations of states for the variables PHYT, TEMP and TS in the learning cases.

The marginal probability $P(FG1_{Good}|HELOPH_{High\_dens}) \approx 0.5$, while the marginal probability for $P(FG1_{Good}|HELOPH_{Low\_dens}) \approx 0.11$. This means that the most likely state for FG1 for low density helophyte is Moderate. Again, the CPT for FG1 shows a counter intuitive response to the state of helophytes. The unexpected probability distributions in the CPTs for FG1 limit our ability to determine the direction of the impact of the explanatory nodes on the target nodes.

FG2 is mainly influenced by its parent nodes SEDI and MAFA and by the variable PHYT. The marginal probability for the soft sediment type $P(FG2_{Good}|SEDI_{Soft}) \approx 0.2$, while for the marginal probability for sandy sediment $P(FG2_{Good}|SEDI_{Sandy})) \approx 0.0$. In both cases, the most likely state for FG2 is Moderate, but the probability $P(FG2_{Good}|SEDI_{Soft})$ is higher than $P(FG2_{Good}|SEDI_{Sandy})$, which corresponds with the assumed structure preferences for FG2. For macrofauna, the marginal probability $P(FG2_{Good}|MAFA_{High\_avail}) \approx 0.08$, while the marginal probability $P(FG2_{Good}|MAFA_{Low\_avail}) \approx 0.12$. These marginal probabilities illustrate that, according to the BN model, the probability for good habitat conditions for FG2 increase for a lower abundancy of macrofauna. Again, the results from the analysis of the CPT for FG2 are counter intuitive.

The model evaluation demonstrates some short-comings of the current model. At first sight, regarding the analysis of the SRR and error rates, the model performance is acceptable, with values for the Quadratic loss and Spherical payoff indicating relatively high model certainties. Though, the results from the sensitivity analysis indicates the magnitude of the impact of the explanatory nodes on the target nodes is limited. Moreover, the analysis of the model CPTs shows that the direction of the influence of some of the the explanatory nodes on the target nodes are ambiguous. In addition, the results from the cross-validation suggest that limited data availability for the current model variables are at the cause of the constrains with regard to model accuracy.

4.2 Assessment of current fish habitat conditions

The research question which will be addressed in this paragraph is:

How can the model determine the impact of local conditions on fish habitat differentiation among the target groups?

The approach for this research question is by probabilistic inference for three aggregated cases in the Rotte basin. The three cases for the sections Moerkapelle, Rottemeren and Noorderkanaal have been selected based on the maximum data availability and the observed data for these cases have been added to the model as findings.

Table 4.3 presents the cases, including the observed state for the FHS for FG1 and FG2:

The cases for Moerkapelle and Rottemeren share the same evidence, especially with regard to the parent nodes for the target nodes FG1 and FG2. The parent nodes for FG1 are helophyte, hydrophyte and phytoplankton. For Rottemeren, the data for hydrophyte is not available, but the conditions for the thickness of the mud layer (MUDL) and WDT ratio result the exact state for hydrophyte compared to Moerkapelle. The state for phytoplankton for both Moerkapelle and Rottemeren have been determined by the variables TEMP and TS.

The parent nodes for FG2 are macrofauna and sediment type. For both cases for Moerkapelle and Rottemeren, this data is available, and, as described in paragraph 3.1, the evidence for the parent nodes create a conditional independence from all other variables in the network.

Section	Moerkapelle	Rottemeren	Noorderkanaal
Location	RB-14-SK-01	RO-20-SK-02	RB-14-EL-22
Date	2014-07-31	2020-07-31	2014-07-31
FG1	Moderate	Good	Moderate
FG2	Good	Good	Moderate
MAFA	Low avail	Low avail	*
PHYT	*	*	*
HYDROPH	Low abund	*	*
HELOPH	Low dens	Low dens	*
SEDI	Soft	Soft	Soft
RIPZONE	Natural	Foreshore	Embankment
MUDL	Below thres	Below thres	*
WDT	Above thres	Above thres	*
TS	Hypertrophic	Hypertrophic	Eutrophic
TEMP	High	High	High
Missing data	1	2	6

Table 4.3: Cases for the analysis of the impact of local conditions via probabilistic inference.

Figure 4.3 illustrates the progress for the beliefs for FG1 and FG2, starting from the prior probability distributions and sequentially adding evidence of the explanatory variables:

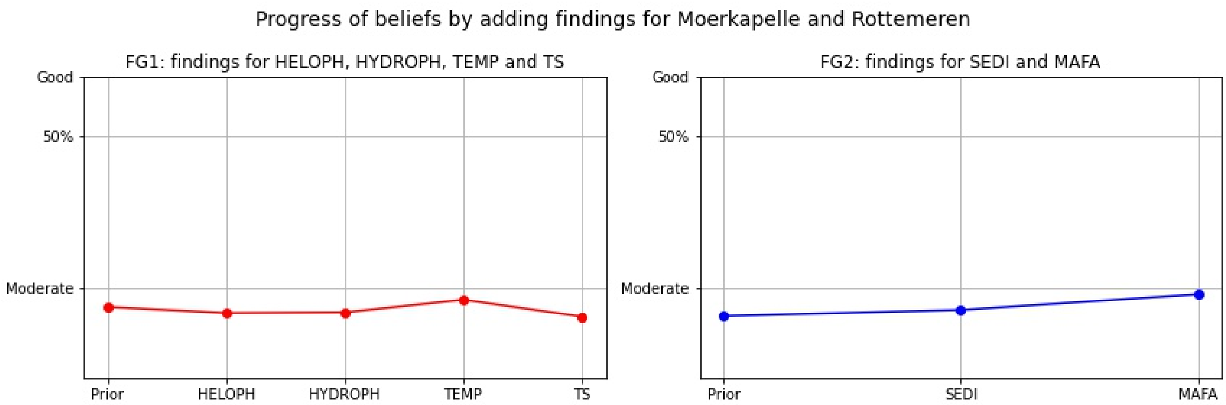


Figure 4.3: Results for the probability inference for Moerkapelle and Rottemeren

The FHS for FG1 in the sections Moerkapelle and Rottemeren diverge from the state $FG1_{Good}$. The high temperature results in an move towards a good habitat suitability, but the hypertrophic state nullifies it. The diverging trend is in line for the actual state for Moerkapelle, but deviates from the actual state for the Rottemeren. The FHS for FG2 converges towards a Good habitat state, which is in line with the actual state. Both the soft sediment state and the low abundance for macrofauna are causing this trend. But, as explained earlier, the impact of low macrofauna availability on FG2 is ambiguous.

The evidence for the Noorderkanaal consists of only four variables and none of these variables are the parent nodes for FHS. Figure 4.4 illustrates the progress for the beliefs for FG1 and FG2 for the Noorderkanaal:

The findings for the Noorderkanaal result in a convergence for the habitat suitability $FG1_{Good}$. This is primarily the effect of the variables TEMP, TS and RIPZONE. The states for TS and TEMP in the Noorderkanaal result in high abundance for phytoplankton, which, in turn, controls the habitat suitability towards $FG1_{Good}$. The state $RIPZONE_{Embankment}$ results in a higher probability $P(HELOPH_{Abscent})$, which in turn, according to the model parameters, controls the habitat suitability for FG1 towards $FG1_{Good}$. Eventually, the convergence towards $FG1_{Good}$ does not correspond with the actual state for FG1 in the Noorderkanaal.

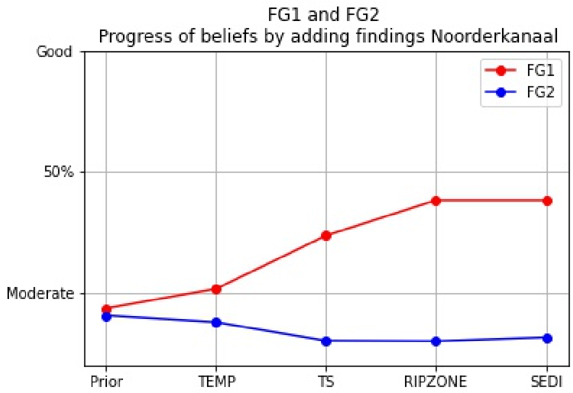


Figure 4.4: Results for the probability inference for Noorderkanaal

The habitat suitability for FG2 is controlled by all variables for the Noorderkanaal. The high probability for $PHYT_{High\_abun}$, which is the result for the findings for TEMP and TS, results in a high probability for $MAFA_{High\_avail}$. Also the increase of $P(HELOPH_{Abscent})$, which is the effect of the state $RIPZONE_{Embankment}$, has a positive impact on $MAFA_{High\_avail}$. As discussed earlier, the high availability for macrofauna has a limiting effect for the habitat suitability for FG2. The state $SEDI_{Soft}$ has a direct impact on the habitat suitability for FG2 and controls FG2 towards $FG2_{Good}$. The net effect is a divergence of FG2 from $FG2_{Good}$, which indicates that the impact of MAFA on the habitat suitability for FG2 dominates over the variable SEDI. The divergence of FG2 from $FG2_{Good}$ corresponds to the actual state for the Noorderkanaal.

Table 4.4 presents the results from the inference for the three sections and the comparison with the actual state according to the monitoring data. The results from the analysis deviate partly from the actual cases, as table shows. But the model correct predictions can be caused for the wrong reasons, and the same accounts for the wrong predictions, since model parametrisation is limited by the data availability and the frequency for combinations of variable states. Nevertheless, the applied methodology based on the probabilistic inference for the observations and the analysis for diverging or converging beliefs illustrates the impact of local conditions for the habitat suitability.

Section	Fish group	Actual state	Results inference
Moerkapelle	FG1	Moderate	Diverging from Good
	FG2	Good	Converging to Good
Rottemeren	FG1	Good	Diverging from Good
	FG2	Good	Converging to Good
Noorderkanaal	FG1	Moderate	Converging to Good
	FG2	Moderate	Diverging from Good

Table 4.4: Results from probabilistic inference for cases Moerkapelle, Rottemeren and Noorderkanaal.

4.3 Scenarios for the management options

The research question which will be addressed in this paragraph is:

How does the model respond to different management options and what does this imply for the FHS and population composition?

The management options that have been examined are:

Scenario 1: Scenario 1: Reduction of nutrient pressures on the Rotte basin.

Scenario 2: Temporarily lowering of water levels.

Scenario 3: Increase of the area of foreshores in the riparian zone at the expense of the areas with natural banks and embankment.

The management options focus on controlling the physio-chemical and hydromorphological variables. These variables are the parentless nodes in the model, for which a PPD has been defined based on the data analyses for the total Rotte basin (see also paragraph 3.2). For modelling the scenarios, the PPDs have been adjusted based on hypothetical *Key Performance Indicatorss (KPIs)* for the management strategies by the water authority.

The hypothetical KPIs were applied, in stead of current policy options, because of the limited time and opportunity to discuss actual management goals and KPIs with the water authority. Besides, the current model parametrisation shows irregularities and the results from the scenario studies with the current model version will have limited value for future decision making. The scenarios discussed in this section are intended for testing the methodology to simulate scenarios by adjusting the prior probability distributions and to illustrate the line of reasoning from the results of the scenarios.

The analyses focus on the impact of the policy options for the balance between the habitat suitability for FG1 and FG2. Political interest for the abundance of the fish groups in the Rotte is, partly, motivated by the WFD. HHSK states that the current poor ecological state of the Rotte is partly caused by the high abundance of benthivorous fish, such as Bream (*Abramis brama*) because the high abundancy of Bream has a negative impact on the Ecological Score Card (ESC) (see also paragraph 2.2). Hence, there is a political interest to reduce the abundance for FG2 in favour of FG1.

Scenario 1: Reduction of nutrient pressures on the Rotte basin.

The reduction of nutrient pressures in the Rotte basin is part of the current programmes of the water authority, for example the programme for the zero-emission greenhouses. Goals for the trophic state do not exist and, most likely, will not come into place in the short or medium term because of the economic and societal positions of the main stakeholders and the expected resistance against progressive objectives, but also due to the complexity of this topic with regard to the chemical-biological process involved. With these limitations in mind, there are no specific actions attached to this scenario. This scenario is motivated from a strategic point of view, to investigate the What if - question: What if we could substantially lower the trophic levels in the Rotte, what could be the possible results for the fish habitat conditions?

The structural reduction of nutrient loads in the scenario increases the frequency of the lower trophic states at the expense of the higher trophic states. The analysis focuses on the alterations of the resulting probability distributions of FG1 and FG2 in the model. Table 4.5 presents the PPD for the variable TS and figure 4.5 presents the results for the fish habitat conditions:

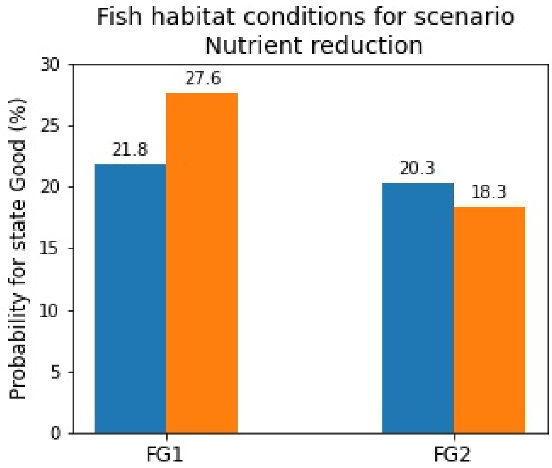


Figure 4.5: Results Nutrient reduction.

TS	Prior	Nutrient reduction
Hypertrophic	68%	25%
Eutrophic	30%	40%
Mesotrophic	2%	30%
Oligotrophic	-	5%

Table 4.5: TS probability distribution

The scenario results in an increase of $P(FG1_{Good})$ and a decrease of $P(FG2_{Good})$. These changes are caused by the changes of the probability distributions (or beliefs) for phytoplankton and macrofauna, since both $P(PHYT_{High\_abund})$ and $P(MAFA_{High\_avail})$ increased in this scenario. The changes in the beliefs of the model variables challenge our understanding of the biological system. First of all, the nutrient reduction in the scenario aims to reduce phytoplankton concentrations in order to create a state of clear water dominated with aquatic vegetation [50]. The current model parametrisation results in an opposite effect, because both $TS_{Eutrophic}$ and $TS_{Mesotrophic}$ increase $P(PHYT_{High\_abund})$ instead of causing a decrease. Another characteristic of the model parametrisation causing this unexpected outcome, is the negative impact of $MAFA_{High\_avail}$ on the probability for a good habitat condition for FG2 $P(FG2_{Good})$.

The model predictions for this scenario are ambiguous, because of the illogical responses of the variables $PHYT$, $MAFA$, $FG1$ and $FG2$. But the scenario demonstrates the different responses of the habitat conditions for FG1 and FG2 on nutrient reductions in the Rotte basin and therefor provides valuable information to evaluate the impact of different scenarios.

Scenario 2: Temporarily lowering of water levels.

HHSK has recently adopted a more flexible water level, which allows temporarily lower water levels without necessarily supplying water to the system. Lowering the water levels alters ecological factors, such as food availability, protection for juvenile fish and water temperature, but for our model the main impact will be the WDT ratio and, as a response, the production of hydrophyte. Lowering the water level during Spring and Summer has been suggested from an ecological perspective (personal communication with Marit Meier, September 26, 2022), but is under debate because of the claims for water supply for agricultural production and the stability of the water shores, among others. The scenario explores the ecological impact of the management option, in order to add the ecological perspective and the targets within WFD to the debate on water level management.

The starting point of this scenario is to allow lower water levels and as well as to allow longer durations of lower water levels. The output of the policy is a change in the classification of the water levels (high, medium and low), with a higher frequency of low water levels in the basin. At the same time the policy results in a lowering of the average water level for each class. The turbidity is assumed to be independent from the water level. Therefor, the prior distribution for turbidity, which followed from the physio-chemical monitoring data, has not been adjusted.

Calculations for the probability distribution for the WDT show that the adjustment of this variable requires adjustments in the water level management which deviate significantly from the current practices. This is caused by the relatively steep bed profile and the limited presence of shallow waters in the Rotte basin. For the hypothetical case which include more extreme weather (e.g. draughts and extreme rainfall) and the implementation of rigorous measures for water level management, the PPD for the WDT have been altered by nearly 4%.

Table 4.6 presents the probability distribution of the water levels and the resulting prior probability distributions of the variable WDT for the current situation and the scenario:

	Prior		Water level reduction	
Water level	Avg. water level (mNAP)	Probability (%)	Avg. water level (mNAP)	Probability (%)
High	-1.03	29%	-1.05	30%
Medium	-1.07	71%	-1.20	10%
Low	-1.19	$> 0\%$	-1.40	60%
	WDT for Prior		WDT for water level reduction	
Above threshold	92%		88%	
Below threshold	8%		12%	

Table 4.6: Water level and WDT probability distributions.

The changes in the WDT result in a small increase for $P(HYDROPH_{Highabund})$, from 0.27% in the prior distribution to 0.4% for the scenario output, but this change is far to small to impact FHS. The results from this simulation imply that, given the current bed profile and turbidity in the Rotte basin, reduction of the water levels alone will have an insignificant effect on the abundance of hydrophyte. Alternative measures will be needed to improve this factor.

Scenario 3: Increase of the area of foreshores in the riparian zone at the expense of the areas with natural banks and embankment.

The creation of foreshores is an ongoing project of HHSK and since 2020, approximately 2.0 ha foreshores has been realised (personal communication Peter Verkaik, 7^{th} of June 2023). The objective of these projects is to stimulate the production of helophyte in the areas where they are abscent. According to HHSK, the increased production of helophytes can reduce the nutrient levels in the water and in the soil. In our model, increasing the availability of, preferably low density, helophyte creates habitat structure and increases food availability for FG1. At the same time, the availability of helophyte stimulates the production of macrofauna and, in doing so, increases food availability for FG2. Thus, the expected result for this measure is a positive effect on the habitat suitability for FG1 and FG2.

Table 4.7 and 4.6, both on page 35, present the distribution for the classification of the riparian zone and the results for the habitat suitability for FG1 and FG2.

RIPZONE	Prior	Realisation Foreshores
Natural	51%	30%
Embankment	44%	40%
Foreshore	5%	30%

Table 4.7: RIPZONE probability distribution.

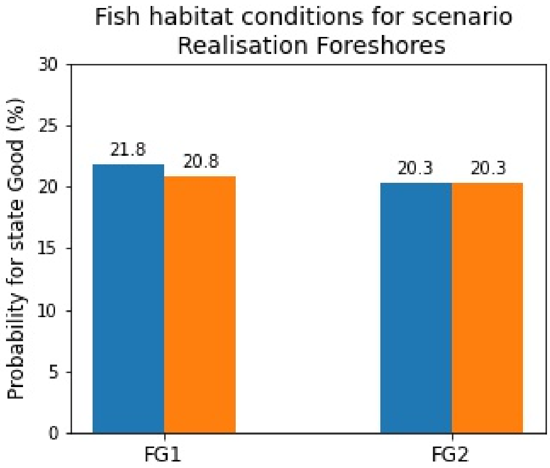


Figure 4.6: Results Realisation Foreshores.

The realisation of foreshores result in an increase of $P(HELOPH_{High\_dens})$ at the expence of $P(HELOPH_{Abscent})$. The effect of the riparian zone on helophyte has been immediately forwarded towards the target node FG1, which gives results in a decrease of $P(FG1_{Good})$. The increase of $P(HELOPH_{High\_dens})$ also reduces $P(MAFA_{High\_avail})$ compared to the prior distribution, but this effect seems to be smoothened out at the target node FG2.

Again, the model predictions do not correspond to the expected response of FG1 and FG2, but the results illustrate that the BBN model for FHS can give valuable insight in the different responses of the target nodes FG1 and FG2 to scenario inputs.

Chapter 5

Discussion

Choice for the modelling approach

The prior reasons for the development of a BBN model in this research are the ability to use different forms of information, the high level of uncertainties and non-linearity of the ecological system and the limited data requirements of the model. Also, BBN models allow for simplifications of the model, without hampering model results [20]. And finally, BBN models are appreciated for their strong communicative power, because it takes the relation diagram as a starting point and presents the results in terms of likelihood including the uncertainties about the outcomes ([37], [7]).

Although BBN models have often been developed and used for ecological management decisions, the use of this type of models within the WFD is uncommon. Conallin *et al.* [9] don't even mention BBN models as an alternative in his overview of model approaches for European water resource managers. And maybe that is illustrative for the position of these model types in the Directive. Within there is a tendency to develop and expand DSS models with large complexities to represent the ecological systems. Junier [30] criticises this development, because these systems can, in practice, only be used by experts. She also raises the story how, in the Netherlands, an ecological process model (WFDE) has been replaced by a statistical neural network model, because of the increasing complexity of the former. In this respect, the BBN modeling approach is useful for all kinds of stakeholders. Furthermore, BBN models provide an elegant balance between the causal modelling approach and the statistical approach, with the relational diagram as a starting point and data-based machine learning for the parametrisation.

Underlying research illustrates the application of the BBN model for ecological studies. The main steps in the research process have been the development of the model and the development of a methodology for model application and interpretation of the model results. Concerning the development process, the strong communicative power of the BBN model has proven itself during the development process with stakeholders with different expertise, interests and knowledge levels. A limiting factor in the development of the model has been data availability to develop sufficient learning cases. The results of the model evaluation in paragraph 4.1 suggest model uncertainties, because of the existing error rates and the results of the cross-validation. But the analysis of the CPTs revealed the short comings in the parametrisation of the model. The current model presents ambiguous predictions which challenge our understanding of the ecosystem. It is clear that model parametrisation has to be improved in order to enhance the current BBN model.

The methodology to address the research questions with regard to model development and application consist of the features supplied by the NETICA software package, for example the sensitivity to findings and the use of probabilistic inference, and includes additional analyses, such as the marginal probabilities for the CPTs. The results show that the BBN model is very suitable to evaluate the current habitat suitability and the effectiveness of different management options. That is, leaving aside the model uncertainties caused by the current model parametrisation.

Different studies proclaim the use of multiple, alternative, model types parallel to each other, for a better understanding of the uncertainties in ecological modelling ([36], [58], [20]). In this respect, Chen & Pollino [7] mentions that the use of different models can help to detect equifinality. Marcot *et al.* [37] states that using different models gives a better understanding whether the results follow from the model structure and parametrisation or from ecological conditions. For the case of the FHS in the Rotte basin, alternative modelling approaches can be complementary to the BBN model developed in this research. A good alternative could be the *fuzzy logic* model based on learning rules in a decision tree. This model approach requires limited data and is also advised by Conallin *et al.* [9] for WFD water bodies.

Discussion model structure

The choices with regard to the model structure are a balancing act between model parsimony, prediction accuracy and precision of the model representation [37]. Hamilton *et al.* [20] states that, especially, situations with limited data requires simpler BN model structures. These simplifications require more assumptions, but do not necessarily affect the reliability of the model. In this respect, Hamilton *et al.* refers to an interesting feature of the BN models that the uncertainties due to missing explanatory variables or poorly supported classifications of model states will eventually expressed in the probability distribution of the target nodes. Marcot *et al.* [37] and Hamilton *et al.* [20] explain that simpler models with fewer nodes or fewer states per node can also provide more accurate predictions because of a better match of the model structure with the scope of available data.

These views on BBN models are contrary to the approach for the BBN model as described by Teurlincx *et al.* [54]. The objective of the BBN model in this report is to illustrate and quantify the individual factors and the input on the functioning of the ecosystem for still waters [54]. The model structure consists of a myriad of model variables, including water balances, wind speed, suspended solids in the water and dredging frequencies, among others. Furthermore, the parametrisation of the CPTs is based on three resources, namely meta- or process models, statistical models comparable to the learning cases in underlying study and expert judgement. The choice for a complex structure is motivated by the preference to represent the complexity of the ecological processes in still waters in a model.

The function of the BBN model in underlying research is mainly to support decision making for ecological management and is less focused to represent actual ecosystem complexity. Therefor, objectives for this model are more in line with the modelling approach as pictured by the studies by Marcot *et al.* and Hamilton *et al.*, who give preference for the more simpler model structures.

Marcot *et al.* [37] presents some practical guidelines with respect to the model structure. First of all, the model structure should consist a maximum of four layers. Second, the number of parent nodes per child node is preferably less than four. And thirdly, the maximum state per variable is five. The reduction of the number of layers is important to limit intermediate nodes, because these nodes result in an unnecessary propagation of uncertainties throughout the model and will damp out the impact of the input nodes on the output nodes.

The limitations for the number of parent nodes and variable states are important to reduce the size of the CPTs and to keep the CPTs tractable and understandable. Equation 5.1 calculates the number of entries of the CPT given the number of parent nodes, the variable states of the parent nodes and the variable states of the child node. The equation shows that the number of entries, which correspond to a unique combination of all the variable states, increases progressively:

$$Entries = S * \prod_{i=1}^{n} P_i \tag{5.1}$$

In this formula S is the number of states of the child node, P_i is the number of the i_{th} parent node and n is the total number of parent nodes.

The structure of the model satisfies the guidelines discussed above. The BBN model consists of 12 nodes and, in total, 113 entries combined for all CPTs. The side effect of this rather simple model structure is that the impact of the explanatory nodes FG1 and FG2 is limited, as the sensitivity analysis points out (see also paragraph 4.1). The model structure basically comes from the choice to include only the habitat preferences for food and for structure in the model. This choice leaves out many other variables that impact fish abundance. For example, the predation by Cormorants has major impact on fish populations, according to many anglers in the Rotte and the angler associations. The impact of predation is the result of the Cormorant population, the available fish population and the effectiveness of the Cormorants to catch the fish.

Likewise, the model structure can be expanded based on available monitoring data for the Rotte basin, for example by adding oxygen saturation or acidity (in pH values) as explanatory nodes for FHS. For this expansion, understanding of the direct and indirect impact of the variables on the target nodes as well as the interaction with other model variables is required.

The examples for expansion of the model structure show that adding a variable to the model imposes new challenges on data availability. In the case of adding new habitat factors, such as predation by Cormorants, new data is needed required for these specific features, such as the Cormorant population or data on functional underwater structures for protection of the fish. In case of adding model variables based on available data, such as oxygen saturation or pH, model expansion involves mainly new relationships between the existing nodes and the new additions based on our understanding of the ecosystem. In both cases, model expansion requires new data for the specific features and additional combinations of data are required for the parametrisation of the CPTs.

Given the current model prediction accuracy, expansion of the model variables should be discussed, regardless of the increased model complexity and data requirements. New explanatory variables can certainly improve model predictions but they also expresses the concerns for different stakeholders with respect to the FHS. BBN models have the quality that they can be developed in an iterative process. The BBN model can be easily adjusted and expanded in an iterative process, as data and knowledge gaps of the (ecological) system or specific model variables are filled.

Reliability of the model

Bayesian learning uses hypotheses and calculates the probabilities for each hypothesis after observing the evidence (or findings). Based on the hypotheses for a combination of variable states and the weight of their probabilities, the model makes new predictions. The quality of the predictions will improve as the hypothesis space decreases by reducing false hypotheses [48]. If data availability is limited, opportunities to disapprove false hypotheses decreases. This could result in overfitting of the model, which means that the model will respond more to rare combinations of variable states and extreme events (*false hypotheses*) and will be less robust for normal situations. Thus, the risk for overfitting is being reduced when more data is available, because the posterior probabilities for any false hypothesis will eventually vanish due to learning.

For that reason Chen & Pollino [7] have introduced a guideline to use a minimum of 20 cases for each combination of variable states of the parent nodes. In our case, this guideline for 20 cases for each combination has not been reached due to limited data availability. The causes for the limited data in the Rotte basin is two-fold. First of all, biological data is not sufficiently available because of the number of monitoring occasions. The three-annual monitoring has resulted in 4 years of data and the number of monitoring locations, especially for macrofauna and macrophyte, are few in numbers. Furthermore, the monitoring locations for these biological factors are not evenly spread over the Rotte basin. Therefor, biological data are not available for the sections Noorderkanaal, Binnenrotte and Boezem.

The second reason for data limitations is that some conditions, mainly hydromorphological factors, are not widely present in the Rotte basin. For example, $SEDI_{Sandy}$ is only detected at 6 locations out of 64 and also the frequencies for $MUDL_{Below\_thres}$ (7 out of 795) and WDT_{Below_thresh} (19 out of 91) are limited available.

The main options to improve data availability and reduce the risk of overfitting are (1) expansion of the BBN model from the Rotte basin to other water bodies and (2) the use of metamodels for the (ecological) processes based causes and effect.

The expansion of the model should contain all the water bodies within the WFD with similar characteristics, such as heavily modified systems, storage canals, and particularly, water bodies for which fish passages between lentic and lotic waters have been realised. These water bodies are subjected to the same rules for monitoring and data collection as the Rotte basin, therefore similar variable data is available. And, not importantly, the defined fish groups for plant loving and benthivorous fish, including the classification of the fish species, recognisable and applicable for all water bodies. The expansion of the BBN model to other water bodies means that more monitoring data will be available. It also increases representation of different variable states for the learning cases and testing cases, because water bodies have different (hydromorphological) features and therefor the learning cases for the different water bodies will be complementary to each other.

The metamodels can be applied specifically for the model variables for which data availability and uncertainty is the prior limitation. But they can also be used to attach data for locations where no measurements have taken place. In most cases, the metamodels will require proxy data, based on other data resources, to predict the model variables. The application of metamodels requires a concise understanding of the biological and ecosystem processes. Therefore, the use of metamodels impose new challenges for the development of the BBN model, but can be overcome when building upon available resources. It is very likely that metamodels for most of the essential model variables are available from earlier ecological modeling processes within the WFD. Additionally, BBN models have the flexibility to add data via learning cases, but also via (biological) process models or a combination of both.

Chapter 6

Conclusions

The starting point for this research is the increasing political interest in fish habitat suitability in fresh water bodies. This interest is motivated to secure effective investment in fish passages and longitudinal connectivity, but also by the challenge to achieve a good ecological status as part of the WFD. With respect to the latter, the presence of plant-loving fish species (FG1) have a favourable impact on ecological quality, whereas benthivorous fish (FG2) have a negative appreciation in the Dutch approach for the WFD.

The BBN model that has been developed in this research, functions as a DSS for the ecological restoration of the Rotte basin with respect to fish habitat suitability. Starting point for the current model is the distinction between between food preferences and habitat structure preferences between benthivorous and plant-loving fish species. Monitoring data for the Rotte basin is used for parametrisation of the model.

The choice for the BBN modelling approach is primarily motivated by the limited data requirements for this model approach in terms of data availability for parametrisation and the use of different types of data, such as continuous and categorical data. BBN models calculate the probability distribution for variable states and built upon the principle that equal conditions can result in different variable states. In line with this principle, BBN models reflect uncertainties which are inherent to complex ecosystems. However, the evaluation of the BBN model shows that data availability for the Rotte basin is insufficient for proper parametrisation. To illustrate this further, the data set for machine learning does not comply with the guideline of a minimum of 20 cases for each combination of variable states. The result is a model with high predictive uncertainties and outcomes which challenge our understanding of the ecosystem.

The model structure, composed of the model variables and the edges connecting the variables, is the result of the selection of habitat factors (i.e. for food and habitat structure) and the consultation of literature and the review of stakeholders and experts. The accuracy of the resulting model in this research is limited. The error rate varies between $20 - 30\%$, but, more importantly, the parametrisation and model predictions depend for a large part on the sample data for machine learning in stead of the model structure. In this respect, the choice of the model variables impacts model prediction accuracy in two ways:

1. The current data set is not sufficient to phase out the false hypotheses during machine learning. This is main cause for the irregular model prediction results and most likely limits the explanatory power of some of the model variables as well. Data availability can be improved by adding data from other water bodies or by using metamodels for specific variable nodes. The use of data from other water bodies is preferred for three reasons. First, for the designated water bodies in the Netherlands have similar monitoring data is available. Therefore, the use of additional data from other water bodies is easy to perform. Secondly, the addition of learning cases data from other water bodies in the Netherlands can be the first step for the deployment of the BBN model. In this respect, it would be best to start with the water bodies with similar features like the Rotte, which is classified as a heavily modified storage canals with restored longitudinal connections to lotic water systems. Currently there are at least 36 similar water bodies within the Netherlands which are designated to the WFD. Thirdly, the use of metamodels to improve data availability increases the risk that focus on the model validity gets the upperhand, despite the usability of the model as a political decision support tool.

2. The nodes and edges in the model have an explanatory value for the target nodes FG1 and FG2. The sensitivity analysis points out that the explanatory value of some of the parent nodes for FG1 and FG2 is very limited. This could be at the cause of the selected habitat factors, which leave out other explanations for the presence of the target fish groups. By expanding the model structure with relevant habitat factors, including additional parent nodes, model prediction accuracy could be improved. On top of that, BBN models can be easily expanded for additional model variables, but the expansion requires sufficient understanding of the ecosystem to include additional parent nodes and edges as well as sufficient data to cover the possible combinations of variable states.

After model development and evaluation, the BBN model is applied for the assessment of the current habitat conditions and to simulate the impact of different management options on fish habitat suitability.

The assessment of the current habitat conditions has been carried out by probabilistic inference. Data from the learning cases have been added to the explanatory nodes as findings. The sequential addition of these findings updates the beliefs for the states of the target nodes FG1 and FG2. The direction (i.e. convergence towards or divergence from a variable state) and the magnitude of these updates indicate the impact of the local condition on the fish habitat suitability for the target groups.

For example, the analysis in paragraph 4.2 shows that fish habitat suitability for FG1 in the sections Moerkapelle and Rottemeren is primarily influenced by the variables TEMP and TS. The high temperatures pushes the variable state for FG1 towards 'Good', while the hypertrophic state drives the state for FG1 towards 'Moderate'. The variables TEMP and TS also have the largest impact on fish habitat suitability for FG2 in the Noorderkanaal. The high temperature and the eutrophic state both drive the variable state for FG2 towards 'Moderate'. The soft sediment drives the variable state for FG2 towards 'Good', but this impact is less strong.

The management options have been simulated by changing the prior probability distributions for the target variables of the policy. The scenarios in this study focus on the trophic state, WDT ratio and the riparian zone. The scenarios for the reduction of nutrient levels and extension of the foreshores show a clear response of the fish habitat suitability for FG1 and FG2. The reduction of nutrient levels has a positive effect on the fish habitat suitability for FG1 and a negative impact on the fish habitat suitability for FG2. According to the current model, lowering the trophic state will most likely advantage FG1 over FG2. Realisation of foreshores will reduce the habitat suitability for FG1, while the habitat suitability for FG2 remains the same. Thus, in this scenario the intervention will most likely advantage FG2 over FG1. The scenario for temporarily lowering the water levels does not result in a differentiation between the fish habitat suitability for FG1 and FG2, indicating that this management options will most likely have a limited effect to modify fish population compositions.

As stated before, the model predictions of the current BBN model show irregular and illogical outcomes. Therefore, the results from the model application are as yet unusable for policy making. Never the less, the methodologies for the assessment of the local habitat conditions and the simulation of management options are functional and efficient to apply. And also, the BBN model is an excellent tool for participation of stakeholders with different expertise, interests and knowledge levels, as experienced during underlying research. The current BBN model should be considered as a prototype model to built upon and to improve.

The main improvement is the parametrisation of the model by expanding the scope to similar water bodies in the Netherlands. The prediction accuracy of this new model should be tested by using SRR, sensitivity analysis and cross-validation. As a third step, expansion of the model structure should be discussed with ecologists and other stakeholders, to include more habitat factors to determine fish habitat suitability.

Appendix A

Monitoring locations HHSK

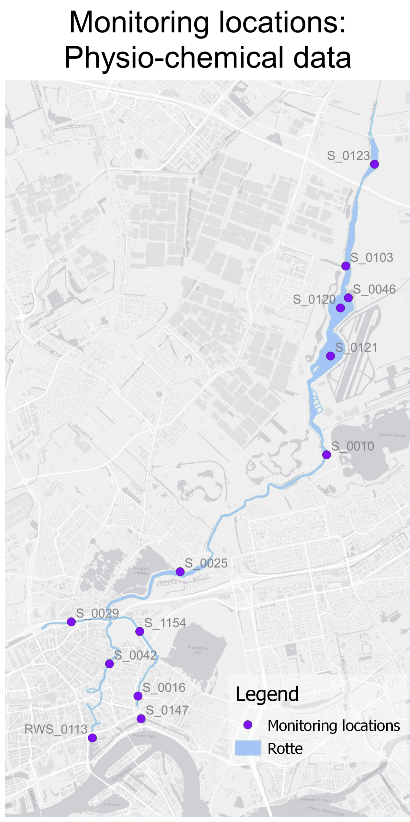


Figure A.1: Monitoring locations for physio-chemical data.

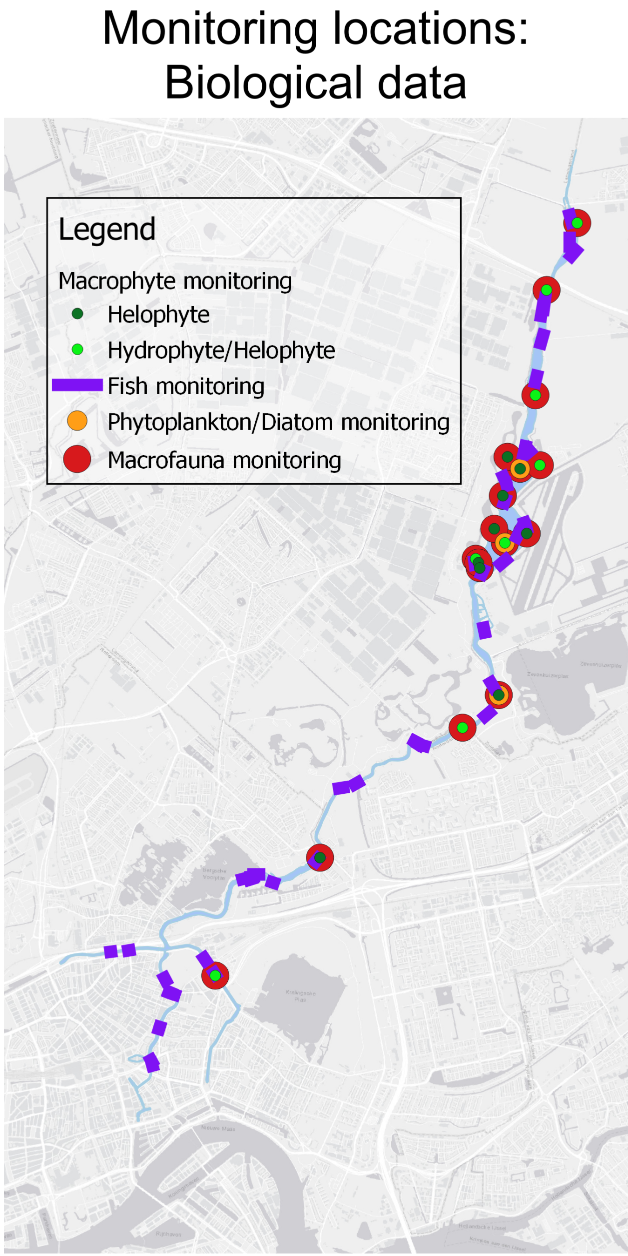


Figure A.2: Monitoring locations for biological data.

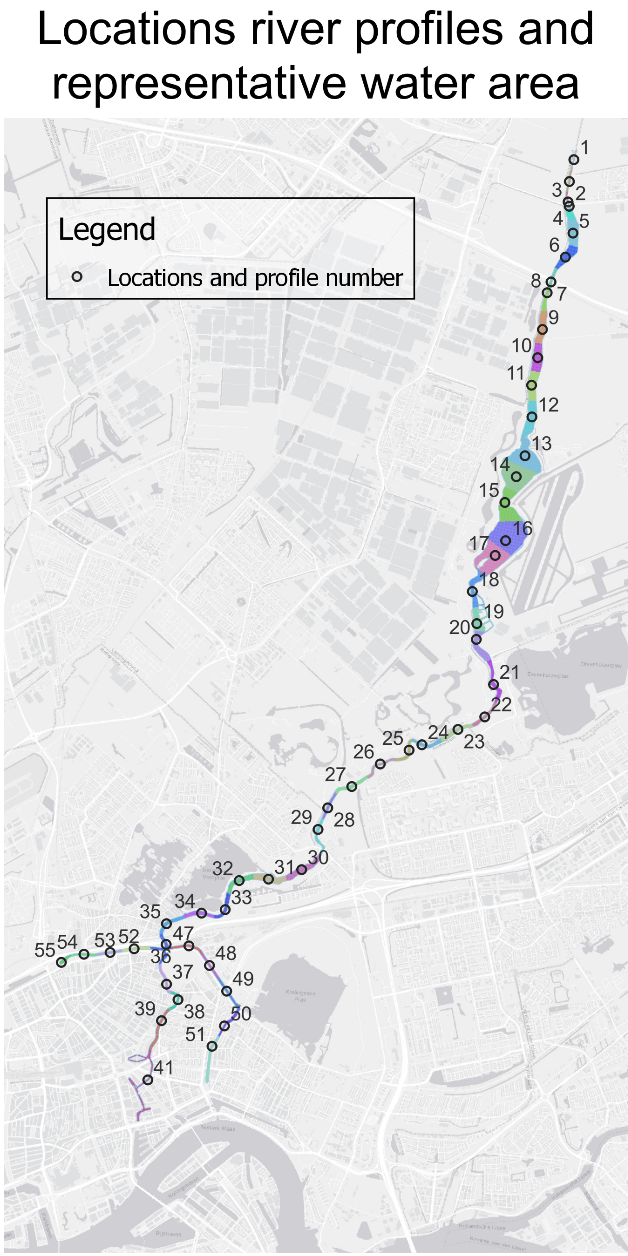


Figure A.3: Locations for bed profiles in the Rotte basin.

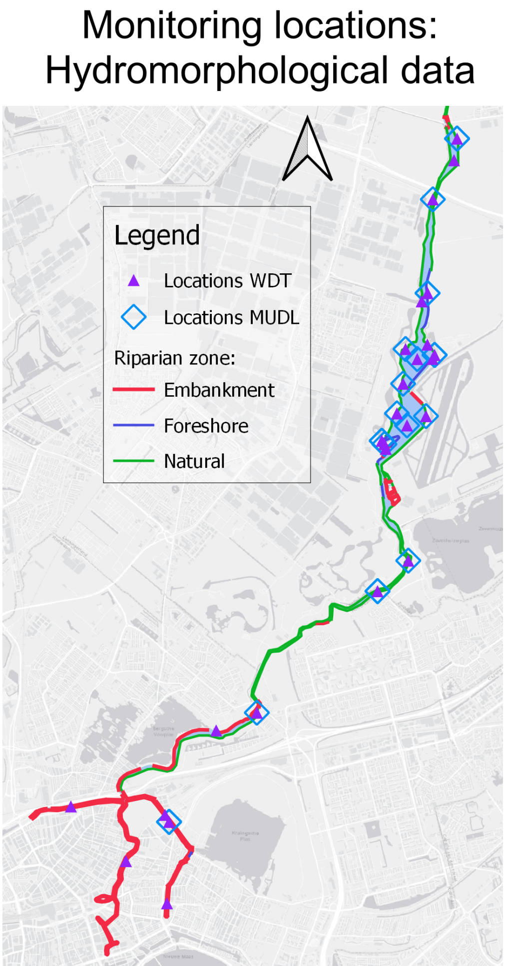


Figure A.4: Monitoring locations for hydromorphological data WDT-ratio, MUDL and Riparian zone.

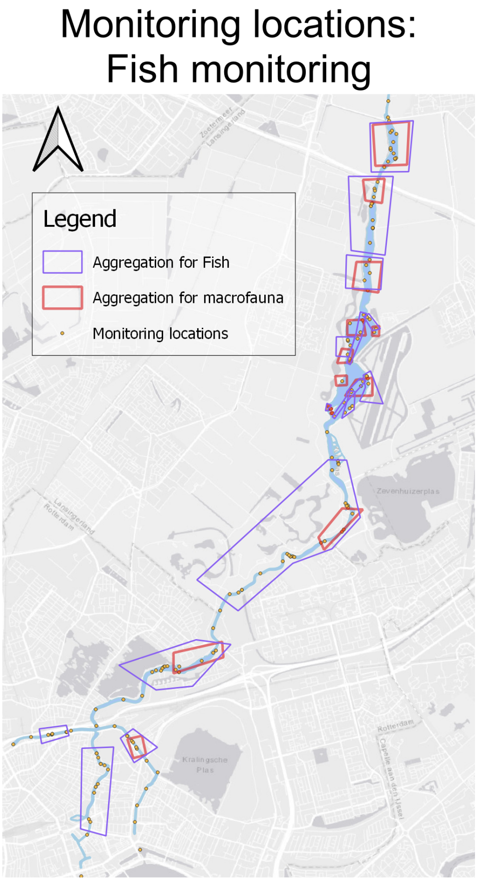


Figure A.5: Spatial aggregation scale for macrofauna and fish groups.

Appendix B

Bed profiles of the Rotte basin

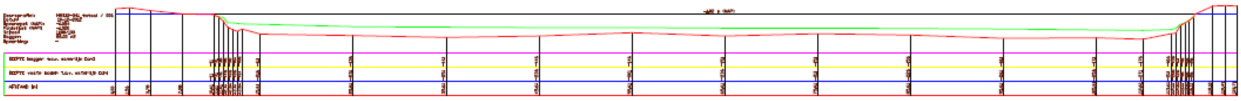


Figure B.1: Profile 6 - Lake-shaped water body with natural shores in section Moerkapelle

Figure B.2: Profile 15 - Lake-shaped water body with natural shores in section Rottemeren

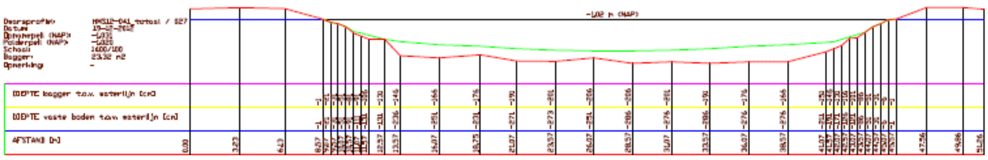


Figure B.3: Profile 27 - Line-shaped water body with natural shores in section Rotte

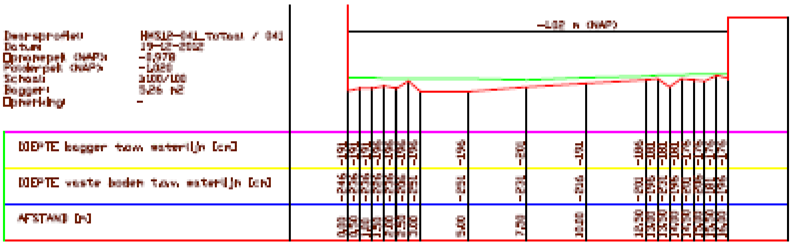


Figure B.4: Profile 41 - Line-shaped water body with vertical shores in section Binnenrotte

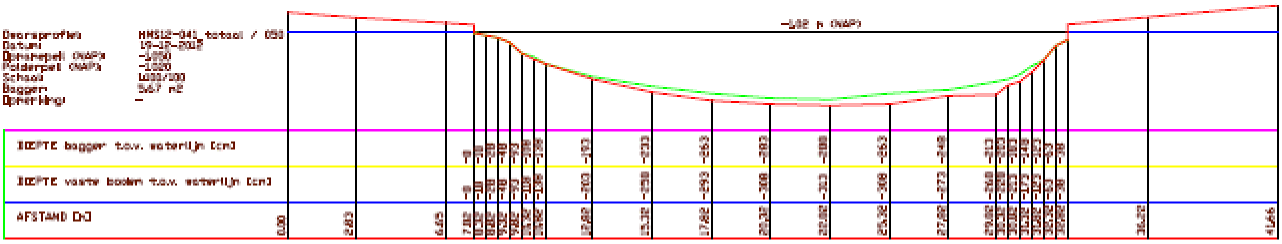


Figure B.5: Profile 50 - Line-shaped water body with vertical shores and flood plains in section Boezem

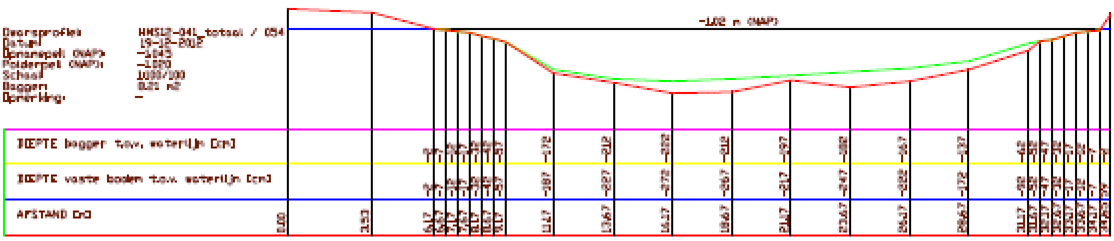


Figure B.6: Profile 54 - Line-shaped water body with shallow littoral zone in section Noorderkanaal

Appendix C

Fish population in the Rotte basin and connected waters

[]

ENG	DUTCH	SCIENT	FLOWGUILD	ECOL	HAB	SUBS	FOOD	SPAWN	FGGroup	AvgRB	AvgBPL	AvgKP	AvgSB
Eel	Aal	Anguilla anguilla	Eurytope	ca	b	f	p,i,j,f	Op	Other	18.3	6.9	2.8	25.3
Bleak	Alver	Alburnus alburnus	Eurytope	fw	p	-	i,j,f	Ov	Other	-	-	-	3.3
Perch	Baars	Perca fluviatilis	Eurytope	fw	p	-	p,i,f	Ov	Other	505.7	4807.9	3230.6	947.0
Barbel	Barbeel	Barbus barbus	Rheophilic	fw	d	s	i,j	Ob	Other	-	-	-	-
Bitterling	Bittervoorn	Rhodeus sericeus	Limnophilic	fw	b	m,v	i,v	Os	FG1	13.3	-	-	0.3
Roach	Blankvoorn	Leuciscus rutilus	Eurytope	fw	p	-	p,i,j,v	Ov	FG1	663.7	1635.7	1669.2	343.3
Topmouth Gudgeon	Blauwband	Pseudorasbora parva	Eurytope	fw	d	-	p,i	?	Other	-	-	-	-
Flounder	Bot	Platichthys flesus	Limnophilic	ca	b	f	i,f	Op	FG2	-	-	-	0.0
Common Goby	Brakwatergrondel	Pomatoschistus microps	Eurytope	er	b	s	i	Ob	Other	-	-	-	-
Bream	Brasem	Abramis brama	Eurytope	fw	d	m,v	p,i	Ov	FG2	3807.0	3938.0	1862.8	1005.3
Sand Goby	Dikkopje	Pomatoschistus minutus	-	er	b	s	i	Ob	Other	-	-	-	-
Three-spinded Stickleback	Driedoornige stekelbaars	Gasterosteus aculeatus	Eurytope	ca	p	-	i,f	Og	FG2	-	-	1.1	11.0
Gibel Carp	Giebel	Carassius auratus gibelio	Eurytope	fw	d	m,v	o	Ov	FG1	-	-	-	2.3
Golden-grey Mullet	Goudharder	Liza aurata	-	ms	p	-	p,i,j,v	Op	FG1	-	-	-	-
Goby spec.	Grondel spec.	Gobiidae spec.	-	er	-	-	-	-	Other	-	-	-	-
Mullet spec.	Harder spec	Mugilidae spec.	-	ms	-	-	-	-	Other	-	-	-	-
Atlantic Herring	Haring	Clupea harengus	-	mj	p	-	i,f	Ob	Other	-	-	-	-
Hybrid	Hybride	Hybrid	-	-	-	-	-	-	Other	11.7	9.8	0.6	1.7
Carp	Karper	Cyprinus carpio	Eurytope	fw	d	m,v	o	Ov	FG2	1.0	52.9	0.0	0.3
Spined Loach	Kleine modderkruiper	Cobitis taenia	Eurytope	fw	-	s	p,j,i	-	FG1	1.0	1.3	-	3.0
White Bream	Kolblei	Blicca bjoerkna	Eurytope	fw	p	-	p,i,v	Ov	FG1	203.0	0.2	-	53.0
Tubenosed Goby	Marmergrondel	Proterorhinus marmoratus	Eurytope	er	-	-	p	-	Other	64.7	54.8	30.8	2.0
Monkey Goby	Pontische stroomgrondel	Neogobius fluviatilis	Eurytope	er	-	f	i,j	-	FG2	-	-	-	-
Ruffe	Pos	Gymnocephalus cernuus	Eurytope	fw	d	f	i,j,v	Ov	FG1	870.7	3448.4	29.9	127.0
Rudd	Rietvoorn	Scardinius erythrophthalmus	Limnophilic	fw	p	-	i,p,v	Ov	FG1	89.3	47.2	139.7	14.0
Bullhead	Rivierdonderpad	Cottus gobio	Rheophilic	fw	b	r	i,f	Og	Other	1.0	2.0	-	1.3
Gudgeon	Riviergrondel	Gobio gobio	Rheophilic	fw	d	s	i	Ov	Other	2.5	-	-	0.3
Asp	Roofblei	Aspius aspius	Eurytope	fw	p	-	i,j,f	Ob	Other	1.3	-	0.0	1.0
Nase	Sneep	Chondrostoma nasus	Rheophilic	fw	d	r	v	Ob	FG1	-	-	-	-
Pike	Snoek	Esox lucius	Eurytope	fw	d	m,v	i,f	Ov	FG1	8.7	29.4	3.4	5.0
Pikeperch	Snoekbaars	Sander lucioperca	Eurytope	fw	d	r	i,f	Ob	Other	168.7	27.5	6.8	30.3
European Sprat	Sprot	Sprattus sprattus	-	ms	p	-	p	Op	Other	-	-	-	-
Ten-spinded Stickleback	Tiendoornige stekelbaars	Pungitius pungitius	Limnophilic	fw	d	f	i	Og	FG1	-	-	-	0.7
Sunbleak	Vetje	Leucaspius delineatus	Limnophilic	fw	-	-	p,i	-	FG1	-	-	-	0.3
Ide	Winde	Leuciscus idus	Rheophilic	fw	p	-	i	Ov	Other	3.0	-	3.8	11.3
White-finned Gudgeon	Witvingrondel	Romanogobio belingii	Rheophilic	fw	?	s	i	?	Other	-	-	-	-
Small Sandeel	Zandspiering	Ammodytes tobianus	-	er	b	s	p	Ob	Other	-	-	-	-
European Seabass	Zeebaars	Dicentrarchus labrax	-	mj	d	m	i,f	Op	Other	-	-	-	-
Tench	Zeelt	Tinca tinca	Limnophilic	fw	p	-	i	Ov	FG1	10.3	13.2	0.4	9.3
Round Goby	Zwartbekgrondel	Neogobius melanostomus	Eurytope	er	-	s,r	i,j,f	-	Other	1.0	-	15.7	150.7

Table C.1: Fish species, guilds and average abundance (n/ha) for the Rotte basin (RB), Bergse Plas (BPL), Kralingse Plas (KP), Schie Boezem (SB) and Nieuwe Maas (NM). Ecological guilds (ECOL): fw = freshwater, ca = catadromous, er = estuarian residents, ms = marine seasonal guest, mj = marine juvenile. Habitat guilds (HAB): b = benthic location (bottom), d = demersal (water column), p = pelagic (water surface). Substrate guilds (SUBS): f = fine sediment (sand, mud, silt), s = sandy sediment, r = hard soil, m = no preference, v = in or above vegetation. Food guilds (FOOD): p = zooplanktivore, i = zoobenthos, j = opportunists, f = piscivore, v = herbivore, d = detrivore, o = omnivore. Spawning guilds (SPAWN): Op = species producing pelagic eggs, Ov = Species that produce adhesive eggs that become attached to substrata and/or vegetation, Ob = Species that produce eggs which settle on the substratum, Os = Species that shed their eggs and then protect them for a period in a part of their body, Og = Species in which one or the other parent guards their eggs externally ([32], [16]). Fish groups (FG Group): FG1 = Fish group 1 for plant-loving fish species with a habitat preference for vegetation, FG2 = Fish group 2 for benthivorous fish with a habitat preference for soft bottoms, Other = Fish group with other preferences for food and habitat.

ENG	DUTCH	SCIENT	FGGroup	BR	NK	B	M	R	RM
Asp	Roofblei	Aspius aspius	Other	-	-	-	3.0	1.0	2.5
Bitterling	Bittervoorn	Rhodeus sericeus	FG1	-	-	-	9.0	6.7	19.9
Bream	Brasem	Abramis brama	FG2	1.0	1.0	2.0	1322.7	7871.0	2331.3
Bullhead	Rivierdonderpad	Cottus gobio	Other	2.0	-	-	0.7	-	1.7
Carp	Karper	Cyprinus carpio	FG2	-	-	-	0.3	2.0	1.0
Eel	Aal	Anguilla anguilla	Other	2.0	3.0	3.7	8.3	12.7	7.0
Gudgeon	Riviergrondel	Gobio gobio	Other	-	-	-	-	5.9	1.0
Hybrid	Hybride	Hybrid	Other	-	-	-	5.6	51.4	2.0
Ide	Winde	Leuciscus idus	Other	-	-	-	0.7	7.0	1.0
Perch	Baars	Perca fluviatilis	Other	7.5	16.1	14.0	196.1	1162.6	110.8
Pike	Snoek	Esox lucius	FG1	1.7	2.0	1.0	8.0	10.3	3.7
Pikeperch	Snoekbaars	Sander lucioperca	Other	-	-	-	56.7	179.5	192.1
Roach	Blankvoorn	Leuciscus rutilus	FG1	-	-	-	291.5	2049.6	308.6
Round Goby	Zwartbekgrondel	Neogobius melanostomus	Other	-	-	1.0	-	-	-
Rudd	Rietvoorn	Scardinius erythrophthalmus	FG1	38.7	4.0	17.0	101.7	80.3	15.8
Ruffe	Pos	Gymnocephalus cernuus	FG1	-	0.5	-	133.1	2436.6	412.5
Spined Loach	Kleine modderkruiper	Cobitis taenia	FG1	2.0	-	-	0.7	-	1.0
Ten-spinded Stickleback	Tiendoornige stekelbaars	Pungitius pungitius	FG1	-	-	-	-	-	3.0
Tench	Zeelt	Tinca tinca	FG1	7.7	3.3	2.5	7.0	6.3	1.0
Three-spinded Stickleback	Driedoornige stekelbaars	Gasterosteus aculeatus	FG2	-	0.5	1.0	0.0	0.3	1.0
Tubenosed Goby	Marmergrondel	Proterorhinus marmoratus	Other	19.0	21.0	23.5	20.3	38.3	28.8
White Bream	Kolblei	Blicca bjoerkna	FG1	-	-	-	94.0	751.3	42.6

Table C.2: Average fish count (n) for the sections in the Rotte basin for monitoring years 2014, 2017 and 2020. The Southern sections are Binnenrotte (BR), Noorderkanaal (NK), Boezem (B). The Northern sections are Moerkapelle (M), Rotte (R) and Rottemeren (RM). Fish groups (FG Group): FG1 = Fish group 1 for plant-loving fish species with a habitat preference for vegetation, FG2 = Fish group 2 for benthivorous fish species with a habitat preference for soft bottoms, Other = Fish group with other preferences for food and habitat.

Appendix D

Participants

Naam	Organisation
Joost Backx	RWS
Marcel van den Berg	RWS
Peter den Boef	Hengelsportvereniging Groot Rotterdam
Marja de Bruyn	PlezierRivier de Rotte
Paul vd Burgt	Roeivereniging Rotterdam
Tom Buyse	WUR/ Deltares
Ryan van der Eijk	HV Groot Rijnmond/TUD
Gerben van Geest	Deltares
Christa Groshart	Voormalig HHSK
Niels Houben	Sportvisserij NL - ZW Ned
Piet Kalkman	Piet Kalkman
Jan Pieter Kalkman	HHSK
Jan Leeuwangh	Hengelsportvereniging Groot Rotterdam
Marit Meier	HHSK
Marjoke Muller	RWS
Panos Panagiotopoulos	WUR
Anne Regtien	HHSK
Jan Roelofs	B-Ware
Tim van Rooijen	HHSK
Joost van den Roovaart	Deltares
Rianne Trompetter	HHDelfland
Erik Venneman	HHSK
Marc Weeber	Deltares
Erwin Winter	WUR

Table D.1: List of participants.

Appendix E

Model variable CPTs

Good	Moderate	HELOPH	HYDROPH	PHYT
0.99	0.01	High dens	High abund	High abund
0.99	0.01	High dens	High abund	Low abund
< 0.01	0.99	High dens	Low abund	High abund
< 0.01	0.99	High dens	Low abund	Low abund
< 0.01	0.99	Low dens	High abund	High abund
< 0.01	0.99	Low dens	High abund	Low abund
0.29	0.71	Low dens	Low abund	High abund
0.17	0.83	Low dens	Low abund	Low abund
0.50	0.50	Abscent	High abund	High abund
0.49	0.51	Abscent	High abund	Low abund
0.99	< 0.01	Abscent	Low abund	High abund
< 0.01	0.99	Abscent	Low abund	Low abund

Table E.1: CPT for FG1

Good	Moderate	SEDI	MAFA
0.1	0.84	Soft	High avail
0.24	0.76	Soft	Low avail
< 0.01	0.99	Sandy	High avail
< 0.01	0.99	Sandy	Low avail

Table E.2: CPT for FG2

High avail	Low avail	PHYT	HYDROPH	HELOPH
0.99	< 0.01	High abund	High abund	High dens
< 0.01	0.99	High abund	High abund	Low dens
0.50	0.50	High abund	High abund	Abscent
0.46	0.54	High abund	Low abund	High dens
0.99	< 0.01	High abund	Low abund	Low dens
0.99	< 0.01	High abund	Low abund	Abscent
0.99	< 0.01	Low abund	High abund	High dens
< 0.01	0.99	Low abund	High abund	Low dens
0.491	0.5	Low abund	High abund	Abscent
< 0.01	0.99	Low abund	Low abund	High dens
< 0.01	0.99	Low abund	Low abund	Low dens
< 0.01	0.99	Low abund	Low abund	Abscent

Table E.3: CPT for MAFA

High abund	Low abund	TS	TEMP
0.26	0.74	Hypertrophic	High
0.16	0.84	Hypertrophic	Low
0.99	< 0.01	Eutrophic	High
< 0.01	0.99	Eutrophic	Low
0.99	< 0.01	Mesotrophic	High
0.99	< 0.01	Mesotrophic	Low
0.5	0.5	Oligotrophic	High
0.5	0.5	Oligotrophic	Low

Table E.4: CPT for PHYT

High abund	Low abund	WDT	MUDL
< 0.01	0.99	Above thres	Above thres
< 0.01	0.99	Above thres	Below thres
< 0.01	0.99	Below thres	Above thres
0.47	0.53	Below thres	Below thres

Table E.5: CPT for HYDROPH

High dens	Low dens	Abscent	MUDL	RIPZONE
< 0.01	0.95	0.05	Above thres	Natural
< 0.01	0.77	0.23	Above thres	Embankment
0.17	0.83	< 0.01	Above thres	Foreshore
0.23	0.77	< 0.01	Below thres	Natural
< 0.01	0.99	< 0.01	Below thres	Embankment
< 0.01	0.99	< 0.01	Below thres	Foreshore

Table E.6: CPT for HELOPH

High	Low
0.51	0.49

Table E.7: PPD for TEMP

Hypertrophic	Eutrophic	Mesotrophic	Oligotrophic
0.68	0.3	0.02	0

Table E.8: PPD for TS

Above thres	Below thres
0.92	0.08

Table E.9: PPD for WDT

Above thres	Below thres
0.93	0.07

Table E.10: PPD for MUDL

Natural	Embankment	Foreshore
0.51	0.44	0.05

Table E.11: PPD for RIPZONE

Soft	Sandy
0.95	0.05

Table E.12: PPD for SEDI

Glossary

catadromous Migration from freshwater to sea for spawning. 6

chlorophyl-a Leaf green to obtain energy from light due to photosynthesis.. 6

confusion matrix A matrix which summarizes the correct and incorrect model predictions by count values for each class. *True positives* and *False negatives* are the classes for correct predictions, whereas *False positives* and *True negatives* are classes for incorrect predictions. Source: https://machinelearningmastery.com/. 24

diadromous Migration between sea to freshwater and vice versa. 2, 6, 11

diatom Specific algal class attached to algal class found in sediments or attached to solid substances.. 8

dispersal Fishes moving to another place for Y. 1

electro fishing Fishing technique using an electric filed to temporarily paralyse nearby fishes.. 8

eurytope Living in both flowing and standing water. 2, 6, 11

foreshore Barrier between shore and main waterway, typically built up of natural material, to create shallow waters near shores and stimulate development of aquatic plants.. 10

guild Group of species that exploits the same kinds of resources in comparable ways and form a functional unit for ecological analyses.. 11

helophyte Perennial littoral plants that are rooted under water and than emerges. It bears its overwintering buds in the mud below the surface. Source: Collins Dictionary. 5, 17

hybrid Offspring of parents that differ in genetically determined traits and different species, genera, or (rarely) families. 11, 12

hydromorphological status Conditions of form and structure of water bodies and substrate and the interaction between them.. 6

hydrophyte Aquatic plants growing in the water. The plants can have submerged, floating and emerged forms of growth or a combination.. 17

ichthyofauna The fish life of a region. Source: Merriam Webster Dictionary. 1

lentic Water system with non-flowing or slow flowing water or stagnant water system such as lakes and canals. 1

limnophilic Living in / preference for standing water. 11

littoral zone Zone in or near the water body and with direct experience of the surrounding water.. 5

lotic Water system with flowing water or fluvial water system such as rivers and transitional waters. 1

macrofauna Animals large enough to be seen by the naked eye, especially invertebrates and insects. Source: Merriam Webster Dictionary. 8

macrophyte Member of plant life, especially aquatic plants, which is observable with the naked eye. Source: Merriam Webster Dictionary. 8

migration Fishes moving to another place for X. 1

natural bank Typical clay levees covered with grass.. 10

phytoplankton Plankton which contain chlorophyl and obtain energy from photosynthesis.. 8

polder Tract of lowland reclaimed from a body of water, often the sea, by the construction of dikes roughly parallel to the shoreline followed by drainage of the area between the dikes and the natural coastline.(Source: Britannica). 4, 5

quail fishing Fishing technique using a dragnet from water top to bottom.. 8

rheophilic Living in / preference for flowing water. 2, 11

riparian zone Zone located on the bank of a natural watercourse (such as a river) or sometimes of a lake or a tidewater. Source: Merrian-Webster. 4

storage canal Canal for drainage and inlet of water for water level management in polders. 2, 4

trait A quality that makes a specie different from another.. 15

vertical embankment Vertical shores typically built up of wood board, concrete or corrugated steel.. 10

Acronyms

AUC Area Under Curve. 24

BBN model Bayesian Belief Network model. 15

CPT Conditional Probability Table. 16

DAG Directed Acyclic Graph. 15
DSS decission support system. 2

ECC Eco-colour Course. 7
EKF Ecological Key Factor. 6
EM Expectation Maximization. 23
ESC Ecological Score Card. 32
EU European Union. 2

FG1 target group for plant-loving fish species. 18
FG2 target group for benthivorous fish species. 18
FHS fish habitat suitability. 2

GAM Generalized Additive Models. 15
GD Gradient Descent. 23
GLM Generalized Linear Models. 15

HELOPH model variable for helophyte. 18
HHDL Hoogheemraadschap Delfland. 11
HHSK Hoogheemraadschap Schieland en Krimpenerwaard. 2, 5–7, 10, 11
HYDROPH model variable for hydrophyte. 18

KPI Key Performance Indicators. 32

MAFA model variable for macrofauna. 17
MUDL model variable for the thickness of the muddy layer. 20

PHYT model variable for phytoplankton. 17
PPD Prior Probability Distribution. 23

RIPZONE model variable for the design of the riparian zone. 21
ROC Receiver Operating Characteristics. 24
RWS Rijkswaterstaat. 11

SEDI model variable for sediment type. 21
SRR Scoring Rule Results. 24

TEMP model variable for temperature. 18
TS model variable for the trophic state. 18

WDT model variable for WDT-ratio. 18
WDT-ratio Water depth-Turbidity ratio. 10
WFD EU Water Framework Directive. 2, 5–9

Bibliography

1. Adriaenssens, V., Goethals, P. L. & De Pauw, N. Fuzzy knowledge-based models for prediction of Asellus and Gammarus in watercourses in Flanders (Belgium). *Ecological Modelling* **195.** Selected Papers from the Third Conference of the International Society for Ecological Informatics (ISEI), August 26–30, 2002, Grottaferrata, Rome, Italy, 3–10. ISSN: 0304-3800. `https://www.sciencedirect.com/science/article/pii/S0304380005005703` (2006).
2. Arntz, J. *KRW Visstandbemonstering 2018 Hoogheemraadschap van Delfland* Dutch (Arcadis, 's Hertogenbosch, 2019).
3. Belletti, B. *et al.* More than one million barriers fragment Europe's rivers. *Nature* **588,** 436–441. ISSN: 1476-4687. `https://doi.org/10.1038/s41586-020-3005-2` (2020).
4. Bijkerk, R. *Handboek Hydrobiologie. Biologisch onderzoek voor de ecologische beoordeling van Nederlandse zoete en brakke oppervlaktewateren.* Dutch (STOWA, Amersfoort, 2014).
5. Brachet, C., Thalmeinerova, D. & Magnier, J. *The handbook for management and restoration of aquatic ecosystems in river and lake basins* 96. ISBN: 978-91-87823-15-2. `https://www.inbo-news.org/en/documents/handbook-management-and-restoration-aquatic-ecosystems-river-and-lake-basins` (Mar. 2015).
6. Brevé, N., Buijse, A. D., Kroes, M. J., Wanningen, H. & Vriese, F. T. Supporting decision-making for improving longitudinal connectivity for diadromous and potamodromous fishes in complex catchments. *Science of the Total Environment* **496,** 206 –208. `https://www.sciencedirect.com/science/article/pii/S0048969714010729` (2014).
7. Chen, S. H. & Pollino, C. A. Good practice in Bayesian network modelling. *Environmental Modelling Software* **37,** 134–145. ISSN: 1364-8152. `https://www.sciencedirect.com/science/article/pii/S1364815212001041` (2012).
8. Commission, T. E. "DIRECTIVE 2000/60/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 October 2000 establishing a framework for Community action in the field of water policy" or, in short, the EU Water Framework Directive. *Official Journal of the European Communities* **L 327,** p. 1–72 (2000).
9. Conallin, J., Boegh, E. & Jensen, J. K. Instream physical habitat modelling types: an analysis as stream hydromorphological modelling tools for EU water resource managers. *International Journal of River Basin Management* **8,** 93–107. eprint: `https://doi.org/10.1080/15715121003715123`. `https://doi.org/10.1080/15715121003715123` (2010).
10. De Leeuw, C. & Backx, J. *Naar een herstel van estuariene gradiënten in Nederland : een literatuurstudie naar de algemene ecologische principes van estuariene gradiënten ten behoeve van herstelmaatregelen langs de Nederlandse kust* (Lelystad, 2000). `https://puc.overheid.nl/rijkswaterstaat/doc/PUC_62358_31/1/`.
11. Directorate-General for Environment. *Biodiversity strategy for 2030 : Barrier removal for river restoration* (Publications Office of the European Union, 2022).
12. Doef, L. & Van Giels, J. *KRW Visstandonderzoek HHSK 2020* Dutch (ATKB, Wageningen, 2021).
13. EEA. *Surface water ecological status* `https://tableau.discomap.eea.europa.eu/t/Wateronline/views/WISE_SOW_Status/SWB_Status_Category?:embed=y&:showAppBanner=false&:showShareOptions=true&:display_count=no&:showVizHome=no`. (accessed: 20.12.2022).
14. EEA. *European Waters. Assessment of Status and Pressures 2018* English. Annual report (European Environment Agency, Luxemburg, 2018).

15. Elbersen, J., Verdonschot, P., Roels, B. & Hartholt, J. *Definitiestudie KaderRichtlijn Water (KRW); 1. Typologie Nederlanse Oppervlaktewateren* Dutch (Alterra, Research Instituut voor de Groene Ruimte, Wageningen, 2003). `https://waterkwaliteitsportaal.overheidsbestanden.nl/factsheets/`.

16. Elliott, M. *et al.* The guild approach to categorizing estuarine fish assemblages: a global review. *Fish and Fisheries* **8,** 241–268. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-2679.2007.00253.x`. `https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-2679.2007.00253.x` (2007).

17. FISRWG. *Stream Corridor restoration: Principles, Processes and Practices* (FISRWG, Oct. 1998).

18. Geist, J. & Hawkins, S. J. Habitat recovery and restoration in aquatic ecosystems: current progress and future challenges. *Aquatic Conservation: Marine and Freshwater Ecosystems* **26,** 942–962. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/aqc.2702`. `https://onlinelibrary.wiley.com/doi/abs/10.1002/aqc.2702` (2016).

19. Grill, G. *et al.* Mapping the world's free-flowing rivers. *Nature* **569,** 215–221. ISSN: 1476-4687. `https://doi.org/10.1038/s41586-019-1111-9` (May 2019).

20. Hamilton, S. H., Pollino, C. A. & Jakeman, A. J. Habitat suitability modelling of rare species using Bayesian networks: Model evaluation under limited data. *Ecological Modelling* **299,** 64–78. ISSN: 0304-3800. `https://www.sciencedirect.com/science/article/pii/S0304380014006103` (2015).

21. HHSK. *Ecologisch maaien en schouwen: de ecokleurenkoers* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2009).

22. HHSK. *Toelichting peilbesluit de Rotte* Dutch, 78. `https://www.schielandendekrimpenerwaard.nl/wat-doen-we/het-hoogheemraadschap-beheert-het-water-de-dijken-de-wegen-in-onze-regio/zorg-voor-voldoende-water/peilbesluiten-schieland-en-de-krimpenerwaard/peilbesluit-rotte/` (May 2012).

23. HHSK. *Nota Vis 2016-2021 (Verlenging Kadernota vis)* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2016).

24. HHSK. *Vis plan 2004 - 2014* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2016).

25. HHSK. *Beleidsuitwerking VIS (Verlenging Nota VIS)* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2022).

26. HHSK. *Factsheet OW39 Hoogheemraadschap van Schieland en de Krimpenerwaard (2022)* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2022). `https://www.waterkwaliteitsportaal.nl/krw-factsheets/`.

27. HHSK. *KRW-plan 2022-2027* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2022).

28. HHSK. *Rapportage gewasbeschermingsmiddelen HHSK* Dutch (Hoogheemraadschap van Schieland en de Krimpenerwaard, Rotterdam, 2022). `https://www.schielandendekrimpenerwaard.nl/kaart/waterkwaliteit/waterkwaliteitsrapportages/GBM_rapportage/`. accessed: 07.04.2023.

29. Hu, F. *et al.* FABM-PCLake - Linking aquatic ecology with hydrodynamics. *Geoscientific Model Development* **9,** 2271–2278 (July 2016).

30. Junier, S. *Modelling expertise: Experts and expertise in the implementation of the Water Framework* (Sept. 2017).

31. Koller, D. & Friedman, N. *Probabilistic Graphical Models: Principles and Techniques* ISBN: 9780262013192 (MIT Press, 2009).

32. Kooiman, M., Ploegaert, S. & Vos, M. *Een zegen in de Delta - 2018 - 2020: Een onderzoek naar de krsaamkamerfunctie van de Zuid-Hollandse Delta* Dutch (RAVON, Nijmegen, 2022).

33. Kornis, M. S. *et al.* Fish community dynamics following dam removal in a fragmented agricultural stream. *Aquatic Sciences* **77,** 465–480. ISSN: 1420-9055. `https://doi.org/10.1007/s00027-014-0391-2` (2015).

34. Kroon, J. & Van Wijk, B. *KRW Visstandonderzoek 2021: Bemonstering vier waterlichamen Hoogheemraadschap van Delfland* Dutch (ATKB, Wageningen, 2021).

35. Lek, S. & Guégan, J. Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling* **120,** 65–73. ISSN: 0304-3800. https://www.sciencedirect.com/science/article/pii/S0304380099000927 (1999).

36. Lin, Y.-P., Lin, W.-C. & Wu, W.-Y. Uncertainty in Various Habitat Suitability Models and Its Impact on Habitat Suitability Estimates for Fish. *Water* **7,** 4088–4107 (July 2015).

37. Marcot, B., Steventon, J., Sutherland, G. & Mccann, R. Guidelines for Developing and Updating Bayesian Belief Networks Applied to Ecological Modeling and Conservation. *Canadian Journal of Forest Research* **36** (Dec. 2006).

38. Marcot, B. G. Metrics for evaluating performance and uncertainty of Bayesian network models. *Ecological Modelling* **230,** 50–62. ISSN: 0304-3800. https://www.sciencedirect.com/science/article/pii/S0304380012000245 (2012).

39. McGarvey, D. J. *et al.* Do fishes enjoy the view? A MaxEnt assessment of fish habitat suitability within scenic rivers. *Biological Conservation* **263,** 109357. ISSN: 0006-3207. https://www.sciencedirect.com/science/article/pii/S0006320721004092 (2021).

40. McLaughlin, R. L. *et al.* Unintended consequences and trade-offs of fish passage. *Fish and Fisheries* **14,** 580–604. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/faf.12003. https://onlinelibrary.wiley.com/doi/abs/10.1111/faf.12003 (2013).

41. Mellor, H., Verbeek, S. & Van de Wijngaart, T. *Ecological key factors* English (STOWA, Amersfoort, 2017). https://www.stowa.nl/onderwerpen/waterkwaliteit/ecologische-krw-doelen.

42. Ministry of Infrastructure and Water Management. *National fishroute map* https://storymaps.arcgis.com/stories/784f89c209bb4362b6453e6ad8f733be. (accessed: 09.09.2022).

43. Molen, D. *et al. Referenties en maatlatten voor natuurlijke watertypen voor de kaderrichtlijn water 2021-2027* (Jan. 2019).

44. Mouton, A. M. *et al.* Optimisation of a fuzzy physical habitat model for spawning European grayling (Thymallus thymallus L.) in the Aare river (Thun, Switzerland). *Ecological Modelling* **215.** Selected Papers from the International Conference on Ecological Modelling, 28 August – 1 September 2006, Yamaguchi, Japan, 122–132. ISSN: 0304-3800. https://www.sciencedirect.com/science/article/pii/S0304380008000859 (2008).

45. NETICA Corp. *NETICA website Sensitivity Equations* https://www.norsys.com/. (accessed: 17.06.2023).

46. Niemeijer, B., Mies, J., & Van Giels, J. *KRW Visstandonderzoek HHSK 2017* Dutch (ATKB, Wageningen, 2018).

47. Puts, T. & Kruitwagen, G. *KRW Visstandonderzoek 2015: Polder Berkel en Oostboezem* Dutch (Witteveen + Bos, Deventer, 2016).

48. Russel, S. & Norvig, P. *Artificial Intelligence: A Modern Approach* Fourth. ISBN: 9781292024202 (Prentice Hall, 2009).

49. RWS. *Protocol monitoring en toestandsbeoordeling oppervlaktewaterlichamen KRW* Dutch (RWS, The Hague, 2020).

50. Scheffer, M., Hosper, S., Meijer, M.-L., Moss, B. & Jeppesen, E. Alternative equilibria in shallow lakes. *Trends in Ecology Evolution* **8,** 275–279. ISSN: 0169-5347. https://www.sciencedirect.com/science/article/pii/016953479390254M (1993).

51. Silva, A. T. *et al.* The future of fish passage science, engineering, and practice. *Fish and Fisheries* **19,** 340–362. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/faf.12258. https://onlinelibrary.wiley.com/doi/abs/10.1111/faf.12258 (2018).

52. Smith, V., Tilman, G. & Nekola, J. Eutrophication: impacts of excess nutrient inputs on freshwater, marine, and terrestrial ecosystems. *Environmental Pollution* **100,** 179–196. ISSN: 0269-7491. https://www.sciencedirect.com/science/article/pii/S0269749199000913 (1999).

53. Stoffers, T. *et al.* Freshwater fish biodiversity restoration in floodplain rivers requires connectivity and habitat heterogeneity at multiple spatial scales. *Science of The Total Environment* **838,** 156509. ISSN: 0048-9697. https://www.sciencedirect.com/science/article/pii/S0048969722036063 (2022).

54. Teurlincx, S. *et al. Linking ESF - Het begrijpen van de samenhang tussen de ecologische sleutelfacotoren voor stilstaande wateren.* Dutch (STOWA, Amersfoort, 2018).

55. The European Commission. COUNCIL REGULATION (EC) No 1100/2007 of 18 September 2007 establishing measures for the recovery of the stock of European eel. *Official Journal of the European Communities* **L 248/17,** p. 17–23 (2007).

56. Van Gerven, L. & Evers, N. *KRW-Watersysteemanalyse Boezem Schieland* Dutch (RHDHV, Amersfoort, 2017).

57. Van Gerven, L. & Holstein, A. *Systeemanalyse Schieland: Effect van nutriëntenbelasting op de ecologie van sloten* Dutch (RHDHV, Amersfoort, 2017). `https://www.waterkwaliteitsportaal.nl/sgbp-achtergronddocumenten`.

58. Özesmi, S. L. & Özesmi, U. An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling* **116,** 15–31. ISSN: 0304-3800. `https://www.sciencedirect.com/science/article/pii/S0304380098001495` (1999).