International evaluation of the Danish marine models

Performed by the Panel of international experts

10. oktober 2017

Implement Consulting Group
Strandvejen 54
2900 Hellerup

Tel +45 4586 7900
Email info@implement.dk
Implementconsultinggroup.com

CVR 32767788
Bank 4845-3450018236
SWIFT DABADKKK
Iban DK3030003450018236

Table of contents

1. Introduction

This report presents a scientific review of the Danish management approach regarding coastal waters in relation to the implementation of the European Water Framework Directive (WFD) in Denmark. The parties to the Agreement on Food and Agriculture Package (22 December 2015) have decided to evaluate the modelling tools (pressure-impact models) used to calculate the mitigation demands for nitrogen (N) runoff from land in the Danish River Basin Management Plans. The results of the evaluation will be utilised towards the development and application of models in the 3rd generation water plans valid for 2021-2027.

Task description by the Ministry of Food and Agriculture

In agreement with the EU Water Framework Directive, Denmark has produced the River Basin Management Plans devising a strategy for improving and securing that coastal waters, lakes, streams and ground waters fulfil the demand for Good Ecological Status as stated in the directive. For Danish coastal waters, it has been estimated that reductions in N runoff from land are the primary concern if goals of Good Ecological Status in coastal waters are to be fulfilled. On this background, mitigation measures have been implemented in the 2015-2021 River Basin Management Plans to additionally reduce the N runoff to coastal waters, corresponding to roughly half the total estimated reduction needs.

The task of the evaluation panel is to perform a thorough evaluation of the marine modelling tools that form the basis for the mitigation demands for land-based nitrogen (N) runoff in the Danish River Basin Management Plans with regards to the importance of N as well as other relevant pressures such as phosphorous, fisheries etc. In particular, the evaluation panel has to:

i. Evaluate the use of models for determination of type-specific reference values (according to the Water Framework Directive, Annex 2) for the water quality element phytoplankton (chlorophyll).

ii. Evaluate the use of models to determine environmental targets (Maximum Allowable Inputs (MAI) of nitrogen)) and mitigation needs to achieve good environmental status and evaluate differences and similarities between the use of different methods and model types for coastal waters with different typology.

iii. Evaluate the estimated nitrogen target loads and mitigation needs in the Danish River Basin Management Plans and evaluate the method for determining the Danish proportion of total mitigation needs. How is the current environmental status in Danish coastal waters determined by N runoff from Danish land areas in relation to other pressures such as N released from sediments and N loads from catchments in neighbouring countries and airborne N deposition (the Danish share of the total mitigation needs related N)?

Further, the Panel is expected to address the technical questions and comments from the stakeholders.

Recruitment of experts

The Danish Ministry of Environment and Food has been responsible for the recruitment of an international panel of five experts to carry out the evaluation. The recruitment of experts has been conducted by a nomination process where the Danish Ministry of Environment and Food has requested water management authorities in other countries (Sweden, Finland, Poland, Germany, The Netherlands and England) and the European

Environment Agency, Joint Research Centre (JRC) and the European Commission (DG Environment) to nominate experts to conduct the evaluation. It has been stated in the request that the nominees should have expert knowledge in the following areas: marine ecology, marine ecosystem models, statistical methods and experience in marine water management in relation to the Water Framework Directive.

The request by the Ministry resulted in the nomination of 14 experts of which 9 experts subsequently indicated that they were interested in being part of an expert panel. Of these, the Ministry has selected the following five experts to conduct the evaluation:

- Professor Peter Herman, Deltares, Institute for applied research in the field of water and subsurface, the Netherlands.
- Professor Alice Newton, NILU – Norwegian Institute for Air Research
- Professor Gerald Schernewski, Leibniz Institute for Baltic Sea Research, Warnemunde
- Director Bo Gustafsson, Baltic Nest Institute (BNI), Stockholm University, Sweden
- Senior Researcher Olli Malve, Finnish Environment Institute SYKE
- Professor Peter Herman was chosen as chairman of the Panel

The five experts were chosen according to an assessment of their qualifications with regards to experience with and competences in the following fields of study: *marine ecology/coastal ecology*, c*oastal ecosystem modelling*, *use of statistics in environmental science* and *marine management experience related to the implementation of the Water Framework Directive*.

1.1 Aim and focus of the evaluation

This section presents the aim and focus of the evaluation according to the international panel (hereafter referred to as the Panel) and should therefore be seen as the Panel's further operationalisation of the task description in section 1.1.

The main aim of the evaluation

The main aim of the evaluation is to review whether the marine models – as presented in the Scientific Documentation Report and as commented by the researches and stakeholders – *provide solid and robust scientific evidence that the proposed reductions in land-based N runoff will be both necessary and sufficient to reach Good Ecological Status as defined in the Water Framework Directive*.

- By "solid", the Panel means well based in international scientific literature, well performed, credible
- By "robust", the Panel means not unduly dependent on arbitrary details, reliable with acceptable precision
- By "necessary", the Panel means that by doing less the goals would not be reached
- By "sufficient", the Panel means that by executing the plans, there is a high probability of reaching the goals

The evaluation concerns the modelling tools (pressure/impact models) forming the basis for the mitigation demands for land-based nitrogen (N) runoff in the Danish River Basin Management Plans. The evaluation results will enter into the calculation of N mitigation demands for coastal marine areas in the 3rd generation water plans valid in 2021-2027.

The evaluation will answer questions related to points (i)-(iii) in the task description above and is therefore focused on the scientific underpinning of the plans, in particular the modelling tools. The evaluation must take into account the internationally agreed goals of achieving Good Ecological Status in the Water Framework Directive. One that basis, the Panel has defined the aim and focus of the evaluation as stated in the Text Box shown above.

The scope of the evaluation does not include other models than the marine model and other environmental targets than those applying to coastal areas. The scope of the evaluation does not include the societal costs and benefits of the measures that would be needed to fulfil the environmental targets.

1.2 The basis and process of the evaluation

The basis for the evaluation

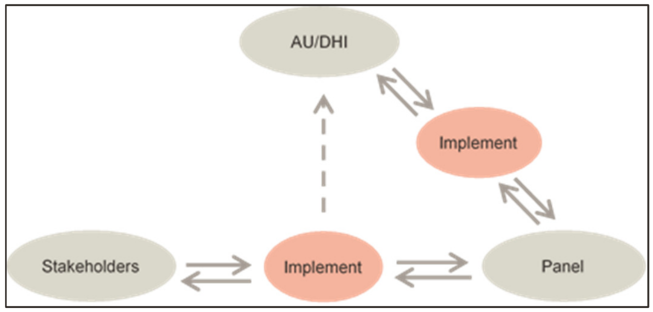
The basis on which the Panel has made the final evaluation consists of the following materials:

- The Scientific Documentation Report written by Aarhus University (DCE) and DHI in June 2017, which documents the model tools and calculated MAI that were developed for the Ministry over the period 2013-2015.
- Questions and comments from the stakeholders to the Scientific Documentation Report (see Annex 1 of the evaluation report)
- Answers from the researchers to questions and comments which were formulated by the Panel after the members of the Panel read and considered the report as well as the questions and comments from the stakeholders (see Annex 2a and 2b of the evaluation report).
- Answers from the Panel on how they took into account each of the technical questions and comments from the stakeholders (see Annex 3 of the evaluation report).
- Selected background materials cited by the researchers, the stakeholders and the Panel

The means for ensuring independence during the process

It is considered crucial that the evaluation of the Danish marine models be performed by independent scientists. In order to guarantee independence, it was decided that the Ministry of Environment and Food, the scientists from AU and DHI and the stakeholders should keep arm’s length to the Panel throughout the process of the evaluation. Implement Consulting Group (Implement) was engaged by the Ministry to facilitate the process.

Figure 1. Communication model



As illustrated above, the communication model was designed to facilitate a dialogue that ensured arm's length between the involved parties and to promote a transparent flow of communication. Implement has been the link between the Panel, the stakeholders and the scientists. Besides facilitating the final writing workshop and the preparations leading up to that, the main role of Implement has therefore been to ensure timely communication and convey relevant material and information between the parties.

The evaluation process

The evaluation process started in June 2017. It resulted in an evaluation report on 19 September, which was finalised after a writing workshop in Helsingør which took place between 11-15 September. After the hearing process between 19. September and 2 October some minor corrections were made to the final report which was completed on 10 October.

The text and the activity plan below provide a more detailed overview of the evaluation process.

Initially, the stakeholders from Blåt Fremdriftsforum, the scientists from AU and DHI and the Panel were invited to participate in separate meetings where Implement explained the process of the evaluation. At the meetings, the activity plan and a communication model were presented to make sure that all parties were properly informed about the practical aspects, important deadlines and rules of communication. The process leading up to the final evaluation workshop was thereafter as follows with respect to each of the parties:

- The stakeholders received the Scientific Documentation Report written by the scientists from AU and DHI on 6 June and had until 4 July to formulate questions and comments to the report. The comments and questions had to be submitted in a table – made specifically for that purpose – that followed the structure of the report. A "hotline" for questions regarding the practical aspects of formulating and submitting the questions and comments was established by Implement. The stakeholders submitted their questions and comments on 4 July, and they were all forwarded by Implement to the Panel on 6 July. In order to sum up their main points of view in front of the Panel, the stakeholders were invited to participate in a physical meeting with the Panel hosted by the Ministry in Helsingør on 11 September during the writing workshop.
- The Panel received the Scientific Documentation Report at the same time as the stakeholders – on 6 June. Implement held a couple of status meetings with the Panel in June and the beginning of July when the comments and questions from the stakeholders were forwarded to the Panel on 6 July. Based on the reading of the Scientific Documentation Report and the questions and comments from the stakeholders, the Panel has jointly formulated questions to

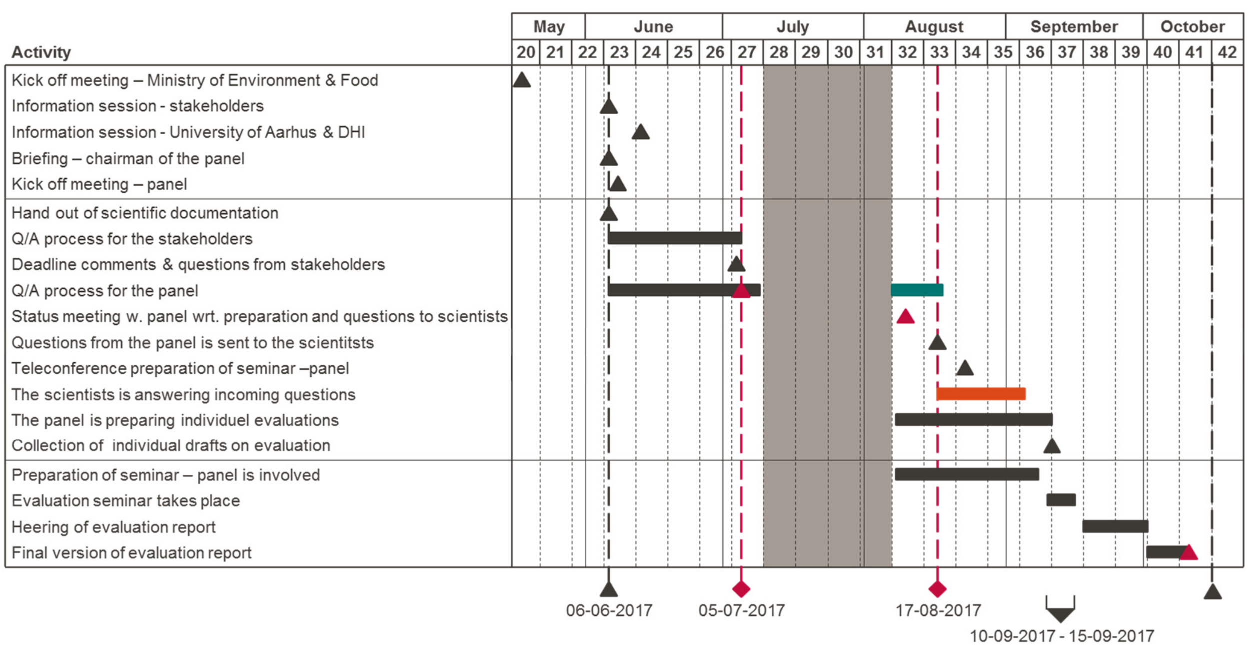
the Danish scientists that were forwarded by Implement on 15 August. The scientists from AU and DHI answered these questions on 4 September in order for the Panel to take it into account in the evaluation. Throughout August, Implement held some status meetings with the Panel to monitor the progression and to prepare the writing workshop in Helsingør.

- The scientists from AU and DHI worked out the Scientific Documentation Report that was distributed by Implement to the stakeholders and the Panel. The scientists received the comments and questions by the stakeholders on 6 July as an orientation. As stated above, the scientists received questions from the Panel and replied to these on 4 September. During the writing workshop from 11-15 September, the scientists have answered a limited amount of additional questions from the Panel.
- The Ministry of Environment and Food was not directly involved in the evaluation process due to the principle of arm's length. Implement has occasionally informed the Ministry about the progress in the evaluation, and the parties had a dialogue about the practicalities of the writing seminar in Helsingør. Representatives from the Ministry were present at the meeting with the stakeholders and the Panel in Helsingør on 11 September.

The writing of the evaluation report was thereafter carried out by the Panel and facilitated by Implement in a final writing workshop in Helsingør between 11-15 September. The evaluation report was edited and submitted for hearing on 19 September.

The hearing of the evaluation report among stakeholders from Blåt Fremdriftsforum and the scientists from AU and DHI took place between 19 September and 2 October.

Figure 2. Activity plan of the evaluation process



After the hearing process, the evaluation report will be published by the Ministry of Environment and Food along with annexes containing the hearing comments and answers by the Panel. The activity plan above illustrates the entire process of the evaluation.

1.3 Content and structure of the evaluation report

The evaluation report is divided into a number of themes which the evaluation panel found to be the most important in order to cover the topics in the terms of reference and pursue the aim of the evaluation. According to the Panel, the main themes are those covered in Chapters 2-9 in the evaluation report which has the following structure:

- Introduction (Chapter 1)
- Compliance with the Water Framework Directive (Chapter 2)
- Coastal water typology (Chapter 3)
- The use of seagrass and Kd as environmental indicators (Chapter 4)
- Emphasis on nitrogen versus phosphorus (Chapter 5)
- Statistical modelling (Chapter 6)
- Mechanistic modelling (Chapter 7)
- Calculation procedures to estimate Maximum Allowable Inputs from model results (Chapter 8)
- Evaluation of Maximum Allowable Input results (Chapter 9)
- Overall assessment and conclusions (Chapter 10)
- Recommendations for going further (chapter 11)

By going through the most important themes and discussing the main problems within each theme, the review by the Panel focuses on whether these problems have been adequately solved in the Scientific Documentation Report – rather than going through the details in the report chapter by chapter.

This means that the review text by the Panel mainly concentrates on investigating possible weaknesses in the overall modelling approach followed by the researchers from Aarhus University (DCE) and DHI. However, the review contains conclusions with respect to both the strengths and weaknesses of the approach, and critical remarks should be viewed in the context of the overall assessment as presented in Chapter 10.

After the thematic chapters, the evaluation contains an overall assessment of the marine modelling approach and report. The final assessment will provide an answer to the central question on whether the modelling approach and report provides solid and robust scientific evidence that the proposed reductions in land-based N runoff will be both necessary and sufficient to reach Good Ecological Status as defined in the Water Framework Directive. Moreover, the assessment will answer other related questions to cover the terms of reference.

Finally, recommendations are given as to how the Danish marine models might be improved in the future. Focus is on improvement that can be made within a reasonable time frame and without investing excessive resources.

1.4 Future development following the evaluation

Once the researchers have made the adjustments to the modelling, they are encouraged to publish their work in peer-reviewed journals to showcase Danish leadership in this field.

The work of the researchers has been performed over decades and several administrations. This “organic” process has given rise to numerous interactions between the scientists and authorities. In order to avoid confusion and misunderstandings, terms of reference, the scope of the missions set by the Ministry and agreements on choices, e.g. indicators to be used, should be well-defined. This can be important for the further political process, but has not been subject to examination by the Panel.

The Panel hopes that the attention given to the views of the stakeholders and the responses of the researchers during the scientific scrutiny of the Scientific Documentation Report will help to build trust between the parties and contribute to a successful outcome.

2. Compliance with the Water Framework Directive

This chapter examines whether the Scientific Documentation Report complies with Directive 2000/60/EC, commonly referred to as the Water Framework Directive (WFD). It also addresses some of the concerns and questions of the stakeholders. In this chapter, we focus on the following general questions relating to the compliance with the Water Framework Directive:

- Is the procedure for setting the type-specific reference condition WFD compliant?
- Is the choice of indicators selected WFD compliant?
- Have the indicators been intercalibrated?
- Has the “one-out, all-out” principle been respected?

The questions are the basis for the subsections of the chapter. In addition, the chapter devotes attention to the question whether all relevant stressors have sufficiently been taken into account.

The community policy on water was adopted by the European Parliament and Council on 23 October 2000 as an integrated Community Directive 2000/60/EC, commonly referred to as the Water Framework Directive (WFD). It was published in the Official Journal (OJ L 327) on 22 December 2000 and was also adopted by member states (MS). In Denmark, it was adopted as national legislation in 2003.

Article 1 states: “The purpose of this Directive is to establish a framework for the protection of inland surface waters, transitional waters, coastal waters and groundwater”. Preamble 26 of the WFD states that “member states should aim to achieve the objective of at least good water status by defining and implementing the necessary measures within integrated programmes of measures, taking into account existing community requirements”. Article 4 introduces the concept of the River Basin Management Plans (RBMP) as fundamental to “making operational the programmes of measures”, and these are detailed in article 13. RBMP are a single system of water management by river basin, which are the natural geographical and hydrological units, instead of according to administrative or political boundaries.

The report “Development of models and methods to support the establishment of the Danish River Management Plans”, which we refer to as the Scientific Documentation Report, contributes to the implementation of the WFD to maintain or achieve Good Ecological Status in Danish Coastal Waters (CW). Therefore, an evaluation of the WFD compliance of the methodology and results is valuable and important.

2.1 Reference conditions and boundary setting

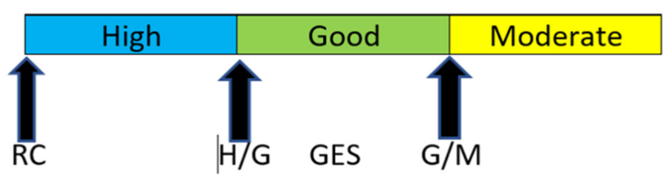
Annex 2 of the WFD (section 1.1) addresses the characterisation of surface water body types. First, the water bodies must be placed in one of the surface water categories: rivers, lakes, transitional waters or coastal waters. Another possible category is artificial and heavily modified bodies of water.

The WFD then specifies that the water bodies are to be differentiated according to their type. Denmark merged transitional waters with coastal waters, therefore a Fish BQE is not necessary. Denmark has 119 marine water bodies[1]. These are categorised into six open water body types and 12 estuarine water body types, all included in coastal waters, according to a report by Dahl et al (2005). In Chapter 3, we discuss the further implications and consequences of this typology. We note that in their answers to the Panel, the researchers state that there is a project proposal on an “update of the typology applied towards the RBMP 2021-2027”.

Annex 2 of the WFD (section 1.3) specifies the procedure for the “establishment of type-specific reference conditions for surface water body types”. Type-specific reference conditions (RC) may be either spatially based or based on modelling or may be derived using a combination of these methods. Where it is not possible to use these methods, member states may use expert judgement to establish such conditions. The Danish approach relies on modelling and a 1900 baseline, since there are no pristine systems that can be used as a reference. This approach is appropriate, WFD compliant and better than only using expert judgement.

Good Ecological Status (GES) falls between the “High/Good” boundary and the “Good-Moderate” status. The relationship between reference condition, the boundaries and GES is shown in Figure 3. The setting of the reference conditions and the boundaries, especially the G/M boundary, is important. This determines whether management measures are necessary. Classification below the G/M boundary requires management measures to be adopted.

Figure 3. Relationship between reference condition, the boundaries and GES



Target values must fall in the green (GES) range.

2.2 Choice of indicators

Annex 5 of the WFD specifies the quality elements (QE) for the classification of ecological status of coastal waters (1.1.4). Good Ecological Status is an assessment based on a combination of biological quality elements (e.g. phytoplankton, other aquatic flora and benthic invertebrates); Hydro-morphological elements (e.g. structure and substrate of the seabed, tidal regime); chemical and physico-chemical elements (e.g. transparency, oxygenation conditions, nutrient conditions).

The indicators chosen in the Danish RBMP report are Chlorophyll a, Kd and a benthic index as well as some secondary indicators in statistical modelling approaches (see Chapter 4 of this evaluation report). Denmark has not adopted transitional waters as a separate category, and therefore there is no need to include a fish biological quality element. Chlorophyll a is a proxy for phytoplankton biomass and has been intercalibrated (see 2.3 below). Kd is a measure of attenuation, hence an indirect measure of growth conditions for benthic plants and algae. Thus, it is not a direct indicator of aquatic flora (eelgrass), but rather a light control on the distribution of eelgrass. Furthermore, Kd is not independent of Chlorophyll a, since phytoplankton cells contribute to light attenuation and a loss in transparency. Kd has not been

[1] Scientific Documentation Report, section 3.1, p. 14, Bek.nr. 837 2016

intercalibrated, although eelgrass depth limit, for which it is a proxy, has (see 2.3 below). The Danish Benthic index addresses the benthic invertebrates' biological quality element. Chlorophyll a was the indicator chosen for the intercalibration of the WFD, which Denmark participated in. The marine models under evaluation only considered indicators for the physico-chemical elements' oxygenation condition and nutrient condition in the statistical modelling.

- Most of the calculations in the modelling are based on only Chlorophyll a and Kd to derive the targets for N
- Furthermore, the choice of Kd as an indicator for submerged aquatic vegetation (eelgrass) may be insufficient (Chapter 4)
- The inclusion of other indicators throughout the process and modelling (oxygenation condition and nutrient limitation) are also discussed in Chapter 4

2.3 Intercalibration

A Common Implementation Strategy (CIS) has been in operation since 2001, bringing together national experts, stakeholders and the Commission involved in the implementation of the WFD. During the process, a series of Guidance Documents and CIS Thematic Information Sheets were produced. These are not legally binding but give mainly technical advice about the implementation process.

The WFD requires the national classifications of Good Ecological Status to be harmonised through an intercalibration exercise, Birk et al (2013). This is to avoid adjacent water bodies being classified in a different way. Intercalibration was carried out by member states that share typologies and transboundary water bodies. In the case of Denmark, the shared water types were NEA 1/26C: NEA 8B and BC 6. These are explained in Table 1.

Table 1. Common typologies intercalibrated for Chlorophyll a with Germany and Sweden

Code	Water type	Shared with	Intercalibration
NEA 1/26C	The North-East Atlantic, enclosed seas, exposed or sheltered, partly stratified	DE	Intercalibrated
NEA 8B	The North-East Atlantic Kattegat coastal waters	SE	Intercalibrated
BC 6	Baltic Coast (SW)	SE	Intercalibrated

As mentioned in 2.2, Chlorophyll a was the indicator chosen for the intercalibration of the WFD, which Denmark participated in. The overall status with respect to intercalibration of the indicators used in the Danish marine models is as follows:

- Chlorophyll a has been successfully intercalibrated with SE and DE
- Kd has not been intercalibrated (as confirmed by the researchers from Aarhus University (DCE) and DHI and the European Commission's Joint Research Centre).
- Eelgrass depth limit has been intercalibrated

2.4 "One out, all out"

The WFD preamble 11 specifies that it is "based on the precautionary principle and on the principles that preventive action should be taken". The "one-out, all-out" principle is a key principle that reflects the WFD integrated approach for the protection of water resources and associated aquatic ecosystems. Quality elements comprised in the definition of ecological status provide a holistic picture of the health of the aquatic environment. The overall status would only be "good" if all the elements comprised are at least considered "good". This ensures that all pressures capable of degrading the water status are addressed and are a guarantee of the environmental integrity of the objectives of the directive.

Progress achieved towards "good" status of water bodies can be reported using indicators at individual quality element level. However, this does not preclude the "one-out, all-out" principle. The WFD will be reviewed by 2019, taking into account the results of the second RBMPs. The proper implementation of the Nitrates Directive, which is a basic measure under the WFD, is necessary for the achievement of the WFD objectives. However, in many cases, this will not be sufficient, and additional measures will have to be taken by member states to ensure that the WFD objectives are reached.

Based on the "one-out, all-out" principle, indicators for different quality elements should be considered individually. If one is classified as below the G/M boundary, then management measures must be applied. This was not applied in the Scientific Documentation Report as confirmed by the researchers in their answers to the panel questions. The different methods of aggregation and their implications in both the WFD and MSFD are discussed in Borja et al (2014). We further discuss the "one-out, all-out" principle in relation to indicators in Chapter 4 and in relation to the calculation procedures in Chapter 8.

2.5 Other stressors on the ecosystem

In his classic paper on eutrophication problems, Cloern (2001) describes how the vision on eutrophication problems has evolved from viewing nutrient enrichment as a single isolated issue, towards a vision that emphasises the interactions between multiple stressors, the physics and hydrography of the systems, and eutrophication. He makes a plea for integrated models and tools that describe how nutrient enrichment modulates the response of ecosystems to other stressors, such as chemical pollution, introduction of invasive species, habitat modifications, fishing pressure and others, in the physical setting of a water body. The different stressors should not be viewed as additive factors with – from a management perspective – the option to choose reduction of any of these stressors to obtain a similar percentage of improvement in the ecosystem response. Restoring physical habitat quality, as an example, will have very little effect if eutrophication leads to oxygen problems, low transparency of the water or low phytoplankton quality due to the interactions of the causal factor "physical structure" with eutrophication. Reversely, remediation of eutrophication problems may not suffice to improve ecological quality, if additional action on other stressors is needed.

In many questions and comments of the stakeholders, reference was made to the report by Andersen et al (2017) that lists many stressors on the marine ecosystem and, using a particular weighting, concludes on an overall percentage of stress due to nutrient loading. One could try and argue that this provides evidence that similar improvements of ecological status could be obtained by working on other stressors than nutrient loading, but in so doing would miss the essential point that the effect of the different stressors is not additive and that the final ecosystem response is modulated by the interaction between the stressors, not their individual additive effect. The Panel endorses the fundamental view on interaction between stressors, and on the key role of nutrient loading and eutrophication in modulating the ecological response of Danish coastal waters, that is expressed in the Scientific Documentation Report and in the models (especially the mechanistic models) underlying the analyses.

This fundamental view on the importance of water quality as the main modulator in promoting Good Ecological Status is fully in line with the Water Framework Directive implementation and with the use of intercalibrated indicators such as Chlorophyll a and measures of chemical pollution as the prime measures of ecological status. Inclusion in the WFD was based on extensive and in-depth reviews of ample scientific evidence.

Other legal instruments, e.g. the Marine Strategy Framework Directive, take a broader view and also include more explicitly other stressors such as invasive species, shipping, fishing and physical modification. The Panel is convinced that these aspects fully merit inclusion in a holistic view on restoring Good Ecological Status but in no way decrease the importance that has to be attached to controlling nutrient loadings as a necessary condition for restoration of Good Ecological Status.

3. Coastal water typology

The Scientific Documentation Report uses a modified Danish coastal water typology as the basis to calculate reference conditions and targets for coastal waters as well as Maximum Allowable Inputs (MAI). The typology is a crucial element for all following steps. Therefore, in this chapter, the Panel evaluates its suitability, analyses its shortcomings and provides suggestions.

3.1 Basic idea behind typology

Annex 2 of the WFD gives instructions on how typology should be carried out and lists the obligatory and optional factors that can be used (see Chapter 2 of this evaluation). Most European Union member states applied the most specific system B. In this approach, the physical and chemical factors that determine the characteristics of coastal and transitional waters are latitude, longitude, tidal range and salinity as obligatory factors. Optional factors are current velocity, wave exposure, mean water temperature, mixing characteristics, turbidity, retention time, mean substratum composition and water temperature range.

The Common Implementation Strategy (see Chapter 2) for the WFD (2000/60/EC): “Guidance Document No 5” on “Transitional and Coastal Waters – Typology, Reference Conditions and Classification Systems” provides a detailed guideline for carrying out a characterisation of all water bodies, referred to as typology. The aim is to produce a simple physical typology that is both ecologically relevant and practical to implement. It aims at linking similar water bodies under one type to enable the establishment of type-specific reference conditions. Guidance Document No 5 suggests ranges for several factors that could be used in the typology.

Once the water has been characterised as transitional (TW) or coastal (CW), a typology for each is developed by the member states. Denmark, like Germany, chose to include estuarine waters within coastal waters because all the parameters, except depth, are the same. Whether a typology separates transitional and coastal waters or combines both under coastal waters does not make a difference for the calculation of reference conditions, targets and Maximum Allowable Inputs (MAI). It does not necessarily affect the number of types nor the number of water bodies in a country. Because of the narrow guidelines, most countries considered the typology development as a largely technical task.

3.2 The Danish typology

According to the Common Implementation Strategy for the WFD (2000/60/EC), Dahl et al (2005) divided the Danish coastal waters into 15 different types: 5 open water types and 10 estuary types. Transitional waters were included in the typology. Dahl et al (2005) state that “the large number of types reflects the strong salinity gradient present in the Danish coastal waters, but also that the physical factors that are relevant for defining a type, vary greatly among the Danish estuaries”. The national typologies in the Baltic Region show many similarities, and in several cases coastal and transitional waters were merged into one system to reduce complexity. In the Scientific Documentation Report, the Danish typology is further simplified and types are merged. The aim of this simplification is that less Chlorophyll a reference and threshold values for a Good Ecological Status (target values that fall between the High-Good and Good-Moderate boundaries) have to be defined.

The Common Implementation Strategy for the WFD (2000/60/EC) reminds member states that, when developing a typology, they should keep the major objective of the Directive in mind, namely to establish a framework for the protection of both water quality and water resources preventing further deterioration and protecting and enhancing ecosystems. It is pointed out that typology is a tool to assist this process, and it is recognised “that a simple typology system needs to be complemented by more complex reference conditions that cover ranges of biological conditions” (p. 28). It means that every country has the freedom to adjust the typology to its own needs and to refine it to the required degree.

3.3 Suitability of the Danish typology

The major question is whether the typology in the Scientific Documentation Report is sufficiently detailed to allow the definition of reliable reference and target values for Chlorophyll a and the other indicators in all coastal waters. Reliable means that these values well reflect the ecological conditions and properties of all coastal waters. This is a precondition for defining target values that allow to derive to derive reliable MAI for each water body. The general impression is that the typology allows the derivation of suitable target values and MAI for water bodies in the sea with strong water mixing. An indication is that the inter-calibrated values for Chlorophyll a with Germany and Sweden for the sea and outer coastal waters are well in agreement with the results of the Scientific Documentation Report (Schernewski et al, 2015). In general, the comparable German Chlorophyll a target values for the open sea are slightly lower, but would allow a cross-border harmonisation.

Many fjords and coastal bays share a similar Chlorophyll a target concentration of 3.6 mg/m³, namely Norsminde Fjord, Mariager Fjord (outer), Nissum Bredning, Randers Fjord (outer), Horsens Fjord, Kolding Fjord, Vejle Fjord, Odense Fjord, Nyborg Fjord, Kerteminde Fjord, Holckenhavn Fjord, Bredningen, Emtekær Nor, Nærå Strand, Nakkebølle Fjord, Dalby Bugt, Karrebæk Fjord and Roskilde Fjord.

The Scientific Documentation Report and the additional data tables provided by the authors of the report show that water bodies with diverse properties are represented by only one target value. The typology is too simplified to reflect the specific characteristics of the individual fjordic water bodies. The consequence is a large and not sufficiently justified variation in the required load reduction for each water body. In the understanding of the Panel, the Danish typology does not sufficiently reflect the individual properties of the many Danish fjords and inner coastal waters. The solution could be either to subdivide the typology for these systems, taking into account especially water exchange rate and fresh water discharge, or to develop individual Chlorophyll a target values for every single water body. The statistical modelling, especially when carried out across water bodies, could be an excellent basis for this.

3.4 Suitability of the Danish monitoring programme

A precondition for a refined typology for fjords and inner coastal waters is the existence of a suitable and comprehensive monitoring programme. The present Danish national monitoring programme includes more than 90 stations along the coast and in the sea. It is very comprehensive and seems to be well-adjusted to the WFD requirements. Altogether, 119 water bodies are separated in Denmark. In some cases, fjord systems are divided into two or more water bodies and are represented only by one monitoring station. Examples are Mariager Fjord, Randers Fjord, Vejle Fjord and Flensborg Fjord. It means that practically every water body or spatially linked group of water bodies (like a fjord) are represented by one monitoring station. This is important, because only the existence of a monitoring station and regular data collection allows assessing whether the target is reached or not.

3.5 Suggestions towards a modified approach

Such a comprehensive monitoring programme not only allows a refinement of the typology, but would allow the definition of individual Chlorophyll a reference and target values for every water body, respectively some spatially linked group of water bodies. We strongly suggest considering this approach, especially when the aim is to calculate as precise and water-body specific MAI as possible. Denmark is one of the few countries in Europe, where the necessary data, expertise and models are available for such a comprehensive approach. In detail, it has to be assessed if additional monitoring stations, temporary data collection at some locations or complementation of the monitoring programme with remote sensing might be necessary. Neither the application of the meta-modelling nor monitoring of the success of the proposed measures are possible without a minimal set of follow-up actions in the field.

3.6 The look abroad

Similar discussions took place in Germany as well, and the results are reflected in a national report and an international publication (Schernewski et al, 2015). Germany carried out a comprehensive revision of all German Baltic reference and target values for nutrients and Chlorophyll a. The discussion process within the accompanying official national working group came to the conclusion that especially the different estuaries and lagoons have so specific properties and behaviours, and that type-specific Chlorophyll a and nutrient reference and target values would be too general. As a consequence, specific Chlorophyll a and nutrient reference and target values were developed for every single water body, resulting in 35 major Chlorophyll a reference and target values for the German Baltic waters alone.

4. The use of seagrass and Kd as environmental indicators

The use of Chlorophyll a as an indicator for phytoplankton is widespread and accepted in the WFD. The indicator is intercalibrated between Denmark and neighbouring countries and is fully endorsed by the Panel. In this chapter, we will discuss the appropriateness of other indicators. We discuss the use of Kd as an indicator for aquatic macrophytes and angiosperms, and subsequently devote a discussion to the indicators for hypoxia and nutrient limitation, used in the statistical modelling approach. The main questions are whether these indicators are appropriate for demonstrating important ecological quality aspects, whether they can be related to nutrient inputs, and whether they should be maintained in the scientific modelling approaches.

4.1 Kd as an indicator for the biological element “benthic vegetation, macroalgae and angiosperms”

One of the three main indicators used in the WFD as measures of Good Ecological Status is the condition for aquatic macrophytes and angiosperms. In most Danish estuarine and marine waters, this concerns eelgrass (*Zostera marina*), even though this is not the only species of angiosperm that occurs. Pondweed (*Potamogeton* species) and *Ruppia* may cover extensive parts of some systems and should be taken into account as “angiosperm vegetation”. However, in systems where this occurs (e.g. Odense Fjord), it is still eelgrass that dominates the deeper (>1.5 m) parts and thus remains the most critical indicator. The Panel has no complete overview of the situation in all the different water bodies, but stresses the generality of the required “angiosperm vegetation” indicator, so that it may occasionally differ from the single “*Zostera* maximum depth” indicator, at least in principle.

In the scientific documentation report and the underlying model work, water transparency, expressed as the light extinction coefficient Kd (m^{-1}) is used as a proxy for the depth limit for eelgrass. This is based on solid scientific evidence that eelgrass needs a light intensity at the bottom of between 10-20% of the incident light. The choice of 14% is based on this literature, and on area-specific experiments for Danish waters, and is well justified. However, the Panel points out that, due to the non-linear interaction between light intensity and Kd and in the presence of temporal variability in Kd, the average light intensity reaching the bottom in a water system at a particular depth may differ significantly from the light intensity calculated at this depth using the average Kd.[2]

As clearly stated in the report, water transparency is a necessary but insufficient condition for seagrass to re-establish in these estuarine systems. Recent studies of seagrass reproduction, as well as adult seagrass survival, have pointed to factors such as disturbance by floating algae, resuspension of fine material, disturbance by lugworms, herbicides and others to influence recovery (Flindt et al, 2016; Kuusemäe et al, 2016; Canal-Verges et al, 2016). It is likely that the presence of eelgrass itself plays a role in these circumstances, not only as a seed source but also by collecting and fixing fine sediment material. It has been shown in general (van der Heide et al, 2011) and in a specific restoration case in North America (Orth et al, 2012) that this may lead to alternative stable states and strong non-linear behaviour: once extensive seagrass

---

[2] As an example, if Kd has a lognormal distribution with a logarithmic mean of -1.0 (mean approx. 0.4 m^{-1}), and a logarithmic standard deviation of 0.5, the depth was on average 14% of incident light reaches the bottom is 6.15 m, whereas the depth limit calculated from the average Kd is 4.85 m. Note that the difference depends on the statistical distribution of Kd in the field, which the Panel cannot evaluate.

meadows are present, they contribute to keeping the water clear and extend their range to deeper waters, but in the absence of meadows the water remains turbid and prevents the development of meadows. As a consequence, restoration of seagrass meadows (i.e. transition from unvegetated to vegetated state) may require more stringent conditions to be fulfilled than what is needed to maintain an existing vegetation. These more stringent conditions may relate to lower nutrient loadings, but may also relate to the exclusion of other disturbing factors. Therefore, it is unlikely that Kd as a sole indicator covers the entire range of conditions needed for eelgrass restoration, but it is even more unlikely that restoration will succeed without at least restoring Kd to the levels needed for the Good-Moderate boundary conditions.

The time course of Kd in the water bodies studied by statistical modelling is shown in the Annexes to this evaluation report. In most cases, it is very difficult or impossible to detect a significant downward trend in the values. Even though a significant correlation of summer (June to August) averages of Kd with N load is reported for 16 out of 22 stations [p 94], the slopes of these relations are very low [p 94], and no material changes in yearly averages are observed over time despite changes in N loading. Similarly, in the mechanistic modelling, slopes for change of Kd as a function of N load are usually small, and the model is not able to reproduce the reference (observed around 1900) Kd values by modelling reference loads of 1900. The total range in Kd across all systems at reference loadings is approximately 0.15 m^{-1}, whereas the total range in the observations is approximately 0.35 m^{-1}. The model is calibrated to reproduce the mean reasonably, but is not able to capture the full (temporal and cross-system) variation very well.

Based on a statistical analysis of all Danish systems since the 1980s, Riemann et al (2016) report a significant increase in Secchi depth between 1980 and 1990, when many systems made the transition from hypertrophic to eutrophic state, but absence of any systematic trend afterwards. In addition, they remark that Secchi disc depth data even overestimate the effect on Kd, because of a shift from scattering to absorption as the main light extinction mechanism between 1980 and 2010. In summary, none of the within-system statistical analyses or models seem to be able to demonstrate a strong dependence of Kd on nutrient loading in the period 1990-2013.

However, when viewed across systems, the data shown in annex B of the Scientific Documentation Report for Chlorophyll a and Kd in the systems studied with the statistical modelling strongly suggest a close correlation between average Chlorophyll a concentration and average Kd over the study period (see Figure 4 in Chapter 8). It is likely that a common cause – most probably the relative influence of the freshwater end member in the water of the estuary – determines both. However, within each of the systems, we do not observe a correlated evolution in time of the two indicators over the 1990-2012 period. A further interesting observation is that the targets for Kd and Chlorophyll a, as derived in the Scientific Documentation report, show exactly the same correlation but in a narrower range of both indicators. As these reference values represent historic conditions, the suggestion is that on a centennial time scale, Kd and Chlorophyll a covary in time. Therefore, it is possible that Kd does respond to nutrient loading, but with significant delay and only on long time scales.

From these considerations, the Panel concludes that both indicators represent eutrophication effects, but that the estimation of the effect of nutrient reduction on Chlorophyll a is more reliable than the estimation of this effect on Kd.

The Scientific Documentation Report suggests that other causes, in particular the influx of dissolved and particulate organic matter from freshwater, as well as long-term storage of fine and fluffy sediment material, influences the transparency of the water. Trends in benthic filter feeders (that have decreased significantly in biomass between 1990 and

2017 – see Riemann et al, 2016) may also be a causal factor. It can be assumed that filter feeders decrease in biomass as a consequence of the decrease in phytoplankton primary production, which may result in less filtration and fixation of fine particles in the sediment. It is difficult to evaluate each of these hypotheses, but the very good correlation between average chlorophyll and average Kd across systems suggests that the influx of some substance with the freshwater, that may be higher nowadays than in 1900, plays a dominant role. In their answer to the questions of the Panel, the researchers have thoroughly analysed and dismissed the possibility that herbicides play a significant role in this freshwater influence. The most probable hypothesis is that the influx of coloured organic substances has increased between 1900 and now.

The consequences of incorporating Kd as an important indicator for water quality, in the absence of strong slopes between nutrient loading and Kd, are different for the mechanistic modelling and the statistical modelling exercises. The mechanistic modelling estimates which part of the distance between target and status can be bridged by reducing Danish land-based N sources. It corrects for this fraction in the calculation of the effort required. The Panel finds this approach appropriate and does not think it leads to unjustifiable overestimation of efforts needed.

The statistical modelling approach does not follow the same reasoning as the mechanistic modelling. For some water bodies, N load reductions of well above 100% are calculated to be needed in order to bring Kd down to target levels. This is of course physically impossible. The problem is solved by “translating” the required very high efforts into realisable efforts [25% when the calculation is 25%-100%, 50% for calculation 100%-200%, 75% for calculation >200%). Despite questions to the researchers, the Panel has not been able to discover the logic behind this translation. The researchers argue that this is basically expert judgement and further argue that 25% is the order of magnitude of interannual variation of the N load, therefore when an effort is estimated to be “large”, it should be above this level but not too much. In the opinion of the Panel, the “translation” introduces an unnecessary element of arbitrariness into the whole procedure that is in contrast with the general evidence-based approach and that therefore exposes the entire procedure to unproductive criticism. The Panel furthermore observes that the situation here is analogous to the situation treated in the chapter on mechanistic modelling, where very often the target value cannot entirely be reached with reduction of Danish land-based N. Therefore, the Panel suggests harmonising the approach across the two modelling lines and adopting the approach of the mechanistic modelling also in the statistical modelling.

In further work, the Panel recommends reviewing the approach for this WFD indicator by starting from the basic observation that not Kd, but survival and restoration of aquatic angiosperm vegetation is the real criterion. In some systems, this criterion may actually be fulfilled by other species than eelgrass (e.g. *Ruppia* or *Potamogeton* species), in which case the criterion could also be considered as generally fulfilled. However, in most cases, eelgrass will be the species of interest. As mentioned above, recent modelling work of Kuusemäe et al (2016) and Flindt et al (2016) has taken a more comprehensive view on restoration of eelgrass, and the influence of nutrient loading on the process. This work is actually built into the mechanistic models used in the present study, but the results have not been directly used in order to estimate the influence of nutrient reduction on seagrass restoration. The Panel proposes to make better use of these models, probably after more extensive validation, to more directly estimate the effect of nutrient reductions on seagrass development possibilities.

In view of the apparent difficulties in estimating the effect of nutrient reductions on Kd at short time scales, the insufficiency of Kd as a representation of all factors needed for restoration of seagrass, and the high correlation between Kd and Chlorophyll a both in status and targets at longer time scales, the Panel suggests to relatively downweigh the importance of Kd in the final calculations of reductions needed. It further recommends pursuing studies attempting to estimate conditions for seagrass restoration based on already developed more comprehensive models. In the absence of the latter, and given the correlation between Kd and Chlorophyll a, the Panel is of the opinion that adherence to the “one-out, all-out” principle with respect to Kd and Chlorophyll a, is not imperative. A weighted average of reduction needs for both indicators might be preferable.

4.2 Other indicators used in the statistical modelling

In contrast to the mechanistic modelling, the statistical modelling bases its conclusions on three other indicators: (1) the occurrence of hypoxia, (2) ecological signs of hypoxia from nutrients and (3) chlorophyll and (4) the number of days of N limitation of phytoplankton growth. Indicator (2) and (3) are given half weight as they estimate one element together. Compared to Chlorophyll a and Kd, the combination (2) - (3) and the indicators (1) and (4) are given half the weight.

The Panel is surprised by the inclusion of these indicators in only one line of modelling, as it could also have been done in the mechanistic modelling. The latter contains all the variables needed to estimate hypoxic/anoxic conditions as well as direct estimates of nutrient limitation of phytoplankton growth. The asymmetric situation leads to a decrease of comparability of the two models and decreases credibility of the procedure of averaging both approaches, e.g. in meta-modelling. The Panel further notes that in meta-modelling based on the statistical approach, the ancillary indicators are sometimes included and sometimes not, depending on data availability.

With respect to the occurrence of hypoxia, the researchers note in the Scientific Documentation Report that:

“There is direct evidence for a relationship between nutrient loadings and oxygen concentrations in bottom water (Markager et al, 2006) and the size of hypoxic/anoxic areas (Scavia et al, 2003; Christensen et al submitted). However, these relationships are complicated by a considerable time lag and a high sensitivity to climate variables like water temperature and wind stress.”

With respect to the number of days with nutrient limitation, figure 8.7 of the Scientific Documentation Report shows a direct correlation with Chlorophyll a concentrations, but with considerable scatter (considering the log scale of the y-axis). Two questions can thus be posed: Do the additional indicators measure a significantly different indicator compared to Chlorophyll a and Kd, and can the effects of nutrient reduction on both indicators be estimated reliably?

The Panel is of the opinion that both criteria lead to doubt about the usefulness of these indicators. It is clear that phytoplankton production and biomass is related to the amount of organic matter sinking to the bottom and fuelling oxygen consumption. This could be a reason to include the spring in the Chlorophyll a indicator, but anyhow a correlation between Chlorophyll a and the probability of occurrence of hypoxia can be expected. Note, moreover, that excessive summer Chlorophyll a is the ecological sign of hypoxia used. The high dependence of occurrence of hypoxia on weather conditions induces considerable variability that obscures effects of nutrient reductions on the indicators. There is essentially only one (on/off) observation per year. In addition to these difficulties, a rather arbitrary look-up table approach has to be used in order to estimate the required nutrient reduction for improvement in the hypoxia indicators.

For the indicator “days with nutrient limitation”, it can be expected that nutrient reduction, if effective on Chlorophyll a at all, can only be effective through the increase of the time duration of nutrient limitation. It is hard to see how the response of this indicator could differ from the response of Chlorophyll a concentration. The latter, however, can be measured more easily and more reliably. We note that there is considerable disagreement in the literature on the correct value of Km, the Monod limitation parameter, and that it differs considerably between different phytoplankton species and groups. Also for this indicator, a look-up table has to be used to estimate the required load reductions.

In summary, even though the ancillary indicators aim at describing important ecological phenomena, it is not easy to translate them into required load reductions (expert judgment and look-up tables are needed) and their added value compared to Chlorophyll a and Kd is limited. Therefore, the Panel is of the opinion that these indicators do not bring a substantial improvement of the approach. The Panel recommends using the mechanistic models to better study how the important phenomenon of oxygen depletion can be linked directly to required nutrient reductions before using it in practice to estimate required nutrient reduction. If, based on these studies, it can be decided to use these additional indicators, they should be introduced in both statistical and mechanistic modelling approaches for consistency of the approach.

5. Emphasis on nitrogen versus phosphorus

In this chapter, it is evaluated to what extent the Scientific Documentation Report *a priori* focused exclusively on nitrogen (N) reductions as measures to reach Good Ecological Status, or if evidence is given that rules out positive effects from phosphorus (P) load reductions. We thus address the question whether management options were unnecessarily limited by focus on reduction of the yearly N load only.

5.1 Phosphorus limitation

The nutrient emissions from large point sources were dramatically reduced already during the 1980s, causing a large and relatively sudden decrease in the P loads. After that first effort, focus has been directed to mitigation of N loads, primarily through measures in agriculture. This has resulted in a smoother, but substantial, decline in the nutrient inputs (mostly N inputs) thereafter. Currently, the overall inputs of N and P are roughly about 4.2 and 3.4 times higher, respectively, than estimated reference inputs for the year 1900 (Riemann et al, 2016). This indicates that the N/P ratio of nutrient inputs is not exceptionally deviating from the historic inputs.

Previous studies have shown substantial and significant reductions of primary production (e.g. Timmermann et al, 2014) and Chlorophyll a concentrations (e.g., Riemann et al, 2016) in response to the early P load reductions. Thus, there is no doubt that reduction of P loads can, in principle, lead to improvement of water quality in terms of WFD indicators. However, it is uncertain to what degree these historical responses are transferrable to present-day conditions, because emissions from point sources did not have the annual cycle of the diffuse sources and, moreover, they were observed in generally hypertrophic situations that are not comparable to the present state.

Traditionally, marine coastal waters have been regarded as N limited, but in the past decades, scientists have become increasingly aware of complicated co-limitation patterns and intricate nutrient dynamics. Processes such as N fixation and sediment P release can modify long-term response compared to the direct response of phytoplankton to nutrient additions on short time scales. A number of studies from Danish waters confirm that N is in general limiting algal production during summer time, and P is often limiting in spring, but there are seasonal and spatial variations of nutrient limitation. These field studies suggest that at least in a number of systems, regulation of annual primary production by P load reduction could be feasible.

5.2 Treatment of nitrogen and phosphorus in the Scientific Documentation Report

There are several elements that have contributed to the large emphasis on N load reduction in the Scientific Documentation Report. In particular, we discuss the nature of the indicators used, the selection of the study period, the procedures of the statistical modelling and the characteristics of the mechanistic model.

The basis for all calculations are the indicators Chlorophyll a and Kd during summer. This has potential implications for the exclusive focus on nitrogen load reductions. Summer phytoplankton in most Danish water bodies is predominantly nitrogen limited. The choice of summer Chlorophyll a as an indicator may have focused the attention primarily on processes that are dominant in summer and on nitrogen loads as a primary factor responsible for eutrophication. This is also pointed out in the Scientific Documentation Report, where it is suggested that developing new indicators focusing on other parts of the season would give a more diverse focus on both nitrogen and phosphorus.

In general, the Panel is of the opinion that the selection of indicators only representing summer conditions could be too restrictive. In waters with some degree of stratification, the spring bloom has the highest contribution to export production, fuelling the organic matter on the sediment and largely determining the oxygen demand in the rest of the season that could lead to P release from the sediments. The Scientific Documentation Report suggests that limitation of the spring bloom by P occurs in a number of water bodies, thus suggesting that the effectiveness of P load reduction on an indicator representing the full growing season could be significant.

Another factor potentially excluding possible influence of P in the analysis is the selected period for the statistical model (1990-2013). This period excludes most of the period of major development of efficient sewage treatment in the 1980s that caused a major decrease in point source P loads. For most water bodies, the P load trends that are now dominated by diffuse sources are less significant than N load trends, and thereby it is naturally more difficult to find significant effects. However, the Panel endorses the choice of period, because the seasonality and mechanisms of P limitation in current situations may differ from the historical, point-source dominated situation, as argued above.

In the statistical modelling approach, the variable selection procedure may have masked the potential role of phosphorus load reduction. There is a bias in the selection of variables towards regressions with N. This occurs first through the automated variable selection process. Whenever N load is selected as the dominant variable, possible P dependence is disregarded because P load is no longer considered as a secondary independent variable. If, on the other hand, P is selected as the dominant controlling variable, that regression model is not used. Thus, potential influence of P load reductions, or combinations of N and P load reductions, are not investigated further.

The mechanistic models include all relevant processes for modelling effects from both N and P and combinations of them both. However, major focus in the formulations of scenarios is on N, and the few scenarios, including also P reductions, are not detailed and perhaps not optimal for exploring the influence from P load reduction. In addition, we observe that for a significant portion of the water bodies, the models seem to overestimate P concentrations during summer. This can have eliminated the potential impact from P load reductions on the indicators.

5.3 Possible implications for management

Based on the different factors leading to a focus on N load reduction, the Panel concludes that the study does not demonstrate significant contributions from P loads on the summer indicators, but the evidence is not strong enough to exclude that P reductions or combined N and P reductions could be effective in reducing year-averaged chlorophyll levels as well as sediment oxygen demand.

Keeping the option for combined N/P reduction open may have significant management implications in regions where very large N load reductions are demanded. Focused studies resulting in an envelope of combinations of Maximum Allowable Inputs of N and P would probably lead to greater flexibility and more cost-efficient nutrient reduction management in these areas. In making this recommendation, the Panel acknowledges that great efforts have already been made to reduce the P load from urban waste waters, and that little gain has to be expected from intensifying those efforts, following the law of diminishing returns. However, any innovative approach to reducing remaining P loads, including the P load from agriculture, could significantly enlarge the portfolio of potential measures. The Panel recommends using basin load models in combination with the mechanistic models used in the Scientific Documentation Report to investigate these possibilities.

5.4 Seasonality

The exclusive focus on summer indicators in combination with water bodies with short residence times implies a direct link between summer loads and the indicator. Typical residence times in Danish estuaries are short in many cases, ranging from a few days to about 3 months (Rasmussen and Josefsson, 2002). Even if the indicators would include the spring phytoplankton bloom, regulation by N loads would mostly focus on the summer period in water bodies where P limits the spring bloom. There seems to be a possibility to regulate Good Ecological Status by focusing on the summer loads, rather than on the yearly integrated loads. The Panel recognises that the problem is complicated by N retention in the system in the form of organic N stocks accumulating over the season and even years, so that the calculation is not straightforward. Moreover, spatial displacement of problems to other systems as a consequence of flushing winter nutrient loads has to be taken into account. Even so, the Panel estimates that the modelling tools developed, especially the mechanistic modelling, are able to investigate scenarios with seasonal regulation of the N (and P) input into the system. Therefore, nutrient load management could be focused on optimising the effect in the coastal estuaries. The Panel does not have a complete overview of the potential, in agricultural practice, to focus in particular on summer N load. However, it recommends exploring the possibilities to do so and use the mechanistic models to estimate how this would affect the GES indicators.

6. Statistical modelling

In this chapter, the Panel reviews and sums up the objectives and basic setup of statistical modelling and their usage in defining Maximum Allowable Inputs (MAI). The main issues are averaging of statistical and mechanistic models, the analysis of within and cross-system variability in Chlorophyll a and Kd responses, collinearity of phosphorus (P) and nitrogen (N) loading, filtering out the effect of flushing and the uncertainty and resulting risk of over- and under-dimensioning of MAI.

6.1 Setup

The statistical model approach as presented in the Scientific Documentation Report aims at demonstrating the dependence of the indicators Chlorophyll a, Kd, hypoxia, anoxia, number of days with nutrient limitation on the N and P loading of the system as well as on some other physical and chemical characteristics of the system. The statistical models (there is one model per sufficiently monitored water body) also estimate how concentrations of Total N and Total P depend on the nutrient loadings and physico-chemical characteristics. These latter analyses are informative on the functioning of the systems but are not really used any further in the overall modelling procedure.

A few basic choices for the setup of the statistical modelling have been made at the start of the study. The most important choices were:

- Restrict the database analysed by the statistical modelling to the period 1991-2012. This implies that the major decrease in P input, as well as the ecological consequences of this decrease, in general are not part of the analysed database.
- Restrict the construction of statistical models to those systems where sufficient data are available. What is “sufficient” is always open to discussion, but the Panel is of the opinion that the choices are reasonable and have been well justified.
- Use annual averages of nutrient loads, concentrations and other variables as the basis for modelling.
- Construct one statistical model per water body without cross-system model building.
- Perform a variable selection method for significant independent variables, where (due to collinearity problems) only one type of nutrient loading (either N or P) was entered into the set of independent variables.

The most important results of the statistical models are the slopes of the relation between N loads and the indicators Chlorophyll a and Kd. These slopes are only determined, if N load was selected as the most important independent variable and thus entered into the statistical model. When this was not the case, substitute solutions have been used. The statistical model that was used to estimate the slopes (Partial Least Squares) is different from the model used to select the variables (Multiple Linear Regression).

The construction and use of the statistical model are well explained in the Scientific Documentation Report. Measures of goodness-of-fit are given at different stages in the description. No formal uncertainty analysis of the model as a whole, nor variance estimation of the estimated parameters (in particular the N load – indicator slopes) have been given.

6.2 Panel evaluation of basic model setup

The Panel distinguishes three major uses of the results of the statistical modelling:

- Estimation of the relation between nutrient loading and indicators at relatively long time scales (5-10 years), as a basis for estimation of reference conditions of Chlorophyll a, and of the effectiveness of load reductions in reaching the target conditions.
- Provide insight into the water body characteristics that explain the differences between water bodies in status or slopes.
- Provide an independent, evidence-based check on the accuracy of the mechanistic modelling approach.

The Panel remarks that the statistical models are not needed to ascertain that nutrients, both N and P, are important for phytoplankton. This point was also made by the researchers, stressing that the body of scientific evidence showing these relations is massive.

In contrast to the researchers, however, the Panel questions if the step of variable selection was needed at all. It involves mixing of two methods (MLR and PLS). It also leads to the suggestion that in some systems, N load was not involved at all in determining Chlorophyll a and Kd. In systems where N load was selected as the most important determining factor, possible secondary effects of P load cannot be shown and are obliterated. The most important consequence of this option, however, is that it may lead to biased estimates of the slopes and MAI. If, in a particular water body, the slope is very small (close to zero), it is very likely that N load will not be selected as the most important independent variable in the variable selection procedure. Subsequently, for this system, the slope will be estimated as the average type-specific slope, almost inevitably leading to a higher slope than shown by the data. This will then lead to a lower reference and target value for the system than the one suggested by the data. As these reference values will enter into a type-specific averaging afterwards, the final consequences of these choices become difficult to assess, but likely affect the targets for all systems in the type.

Moreover, the Panel is of the opinion that there is no real reason for estimating the short-term response of the indicators on year-to-year variations in nutrient loads, with or without time lags of a few months. Both nutrient loads and concentrations of nutrients and chlorophyll are known to vary considerably with freshwater discharge, which is variable from year to year. Short-term (i.e. year-to-year) responses of indicators to short-term variations in nutrient loads will not necessarily be the same as the decadal-scale responses that the study really wants to estimate. For instance, high discharge will not only increase the total load of nutrients to a system, but simultaneously also decrease the freshwater residence time and thus the ability of the ecosystem to take up and use these nutrients. This may contrast with a decadal-scale increase in nutrient load, where clearer and possibly also different ecological responses might be expected.

Therefore, the Panel is of the opinion that a clearer focus on the long-term slopes and the cross-system variability is needed. Through the use of mixed or Bayesian hierarchical models, short-term and long-term variations can be separated and collinearity between variables can be built in as part of the model (Malve & Qian, 2006).

Danish water systems differ in a number of morphological and hydrographical characteristics, leading to a diversity of systems that is not very well captured by the few types used in the typology (see Chapter 3 in this evaluation report). However, there are a few characteristics that presumably dominate the differences in nutrient, chlorophyll and Kd status between systems. The relative influence of freshwater in the water, dependent on discharge rates, flushing rates and exchange rates with the coastal system, will most probably be a key parameter. Nutrient concentrations in seawater are relatively stable and do not differ very much between the reference conditions and now. In contrast, nutrient concentrations in freshwater are much higher and are obviously much more directly influenced by nutrient loads. As a consequence, it may be expected that much of the variation in status and slopes between systems may be explained with a cross-system statistical model as suggested above. The main purpose of this setup is to improve within-system estimates of slopes with information coming from similar systems elsewhere, and to improve meta-modelling applications. It should lead to a model that estimates the slopes (which are the results of primary importance) based on independent variables summarising the water body characteristics, while simultaneously estimating (and evaluating) system-specific deviations. Such an approach could constitute an improvement with respect to the current within-system modelling approach.

In order for the statistical model to provide an independent, evidence-based check on the results of the mechanistic modelling, two requirements must be fulfilled. First, the procedures of the statistical and mechanistic modelling should not be unduly mixed at early stages (see comments in Chapter 8 in this evaluation report). Second, the statistical model should contain a formal estimation of variances of the estimated parameters. Statistical modelling techniques have much better formal methods to estimate uncertainty than mechanistic models, and this opportunity should be taken in order to better formalise both uncertainty resulting from modelling and from data uncertainty. For this evaluation to be effective, the setup of a single cross-system statistical model is better suited than the current set of separate within-system models.

6.3 Panel evaluation of statistical model results

Even though the simultaneous development of two model lines, statistical and mechanistic, may seem redundant at first sight, the Panel endorses the continuation of this approach. The richness of the Danish database is an internationally exceptional asset that provides the opportunity of an evidence-based check on mechanistic model outcomes. This asset should be used, and the two modelling lines are a very good way to do so. However, the Panel recommends strengthening this aspect, e.g. by keeping the two model lines more separate and independent throughout the modelling procedure, so that the check becomes clearer and more explicit in the final stages of result interpretation. Furthermore, as specified above, the Panel is of the opinion that a cross-system approach in the statistical modelling would strengthen the possibilities of obtaining insight into possible causes for model divergence and would assist better in choosing final management strategies based on the model comparison. In addition, a formal uncertainty analysis of the statistical model would contribute to this goal.

With respect to the present outcomes of the statistical modelling, the Panel sees reasons to suspect bias in estimated slopes and reference values due to the variable selection procedure, as specified above. The Panel suspects that the slope estimates, being a mixture of short-term and long-term ecological responses, might be biased as estimators of the long-term response. However, the Panel does not consider these remarks as a reason to entirely dismiss the statistical model results as unreliable. The mentioned discrepancies are probably minor in comparison with the overall range of the results and in comparison with the inevitable variability in the observations. The within-systems PLS regression approach used is robust and not expected to be overly influenced by the mixture of short and long time scales. The variable selection procedures have led to the replacement of slopes with type-averaged slopes, but mostly in types with small slopes. Nonetheless, there is enough reason to improve the statistical model and the slope estimates that follow from it.

7. Mechanistic modelling

The mechanistic models are evaluated in this chapter with respect to included processes and technical implementation, performance and the different scenarios that are used.

7.1 The models

The mechanistic modelling is based on the DHI systems MIKE 3 combined with ECOLAB. Four models are set up: a large-scale model encompassing the whole Baltic Sea up to Skagerrak (IDW model) and three models of specific estuaries; Limfjorden, Roskilde Fjord and Odense Fjord (estuary models). In all, 45 of the 119 Danish water bodies are covered by the mechanistic models. The IDW and estuary models differ in some specific ways, adapting them to the circumstances. However, the three implementations of the estuary model are identical in terms of processes, but needed somewhat different calibration.

The pelagic dynamics in both models follow classic NPZ concepts similar to other models, and the bacterial loop is not explicitly resolved. An addition to many other similar models is that internal nutrient pools are explicitly modelled using the Droop equations (Droop, 1968). Both models also feature explicit benthic vegetation state-variables, but not benthic fauna.

The estuary models are quite comprehensive in terms of processes, including sophisticated representation of benthic vegetation and elaborate description of resuspension coupled to dynamic wave-shear processes from the hydrodynamic model. Spatial sediment characteristics are taken into account both for sediment-water interaction and as controlling the benthic vegetation.

Specifically, the IDW includes three autotrophic groups to take into account the seasonal succession and nitrogen fixation typical for the open sea areas of the Baltic Sea. Further, the representation of the sediments does not include explicit representation of inorganic particles and instead an empirical direct relationship between shear stress and turbidity is used. Simplification of the sediment module was necessary because of lack of detailed information from the wider area and because of computational constraints.

In many biogeochemical models, Chlorophyll a is estimated from the autotroph biomass in retrospect using a specific ratio. The models used in the Scientific Documentation Report are more advanced in this aspect in that Chlorophyll a is dynamically calculated based on fitness of the autotrophs and light conditions. In the IDW model, where there are three autotrophic functional groups, the weighted average contributions from all groups are taken into account in calculating the production and removal of Chlorophyll a.

The water transparency, Kd, is computed from a relationship that includes Chlorophyll a concentration, detritus carbon, (coloured) dissolved organic carbon and inorganic matter. All these components are explicitly modelled, although the inorganic matter representation in the IDW is a less complicated empirical relationship than in the estuary model.

In summary, the models are quite comprehensive and include all processes that we think are relevant for the problem at hand. The MIKE system, with its sub-components, is a mature system, although it is not so frequently used by research scientists, and, therefore, there are not that many peer-reviewed articles with applications as there are for some open access model systems. Despite this, we have no reason to question the model system capabilities.

7.2 Model setup, calibration and validation

All model setups have high resolution, both in horizontal and vertical. The IDW resolution is sufficient to resolve the internal physical dynamics, both of the narrow straits and geostrophically balanced Kattegat-Skagerrak front. The computing cost of the high resolution and high degree of complexity is significant, leading to a trade-off in the execution of calibration and experiment simulations. All relevant forcing functions are taken into account in a sensible way. The time period of simulation was 2002-2011. A critical part of the riverine inputs is the division of whatever carbon data available into the different categories of organic carbon in the model, especially the CDOC (coloured dissolved organic carbon) that influences Kd and the refractory and labile fractions of organic nutrients. This has been handled to the extent possible according to the Scientific Documentation Report.

At least the hydrodynamics of the IDW model have been used previously and were set up as a part of the EIA for the Fehmarnbelt fixed link project. The models for Odense Fjord and Roskilde Fjord are applied to vegetation modelling applications in Kuusemäe et al (2016) and Flindt et al (2016). Only the Limfjord model is newly developed. Thus, there is some history behind three out of four implementations.

All four model implementations are calibrated independently. That resulted in somewhat different parameter setting, also of the structurally identical estuary models. According to the researchers, there are only about 10 parameters that differ, and all of these are within the sediment module. The actual calibration procedure is not described in detail, but for the three models that have a past history, it can be expected that this has been an iterative process over some time.

7.3 Validation

The hydrodynamics are evaluated quantitatively with respect to salinity and temperature. Salinity is important since it indicates whether circulation is correct and gives the right mixing between the riverine water and open sea water in the estuaries of different sizes. Temperature is of less importance for the circulation, but of imperative importance for the biogeochemical processes. The quantitative comparison shows that the model results are well within the criteria. Upon request from the Panel, the researchers supplied direct time-series comparisons between observations and model results for all four models, and inspection of these shows excellent agreement between model and data for both salinity and temperature. The Panel is convinced that the models give a quite accurate representation of the physical processes.

The validation of the biogeochemical models is done primarily through comparison with observations of Chlorophyll a, Kd and nutrient concentrations. To simplify presentation of the validation of the biogeochemical processes in the models, results are aggregated per water type and month. This presentation may hinder interpretation of the magnitude of the difference between modelled and observed annual cycles. Quantitative skill assessment was performed by computing a cost function (measuring mean deviation scaled by variation of the variable) and correlation from simultaneous model results and observations. Upon request from the Panel, the researchers also supplied example time-series of concurrent observations and model results from selected locations for the standard measured variables.

The model separates well the differences in Chlorophyll a, Kd and nutrients between the water types, and mean values are well captured for all variables.

The seasonal cycle of Chlorophyll a is well captured, although levels are somewhat low during late spring – early summer in type 2 and 3 water bodies, and the autumn bloom seems to be underestimated in type 1 and 2 water bodies. The seasonal cycle of Kd is quite weak in especially type 1 and 2 water bodies, so it is difficult to value the accuracy from the seasonal averages in these water bodies. The tendency for all types 1-3 is, however, that Kd in summer is less than during winter, indicating some influence from an early spring bloom, but probably more from winter river runoff and turbidity from resuspended material. The time-series plots of Kd supplied by the researcher confirm the complications. The two open sea stations show seemingly random variations in time of observed Kd due to short-term variability, and no visual seasonal cycle or trend can be identified. There is no annual cycle (and only small variation) to be seen in the time-series supplied for Odense Fjord and central Limfjorden, neither in observations nor in model results. In Roskilde Fjord, there is significant variation in Kd with the seasons, but from visual inspection the pattern is irregular and not very well captured by the model, and it is not obvious what causes the variations. The time-series with clear seasonal cycle are from the inner part of Skive Fjord, and here the model accurately simulates the low Kd in winter time and high Kd in summer time.

The seasonal TN is modelled accurately for all water types. Winter DIN is somewhat overestimated in type 1 and 2 water bodies, and DIN is somewhat overestimated in late spring – early summer in type 3 waters. The overestimate of winter DIN in type 1 waters is confirmed for the time-series examples supplied by the researchers. However, overall, the model performs well on the nitrogen cycles.

There seems to be a consistent overestimation of DIP in the summer in type 1 and 2 waters, although somewhat later in type 1 than in type 2 waters. From inspection of the time-series, it seems that the problem is larger in the IDW, smaller in the Limfjorden model, while in the Odense Fjord and Roskilde Fjord, the seasonal cycle is quite correct. Winter DIP and TP concentrations are accurately modelled for all water types.

The quantitative validation in terms of cost function and correlation confirms the qualitative validation discussed above. Overall, the model is low in bias (cost function) indicating that the levels are modelled accurately, with exception of DIP in type 3 and to some extent type 1 waters and Kd in the type 5 waters. However, correlation is absent for type 1 water bodies and weak for type 3 water bodies for Kd.

The comparison between modelled and observed primary production indicates that the model performs well in this respect.

7.4 Reference conditions simulation

A hindcast simulation representing conditions around 1900 was performed. Forcing in general was kept as for the 2002-2011 period, but loads and nutrient boundary conditions needed adjustments. Appropriate waterborne and airborne loads were obtained from existing well-established data sets, and boundary concentrations in Skagerrak were adjusted according to previously published methodology. To overcome the computational challenge of running the whole of the Baltic Sea to steady state, initial conditions were adjusted in the IDW model according to literature values. It is the Panel's opinion that the setup of the simulation of reference conditions with the mechanistic model is sound and based on current published scientific knowledge on the nutrient loads around 1900.

Reference Chlorophyll a concentrations for all water bodies were extracted as average of the last 5 years of the simulation. In a few cases, simulations were repeated in order to be sure that average conditions were in equilibrium with the reference loads. It is unclear whether Chlorophyll a concentrations were spatially averaged over the water bodies or not.

7.5 Scenarios and establishment of cause-effect relationships

A prerequisite in construction of load reduction scenarios is implementation of BSAP for other countries than Denmark. That implies major reductions of primarily phosphorus to Baltic Proper, Gulf of Finland and Gulf of Riga, but also nitrogen to Baltic Proper, Kattegat and Gulf of Finland. The response time to the load reductions to Baltic Proper and the Gulfs is very long. Estimations show that during the first decade after implementation, all changes are within natural variability, but significant reduction in winter nutrient (primarily phosphorus) concentrations will be seen between one and two decades after implementation (HELCOM, 2013). The response time-scale has been shown to vary between models (Eilola et al, 2011), but is long in all cases. This means that the influence from load reductions to the Baltic Proper and the Gulf is limited during the decade considered here. It could be noted that in the longer perspective, nutrient concentrations would continue to decrease, and according to the underlying calculations in the BSAP, load reductions to the Baltic Proper was a prerequisite for obtaining GES in the Danish straits.

Three nitrogen reduction scenarios for the Danish loads were constructed by reducing proportionally all waterborne loads by 15, 30 and 60%, respectively.

There is also a set of scenarios where the three nitrogen scenarios are combined with a spatially distributed phosphorus reduction scenario according to reductions specified by the Danish EPA. It is mentioned that no significant effect could be detected from the P load scenarios, but there is no further elaboration on these scenarios. If the distribution is such that most of the reduction occurs to relatively few water bodies, there could potentially be an effect in these that would not be seen overall.

Indicator values from model results are calculated as water body spatial means, and these are corrected to match the mean observational value at the measurement station.

The scenarios without phosphorus load reduction are used to estimate parameters for a simplified surrogate model, built on temporal averages over 2007-2011. The three scenarios are used to establish the linear response function. The extrapolated value at present day Danish loads will represent the indicator value, given only reductions by other countries. In the Scientific Documentation Report, also an average indicator value from the reference scenario is included to indicate how much higher the value will be because of higher loads from other countries. For most water bodies and indicators, the linear approximation is appropriate. It should be remembered though that only a proportion of the full effect from BSAP reductions has had time to develop in the scenarios, and one would expect that for open sea water bodies, especially in the south, water quality will continue to improve as time goes by.

There are implications from the approach of running scenarios with a constant proportional load decrease in the scenarios. Some water bodies will be subject to change due to load reductions to adjacent water bodies. Therefore, one cannot directly sub-divide MAI to individual water bodies, if there is a risk that reduction is necessary also in adjacent basins to obtain GES. To fully disentangle the individual contribution spatially between all water bodies, one would need to test sensitivity to load reductions to each individual water body by itself, and perhaps, if the effect is non-linear, even combinations of water bodies. This would be a major computational challenge, and the improvement in the results would most probably be minor. The reason for the latter is that the problem mostly applies to open sea water bodies that would in any case integrate the load reduction for a relatively large region, while enclosed water bodies still are mostly dominated by local reductions.

7.6 Conclusion on the mechanistic models

Having evaluated the mechanistic models, the Panel comes to the following conclusions:

- The models are clearly state-of-the-art, both in terms of numerical techniques and processes included. The quality of the results follows a high standard and is as good as, or better than, other similar coupled physical-biogeochemical model systems.
- The hydrodynamics seem to perform excellently.
- Levels of Chlorophyll a, Kd and nutrients are accurately modelled across water body types.
- The biogeochemistry seems to perform overall somewhat better for nitrogen than for phosphorus, although in the models for Roskilde Fjord and Odense Fjord, also phosphorus performs excellently. Weakest is the performance of nutrients in the IDW model, where relatively frequently DIP seems to be overestimated during summer or early autumn and nitrogen during winter. Observed short-term variability in Kd in open waters is such that it seems impossible to model.
- Long-term response to large changes in nutrient loads has not been validated.
- The nitrogen reduction scenarios are appropriately set up and relevant.
- The scenario for P reduction is not extensively described, and it cannot be judged whether it forms sufficient basis for exclusive focus on N.
- It should be noted that in a longer time perspective, >10-20 years, the effect from BSAP load reductions will influence the open sea water bodies, especially in the southern part of the region.
- It would be extremely valuable to extend the mechanistic modelling system to as many water bodies as possible.

8. Calculation procedures to estimate Maximum Allowable Inputs from model results

In this chapter, we discuss the general build-up of the procedure to estimate reference conditions, Good-Moderate boundary targets and the required N load reductions to reach the target conditions. These procedures are based on the statistical and mechanistic model results, but use and interpret them in a diversity of ways. In our discussion, we focus on how the different models interact and on the different steps taken to arrive at the final MAI per water body.

8.1 Steps in the calculation of targets and MAI

Despite the general logical nature of the procedure, and even though the Scientific Documentation Report gives extensive explanations of the detailed procedures followed, it is not easy to follow and weight the different steps used in deriving the Maximum Allowable Inputs (MAI) for the water bodies. The essential steps, as the Panel understands them, are summarised in the diagram shown in Table 2. The left column refers to the procedures followed in the statistical modelling, and the right column to the mechanistic modelling. Joint cells point to steps where both approaches are joined.

Table 2. Essential steps in the calculation procedure of targets and MAI in the Scientific Documentation. Steps with averaging have red boxes.

Statistical modelling	Mechanistic modelling
Estimate the slope of the Chlorophyll a/N load relation for those systems where N load was selected as a significant independent variable in the regressions. For 8 water bodies where this was not the case, the average type-specific slope was used.	
Estimate the slope of the Kd/N load relationship, and substitute with average type-specific slopes where no significant relations could be found (6 water bodies).	
Estimate 1900 reference Chlorophyll a levels, using 1900 N loads and the slopes as input.	Estimate 1900 reference Chlorophyll a levels, using a 1900 scenario with adjusted nutrient inputs (N, P), adjusted benthic stocks etc.
Do not estimate 1900 Kd from the models; use historic observations instead. Where no direct observations were available, use observations from nearby similar water bodies.	Use the same historic data on Kd as 1900 reference as the statistical modelling.
Estimate Chlorophyll a reference levels per water body type by averaging the reference levels coming from the statistical and the mechanistic models of all water bodies in the type. Notes: For type 1, some subtypes are defined; a few systems have a *status aparte.* The slopes are not averaged, but kept per water body and model type.	
The same procedure is NOT followed for Kd. Historic references are used in both approaches.	
Estimate the required N load reduction to reach the target values for Chlorophyll a, Kd, hypoxia, anoxia, days with N limitation. Where logical inconsistencies may exist (reductions >100%), use a look-up table to substitute calculations. Unclear how this is done if	Estimate the required N load reduction to reach the target values for Chlorophyll a and Kd, taking into account the fraction due to Danish land-based sources. Based on scenarios with varying degree of

>100% is needed for Chlorophyll a.	overall reductions of N input.
Calculate the required load reduction as a weighted average of the results in previous step.	Calculate the required load reduction as a simple average of the results in previous step.
	Smooth the variability in required N load reduction by regional averaging. Unclear what was the basis for delineating the regions.
Meta-model systems without a model. IF status information is available, use type-averaged slopes for N Chla, N Kd and N other indicators (latter only if their status is known). Calculate weighted average required N reduction. ELSE use type-averaged required reduction.	Meta-model systems without a model. IF status information is available, use type-averaged slopes for N Chla, N Kd. Calculate average required N reduction. ELSE use regionally averaged required reduction.
Average required N reduction for meta-modelled systems across statistical and mechanistic approaches.	
IF mechanistic model exists for system: Drop information from statistical model and only use mechanistic model result. ELSE use statistical model result.	
Apply upstream-downstream rules.	
All done!	

8.2 Averaging and "ensemble modelling" aspects in the procedure

In this procedure, both modelling approaches are largely independent and focused on individual water bodies. However, four critical averaging steps intervene:

- The Chlorophyll a targets are averaged per type, over the statistical and mechanistic models. As far as the Panel understands it, this does not apply to the slopes, which remain water body-specific. In the Scientific Documentation Report, the averaging is justified as a means of reducing variability. As a consequence of the averaging, there is a possibility of incompatibility between slopes and targets: More than 100% of N load should be reduced. It remains unclear how this is solved. Presumably the reductions >100% enter the weighted average over indicators as such and are corrected by the averaging over indicators.

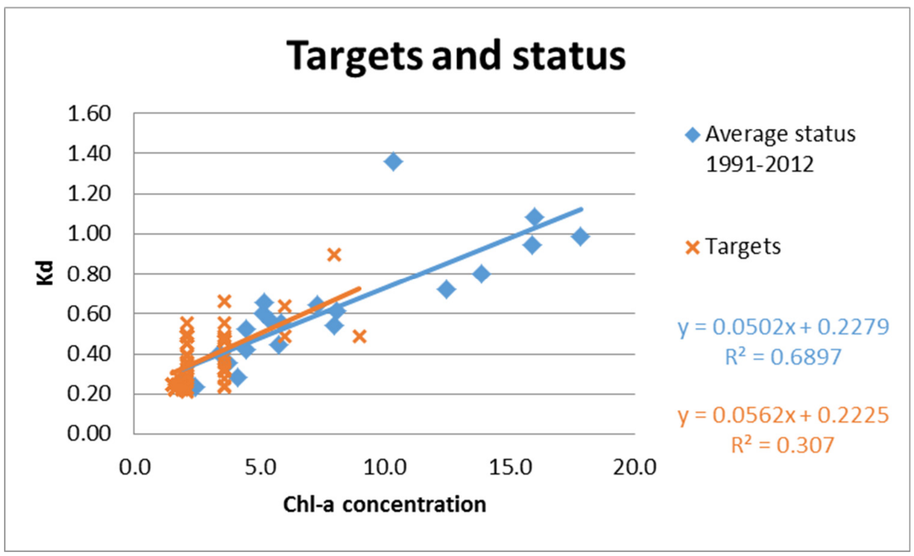
- The required load reduction for the different indicators (most importantly Chlorophyll a and Kd) is averaged quite early in the procedure. This averaging violates the "one-out-all-out" principle, as clearly stated by the researchers, but is justified in the Scientific Documentation Report based on arguments of reducing random variation in the required load reductions.

- The mechanistic model results, in terms of required percentage load reduction, are spatially averaged before the final meta-modelling step is applied to systems without monitoring data. This step obscures differences between water systems in regions. In the Scientific Documentation Report, it is very shortly discussed, and justified based on observed variability between systems in the regions that is – without proper argumentation – attributed to variability in the data on system status. It is not clear to the Panel why this would be the justification for a relatively drastic step in reducing the spatial differentiation in the model results. The Panel feels that this is an important yet poorly argumented step in the entire procedure.

- For meta-modelled systems, statistical and mechanistic meta-model results are averaged. This is not the case, however, for individually modelled systems where the mechanistic model has prevalence in the cases where both models are available.

In the opinion of the Panel, the most problematic aspect of the procedure is the averaging of Chlorophyll a reference (and GM boundary target) values across model types and within water body types. As the types are relatively broad and contain water bodies with quite dissimilar characteristics, this loss of local detail can easily lead to situations where too much effort is spent in one system and too little in another. In that case, there will be both economic and ecological loss of efficiency. In this respect, it is informative to compare the targets used for Kd and for Chlorophyll a across all systems, as is shown in Figure 4.

Figure 4. Comparison of target values for Kd and Chlorophyll a across all water bodies. Target values are shown as orange crosses. For comparison, the average status values 1990-2012 are shown as blue diamonds. Regression lines of the two sets are remarkably similar.



The figure shows that the target values of Kd, based on historic observations, are quite variable within each of the water body types and largely overlapping between types. There is much more discrimination between water bodies based on Kd than based on the (uniformed) Chlorophyll a targets, implying that the latter are not optimised for the water bodies.

The averaging of the required load reductions across the indicators (primarily Chlorophyll a and Kd) early on in the procedure renders it impossible to judge whether, and how much, the final results in terms of MAI depend on this violation of the one-out-all-out principle. If, for reasons of compliance to the WFD procedures, it would be decided that this is unacceptable, the present results cannot be used to make the recalculation. Also, no elements are offered to evaluate the importance of this decision.

The adoption of common Chlorophyll a target values across the statistical and mechanistic models also leads to a loss of independence of the two modelling approaches. As a consequence, comparison of the required N load reduction obtained by the two different methods, as a check on the methodology, is not really possible anymore. The Panel is of the opinion that it would be better to keep both methods separated up to the last stage and then do an in-depth comparison, taking into account water body characteristics to explain or understand any discrepancies. Averaging two independent model results as “ensemble modelling” is an option that can be taken in order to reduce variation in results, but it does not necessarily lead to a better solution. If one of the methods is biased (e.g. is clearly unable to make reliable estimates in particular types of systems), averaging is a worse solution than dropping the bad prediction. The availability of extensive databases on almost all systems should allow the model comparison to be evidence-based (e.g. model comparisons with data can be made for systems where predictions differ substantially), so that well-justified choices can be made.

Very little justification is given for the choice to give prevalence to the mechanistic models where both models are available. Even if the choice could be well justified (which is questionable since an independent comparison is impossible), it contrasts with the meta-modelling approach where both are averaged. Consistency in the choice would improve the overall approach.

In the opinion of the Panel, the spatial averaging of the mechanistic model results is not necessary nor justified. There can be good reasons, from a management point of view, to smoothen required load reductions regionally so that no abrupt changes in requirements occur at too small scales where control would be virtually impossible. However, such decisions could better be made on the basis of a map showing the original model results, allowing one to judge whether a management problem is posed or not. As it is performed now, it remains unclear to what degree the spatial averaging leads to under- or overdoing in particular watersheds, leading to lack of transparency in the management rules.

Summarising, the Panel recommends postponing the averaging operations to the very last stages of the procedure. This will keep the two modelling approaches independent, it will allow estimating the consequences of violating the one-out-all-out principle and avoid confusion due to regional averaging. As a consequence, the effects of different modelling strategies and different indicators will remain clear in the different results on nutrient reduction. A close examination and comparison of these differences will allow making informed decisions on the choice of strategy.

8.3 Conceptual differences between modelling approaches

There are a few points where the statistical and mechanistic modelling approaches are not conceptually consistent. The most important point is that the statistical modelling takes into account additional indicators, while the mechanistic models (although capable of doing the same, as all relevant variables are calculated in the model) only focus on Chlorophyll a and Kd. Even if this would not lead to large differences (this cannot be controlled based on the report), the Panel feels that it leads to a different treatment of water bodies, depending on the model applied to it, that is difficult to justify. Based on the observation that the ancillary indicators are strongly correlated with Chlorophyll a and can hardly be justified as probing independent characteristics of the ecosystem (see Chapter 4), the Panel suggests dropping these ancillary indicators from the procedure. It will make the two modelling approaches more comparable without apparent loss of information on the ecosystem. The ancillary variables could better be used as corroborating evidence for the need to take action (or not) in the water bodies concerned and as documentation of the range of ecological results to be expected from sanitation of the nutrient input.

A second potential inconsistency between the two models is that the mechanistic modelling explicitly separates how much of the distance to target can be reached by reducing Danish land-based N loading alone, while such separation is not done for the statistical model. However, the latter bases its regression approach on Danish (in fact: local) land-based N loads, so that, implicitly, the reasoning is probably more similar than it may appear when written out. Overall, the consequences of this difference are extremely difficult to trace, but intuitively the Panel does not estimate this to be a major conceptual difficulty. It may, however, have consequences in practice. A comparison of independent model results of the two approaches would also inform better about this aspect.

8.4 Meta-modelling

With respect to meta-modelling, the Panel remarks that the coarseness of the typology also has potential impact on the meta-modelling. It is clear that for the meta-modelling, some knowledge on the (un-sampled or under-sampled) systems must be used in order to set the best targets and use the appropriate slopes. As the typology used is rather coarse, the current choices may not be optimal for these systems. The Panel sees a role for statistical modelling here, provided that the statistical modelling would also focus on understanding and modelling cross-system differences (in slopes and consequently in targets) as a function of hydrographic and morphological characteristics of the systems. Particularly the importance of freshwater influence in the systems and the flushing rates may be overarching determining characteristics. The Panel estimates that a regression-based approach could be better than a classification approach.

In the opinion of the Panel, the meta-modelling of the North Sea water bodies is less reliable than that of the other water bodies. For the North Sea coastal systems, the background modelling, which has focused on the Baltic systems instead, is not very strong, and the meta-modelling is based on daring extrapolations from systems with quite different ecological characteristics. The Panel recommends that more study is made of the North Sea estuaries in order to improve the estimates of the nutrient load reduction requirements, based on their very different physical and ecological characteristics as well as on the very different basis (OSPAR intercalibration) for the references and targets.

9. Evaluation of Maximum Allowable Inputs results

Maximum Allowable Inputs (MAI) define the annual load of nutrients, in this context of nitrogen, that are acceptable to keep a coastal water body in a Good Ecological Status according to the WFD or allow a water body to return to this status. Since nutrient load management is a complex task and nutrient load reductions are associated with high costs, reliable overall and water body-specific MAI are of outstanding importance. In this chapter, the Panel reflects to what extent the suggested MAI can be regarded as reliable enough to form the basis for policy and management actions.

9.1 The overall Danish MAI in an international framework

The nutrient reduction scheme of the HELCOM Baltic Sea Action Plan was revised in the 2013 HELCOM Ministerial Meeting, based on a new and more complete dataset as well as an improved modelling approach. The new MAI, compared to the reference inputs of 1997-2003 for the Baltic Sea sub-basins Kattegat and Danish straits, demands only a minor load reduction requirement of about 3%. In this revision, Denmark agreed to reduce N loads to the Baltic Sea (from both land and air) by 2,890 t/a and P loads by 38 t/a. The Scientific Documentation Report suggests low N load reductions (>10%) for Western Jutland and most parts of Zealand as well as Lolland and Falster (Figure 8.23, p. 127). This seems reasonable and is well in agreement with international requirements.

However, to meet the targets for a good status, the Scientific Documentation Report demands much higher load reduction, especially on Funen and Jutland. Here, Denmark faces a situation similar to the Baltic coastal waters of Germany. Especially the inner coastal waters, estuaries and bays in Germany require higher N load reduction than demanded in the HELCOM Baltic Sea Action Plan to reach the GES. According to the German plans, the N load from German Baltic river basins has to be reduced by 21,500 tTN/a, with an average maximum allowed total N concentration in rivers of 2.5 mg/l, resulting in an overall reduction of 34%.

For Denmark, depending on the model approach, an average overall reduction between 29% and 34% is suggested. There are many similarities with respect to geomorphology, land use pattern and intensity as well as population and state of sewage purification between the German and Danish Baltic catchments, and the coastal waters share many similarities too. Therefore, the very good agreement in the assumed relative reduction requirements between both countries indicates that the values meet the right order of magnitude and seem reasonable.

However, the reliability of water body-specific MAI depends on the approach for calculating reference conditions and subsequent target conditions, the typology and type-specific targets, the considered indicators, the applied weighting, the model and meta-model approach as well as the data processing and aggregation. The major question is if all these aspects are sufficiently taken into account and if the application has a sufficient quality to determine reliable water body-specific MAI and mitigation needs to achieve the GES in Danish coastal waters.

9.2 Historic conditions as basis for target setting

The process in the Scientific Documentation Report follows the implementation guidelines of the WFD. It means that it is based on historic reference conditions and assumes that these conditions can serve as a basis for the definition of present and future targets. The reference conditions describe the status of biological quality elements that would exist in a situation with no or very minor disturbance from human activities. Reference conditions are therefore not pristine conditions. The WFD allows different methods to calculate reference conditions. In countries with long monitoring data records and the availability of suitable models, historic conditions are usually used as reference state. Because of data availability, this period often refers to a period around 1900, being aware that this period not always reflects a state with very minor disturbance from human activities. Similar to Germany, the Scientific Documentation Report uses the years around 1900 as reference. The Panel finds this approach well justified and the data basis sufficient and suitable.

However, it is obvious that between 1900 and today, land-use pattern and population densities have changed and different regions in Denmark developed differently until today. Further, the year 1900 is well suitable to reflect a high ecological status in rural areas, while cities already at that time emitted significant amounts of untreated sewage and caused pollution in their surroundings beyond the thresholds for a high ecological status.

For the definition of reliable targets, the question is less how did it look like in 1900, but rather how would reference conditions in a region look like, assuming present land-use and population pattern. This means that targets and water body-specific MAI based on historic conditions around 1900 bear uncertainties and for some water bodies may require a deeper analysis. This is especially true for areas with known strong changes between 1900 and today. However, the Panel agrees that this approach is the best choice that still ensures full compliance with technical WFD implementation guidelines.

In Germany, the official national working group on targets and MAI discussed if reference conditions should be calculated for and translated into the present situation. The approach was to use combined river basin and marine models and present population density and land-use pattern as well as the historic specific emissions per hectare and capita to calculate resulting regionalised Chlorophyll a and nutrients concentrations. The idea was to use the values as reference conditions to account for the fact that different regions developed differently during the last 120 years and to be able to provide even more reliable water body-specific targets. However, the majority of the working group declined this approach for containing too many assumptions and for not fully following the technical WFD implementation guidelines. Denmark would face similar problems with this alternative approach.

9.3 Effects of climate change on targets and MAI

Climate change shows its effects only gradually on a time horizon of decades, while the implementation of the WFD and measures to reach GES must take place within a decade. Further, depending on the emission scenario, climate change effects on countries and on regions within a country are uncertain, they show a large variability and are hard to predict. The Scientific Documentation Report addresses this topic and, in our opinion, provides sufficient evidence and reasons why climate change has not been taken into account in the definition of targets and in calculating MAI in Denmark.

However, several nutrient load reduction measures in river basins show the full effect only after decades. Major effects of climate change on Danish coastal waters, very likely, will result from changed nutrient loads as a result of altered spatial and seasonal precipitation and discharge patterns. Therefore, linked river basin – coastal water – sea models used for the assessment of the effectiveness of measures in the river basin should take into account climate change effects on river basin loads and shifts between nutrients. However, climate change can affect internal processes in coastal waters as well. Riemann et al (2016), for example, point out that more frequent stratification and higher water temperatures presumably hampered the improvement of bottom water oxygen conditions and counteracted the expected positive effects of reduced nutrient inputs in Denmark.

9.4 Relevance of typology on MAI

As indicated in Chapter 3, the Panel has the opinion that the Danish typology used in the Scientific Documentation Report does not sufficiently reflect the individual properties of the many Danish fjords and inner coastal waters. This is also true for the typology reported in Dahl et al (2005). Type-specific targets for the indicators, especially Chlorophyll a, that are applied to a wide range of significantly different water bodies do not sufficiently reflect their properties and behaviour to loads reductions. Consequences are less reliable water body-specific MAI. This may cause an underestimation of the required load reduction for some water bodies and an overestimation for others.

9.5 Relevance of indicator choice on MAI

The Panel agrees that Chlorophyll a is a core indicator, and coastal water body-specific Chlorophyll a concentrations are a sound basis for calculating water body-specific MAI. Further, the Panel agrees that water transparency has to be restored as one necessary condition to enable the recovery of eelgrass in coastal waters. Potentially, Kd can serve as an indicator for describing suitable growing conditions for eelgrass. Eelgrass can serve to indicate the status of macrophytes, a biological element in the WFD. Therefore, Kd has the potential to be an important parameter for calculating MAI.

However, as pointed out in Chapter 4, the relationship between Kd in coastal waters and external nutrient loading is sometimes very weak. Further, Kd and the insufficient relationship have different consequences for and are differently treated in the mechanistic and the statistical modelling exercises. In the statistical modelling approach, for example, the use of Kd in some cases causes impossible N load reduction requirements of above 100%. Further, Kd shows only a slow response to load reduction, the data are subject to high variability, and it shows a correlation to Chlorophyll a. Altogether, we consider Kd as a less suitable indicator in many Danish coastal water bodies. A strong weight of Kd in the calculation of MAI should be avoided and would add uncertainty to water body-specific MAI. Chapter 4 outlines possible solutions to overcome or at least to deal with some of these problems. In the Scientific Documentation Report, other indicators are sometimes mentioned and used in the statistical model. We do not see a major advantage of these indicators for the calculation of MAI, because they do not provide significant new information or show correlations to the exiting indicators.

9.6 Relevance of model quality and approach for MAI

In general, the mechanic model has a very good potential for calculating water body-specific MAI, but in the present state it does not cover all water bodies. The statistical modelling is based on real monitoring data, and in most coastal water bodies it can serve as a valuable tool to assess long-term trends as well as the mechanistic model performance. As indicated in Chapter 8, the model application and the process of calculating water body-specific MAI are complex and not entirely convincing. Most problematic is the averaging of Chlorophyll a reference values across both models and within coastal water types. This has negative consequences for the meta-modelled water bodies as well.

9.7 Conclusion and perspectives

Many of these aspects and shortcomings were mentioned and pointed out by several stakeholders as well. The Panel picked up the stakeholder comments and examined in some detail the MAI for specific areas with very high nutrient load reduction demands. Altogether, the Panel largely shares the stakeholder concerns.

The calculation of water body-specific MAI is a challenging task, but potentially has one major advantage: It allows the development of water body-specific management options and solutions. For this purpose, the coastal water and sea models should be combined with river basin models providing information about the quantitative potential and efficiency of single (or sets of) measures and providing load reduction scenarios for coastal models. If river basin models are able to provide nutrient load data on a monthly basis, this would allow the development of scenarios that take into account the seasonality of emissions. Assessing how seasonally differentiated emissions affect the status of coastal water bodies could lead to optimised, cost-effective management.

Taking into account all aspects and associated problems, the Panel has the impression that the water body-specific MAI are not sufficiently reliable to serve as a basis for decision-making and planning of load reduction measures. Further, the MAI are only addressing nitrogen load reductions and leaving out the possibility of potentially managing water bodies via phosphorus load reduction. However, models, competences and data are available in Denmark to meet the challenge to calculate water body-specific MAI. Even a modified processing of the existing model results might lead to much more reliable MAI.

10. Overall assessment and conclusions

The Water Framework Directive aims at restoring Good Ecological Status in surface waters in Europe. The Scientific Documentation Report proposes measures of nutrient load reduction to reach this Good Ecological Status in Danish transitional and coastal waters. The Panel fully endorses the importance attached to nutrient reductions as a necessary requirement to reach this Good Ecological Status and stresses the importance of nutrient conditions as a modulating factor interacting with any additional measures taken to improve the state of the system.

In comparison with many other European countries, Denmark has excellent databases, models and scientific expertise as a basis for the implementation of the Water Framework Directive. The Panel was delighted to see that these resources have been mobilised to achieve a leading position at the European scale. The Panel was impressed by the openness and transparency of the interaction between government, researchers and stakeholders as well as by the high intellectual level of the discussions. This open exchange of ideas and opinions is a perfect basis for a further improvement of the scientific basis for the WFD implementation.

The Panel has reviewed the choice of indicators and procedures, in the context of the WFD requirements and specifications, and found that the indicators, the methods to determine reference conditions and the methods to determine required actions were WFD compliant. The Danish implementation is based on either direct historical observation or model determination of reference conditions. Little or no uncontrollable “expert judgement” is involved. In that respect, the Danish models are attaining the highest possible standard of WFD implementation.

The Panel has analysed the consequences of using a relatively coarse typology of coastal waters for calculating reference conditions, targets and Maximum Allowable Inputs of nitrogen. The Panel concludes that the use of a coarse typology has led to reduction requirements that are not optimal for each of the individual water bodies. The Panel is convinced that the full use of available data and models would allow Denmark to forego the typology and develop advanced, specific reduction targets for each water body. The Panel recommends focusing on the water body scale of resolution throughout the scientific process. The regional grouping of reduction measures should be decided upon only at the stage of translating scientific advice into management action plans.

The Panel has analysed the indicators used and concluded that Chlorophyll a is a useful intercalibrated indicator of phytoplankton, while Kd is less optimal as an indicator of benthic angiosperms and macrophytes. The other indicators, used in the statistical modelling only, currently present methodological problems and are not yet mature enough for inclusion in the management plans. The Panel has identified promising developments in the modelling with respect to angiosperm and macrophyte indicators and made recommendations on how to extend and develop the indicator set in the future.

In view of the large efforts in the past to remove P load from point sources, the Panel endorses the emphasis placed in the Scientific Documentation Report on reducing N loads from diffuse sources. However, at least in principle, there could be an additional role for P load reduction and for seasonal regulation of the N load. The Panel is of the opinion that these options merit further scientific exploration, especially in watersheds where high efforts for N load reduction are required.

Although the maintenance of two parallel modelling lines (statistical and mechanistic) may seem redundant at first sight, the Panel strongly endorses maintaining these lines. Given the wealth of data available, it provides unique possibilities for evidence-based checking of mechanistic model results. The Panel assesses the mechanistic model as a state-of-the-art, very comprehensive tool, but emphasises that independent checking on data as well as uncertainty analysis remain necessary and can be performed by the statistical approach. This coherence can be optimised by improving the approach and methods of the statistical modelling.

The Panel endorses the general logic of the methodology to derive reference and target values from the models and to calculate the required N load reduction to reach the targets. The Panel has identified several points in the workflow where averaging is performed. This results in interdependence of model types, loss of indicator resolution and loss of spatial resolution. It also adds complexity to the procedure and makes it very difficult to understand. None of these losses are necessary since the model results and database do permit a fully transparent derivation of water body-specific required nutrient reduction.

Summing up these different aspects of the work, the Panel positively evaluates that nutrient load reductions are based on **solid** scientific evidence and generally high-level modelling approaches. The Panel is very positive about the near lack of expert judgment in the work and is of the opinion that in the few places where it does occur, it is not necessary and can be removed. The general (country-averaged) level of required nutrient load reduction compares favourably with independent efforts in similar areas and seems a **robust** measure of what is needed. At the same time, the Panel assesses the spatial resolution of the required efforts as **unnecessarily coarse**. The Panel is convinced that the rich database, combined with an **improved statistical approach** and the high-resolution mechanistic modelling tools, are able to derive improved, water body-specific MAI values. Current scientific insight endorses the view that the overall reductions proposed are **necessary**, but cannot guarantee that they will be **sufficient**. Especially for benthic angiosperms and macrophytes, additional measures may be needed.

11. Recommendations for going further

Monitoring: The Danish national monitoring programme used in the Scientific Documentation Report includes more than 90 stations along the coast and in the sea. It is very comprehensive and is generally well adjusted to the WFD requirements. It forms the basis for the further development of models, for most calculations and is required to evaluate the success of measures and whether the targets of the WFD are met. The Panel recommends maintaining this monitoring system at full strength and assessing if additional monitoring stations will be required for a water body-specific management.

Typology: The typology has weaknesses in reflecting the individual properties of fjordic water bodies. Instead of suggesting a refinement of the existing typology, we recommend calculating reference conditions and targets for each of the 119 water bodies in Denmark. Denmark is one of the few countries in Europe, where the necessary data, expertise and models are available for such a comprehensive approach. By taking specific conditions and individuality of every water body into account, the calculated targets and water body-specific Maximum Allowable Inputs will be optimised and lead to minimal waste of resources. For purposes of intercalibration, a robust typology can be based on the results of the water body-specific analyses.

Choice of indicators: Chlorophyll a is a generally accepted and intercalibrated indicator of phytoplankton. Kd, as a measure for macrophytes and angiosperms, has certain limitations. The Panel recommends building on recent efforts towards comprehensive modelling of eelgrass in order to derive a better indicator of macrophytes, but to keep Kd as a proxy meanwhile. The other indicators used in the statistical modelling address important ecological questions, but are not mature in the sense that they lack a clear quantitative relation with nutrient loading. The Panel recommends leaving them out of the present modelling and developing targeted modelling directed at their incorporation into the indicator system.

Statistical modelling: The Panel sees great merit in the strategy to maintain two independent lines of modelling, one based on statistical data analysis and the other based on mechanistic modelling. The Panel recommends reorienting the statistical modelling towards optimal estimation of the long-term slopes of the indicators on nutrient loading in a cross-systems analysis way and keeping in principle both N and P loading as explanatory variables. The Panel recommends elaborating the uncertainty analysis in the statistical modelling and suggests that this will be facilitated when a single cross-system advanced modelling approach is chosen.

Mechanistic models: The mechanistic models are state-of-the-art, both in terms of numerical technique and included processes. They are powerful tools for providing a sound scientific basis for the implementation of the WFD in Denmark. A shortcoming is that they do not cover all water bodies. As a consequence, different approaches were used for the definition of reference conditions, targets and MAI in different water bodies. We recommend extending a mechanistic modelling approach to as many water bodies as possible to ensure that, in future, a uniform methodology can be used for the definition of water body-specific MAI.

Methods to derive targets and MAI from the models: The Panel recommends simplifying the calculation procedure by removing the averaging steps between models, between indicators, between water bodies within types and between water bodies on a regional basis. In this way, the differences and correspondences between modelling approaches, indicators and water bodies will become clear and can be further analysed. Cross-checking of results of the statistical and mechanistic model approaches in systems, where both are available, will form a basis for extrapolation to all systems. The Panel recommends deriving one MAI per water body in this way and only deciding in a later phase on regional averaging or lumping, when scientific results are translated into management actions.

River basin interactions: River basin models allow calculating the load reduction potential of nitrogen and phosphorus for each river basin, the development of water body-specific nitrogen and phosphorus load reduction scenarios and cost estimates. Further, they allow addressing seasonal load and limitation patterns. The Panel recommends a combination of river basin and coastal water models to enable the development of water body-specific optimised management concepts that consider both nitrogen and phosphorus.

International approach: The technical WFD implementation guidelines force similar approaches in all member states. As a consequence, requirements, modelling and challenges are similar in different countries. Further, the WFD asks for an intercalibration and harmonisation of targets with neighbouring countries. Therefore, the Panel recommends a co-ordinated joint scientific approach, especially between Denmark, Germany and Sweden.

12. List of references

Andersen, J.H., Harvey, T., Kallenbach, E., Murray, C., Al-Hamdani, Z., Stock, A. (2017). Under the Surface. Report L.NR. 7128-2017 DK6 by NIVA Denmark.

Birk, S., Willby, N.J. , Kelly, M.G., Bonne, W., Borja, A., Poikane, S. , van de Bund, W. (2013). Intercalibrating Classifications of Ecological Status: Europe's Quest for Common Management Objectives for Aquatic Ecosystems. Sci Total Environ 454-455, 490-499.

Borja, A., Prins, T.C., Simboura, N., Andersen, J.H., Berg, T., Marques, J.C., Neto, J.M., Papadopoulou, N., Reker, J., Teixeira, H. and Uusitalo, L., 2014. Tales from a thousand and one ways to integrate marine ecosystem components when assessing the environmental status. Frontiers in Marine Science, 1, p.72.

Canal-Vergés P, Petersen JK, Rasmussen EK, Erichsen AC, Flindt MR (2016) Validating GIS tool to assess eelgrass potential recovery in the Limfjorden (Denmark). Ecological Modelling 338:135-148. http://dx.doi.org/10.1016/j.ecolmodel.2016.04.023.

Christensen et al. (subm.) *Referred to in the Scientific Documentation Report, but not found back*

Cloern, J.E. (2001). Our evolving conceptual model of the coastal eutrophication problem. Marine Ecology Progress Series, 210, 223-253.

Dahl K (red.), Andersen JH (red.), Riemann B (red.), Carstensen J, Christiansen T, Krause-Jensen D, Josefson AB, Larsen MM, Petersen JK, Rasmussen MB, Stand J (2005) Redskaber til vurdering af miljø- og naturkvalitet i de danske farvande. Typeinddeling, udvalgte indikatorer og eksempler på klassifikation. Danmarks Miljøundersøgelser. 158 s. - Faglig rapport fra DMU nr. 535.

Droop, M.R. (1968). Vitamin B12 and marine ecology IV: the kinetics of uptake, growth and inhibition in Monochrysis lutheri. J. Mar Biol. Ass UK, 48, 689-733.

Eilola, K., Gustafsson, B. G., Kuznetsov, I., Meier, H. M., Neumann, T., & Savchuk, O. P. (2011). Evaluation of biogeochemical cycles in an ensemble of three state-of-the-art numerical models of the Baltic Sea. Journal of Marine Systems, 88, 267-284.

Flindt, M. R., Rasmussen, E. K., Valdemarsen, T., Erichsen, A., Kaas, H., & Canal-Vergés, P. (2016). Using a GIS-tool to evaluate potential eelgrass reestablishment in estuaries. Ecological Modelling, 338, 122-134. http://doi.org/10.1016/j.ecolmodel.2016.07.005

HELCOM. (2013). Summary report on the development of revised Maximum Allowable Inputs (MAI) and updated Country Allocated Reduction Targets (CART) of the Baltic Sea Action Plan. Background Document for the 2013 HELCOM Ministerial Meeting, 1-23.

Kuusemäe, K., Rasmussen, E. K., Canal-Vergés, P., & Flindt, M. R. (2016). Modelling stressors on the eelgrass recovery process in two Danish estuaries. Ecological Modelling, 333, 11-42. http://doi.org/10.1016/j.ecolmodel.2016.04.008

Malve, O., Qian, S. 2006. Estimating nutrients and chlorophyll a relationships in Finnish Lakes. Environmental Science & Technology, 40(24),pp. 7848-7853. DOI: 10.1021/es061359b.

Markager S, Storm L, Stedmon C (2006) Limfjordens miljøtilstand 1985 til 2003. Sammenhæng mellem næringsstoftilførsler, klima og hydrografi belyst ved hjælp af emperiske modeller. Danmarks Miljøundersøgelser. 219 s. - Faglig rapport fra DMU, nr. 577.

Orth RJ, Moore KA, Marion SR, Wilcox DJ, Parrish DB (2012) Seed addition facilitates eelgrass recovery in a coastal bay system. Marine Ecology Progress Series 448:177-195.

Rasmussen, B., & Josefson, A. B. (2002). Consistent Estimates for the Residence Time of Micro-tidal Estuaries. *Estuarine, Coastal and Shelf Science*, *54*(1), 65-73. http://doi.org/10.1006/ecss.2001.0836

Riemann, B., Carstensen, J., Dahl, K., Fossing, H., Hansen, J. W., Jakobsen, H. H., et al (2016). Recovery of Danish Coastal Ecosystems After Reductions in Nutrient Loading: A Holistic Ecosystem Approach. *Estuaries and Coasts*, *39*(1), 82-97. http://doi.org/10.1007/s12237-015-9980-0

Scavia D, Rabalais NN, Turner RE, Justić D, Wiseman WJ (2003) Predicting the response of Gulf of Mexico hypoxia to variations in Mississippi River nitrogen load. Limnology and Oceanography 48:951-956. doi:10.4319/lo.2003.48.3.0951

Schernewski, G., Friedland, R., Carstens, M., Hirt, U., Leujak, W., Nausch, G., Neumann, T., Petenati, T., Sagert, S., Wasmund, N., von Weber, M. (2015). Implementation of European marine policy: New water quality targets for German Baltic waters. *Marine Policy*, 51, 305-321. https://doi.org/10.1016/j.marpol.2014.09.002.

Timmermann, K., Dinesen, G. E., Markager, S., Ravn-Jonsen, L., Bassompierre, M., Roth, E., & Støttrup, J. G. (2014). Development and Use of a Bioeconomic Model for Management of Mussel Fisheries under Different Nutrient Regimes in the Temperate Estuary of the Limfjord, Denmark. *Ecology and Society*, *19*(1), art 14-11. http://doi.org/10.5751/ES-06041-190114

van der Heide T, van Nes E., van Katwijk MM, Olff H, Smolders AJ (2011) Positive Feedbacks in Seagrass Ecosystems – Evidence from Large-Scale Empirical Data. PLoS ONE 6(1): e16504.