General response.

When studying the reactions of researchers and stakeholders, the Panel has realized that a few factual errors occurred in its report, and that furthermore the choice of wording was not always entirely precise and consistent, with respect to the difference between water body specific, type-specific, and regionalized references, targets and MAIs. In order to avoid further confusion with other readers of the report, the Panel has decided to slightly edit the report. The changes, compared to the version commented upon by the stakeholders and researchers, are summarized in the table below.

Page	Original text	New text
9	Target values must fall in the green (GES) range. However, the error bars of the calculated target value should also fall in the green or blue area. Therefore, the longer the error bars, the more the target moves to the left and is consequently more stringent.	Sentence dropped, as it gave rise to discussions that cannot be finally settled
10		The fact that eelgrass depth limit (not Kd) has been intercalibrated has been added to the text.
14	To derive reliable regionalized MAI	to derive reliable MAI for each water body.
15	when the aim is to calculate regionalised MAI.	.. when the aim is to calculate as precise and water-body specific MAI as possible
17	Reference link lost	Proper reference to figure restored
23	One sentence had dropped out from the text	Add: “Moreover, spatial displacement of problems to other systems as a consequence of flushing winter nutrient loads has to be taken into account.”
27	Further, the representation of the sediments does not include redox-dependent inorganic phosphorus (iron-oxide bound) dynamics and an empirical direct relationship between shear stress and turbidity.	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.
32	Reference link lost	Proper reference to table restored
34	Reference link lost	Proper reference to figure restored
37	Regionalized MAI (3 times)	Water-body specific MAI (3 times)
38	Regionalized MAI	Water-body specific MAI
39	Regionalized MAI (5 times)	Water-body specific MAI (5 times)
40	Regionalized MAI (2 times)	Water-body specific MAI (2 times)
43	Regionalized MAI (2 times)	Water-body specific MAI (2 times)
45	List of references	One reference corrected, several missing references added

Specific response.

In the table below we respond to the questions and comments brought forward during the hearing process.

Organization Stakeholder	Ch/ sect	Comments	Question	Response to question
Landbrug & Fødevarer	General	We would like to thank the Panel for the thorough and skilled work that has been put in this evaluation. We are happy to note that comments from stakeholders have generally been answered and, where necessary, investigated further. This enquiring and persistent attitude towards understanding the scientific matter is much appreciated.		Thanks for the compliments
Landbrug & Fødevarer	2.5	The importance of looking at all relevant stressors to recover degraded marine ecosystems is supported by overwhelming scientific evidence. Nutrients are central,but addressing other stressors in addition to nutrients is important, often essential, to reach the environmental objectives. One example of successful recovery is changing sluice practice, and thereby salinity, of Ringkøbing Fjord. This facilitated recruitment of the suspension-feeding soft-shell clam, Mya arenaria, resulting in a regime shift leading to clear water (Petersen et al. (2008)1). A negative example of ecological disturbance is invasion by the round goby (Neogobius melanostomus) that may disturb the ecosystem e.g. by eating benthic filter feeders (e.g. Poos et al. (2009)2 and Kuhns & Berg (1999)3), causing reduced water filtration. The effect of this will be increased eutrophication signal, even though total N load may be unchanged or even decreasing. Cloern (2001)4 specifically notes that "Changes in coastal water quality and living resources are the result of multiple stressors, so a broader view of coastal eutrophication will consider how anthropogenic nutrient enrichment interacts with other stressors such as translocation of species, habitat loss, fishing, inputs of toxic contaminants, manipulation of freshwater flows, aquaculture, and climate change." This classic paper, thus, emphasizes the importance of taking action against other stressors than	Does the Panel agree that other stressors than direct nutrient load, e.g. invasive species and increasing temperatures due to climate change, may increase the negative effects of eutrophication or directly deteriorate the ecosystem? And does the Panel agree that other actions than reduction of nutrient load from land may improve a degraded ecosystem? And may even be necessary as part of a holistic approach to obtain good ecological status? Does the panel agree that more stressors than nutrient load must be addressed in order to obtain good environmental status of Danish coastal waters? Several stressors may affect the same environmental indicators. E.g. chlorophyll a level will increase with increased nutrient load, but also if e.g. the population of benthic filter feeders is strongly reduced or if temperature increases due to climate change. Does	The Panel has primarily reflected on the Scientific Documentation Report, which focuses on nutrients (mainly N) and indicators of water quality (mainly Chlorophyll a and Kd). The discussion brought forward here distracts from that topic, and touches upon elements that are treated by the Marine Strategy Framework Directive rather than by the Water Framework Directive. The Panel has not been able to fully reflect on these issues. However, in general terms the Panel stressed the importance of *interaction* between nutrient loads and many other stressors in section 2.5 of its report. A consequence of interaction is that other measures to improve the ecological status are unlikely to be successful if nutrient loads, as the primary factors affecting the status of the ecosystem, are not well controlled. Non-linearity of ecological reactions is a real possibility, but as the Panel has stressed in some of its responses to questions from the stakeholders, the Panel does not see reasons to suspect that the current models *over*predict the required nutrient reductions as a consequence of this. In fact, the contrary is more probable. The Panel recommends management on the basics, i.e. the nutrient loads, accompanied with good monitoring that can advise on other needed measures where unexpected state changes appear.

		nutrient load alone. It is supported by Andersen et al. (2017)5, who states that nutrients (N and P together) make up only 50 % of the total stress in the Danish WFD areas. The Panel notes that "other legal instruments, e.g. the Marine Strategy Framework Directive […] include more explicitly other stressors […]", and only P is mentioned in the evaluation report as an important, additional stressor to include in the RBMP. Given the strong arguments provided in both Cloern (2001) and references therein, and the specific study by Andersen et al. (2017), it is difficult to understand the logic in the Panel's conclusion that other stressors than nutrients can be taken care of elsewhere. Furthermore, it is a fact that other stressors are not addressed, or sufficiently addressed, by other legal instruments. Examples: No action at all is taken against invasive species in Danish marine waters, coastal or open. Specific actions to assist in eelgrass reestablishment in Danish fjords can hardly be done elsewhere. The situation now is that all action to improve ecological status is taken against nutrient loss from land. 1 Petersen et al. (2008): Regime shift in a coastal marine ecosystem. Ecological Applications 18:497–510 2 Poos et al. (2009): Secondary invasion of the round goby into high diversity Great Lakes tributaries and species at risk hotspots: potential new concerns for endangered freshwater species. Biological Invasions 12: 1269 3 Kuhns & Berg (1999): Benthic Invertebrate Community Responses to Round Goby (Neogobius melanostomus) and Zebra Mussel (Dreissena polymorpha) Invasion in Southern Lake Michigan. Journal of Great Lakes Research 25(4):910–917 4 Cloern (2001): Our evolving conceptual model of the coastal eutrophication problem. Marine Ecology Progress Series 210: 223-253 5 Andersen et al. (2017): Under the Surface. Report by NIVA Denmark. Under preparation for submission.	the Panel agree that if only one of several stressors affecting the same environmental indicator is addressed in the modelling work, this stressor will have to be reduced relatively more in pursuance of the target? I.e. "pay" for the damage done by other stressors?	
Landbrug & Fødevarer	4.1	The Panel describes the feedback mechanism of eelgrass cover on water quality and clarity and mentions other factors of disturbance to eelgrass recovery. The Panel clearly states that there is no strong dependence of Kd on	Does the Panel agree that other measures should be taken along with nutrient load reductions to obtain eelgrass reestablishment most	The Panel has extensively commented on this in its report. It is quite well possible that additional measures will be needed, and the Panel has recommended pursuing the studies in these complex ecological interactions. The Panel

		nutrient loading in the selected time period (p. 17, end of third section). We suggest adding fishery and crabs to the list of eelgrass stressors, e.g.: Baden et al. Limnol. Oceanogr., 55(3), 2010, 1435–1448 Relative importance of trophic interactions and nutrient enrichment in seagrass ecosystems: A broad-scale field experiment in the Baltic–Skagerrak area doi:10.4319/lo.2010.55.3.1435 Baden et al, Mar Ecol Prog Ser 451: 61–73, 2012, Shift in seagrass food web structure over decades is linked tooverfishing. doi: 0.3354/meps09585 Moksnes et al. Oikos 117: 763777, 2008 Trophic cascades in a temperate seagrass community. doi: 10.1111/j.2008.0030-1299.16521.x Infantes E et al. (2016) Seed Predation by the Shore Crab Carcinus maenas: A Positive Feedback Preventing Eelgrass Recovery? PLoS ONE 11(12): e0168128. oi:10.1371/journal.pone.0168128 K. Matheson et al, Mar Ecol Prog Ser 548: 31–45, 2016. Linking eelgrass decline and impacts on associated fish communities to European green crab (Carcinus maenas) invasion. doi: 10.3354/meps11674 Neckles, Loss of Eelgrass in Casco Bay, Maine, Linked to Green Crab Disturbance, Northeastern Naturalist 2015. http://dx.doi.org/10.1656/045.022.0305 Garbary et al. Mar Biol (2014) 161:3–15. Drastic decline of an extensive eelgrass bed in Nova Scotia due to the activity of the invasive green crab (Carcinus maenas). DOI 10.1007/s00227-013-2323-4	efficiently?	also suspects that responses of Kd to nutrient reductions may be substantially slower than responses of Chlorophyll a, and recommends further study into these processes in order to better understand the long-term processes leading to eelgrass restoration. Note that the cited papers on the influence of *Carcinus maenas* are from North America, where the species is invasive and has other ecological interactions and consequences than in Europe, where it is native and has always lived in the same areas where eelgrass occurred.
Landbrug & Fødevarer	6.2	The reference Malve & Qian (2006) is missing in the references. We assume it is "Estimating Nutrients and Chlorophyll a Relationships in Finnish Lakes"		Apologies. The missing reference has been added to the final version of the report
Landbrug & Fødevarer	6	No conclusions are made in chapter 6, on the statistic models, as section 7.6 on the mechanistic models. For further work, concluding remarks on the statistic models to enable comparison with the conclusions regarding mechanistic models will be necessary. We therefore encourage the Panel to include a section 6.4 "Conclusions on the statistic models"		The Panel has tried to homogenize the structure of the different chapters, but slight differences remain. However, the Panel is of the opinion that assessment and recommendations regarding the statistical modeling are clear in the report. The Panel has made slight editorial improvements in the report, but refrains from changing any important parts such as conclusions, in order to respect the hearing process.
Landbrug &	10	In the final conclusions, clear references are made to all		The Panel feels that the most important findings of all

Fødevarer		chapters except from chapters 6 and 9, which include specific comments on setup and results of the statistic modeling approach and on the fitness of MAI's to serve as basis for deciding measures at a regional level, respectively. We find that the concluding chapter should include references to all relevant chapters of the report to ensure a full overview of findings and an appropriate balance of the final conclusions.		chapters are included in the overview chapter
DHI	2.1	p.9: Target values must fall in the green (GES) range. However, the error bars of the calculated target value should also fall in the green or blue area. Therefore, the longer the error bars, the more the target moves to the left and is consequently more stringent.	We acknowledge the logic behind this statement, and would like to know if any reference to this passage can be found in the literature? Are the panel of the opinion that we should always include the maximum reduction target, and add uncertainties, to ensure the one-out-all-out principle?	This argument was based on logic rather than on a strict directive. If a model predicts that good environmental status will only be reached with a probability of 50 % (as half the uncertainty range lies above the limit), this cannot be interpreted as 'ensuring good ecological status'. However, the statement has been removed from the report as it may lead to discussion and is not strictly necessary in this context.
DHI	2.2	p.9: 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 intercalibrated. p.10: Kd has not been intercalibrated (as confirmed by the researchers from Aarhus University (DCE) and DHI and the European Commission's Joint Research Centre) Kd is not intercalibrated – but is derived (including the uncertainties mentioned by the panel) from the intercalibrated eelgrass depth limit. We acknowledge the uncertainties of this parameter, but just want to state, that it is an attempt to transfer the intercalibrated indicator to an operational model indicator.		The Panel acknowledges that Kd is a proxy for eelgrass depth limit, which is an intercalibrated indicator. As such, Kd is a supporting quality element for the Eelgrass biological quality element, but it has not been intercalibrated in itself. For that reason one should discuss whether it should have equal weight as Chlorophyll a. The intercalibrated status of eelgrass depth limit has been added to the text of the report to clarify the point.
DHI	3.3	p.14: The typology is too simplified to reflect the specific characteristics of the individual fjordic water bodies. The		The Panel is delighted to read this

		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.We acknowledge this comment, and are in the process of defining a project that will assess the typology applied for the RBMP 2015-2021. Based on circumstances behind the project some choices were made during development, but we agree that the typology should have been more detailed, and the comments from the panel will be incorporated in the project definition for the coming project update and pros and cons evaluated.		
DHI	3.6	p.15: 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.	We do not disagree, but will like to state that this approach requires high quality models and observations for all water bodies. Even in the German case, some water bodies are covered by parts of the model with less sufficient resolution which, to our opinion impact the reference and targets values. For the Danish water bodies we have evaluated where we have models of sufficient quality to be used for the assessment of reference and targets values, but for some water bodies we do not have models or few/no observations. What does the panel recommend for these water bodies?	The Panel is aware of this situation. We have recommended to revise and, where necessary, extend the monitoring and observation data base, and to extend, where possible, the mechanistic modeling. Furthermore, we have stressed the importance of a cross-systems statistical approach, because we are convinced that essential characteristics, such as slope of Chlorophyll a versus N loading, can be predicted based on appropriate measures of the hydrography, the local loading intensity and other characteristics of the water body. The German mechanistic model approach covered nearly every water body, but because of the problems you mentioned, the average model performance was less good compared to the Danish approach. However, the German scientists and the stakeholder group saw that the model well reflected the relative changes between today and the situation around 1880. This relative change was applied to the concrete present monitoring data to calculate reference conditions. This is a major difference between the German and the Danish approach. It shows that always different solutions are possible to reach an aim.
DHI	3	The typology is too course, and for the work ahead we will explore the possibilities for improving this part		see 3.3

		significantly. See comment above (section 3.3)		
DHI	4.1	p.18: 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. p.19: It further recommends pursuing studies attempting to estimate conditions for seagrass restoration based on already developed more comprehensive models.	This is of course an ideal approach, but still we see plenty of unsuccessful attempts to restore eelgrass meadows. Some examples exists, but still more failures exists. Furthermore,, we did try to translate depth limit into a model output, but as actual depth limit and modelled depth limit does not necessarily coincide this is a very challenging task. How does the panel suggest we proceed if the mechanistic models does not succeed in describing eelgrass depth limit, and should we work with other indicators for aquatic macrophytes and angiosperms than depth limit? And if the panel have suggestions, which ones? See above	The panel is aware that this is not an easy task, but appreciates the considerable efforts already put in the cited models. As stated elsewhere, we suspect that the time constant for adaptation of Kd to nutrient loading is considerably longer than that of Chlorophyll a; we are of the opinion that this aspect may be key to better understanding (and modeling) of the relations. For this reason, the Panel has recommended to pursue the studies, but to reduce at present the importance of Kd in the weighting.
DHI	4.2	p.20: 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.	We are presently working on a project description focusing on supporting parameters like total N, total P and oxygen. These should be applicable for both statistical and mechanistic model approaches. However, none of these will be intercalibrated. How does the panel suggest we introduce none-intercalibrated indicators for setting final reduction targets?	Intercalibration can only follow once the indicator is well worked out in at least one country. The Panel feels that it is important to continually enrich the practice of the WFD with new knowledge and modeling tools, based on the overall aim of the directive to reach good ecological status. For this reason it recommends following a research line into new, ecologically significant indicators, where the possibilities for this exist. The Panel acknowledges, however, that the WFD has legal status with respect to both the goals and the ways to reach them. The Ministry is the Competent Authority and answerable if it does not implement measures when water bodies fall below the G/M boundary. Therefore it is important to maintain an operational line based on intercalibrated indicators that can form the basis for current management, while not arresting scientific development into improved ways of establishing and reaching the targets.
DHI	5.1	p.21: These field studies suggest that at least in a number of systems, regulation of annual primary production by P load reduction could be feasible.	We fully acknowledge that regulating P can affect e.g. yearly primary production, but as this is not an indicator we find it	The Panel has also indicated that a Chlorophyll a indicator that also covers the spring bloom would be more sensitive to P-loading. On the other hand, P fixation and

			difficult to include P reduction targets based on this. However, a research project lead by AU has been initiated focusing on P measures (catchment scale) and initial studies to allow for future introduction of P sensitive targets.	sedimentation in spring and subsequent release in summer may also partly explain the dominance of N limitation in summer. This could mean that reducing P, through reduction of the spring bloom, might also influence summer Chlorophyll a, probably after a latency period of a few years. The models should allow testing this hypothesis, and subsequent action can be based on the outcome of these tests.
DHI	5.3	p.22: 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. p.23: The Panel recommends using basin load models in combination with the mechanistic models used in the Scientific Documentation Report to investigate these possibilities. For sure this could provide some suggestions for specific water bodies and should be associated with an in-depth analysis of observations at first to locate potential areas where this could add value.	We fully agree, but as no year-averaged chlorophyll indicator has been defined (and intercalibrated) it is difficult to conduct any regulation based on this. Still summer chlorophyll-a will not disappear as it is intercalibrated with Sweden and Germany, but could be supplemented with P-reduction targets. Any comments to this?	There is not really a difference of opinion on this. See also previous remarks and section 4.2 of the report.
DHI	7.1	p.27: Further, the representation of the sediments does not include redox-dependent inorganic phosphorus (iron-oxide bound) dynamics Small misunderstanding – iron bound P is included in the IDW model, see page 74 in previous documentation. It's the adsorption to inorganic particles and potential sedimentation/re-suspension that is not included.		The Panel apologizes for this misunderstanding. It was not always easy to fill the gap between the summary description in the Scientific Documentation Report and the full documentation of the process formulations in the underlying reports. The fact that iron binding was included is positive and increases the value of the mechanistic models for investigating hypotheses regarding P. The Panel's report has been corrected by changing this remark.
DHI	8.2	p.35: 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. We acknowledge this opinion and try to analyze this in the work ahead to evaluate the differences and similarities p.35: However, such decisions could better be made on		Nobody can guarantee the quality of a procedure that has not yet been tested, but it is good to see that the suggestion will be part of future investigations

		the basis of a map showing the original model results, allowing one to judge whether a management problem is posed or not. Valid point – and we will take this into account ahead, although we will not conclude if it still will be appropriate to average or not, but the map would help in the analysis Point taken, and we will evaluate weather this will be valid when continuing working with the RBMPs p.36: Summarising, the Panel recommends postponing the averaging operations to the very last stages of the procedure.		
DHI	8.4	p.36: 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 This is also the opinion of the researchers and a project has been launched focusing on developing a mechanistic model covering the North Sea water bodies.		Good!
DHI	9	p.40: Taking into account all aspects and associated problems, the Panel has the impression that the regionalised 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 regionalised MAI. Even a modified processing of the existing model results might lead to much more reliable MAI.	Is it correctly understood, that you are concerned by the MAIs due to the course typology and early averaging, but that you basically acknowledge the models and methods, and suggest updating according to your recommendations to achieve a stronger basis for setting the reduction targets?	We apologize for slightly confusing language in our report. We have adjusted the text to better express our view, which indeed corresponds to your interpretation
DHI	11	We basically acknowledge and agree on your suggestions for future work		Good!
Bæredygtigt Landbrug	Comments to the evaluati	First of all, we wish to acknowledge the great effort that has gone into the International Evaluation of the Danish Marine Models. We consider the work essential and hopefully it will have substantial influence on the Danish River Basin Management Plans.		Thanks for the compliments

	on proc ess			
Bæredygtigt Landbrug	Gen eral	As we understand it, the conclusion of the Evaluation is that on an overall level, you find that Denmark has a very solid data basis and that the statistical and mechanistic model approach is a conceptual example to be followed.		That is right, but at the same time, we think more is possible with these models and this knowledge, and that improvements can and should be made
Bæredygtigt Landbrug	2.1	Reference Year 1900 The evaluation panel states that model calculations are better than expert opinions: "[… [page 9] The Danish approach relies on modelling and a 1900 baseline, since there are no pristine ystems that can be used as a reference. This approach is appropriate, WFD compliant and better than only using expert judgement. [...]" Furthermore, the panel states: "[...] [page 21] Currently, the overall inputs of N and P re roughly about 4.2 and 3.4 times higher, respectively, than estimated reference inputs for the year 1900 (Riemann et al, 2016). [...]" Unfortunately, it has not been possible to find the referral source in the List of References for us to see the background for the numbers mentioned. In fact these "ancient" nitrogen and phosphorus concentrations in watercourses emptying into fjords and coastal waters are estimated by using expert judgement. (No data exist for year 1900). Contemporary nutrient concentrations measured in streams draining uncultivated soil ("naturvandløb") are not suitable as the majority of Danish land was cultivated, and fertilizer (manure) was used 100 years ago. Therefore, one cannot assume that conditions in "naturvandløb" directly can represent streams draining farmed land more than 100 years ago!	Regardless the method, we would advise the panel to consider making a note on the considerable uncertainty of the termination of the reference year. In this connection we refer to the panel's own point that instead of taking only a single stress factor into consideration, it is the combination of a number of factors that is important. Furthermore, the panel has stated that the Danish typology ought to be segmented further and that the present limit of chlorophyll a of 3.6 mg per m3 should be determined individually and under consideration of for instance the influence of phosphorous. If the typology and the determination objective of chlorophyll a should be individualized - should the reference year be likewise? Summing up, we have the following questions: 1 Has the panel considered the uncertainty by determination of reference year 1900, regardless of method? 2 Should the reference year be individual in consistency with the typology? 3 Should the reference year take variation between spring and summer into consideration?	The Panel has not been presented with the full documentation of the methods used in estimating the 1900 nutrient loadings - the details of this were not part of the evaluation. However, we feel that it is exaggerated to call these estimates simply 'expert judgment' as to our knowledge qualified scientific studies have been performed to estimate 1900 loads and that their overall results compare well with other estimates from other Baltic Sea catchments. The use of loadings around 1900 is an international standard. Especially for the countryside and agriculturally dominated areas it is justified by the calculations showing that the N and P surplus in agriculture (difference between nutrient extraction as harvest and addition as manure) was close to zero or negative at that time. We discussed the question of specifying reference conditions per water body in our report, but concluded that this was not feasible. Note, however, that site specific characteristics of the water bodies (averaged over types - a point that we criticize) were taken into account for setting reference values
Bæredygtigt Landbrug	5.1	The Panel states that great efforts have been made in reducing the P load from especially waste water. We agree that this effort must be acknowledged. However, Denmark is still a country with a very large sewerage system with	Can the panel recommend to regulate/evaluate the status based on "summer loads"?	The Panel has not excluded, in its recommendations, that further efforts be done in removing P from urban waste waters. However, we have acknowledged that great efforts have already been done and that this type of

		far from sufficient capacity, causing outflows of untreated waste water directly into the water environment. There is insufficient control of volume and concentration of these outflows. Therefore, we find it difficult to understand that it is accepted without question that nothing further can be done in connection with P loads. It must also be pointed out that for the last 10 to 20 years, water course management has been downprioritized, which has led to increased water levels and phosphorus mobilization in areas drained to the water courses. Furthermore, many wetland sites that are established in previous agricultural areas with phosphorus saturation will also cause a substantial phosphorus mobilization, which has not been taken into consideration. In our opinion, recent measurements of resuspension/fluctuation of P loads from the seabed are insufficient. There is a much higher resuspension of P loads than N loads. Consequently, it is not only the outflow but also the volume and availability of P loads in the water environment hat should be in focus. In this connection, perhaps, the panel should consider specific projects such as bottom trawls that increase the phosphorus availability. The overall picture of phosphor is lacking, in connection with the outflow as well as volume and availability in the water environment.		measures is subject to the law of diminishing returns: any additional effort will be expensive in relation to its effect. We therefore recommended studying also alternative and innovative ways to further reduce P loading. We think that sufficient expertise is available to execute these studies. We are skeptical regarding the suggestion to dredge away P reservoirs in coastal waters and do not recommend it as a promising avenue for further research.
Bæredygtigt Landbrug	5.4	Good Ecological Status The panel suggests that the status (Good Ecological Status) can be regulated according to summer loads: "[...][page 23] There seems to be a possibility to regulate Good Ecological Status by focusing on the summer loads, rather than on the yearly integrated loads. [...]"		We have recommended studying this possibility, taking into account two important aspects: 1) the effect on the neighboring systems must be taken into account and 2) the effect of the residence time of the nutrients in the water body can be much longer than the residence time of the water. We suggested that the mechanistic models should be good tools to investigate these aspects.
Bæredygtigt Landbrug	10	The Evaluation states that: "[...][page 42] Current scientific insight endorses the view that the overall reductions proposed are necessary, but cannot guarantee that they will be sufficient. Especially for benthic ngiosperms and macrophytes, additional measures may be needed.[...]" Focus in the Evaluation has been the models. In the Evaluation it has been stated that there is no direct connection between Kd and N. The statistical model and the mechanistic model are being mixed together in too		We understand that the sentence was not clear enough and have adjusted it in the text revision. We recommend being more specific about the MAIs and trying to base them on a water body specific approach. We further recommend a number of potential improvements of the model approaches. In short, we recommend refining the models further before actual implementation in the plans, but we do not think that this will change the order of magnitude of the efforts required.

		early stages. The P load has not been taken into consideration. The mechanistic model has potential for including P loads in a better way. However, presently the latter model does not include P loads sufficiently. Also, the N:P ratio has not been sufficiently considered. It is stated that other indicators than eelgrass should be taken into account, i.e. other species of angiosperm, and that this could possibly mean that some areas already achieved "Good Ecological Status". A list of improvement points has been described and a number of other issues have been mentioned in the Evaluation. In the analysis, the panel has dealt stringently and in great detail with the models. The focus of the panel has been to evaluate which bolts are included in the models and if the bolts fit together in their mutual dependencies according to the scientific knowledge represented by the panel. At no time has the analysis gone into how much N, P, or N:P 16 ratio, temperature, etc. should be determining "Good Ecological Status". However, the above sentence can be understood in such a way that the N restrictions of the present water management plans are necessary, although not necessarily sufficient. Such a conclusion does not fit with the analysis and the conclusion in 9.7 "[...][page 40] Taking into account all aspects and associated problems, the Panel has the impression that the regionalised MAI are not sufficiently reliable to serve as a basis for decisionmaking and planning of load reduction measures..[...]" We are afraid the panel's statement could be misunderstood to have a different meaning than intended. We would therefore ask for the sentence to be taken out or to be followed up with a detailed explanation based on the analysis.		
Danmarks Naturfredningsforening	General	In Denmark there are too few measuring stations to be able to subdivide in, for example, outer and inner fjords. Furthermore only more continuous and regular data can disclose if the effort is sufficient.	Is the Panel in agreement with the comments, and if so how many in Denmark and where should they be placed in order to subdivide the Danish water bodies and full fill the Panel's recommendation on this subject? What more data is needed according to	The Panel has stressed the importance of good data bases for the management of nutrient loads. The Panel is unable to make precise recommendations about numbers of stations or number of samples per year, but in general has expressed the view that sufficiency of the data base should be a concern for proper management

			the Panel in addition to the data already available to determine Maximum Allowable Inputs (MAI) over the seasons?	
Danmarks Naturfredningsforening	2.5	There are other stressors than N in the marine environment and Danish coastal waters. But they are not additive. They do not allow choosing freely in reductions to achieve the same percentage response. With reference to the report of NIVA, Denmark: 'Under the surface: a gradient study of human impacts in Danish marine waters' (rapport l.nr. 7128-2017 DK6) and the stressor weighting in this report does not provide vidence that one might as well work on any other stressors instead of nutrients to achieve a corresponding improvement in the ecological condition. Focus on nutrients in Danish waters is fully justified and in strict accordance with WFD. The Panel emphasizes the importance of this focus. It is also apparent from the Panel´s response and comments to the stakeholders, saying, that N and P loads are 2 of the most important pressures. The Danish Society for Nature Conservation would like to emphasize that MSFD takes on a different and wider range of stressors, which may be more holistic, but in no way undermines the control of nutrient supplies as crucial to reach Good Ecological Status (GES) in accordance with WRD.	Is the Panel in agreement with these comments?	That is correct
Danmarks Naturfredningsforening	3.2	The Panel refers to, that The Common Implementation Strategy for the WFD 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. And the Panel refers to, 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. It means that every country has the reedom to adjust the typology to its own needs and to refine it to the required degree.	Does the Panel assess, that Denmark has a simpler typology than other member states? Does the Panel believe that Denmark without violating the WFD could set up an even more simple typology than today? With reference to the recommendation that Denmark should calculate reference conditions and targets for each 119 water bodies in Denmark, instead of refine the existing typology (p. 43) to what extend does the Panel find, that efforts decided concerning lowering N-loading to the water bodies should be set on standby till this calculation has been done? And	The Panel in no way recommends postponing action until an endless series of new studies has been completed. However, it is of the opinion that the material that is currently available already allows estimating MAIs with higher resolution than the relatively coarse typology. There seems to be no reason at all to further coarsen the typology, given the wealth of data, models and knowledge available.

			does the Panel assess, that Denmark can do that without violating the WFD? To what extent does the Panel believe that Denmark should keep the major objective of the WFD in mind, considering the Panel's assessment, that targets for lower annual load of P could be feasible in a number of systems (p. 21)?	
Danmarks Naturfredningsforening	5.1	We have noticed that the Panel reports the following: Summer situations are basis for the calculations, why the Panel points out that there is focus on N limitation. Due to that the importance of P in especially the great spring production may be underestimated. This may mean more costly reduction scenarios than necessary, especially in areas with large N reduction requirements. The model period 1990-2013 excludes the time before the large reduction in the release of P, but reflects the current situation of P-loading, as highlighted by the Panel and therefore the Panel endorses the choice of period. However, the Panel acknowledges that P reductions from point sources have yet reached cost-effective limits, but the inclusion of diffuse sources of P can be steadily included in the input scenarios. The Panel recommends testing the calculations with seasonal inputs of N and P as a contribution to a detail of the effort.		That is correct
Danmarks Naturfredningsforening	5.4	The panel emphasizes the models' exclusive focus on summer indicators in combination with water bodies with a short residence times implies a direct link between summer loads of N and the indicator. But even though a larger proportion of the N-loading takes place during the winter and have been abducted from high-flow water bodies during the winter months, the Nloading ends in other waters bodies where it then affects EC. Therefore, the Panel's assessment that MAI may be overestimated in certain water bodies has to be seen in that context. The Panel emphasizes that it is difficult to assess MAI focusing on summer discharge in high-flow water bodies, and as we understand the Panel finds it is possible to investigate	Does the Panel assess that investigations on scenarios with seasonal regulation of N will lead to more expert judgment? Does the Panel assess that there are valid climate data (rainfall, temperature, etc) to calculate regionalised MAI depending on the season?	There are sufficiently detailed data of rainfall, runoff and seasonal nutrient concentrations to allow a study within seasons. Such study does not need to be done for all water bodies. The mechanistic models do resolve seasonal patterns and should be sufficient to calculate realistic scenarios within seasons. An important question, however, is whether agriculture can steer seasonal cycles of nutrient runoffs and what the margin for management is. The Panel cannot judge these aspects but is aware of the difficulties. The option is not a panacea, but it is interesting enough for further analysis.

		scenarios with seasonal regulation of N. Referring to, that the Panel also emphasizes that it is very positive to the near lack of expert judgment in the work, the Panel does not indicate whether such investigations on scenarios with seasonal regulation of N will lead to more expert judgment.		
Danmarks Naturfredningsforening	8	The Panel refers to that nutrient load reductions are associated with high costs. We understand the Panel so that reference is made to direct economic costs and not to costs for nature and the environment. The Panel has not commented on whether these costs will increase or decrease if there is a delay of achieving GES by postponing the effort to reduce the nutrient load. And the Panel has not commented on whether these costs will increase or decrease in order to protect the groundwater if postponing the effort to reduce the nutrient load to the marine waters.	Does the Panel believe that the costs will increase or decrease if the result of regionalised MAI according to the models is not used as basis for policy and planning of load reductions measures?	It is true that costs are associated with nutrient load reduction (direct economic costs, especially in agriculture), but also with a lack of nutrient reduction (loss of ecosystem services). The Panel has not analyzed any of these cost categories, but only made general remarks on optimizing the efforts: one should try to make any measures as cost effective as possible. The Panel has not advised on postponing measures at all, and is of the opinion that adequate plans can be made for the projected period.
Danmarks Naturfredningsforening	9.7	The Panel writes, that taking into account all aspects and associated problems, the Panel has the impression that the regionalised MAI are not sufficiently reliable to serve as a basis for decision making and planning of load reduction measures.	We hope the Panel will comment on this, when the Panel at the same time emphasizes that the overall reductions proposed are necessary, but cannot guarantee that they will be sufficient (p. 42).	The Panel is of the opinion that the models can be improved and be made more water body specific, and that this does not constitute an impossible effort
Danmarks Naturfredningsforening	9	The WFD aims at restoring GES in surface waters in Europe. Furthermore the WRD aims at protecting the groundwater. In Denmark about 20 % of the arable land covers N-sensitive drinking water areas.	Does the Panel endorse, that a less required load reduction than decided today for The Danish water bodies due to it is overestimated because of less reliable regionalised MAI can have an negative influence on N-loading to the groundwater?	The Panel has not considered groundwater problems at all. This was outside of the scope of the review
Danmarks Naturfredningsforening	10	Based on the evaluation of the Danish marine models by the Panel, it can be concluded that: 1) It is necessary to focus on the nutrient reductions to the Danish marine waters in order to achieve GES in accordance with WFD. 2) It is correct to focus on reduction of N to achieve GES in accordance with WFD. 3) The calculated MAI to achieve GES for the Danish water bodies is necessary, but is not necessarily sufficient. 4) The Danish marine models are appropriate, and		The Panel report contains the full summary of conclusions

		Denmark should continue to apply these models to calculate MAI. 5) There is an excellent scientific expertise in Denmark, which on an objective basis and the use of available data calculate MAI. 6) The Danish marine models used to calculate MAI exclude where ever it is possible expert judgments.		
Danmarks Naturfredningsforening	11	Kd is not only depending on phytoplankton in Danish waters bodies but also on e.g. resuspension of fine material (p. 16). The Panel does not evaluate if it could be taken into consideration to look on material discharge from watercourses in order to achieve GES. While chlorophyll a as a generally accepted and intercalibrated indicator, there should be more focus on this indicator in order to calculate MAI. In addition to the need of focusing on reduction of N-loading, there may also be additional value in looking at reduction of P-loading, to the Danish marine waters. The Panel does not comment on the necessity to reduce P-loading to the Danish lakes, too in order to achieve GES in accordance with WFD for the Danish lakes. Taking this into consideration it could be useful to look at the connection between P-loading to lakes and to coastal waters. It may be useful to look at seasonal N-loading. This is especially true where there is a large N-effort requirement. The Panel does not comment or evaluate if it could cause an increase in N-loading to other fjord sections and the coastal waters. And the Panel does not comment or evaluate if it could cause an increase in N-loading to the groundwater.		The Panel report contains the full list of recommendations. The Panel has not reviewed any work on lakes or groundwater and cannot comment on these points. With respect to seasonal N regulation, the Panel has clearly pointed to possible effects on other water bodies.
University of Aarhus	2.1	Page 9: The panel states that: "Target values must fall in the green (GES) range. However, the error bars of the calculated target value should also fall in the green or blue area. Therefore, the longer the error bars, the more the target moves to the left and is consequently more stringent."	We acknowledge and agree on the logic behind this statement and w uld like to know if there is a reference to a WFD document stating that error bars should fall in the green (Good ecol. status) or blue (High ecol. status) area?	This argument was based on logic rather than on a strict directive. If a model predicts that good environmental status will only be reached with a probability of 50 % (as half the uncertainty range lies above the limit), this cannot be interpreted as 'ensuring good ecological status'. However, the statement has been removed from the report as it may lead to discussion and is not strictly necessary in this context.
University of Aarhus	2.3	Page 10: The panel states that: "Kd has not been intercalibrated (as confirmed by the researchers from		The Panel acknowledges that Kd is a proxy for eelgrass depth limit, which is an intercalibrated indicator. As such,

		Aarhus University (DCE) and DHI and the European Commission's Joint Research Centre)" It is correct that the Kd indicator has not been intercalibrated but the Kd indicator is a transformation/translation of the eelgrass depth limit indicator which has been intercalibrated.		Kd is a supporting quality element for the Eelgrass biological quality element, but it has not been intercalibrated in itself. For that reason one should discuss whether it should have equal weight as Chlorophyll a. The intercalibrated status of eelgrass depth limit has been added to the text of the report to clarify the point.
University of Aarhus	2.4	Page 11: The panel states that 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. It is correct that the one-out-all-out principle was not followed, but since indictors which are ecological better than the good-moderate status are set to zero (page 88 in Scientific Documentation Report) in the calculations all water bodies which are classified below G/M (i.e. at least one indicator is below the G/M target) will induce a nutrient reduction requirement in the calculations.		The Panel has further commented on this principle in chapter 4, and has indicated nuances to its position in this debate. From a legal perspective, it is clear that the one-out-all-out principle across different quality elements has to be used, as was also clarified in the European Parliament. In practice, however, independence of the indicators Chlorophyll a and Kd with respect to the categories is not clear, which leaves room for interpretation.
University of Aarhus	3	Page 13-15: The panel finds the typology coarse and the diversity of chlorophyll a reference values low and suggests refinement of the typology especially related to water exchange and fresh water discharge. We agree that the coarse typology result in a low diversity of especially chlorophyll a reference values and that a refined typology and/or cross system analysis would likely improve the estimation of chlorophyll a reference values. An ongoing project is examining the possibilities and strategies to refine the typology and the chlorophyll a reference values in order to reflect the characteristics of Danish coastal water bodies in higher details. Comments and suggestions from the panel will be addressed in the project.		Good!
University of Aarhus	3.4	Page 15: This is important, because only the existence of a monitoring station and regular data collection allows assessing whether the target is reached or not.		agreed

		We agree – and want to stress that a comprehensive monitoring program is crucial for the assessment of the environmental status of Danish coastal waters.		
University of Aarhus	3.6	Page 15: The panel states that: "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" We are aware of the German approach, and we agree that it is a useful approach although it should be acknowledged that due to high uncertainty in simulating a year 1900 situation variations between systems might reflect model variability/uncertainty to a higher degree than actual system variation. This is likely also the case with some of the German targets.		It is true that the standard deviation of the mean decreases as the number of systems that have been averaged increases. However, if there are systematic, and explainable, variations between the systems, then averaging over too broad a group of systems adds this part of the variation to the noise, whereas system-specific targets do not. For this reason the Panel has recommended to set up the statistical analysis as a cross-system analysis, so that the maximum of explainable variation between the systems is not considered as noise but taken into account in the model. We are well aware and indicated that mechanistic model approaches using a historic state are subject to uncertainties and are not the only possible solution to derive reference conditions. The combination of mechanistic and statistical model approaches is appreciated and endorsed by the Panel.
University of Aarhus	3.7	Page 17: The panel states: "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. We agree with the panel on this aspect and we have in fact data showing the same aspects as mentioned above this quotation. On the other hand, it can also be argued that since Kd in Danish water is governed by DOM and POM, and not by chlorophyll a, over time a decrease in organic matter content both in the water and in the sediment and the other factors like oxygen condition will improve as well. Note that resuspension of fine particles, mainly POM, is the most significant parameter governing Kd in Danish waters.		The Panel does not feel that there is a real difference of opinion about the underlying mechanisms determining Kd. We have suggested relative freshwater influence as a common cause in a cross-system perspective, and left it open whether that relates to nutrients or to direct import of organic matter or both. We have also acknowledged the results of the statistical analyses, but stressed that the slopes are unexpectedly low and that most significant relations were found when concentrating on summer values only. We agree that most probably the time delays in the dependence of Kd on nutrients are much longer than for Chlorophyll a. From all this we concluded that Kd is not very responsive in the short run to nutrient reduction, and likely to respond in a similar way as Chlorophyll a in the long run. Although we are fully aware of the importance of Kd as an ecological target, we are of the opinion that it is less suitable as an indicator to base MAI upon than Chlorophyll a. And we recommend this

		Page 17: The panel states: “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” We are not surprised that a significant downward trend in Kd cannot be recognized, The crucial point is that a relationship between light attenuation and nutrient loadings can be found. The slopes are, as noted in the scientific report and by the panel, relatively low therefore it will either require a constant monotonous decrease in loadings, a sharp reduction over time or very long time series to get a significant trend in the development of Kd. Page 17: The panel states: “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”. This conclusion seems somewhat misplaced. As also mentioned in the report we do find significant relationships for 16 out of 22 water bodies. As documented in the report documenting the models, there as a plausible explanation for the low coefficient. We would also like to draw the attention to the paper in Limnology and Oceanography by Lyngsgaard et al. (2014) that documents such a relationship. Page 17 The panel states: 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 Error! Reference source not found. 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. It should be noticed that in Danish waters Chlorophyll a		issue as an important aspect for further study, including better mechanistic modeling

		contributes with about 5 to 20 percent of Kd. This has been documented in several studies. The correlation between Kd and Chlorophyll a is therefore, as correctly suggested by the panel, due to a common factor. However, this common fact is not freshwater or DOM in freshwater, but the nutrients coming with freshwater. Please consult the reference by Markager et al. (2011) which shows that the overwhelming source of DOM in Danish estuaries is not from freshwater, but from in situ production fueled by nutrients coming with freshwater. A similar study for Roskilde Fjord showing the same is currently under review. Based on this, and other evidence, we find it proven beyond reasonable doubt, that the main mechanism governing variation in Kd over time is the accumulation of organic matter, both as DOM and as POM in the water and on the seafloor. Hence, nutrient loadings are the underlying cause for the low Kd and also the key to lower Kd-values. We fully understand that the panel might be surprised about this, as we know the mechanism is different, e.g. in parts of the Baltic Sea, where direct DOC loadings with freshwater, e.g. from peatlands, is governing Kd. However, as shown the ratio of nutrients to DOC is important, and this ratio is much higher in Danish catchments than in most of the Baltic Sea.		
University of Aarhus	4.1	Page 18: The panel recommends: " 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." We fully agree with the panel about this in the future. However, as a suggestion for the next one to two decades it is not feasible. The reason is that spreading of Zostera marina is a slow process so the approach will face the		The Panel is unsure whether this is a real point of discussion. If for some reason another species of angiosperm would take over the eelgrass niche, there would be little chance to get eelgrass back but the ecological function would be fulfilled. If, what is more likely, other species cannot overtake this function in most systems, then eelgrass distribution will remain as a target.

		same problems as we see for Kd – but even more pronounced. The reason is again the time lag. Not only does Kd need to improve, and the other factors mentioned by the panel, but then the populations of Zostera marina has to respond. Comment: We agree with the panel about including other species than Zostera marina. However, that will require a formal decision by the authorities about changing the indicator or develop additional indicators for the quality element "angiosperms". However, in the project we were in principle restricted to use the formally accepted indicator, which at the moment is Zostera marina. On the other hand, since we have used Kd as a proxy this is only an issue if the light requirements are different for Zostera marina and the other species. We don't agree that other species per se could fulfill the "eelgrass criteria" since other species may have different light requirements, eutrophication tolerance etc. and therefore need to be assessed against other reference conditions than those defined for eelgrass. At present eelgrass is the only angiosperm in Denmark for which a reference condition has been defined.		
University of Aarhus	4.1	P18-19: The panel suggest a downweight of the Kd indicator: "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 19 status and targets at longer time scales, the Panel suggests to relatively downweigh the importance of Kd in the final calculations of reductions needed" The relative weight of the indicators can of course be discussed, however, since Kd is highly ecological relevant and since the Kd target values are based on outstanding historical data, the Kd indicator has several advantages. The fact that fulfilment of the Kd target is a necessary but not sufficient criterion for eelgrass restoration and depth		See higher. There is a difference between the ecological significance of water transparency or accuracy of the historical data - these are beyond doubt - and its usefulness as an instrument to determine MAI as accurately as possible. The fact that very often nutrient reductions beyond physical possibilities would be required, casts doubts on the latter aspect.

		limit target do - in our opinion - justify not to down weight Kd.		
University of Aarhus	4.2	Page 20: The panel states that: "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. This is a minor issue since the threshold for "number of days with nutrient limitation" is estimated based on the correlation between the number of days and the chlorophyll a concentration; hence higher Km would have resulted in a larger threshold and vice versa for a lower Km. Page 20: The panel states that: "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" From these statements, and several other places, we get the impression that the panel is much in favor of Chlorophyll a as the most import indicator for eutrophication. We fully acknowledge that Chlorophyll a is a relevant indicator. However, in our experience the use of Chlorophyll a as indicator is hampered by three serious problems. One problem is to get reliable reference values as historical Chlorophyll a data do not exists and modelled reference conditions are subject to uncertainties. Another disadvantage is the poor relationship between Chlorophyll a and phytoplankton biomass due to the systematic changes that occur in the C:Chlorophyll a ratio for phytoplankton (see paper in Limnology and Oceanography by Jakobsen and Markager (2016) for an analysis of this). In essence, changes in Chlorophyll a do not necessarily reflect changes in phytoplankton biomass. Another even more significant problem by using Chlorophyll a as an		The Panel agrees that Km is a minor problem, and that Chl-a is not THE ideal indicator of eutrophication. The Panel also agrees that other indicators may have high ecological significance. In particular oxygen depletion is probably the most important and directly relevant indicator. For this reason the Panel has recommended pursuing modeling studies that will one day allow to establish reliable dose-effect relations between nutrient loading and these direct indicators of eutrophication. This is, however, not presently the case and therefore, again, the Panel stresses the difference between ecological significance and usefulness as an indicator for the establishment of MAI at the present moment.

		indicator for eutrophication is that phytoplankton biomass governed by the balance between primary production and loss where grazing is important. Thus, as decrease in primary production, followed by a decrease in grazer biomass (benthic or pelagic) might result in only small changes in Chlorophyll a concentration. Thus, our thoughts are currently to decrease the weight put on Chlorophyll a as it is far from being the ideal indicator for eutrophication. That said, it is still useful and an intercalibrated indicator, but there is a need to supplement it with other indicators. We do agree with the panel that the translation from a status in the ancillary indicators to a required load reduction can be improved substantially, but we don´t share the panel's opinion that these indicators do not bring substantial improvement to the approach. On the contrary, we believe that all selected indicators are relevant eutrophication measures and that they complement each other well (e.g., the Chlorophyll a		
University of Aarhus	4.2	concentration may, in one place, be low due to high grazing, here the nutrient limitation indicator will compensate for a Chlorophyll a concentration that does not reflect the eutrophication level).		
University of Aarhus	5.2	Page 22 The panel states that: "In general, the Panel is of the opinion that the selection of indicators only representing summer conditions could be too restrictive." We highly agree. In order to assess the ecological status and propose efficient measures a broader suite of indicators is desirable. As stated in the scientific documentation we also propose indicators covering the spring period and we are currently examining potential indicators relevant for identifying P sensitive ecosystems allowing for a future quantification of P target loads. The choice of e.g. May till September as the assessment period for Chlorophyll a is in our opinion not ideal but chosen simply because it is the WFD intercalibrated indicator for phytoplankton.	Does the panel have suggestions for relevant indicators and suggestions for how to get reliable target values without going through the intercalibration process which is extremely time consuming? 1) Does the panel have experience with year-averaged chlorophyll a concentrations as an indicator? 2) Does the panel have suggestions for how to include non-intercalibrated indicators	The Panel unfortunately has no direct suggestions on how to shortcut the intercalibration process. However, it recommends pursuing studies into the subject with the aim of proving clear and useful relations between nutrient loading and ecologically significant indicators. This is likely to facilitate the intercalibration if the studies show promising results that may also be of use for neighboring countries.

		Page 22: The panel states that: 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. We highly agree. There is a strong evidence in the scientific literature that P is important also for coastal marine ecosystems and even thought we could not demonstrate strong P load effects for the intercalibrated indicators we cannot and should not rule out P as an important pressure for marine ecosystems. Focus on yearly- or growth season- Chlorophyll a concentrations should be considered as also stated in the Scientific Documentation Report, however we face the challenges that these indicators are nor intercalibrated and hence from a management perspective are "unimportant".		
University of Aarhus	5.4	Optimizing efforts in both space and time is desirable but requires knowledge/models/data with high spatial and temporal resolution which is often not available. With respect to seasonal optimization it is also necessary to include adjacent water bodies since nutrients (conservatively) flushed out of one water body will affect other water bodies. And as noted by the panel, the retention time of nutrients in the system is most likely much higher than the hydraulic residence time suggest and in addition complicated to quantify. Hence, addressing seasonality is a desirable but also very ambitious task.		We agree, and have stressed both points in our recommendations
University of Aarhus	6.2	Page 25: The panel states that: "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		The Panel is now confused about the method used. We cite from page 43: "The first conclusion would lead to omission of nitrogen loadings as a management tool for that specific area. The other implies that nitrogen loadings might affect the status of the ecosystem. We assumed that the latter is the case and therefore used an average response for nitrogen loadings versus the response variable obtained from similar areas (the so-called meta

		become difficult to assess, but likely affect the targets for all systems in the type" This is based on a misunderstanding since type-specific slopes are not used to calculate reference values.		model approach, see section 8.6)." And from section 8.6: "As the meta model water bodies all belong to Type 2 (semi-enclosed water bodies with low freshwater influence) or Type 3 (semi-enclosed water bodies with high freshwater influence), type-specific cause-effect relationships for these two categories were estimated as an average of slopes derived from statistical models developed for water bodies of the same type." How else should we interpret this than that type-specific average slopes have replaced the non-significant slopes?
University of Aarhus	6.2	Page 25: The panel states that: "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" We agree with this point, but we are also convinced that to solve the decadal response solidly, a more than two decades long data series is required		We agree that data series are always too short when one analyses them, but would you rather wait for another two decades before doing anything about nutrient loading? In addition, we think that cross-systems analysis will help to resolve this issue better.
University of Aarhus	6.2	Page 25: The panel states that: "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. We did see this in a few estuaries with very low residence time (e.g. Randers fjord) but for the majority of the estuaries there was no effect of high discharge.		We have no idea how important this might be, but it requires attention
University of Aarhus	6.2	Page 25-26: The panel states that: "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" We agree on this, and this is basically what the "F-factor" is compiled of, and this was used to categorize the water bodies		It is our feeling that the information in the F-factor has not been exploited to its full potential, as seen by the relatively broad categories in the types
University of Aarhus	6.2	Page. 26: First, the procedures of the statistical and mechanistic modelling should not be unduly mixed at early stages We do not agree that this is the case. We "mix" the		This is a point of disagreement. We think the statistical modeling would have a much clearer contribution to the overall evaluation process if it is kept independent from the mechanistic modeling approach. Besides, inconsistency problems (unattainable targets) would be

		models with the aim to establish model independent chlorophyll a reference values and apply these similar to the way we use the Kd reference values (which are derived from historical eelgrass observations). Otherwise only MAI results are mixed (for meta-models).		much reduced
University of Aarhus	8.2	Page 34: The panel staes that: 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. Regarding "not averaging across model types" The "true" reference (year 1900) value is independent of the method used to determine it. In practice an estimated reference value depends on the method but effort should be used to come as close as possible to the true reference value. Otherwise it will not be possible to compare a reference value with measured data (status values). In that respect we disagree in the suggestion about not averaging across model types. Averaging across model types may lead to apparent inconsistencies when using models to calculate MAI. This is also the case when using historical data as reference values, however this should not be handled by using method/model dependent reference values. Method independent reference and target values should always be the (ambitious) goal. Regarding "not averaging across water body types" For water bodies where site-specific reference values cannot be established, type-specific reference values are necessary and this requires averaging within water body types (unless cross-system models are used). For water bodies where site-specific reference values can be established (and considered reliable) these could be used when considered more reliable than the type-specific value. This is how historical eelgrass data are used. However, based on the relatively low variation in model estimated site-specific reference values within each type we do not expect much higher variation using site-specific		Estimating reference values using a single modeling approach is the default approach that most countries use. Denmark is probably unique in developing two model lines that can be used to investigate the robustness of the approaches. We endorse this approach, but at the same time feel that much of the advantage is lost when the two model approaches are not kept independent until the stage of comparing their final results. In addition, the statistical model loses the ability to perform a decent uncertainty analysis because it incorporates external information. Compared to these disadvantages, we feel that very little is gained by the mixing of the models at a too early stage.

		values relative to type-specific reference values.		
University of Aarhus	8.2	Page 35: The panel states that "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". We agree, but since we have no a priori knowledge about potential model bias, averaging is the best solution.		We think that comparison, and detection of possible bias, is much better possible if the model lines are kept separate and not averaged across types. That will be a better basis for determining the final strategy.
University of Aarhus	8.2	Page 35: 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. As argued earlier we do not agree with the panel on this issue and we have compared the two models from various separate perspectives including slopes, reference values and the targets based on each approach.		You will agree, however, that the ancillary indicators require unclear expert judgment and in that respect contrast with the overall evidence-based approach. The Panel endorses further study of these indicators across the two model lines, including dose-effect relations, before including them
University of Aarhus	8.2	Page 35 The panel states that: "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" We agree and averaging across model types is likely a better choice.		Averaging across model types, or not, would ideally be based on a critical comparison of independent results of the two model lines
University of Aarhus	9	Page. 37: Since nutrient load management is a complex task and nutrient load reductions are associated with high costs, reliable overall and regionalised MAI are of outstanding importance. We are seriously in doubt about what the panel means by the term 'regionalised MAI'. Particular when we read and compare with the sentence on p. 43 (highlighted). The recommendation by the panel is what we have done. The reported results do in fact have one 'MAI' for each of the 119 water bodies. Regional averaging/lumping has only been performed in situations where the apparent variation		We apologize for slightly confusing language in our report. We have adjusted the text to better express our view

		could not be scientifically justified.		
University of Aarhus	9	Page. 43: 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. In our view, the WFD directive requires one MAI per water body to be estimated, so we believe we are in line with the panel in this aspect.		Yes, with that nuance that the Panel also recommends to be as specific as possible in setting reference values, and to forego types as this leads to unnecessary coarsening
University of Aarhus	9.1	Page 37: 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 regionalised 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 regionalised MAI and mitigation needs to achieve the GES in Danish coastal waters. We are pleased to see that our overall estimate is in line with German values, and acknowledge the many similarities in landscape and land-use As stated above, we have not estimated regionalized MAI but one MAI pr water body. We agree with the panel about the list of factors contributing to the uncertainties of MAI values. Given the overall positive conclusions in the report we find that the panel have endorsed that our approach is sufficient, albeit not perfect.		This section is also cluttered by imprecise language in our report. We mean a MAI per water body, based wherever possible on a water body specific target.
University of Aarhus	9.2	Page 38: 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 regionalised MAI based on historic conditions around 1900 bear uncertainties and for some water bodies		We are all aware of the problems, but seem to agree on the conclusions

		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. We are surprised about this statement in this context. We consider this as a political and perhaps juridical issue. Our report was made on the clear assumption given by the authorities that reference conditions reflect the conditions around the year 1900. From a scientific perspective we can share some of the concern if GES is feasible given changes in e.g. land-use and climate, but as we see it, it requires a political decision or a verdict from the commission to change this. We interpret this statement as a support to our approach.		
University of Aarhus	9.3	Page 38 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 Good point. The time scale issue will to some extent also apply to long term nutrient effects, e.g. regime shifts. Page 38: 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. Comment: We are very pleased that the panel agree with our view on this aspect. Page 38 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.		we agree on this point

		Comment: This work is based on the loadings directly to the recipient (not adjusted for discharge). Hence the climate-related variation in loadings are covered in this framework. How the climate change will affect the run-off and how it should be handled are catchment modelling and managing questions. The targets and the slopes will not change due to this issue		
University of Aarhus	9.4	Page 39: 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 regionalised MAI. This may cause an underestimation of the required load reduction for some water bodies and an overestimation for others. We fully agree with the panel on this aspect. However, it is important for us to emphasize that the typology was a condition for the project, and not something we could change as part of the modelling project. We are happy with the recommendation and also that a new project has been initialized by the Danish EPA on this aspect. In principle, we agree with the panel about this statement. Due to e.g. type specific Chlorophyll a and Kd reference values the calculated MAIs for each water body may deviate from the "true" MAI value. However, since 1) Kd reference values are based on historical observations with incredible spatial resolution and 2) the variance in modelled site specific Chlorophyll a reference values within each type is relatively small, potential under/over estimation of water body MAI is not significant. In addition, as mentioned before we have not calculated regionalised MAIs but one MAI pr. water body.		See previous remarks

University of Aarhus	9.5	Page 39: The Panel agrees that Chlorophyll a is a core indicator, and coastal water body-specific Chlorophyll a concentrations are a sound basis for calculating regionalised 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. 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 regionalised MAI. We are pleased that the panel agree with our emphasis on the indicators Chlorophyll a and Kd. Comment: We are somewhat surprised about this conclusion, as we see it the other way. Given the strong data based evidence for the reference conditions for Kd and the important structuring effect on coastal ecosystem of light attenuation, we find that Kd is a very important indicator. However, we acknowledge the problems with time lag for Kd response. In contrast, the use of Chlorophyll a as indicator is hampered by the fact that we do not have strong data supported evidence for a year 1900 reference condition. Chlorophyll a concentrations in year 1900 can be estimated using models, however, these estimates are uncertain as the models are used way outside the range for which data are available and the fact that many drivers, e.g. loadings, and boundary conditions are highly uncertain. Hence both indicators are associated with	• What is the reasoning behind this statement? In our view, based on decades of working with these data, we can see that the Kd has a lower variability than Chlorophyll a, which reflect the fact that Chlorophyll a, and phytoplankton biomass can undergo large short term variations whereas Kd, governed by e.g. DOM concentrations, are quite stable.	This is a repetition of arguments. We have answered the essence of these questions.

		uncertainties.		
University of Aarhus	9.5	Page 39: 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. We are highly surprised about this statement. We find these indicators, in essence hypoxia and nutrient limitation of phytoplankton, to be key indicators for the environmental status of shallow coastal ecosystems. Moreover, they do really add to MAI in the sense that they decrease the weight on Kd (slow response) and Chlorophyll a (uncertain reference values).		
University of Aarhus	9.5	Page 39: In general, the mechanic model has a very good potential for calculating regionalized 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 regionalised 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. We are pleased that the panel support the use of both types of models. As stated above, we are not calculating regionalized MAI but water body specific MAI. In principle, we agree with the statement that averaging of reference conditions for Chlorophyll a is undesirable. However, we do not see an alternative and in essence this underscores the problematic issues about a reference condition for Chlorophyll a and hence the use of this indicator.		You are calculating site-specific MAI, but based on type-specific targets, therefore your statement is too strong. Further, where mechanistic model results are available, your statistical model results are not used, and the mechanistic model results are regionally averaged. Therefore, the resulting map of required reductions is relatively coarse, reflecting these choices.
University of	9.7	Page 40: The panel states that: “The calculation of		It is quite well possible that the use of river basin models

Aarhus		regionalised 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. As stated above, water body specific MAI were calculated – not regionalized MAIs. We fully agree with the desirable in this approach. However, it is important to point out that this was beyond the scope of the task we were given. We have treated exactly this in several research projects over the last decade, but the Danish EPA did not incorporate this in the project aims. River basin models do exist and are used, but the coupling with the marine models was the responsibility of the Danish EPA.		was beyond the scope of the project. We have not analyzed that aspect.
University of Aarhus	9.7	Page 40: The panel states that: 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. The river basin models delivered data for loadings on a seasonal, in fact daily basis. In the statistical models we used monthly values, and the statistical models do in fact take the seasonal distribution into account. This has also been pointed out in our response to questions from the panel. Making scenarios addressing seasonality is however not an easy task since adjacent water bodies have to be taken into account and since nutrient retention in the coastal zone is highly complex and not fully understood. Page 40: The panel states that: "Taking into account all	Questions • What does the panel mean by regional MAIs? • Is the statement related to all estimated MAIs including MAIs for water bodies with site-specific models and where effects of uncertainty in target/status values can be assessed? And if so, would the use of site-specific target values (if they exist) improve the MAI results? • Is the statement mostly related to meta-model areas (which are not necessarily well represented by the typology)? and does the panel agree that a revision of the typology/cross system analysis would reveal if some areas are "misplaced" and likely increase the reliability of the MAI estimates? We find it a political responsibility of the	Our statement is that the knowledge, data and models seem sufficient to estimate water-body specific references, targets and MAIs, that this is an almost unique position and that this should be exploited to the maximum possible extent. While it is obviously a task of politics to decide on what basis to manage WFD, we feel that it is in everybody's interest to maximize the use of available knowledge and models.

		aspects and associated problems, the Panel has the impression that the regionalised 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 regionalised MAI. Even a modified processing of the existing model results might lead to much more reliable MAI." We are somewhat surprised about this statement/sentence since it seems very different from the overall conclusion. We do agree that the estimates of MAI can be improved, but we strongly disagree with a statement saying that they cannot serve as a basis for decision-making. As stated we are not quite sure what is meant by regionalised MAIs. We have estimated one MAI pr. water body and only for open water types applied a kind of regionalisation. For the coastal types we have applied typology derived (might be interpreted as regionalisation) or site specific Chlorophyll a and Kd reference values and we have used either site-specific or typology-derived model slopes. The use of typologies is an integrated part of the WFD and we have used type-specific values whenever these were considered more reliable that site-specific values. Although it is not within our competences to determine the reliability criteria for decision making, the uncertainty estimates and sensitivity tests that we have performed indicate, that the deviations in estimated MAIs are relatively small between model approaches and when including variations in status and target values (whether these are attributed to uncertainties in measured/modelled values or to deviations from a type-specific value)	authorities to decide if a given approach – in this case modelling - is sufficiently reliable to be applied in management. Such a decision will, in the end, depend on other available alternatives and may also be to use mere expert judgements. • What is the panels criteria for sufficiently reliable?	

		To illustrate the impact of using type-specific instead of site specific Chlorophyll a targets for the MAI estimates we have used Roskilde Fjord as an example. This estuary has a high target Kd value (0.66) and a site-specific Chlorophyll a target value of 3.9. Using type-specific and site-specific target values result in MAIs of 398 Tons N/year and 412 Tons N/year, respectively –a difference of 4%. Applying either type-specific or site-specific reference values for Odense yderfjord (which has the largest deviation between type and site-specific Chlorophyll a target values) results in MAIs of 117 Tons N/year and 98 Tons N/year, respectively –a difference of 16%. That said is likely that some of the (meta-model) water bodies are not well represented by the typology and for the meta-model areas is it not possible to perform (true) sensitivity tests/quantify the uncertainty. Page 40: The panel states that: ….. 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 regionalised MAI. Even a modified processing of the existing model results might lead to much more reliable MAI." As stated above, we support the view that phosphorus should be addressed as well and we will follow the suggestions made by the panel regarding re-examination of the effects of P for the intercalibrated indicators and include additional indicators although these most likely have to be linked to the intercalibrated indicators.		