Comments and Responses to Reports on Physico-chemical supporting elements

Martyn Kelly, Heliana Teixeira, Geoff Phillips, Fuensanta Salas, Gabor Várbíró, Agnieszka Kolada, Anne Lyche Solheim, Sandra Poikane, June 2021

1. Compilation of comments to PHC report – inland waters

MS	**Inland reports**	**TRAC report**	**Expert/ESCOSTAT representative sending the comments**
Cyprus	X	X	Gerald Dorflinger
Germany	X	X	Jens Arle
Italy	X		Aldo Marchetto
France	X	X	Sofia Vauclin
Netherlands	X	X	Marcel Van Berg
Lithuania	X		Diana Osadcaja
Romania	X	X	Carmen Hamchevici/Otilia Mihail
Slovenia		X	Natasa Dolinar
Sweden	X		

Country	Comment	Response
Cyprus	*Inland waters* Data mistake: in WISE all our phys-chem standards were reported as ranges (high-good boundary and poor-bad boundary) and not as the required good-moderate boundaries. In addition, in some instances	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.

	the ranges were reported as being valid for wrong water categories. In Cyprus the phys-chem boundary values are used in the framework of an integrated assessment of chemical - physicochemical status, both of rivers and water reservoirs, and not as single parameters that would on its own downgrade a station. This is because each parameter belongs to a pressure-specific parameter group, and these groups are evaluated following the one-out all-out principle	
Cyprus	*Coastal waters* Data mistake: Cyprus has not established any phys-chem boundaries yet, so the relevant reference in the Report is not correct. All boundaries included, concern only river water bodies, and not coastal ones. It seems that this mistake occurred while submitting the WFD reports of the previous cycle, which led to these conclusions that do not reflect the real condition. Due to the above, the results of the report regarding coastal waters do not reflect the real ecological status of Cyprus. Experts considers that the boundaries established from Greece about phys-chem, which refer to the same Intercalibration type with Cyprus (CW-Type_IIIE), would be appropriate for the assessment.	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.
Czech Republic	**Lakes:** In the conditions of the Czech Republic where only HMWB - reservoirs, no natural lakes, are, standards for good potential have been thoroughly set for total phosphorus. Other physico-chemical parameters that are	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.

	monitored in reservoirs - transparency, water temperature, oxygen saturation, pH, are not considered in the conditions of the Czech Republic to be determining for the potential of reservoirs. Their values are closely related to the manifestations of eutrophication (concentration of total phosphorus). Therefore, the standards are set in a **wide range** and are not decisive for assessment the ecological potential of reservoirs. Local anoxies on tributaries are addressed at the level of individual reservoirs. Even the pH values and acidification of lakes are no longer a major problem in the Czech Republic. Therefore, for the lake category, only total phosphorus should be used in the comparison graphs in the report, other parameters (mainly pH and oxygen saturation) should not be used. **Rivers:** It appears that only the standards used in the 2. RBMB were used for analyses of some parameters (oxygen, pH, BOD5). But for 3. RBMP different approved standards were used. We informed you about it by e-mail from 30.10.2020 and together we sent the current boundaries in an attachment (see the current attachment). Nutrients were listed on the first sheet, other parameters on the third sheet. These new standards are stricter. Standards for nutrients nitrate N (green colour), ammonium N, orthophosphate are listed in the analyses in the report correct. Standards for total phosporus should be: 70, 50, 45 and 30.	
Germany	Concerns on broad typology: The coarse "broad typology" approach was developed to allow the EEA an increased possibility for comparative analyses among the MSs concerning the ecological status/potential. For individual MS the "broad typology" remains meaningless, because their national water body types differ from the "broad	Although the broad typology may not be seen as very useful for single MSs, it is a valid framework for comparing results across MSs at European or at regional scale because it captures a lot more of the national types than the common intercalibration types. The broad typology was requested by the European Parliament in 2012 and developed in dialogue with the MSs in ECOSTAT over a period of two years, and Germany agreed with the links between the broad

	types" with regard to the comparability of the biological status or potential classes, but also because they have different and more explicit biocenotic characteristics. Over more the EUA "broad typology" is no legal instrument of the WFD. Experts do not consider appropriate to do this analysis: 1. because differences in G/M boundaries between MSs biological assessment methods (Intercalibration was done in GIG`s and GIG results were not compared whether these are comparable or not) and 2. It is not useful to lump all "good or better macroinvertebrate status samples together" and to prepare a box-plot for dissolved oxygen" on basis of these samples because you will mix different samples from different geographic regions (e.g. Nordic & Mediterranean).	types and your national types given in the Annex 3 of the ETC-ICM report 2015 (https://www.eionet.europa.eu/etcs/etc-icm/products/etc-icm-reports/european-freshwater-ecosystem-assessment-cross-walk-between-the-water-framework-directive-and-habitats-directive-types-status-and-pressures). The broad typology was also published as a paper in 2019, showing that 83% of German river waterbodies could be linked to the broad types (see Supplementary material Table S3 available at: https://www.sciencedirect.com/science/article/pii/S004896971934020 3?via%3Dihub#s0115). The importance of interpreting all comparisons with caution is emphasised in the report (e.g. 1.2.4: “We recommend readers focus on the big picture: a Member State that has national standards that are consistently more lenient than those of near-neighbours or those sharing similar water body types should regard this report as an opportunity to ask questions to ensure that its standards are sufficiently protective”) and was also stressed during the workshop. We also disagree with the comment about comparisons not being possible between GIGs. The intercalibration process involves several stages that were consistent between GIGs (e.g. compliance checking, statistical processes) and several countries were involved in more than one GIG and would have noted at the time if there were substantial differences in the positions of boundaries between GIGs. Finally, we disagree with the comment that the boxplots are not useful. These, again, have limitations (see 1.2.5) but are the only means by which the position of standards relative to actual biological data at a pan-European scale can be assessed. These are works in progress and may be improved over time (the plots for nutrients, for example, are already divided by broad type and some further subdivision by biogeographical region may also be possible). This is a point where constructive suggestions could well make a difference. We need to be sure that regional subsets will still have gradients that are long enough to enable boundaries to be inferred and which are not biased towards particular

		areas of the pressure gradient, in order to ensure that ecologically meaningful results can emerge.
Germany	Data for presentation: Experts consider not appropriated mix mean and median The approach should be be discussed in Ecostat WG A. The resulting 50th percentiles (red lines in the diagrams) are the product of the values of the MSs that have standards that measure the central tendency only. Other MSs that use other approaches for their standards are excluded. Furthermore, the resulting 50th percentile is influenced by the number of different national types (MS with more national types contribute more values to the "overall 50 th percentile" than MSs with a low number of types. Annual average temperatures differ between e.g. the Nordic MSs and the Mediterranean MSs but have direct influence on the solubility of oxygen in the water. At higher temperatures lower amount of dissolved oxygen is present at saturation (100%) than at lower tempeartures. Therefore lumping all together as done here is wrong. The same is true for phytobenthos.	The graphs treat all standards set as a central tendency in the same way and a further separation into those set using means and those which used medians would have been confusing. We recognise the issue that is being raised here but believe that the graphs in the report gain more from clarity of presentation than they lose through overlooking fine details of statistical distributions. The comment about the 50th percentile is valid. There is no perfect solution, but we will consider a more nuanced approach to calculating these percentiles in future reports. Once again, we emphasise that the "big picture" is unlikely to change radically as a result of this. We will however include a comment on the percentile issue in the updated report. The point on oxygen concentration differences between regions is also a valid point that can be included in the discussion of the oxygen results. Here, again, a region specific analysis could be considered. The comment about phytobenthos argument is unclear.
Germany	Linking the standards to sensitive biological quality elements: Experts think that this approach is not representative, and should be deleted from the report because: 1. different MSs have reported / do report a different number of sample sites per year under the SOE - reporting. And 2. the G/M boundaries for macroinvertebrates in rivers were intercalibrated in GIGs (Northern, Central-Baltic, Alpine & Mediterranean) but the results were never compared between these different GIGs. We do not know whether the G/M Boundaries of the biological assessment systems of MSs among these different GIGs are truely comparable. Therefore it is wrong to lump all available data from the SOE reporting together in this bow plots.	The issue with representativity (with particularly few SoE data from Germany) does not mean that the SoE data reported from a large number of European countries cannot be used for this analysis. We have already recognised that these graphs need to be interpreted with caution, and pointed out that these are works in progress for which constructive comments are welcomed. The possibility of using regional models is already being considered but this does not render the plots in the report as useless. We still believe that these are a valid means of presenting the "big picture", irrespective of any remaining issues. As noted above, a problem with regional splitting of the data is that the gradients covered by each region may become too short or too biased towards certain sections of the gradient to allow ecologically meaningful results to emerge.

	The same (as for macroinvertebrates in rivers) is potentially true for most other BQEs. The intercalibration results were never compared between GIGs and you can not assume that the G/M boundaries are comparable between GIGs- As a results you should not lump all these values "into one box-plot".	We strongly refute all comments about the inadequacies of the intercalibration process (which German scientists played a key role in developing!). We refer you to the comment made by the external assessor of the intercalibration process, Susan Davies: ""To an impartial reviewer, Europe's achievements in this arena, and the ambition to attempt them, are an inspiration. […] While the effort under review has not ended in perfection, nevertheless the intercalibration of ecological status classes has launched the European Union on a heuristic path that, with commitment, can be expected to lead to ever improving comparability, and ultimately, it can be hoped, towards improved ecological sustainability."
Germany	Experts ask Why have the water temperature and Nitrite not been included in further analyses although it is being used by 12 (for water temperature in rivers) an 10 (for nitrite in rivers) countries and is ecologically relevant?	Temperature displays strong regional differences which would have complicated comparisons. Although not included in this report, it should not be forgotten as this work progresses. Nitrite is potentially toxic to invertebrates and fish, but we assume that nitrification rates in nature mean that it is rapidly converted to nitrate-N under most circumstances. Both could be included in future work.
Germany	Results- Oxygen Germany does not use VG / G and G/M boundaries for DO (% and mg/l) in lakes because this variable is influenced by too many natural factors (e.g. temperature, season, thermal stratification, turbidity, day-night rhythm of photosynthesis and others) and the monitoring method (measurement depth, day time, and others). All these sources are responsible for the large variability of the data shown in fugure 3.3. To expert's knowledge there is no standardisation of the monitoring methods for oxygen in lakes between the MSs and therefore you should not compare these values directly (as you did in the text). They suggest to discuss these issues in the text or delete this section.	Comments on validity of lake oxygen standards: these points are already covered in 3.1.3. Comment on Fig. 3.7: We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated with the values provided in your comment. Fig. 3.9: we refer you to earlier responses where we refuted suggestions that intercalibration results from different regions are not consistent. As commented in other responses, we do recognise the need to further develop the analyses presented in Fig. 3.9 and others. Oxygen, in particular, is a supporting element where regional subsets are more likely to be informative than a continent-wide overview.

	Mistake in Figure 3.7: The red and blue dots are wrong for Germany and should be deleted. The legally fixed values are 4, 6, 7 & 8 mg/l MinA-EQS. Figure 3.9: It is not appropriate to do this analysis: 1. because differences in G/M boundariers between MSs biological assessment methods (Intercalibration was done in GIG`s and GIG results were not compared whether these are comparable or not) and 2. It is not useful to lump all "good or better macroinverterbate status samples together" and to prepare a box-plot for dissolved oxygen" on basis of these samples because you will mix different samples from different geographic regions (e.g. Nordic & Mediterranean). Annual average temperatures differ between e.g. the Nordic MSs and the Mediterranean MSs but have direct influence on the solubility of oxygen in the water. At higher temperatures lower amount of dissolved oxygen is present at saturation (100%) than at lower tempeartures. Therefore lumping all together as done here is wrong.The same is true for phytobenthos.	
Germany	Results-Secchi depth Figure 3.13: Phytoplankton biomass and Secchi depth correlate very strongly with thermal stratification. Since the broad types in the lowlands contain stratified and non-stratified lakes, there is a corresponding spread, as shown e.g. in figure 3.13. In expert opinion, this effect can mask the differences between Member States.	We do agree that Secchi depth boundaries are likely to differ depending on the stratification of lakes. This is why the broad types are particularly useful, where type **LW-04 are lowland unstratified very shallow lakes** and the other broad lowland types are **stratified, deeper lakes (e.g. LW-03**). Figure 3.14 shows that the GM boundaries for the unstratified broad type (LW-04) are mostly lower than those for the stratified broad types (LW-03). This explains part of the variation between MSs seen in figure 3.13 and is ecologically meaningful.

Germany	Results-pH Mistake figure 3.19: One pH max value at 8.0 is missing in the plot for Germany	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.
Germany	Results-BOD5 Mistake Figure 3.22: The black, red and blue dots are wrong for Germany and should be deleted. The legally fixed values are 3, 4. & 6 mg/l.	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.
Germany	Results-Ammonium-N Mistake Figure 3.25: The red and blue dots are wrong for Germany and should be deleted. The legally fixed values are 0.1, 0.2 and 0,3 mg/l. please change	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.
Germany	Results-Total Nitrogen (lakes): Experts consider difficult the determination of limit values. National research projects have shown the following: • In contrast to the TP, TN is subject to greater annual dynamics in its inorganic and organic components with different effects on the phytoplankton biomass that is relevant to the assessment according to the WFD. • Even with nitrogen limitation, phytoplankton is able to use DON. The extent of the phytoplankton bloom is ultimately determined by P even if there is a temporary nitrogen limitation in the summer half of the year. • N inputs through N2 fixation of nostocal cyanobacteria do not in principle compensate for the N deficiency. In the case of P Limitation, the growth of nostocals is also predominantly limited. • The influence of nitrate on the release of phosphorus from lake sediments depends on the oxygen conditions over the course of the year. The P release also depends on the iron content and the dynamics of the sulfur. These processes are inadequately mapped using limit values for total nitrogen. Here, too, the ratio of P, Fe and S plays a central role in whether the lake acts as a phosphorus sink or not.	These are all useful comments, but we feel that an evaluation of TN standards is still needed. We agree that determining limit values for TN are potentially more challenging than for TP. However, we have included TN as many Member States report these standards. Additionally, TP and TN are positively correlated and there can be cases where N is limiting or both N and P can be considered limiting. We also note that there is growing evidence for nitrogen-limitation of some BQEs (e.g. macrophytes) in some situations. Overall, these comments emphasise the need to consider not just what threshold values are being used by Member States, but also for how they are used in the water body management process. This, again, is something that ECOSTAT may want to pursue in the future.

	• Denitrification plays a prominent role in river lakes and shallow lakes, but cannot be included in the assessment.	
Germany	Results-Orthophosphate (Rivers) Mistake Figure 3.37: The black, red and blue dots are wrong for Germany and should be deleted. The legally fixed values are 50, 70, 100 & 200 µG/l.	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated, including these corrections from Germany.
Germany	Discussion- Issue 2. Which are the most important variables for explaining variation in ecological quality? The approach (BRTs) is not very useful to determine "thresholds" for the management of single parameters because these thresholds will always depend on the size / values of the other parameters in the model. Total P is the best independent variable for phytoplankton and macrophytes in lakes. In Germany there was a close correlation between P and the good ecological status with regard to phytoplankton. However, the thermal stratification and the size of the catchment area must be taken into account. Regarding the model Phytoplankton EQR: Such high values should be in every case observed with caution. In some cases this might be o.k. But not in most cases. We analyse ecological communities which are commonly influenced by a multitude of natural factors and anthropogenic stressors. If the ecological status, as a "measure for the status of a freshwater ecosystem", is explain to e.g. 80 % by only two or three variables we should really think about whether there is something wrong with our approach. In such cases we should at least intensively discuss what other variable are potentially co-correlated with the selected response	BRTs: We did not suggest that BRTs should be used to set threshold values. BRTs were used to unpack the influence of multiple pressures and, in this respect, they were extremely effective. We have added a sentence of clarification to the relevant section of 4.3. Total P: no change to report needed in response to this comment. It is uncontroversial, but detailed considerations of individual countries is beyond the scope of this discussion. Regarding the model Phytoplankton EQR: No change to the report needed. We agree with the sentiments expressed. The issue of how standards are applied to solve complex management problems is, again, beyond our remit.

	variables. If doing not so, wrong management conclusions will follow.	
Germany	Discussion-Issue 4. Which BQEs should be selected when developing pressure response relationships? As a result of this approach we will become hundreds of context dependent G/M Boundary values (at least more than now). Is this really a senseful goal?. Additionally it will to my opinion become impossible to communicate the complex statistical approach (BRT) to the public and to the politics. Less is sometimes more...	This is a misunderstanding of the text. In order to clarify our position, an extra sentence has been inserted, so that the paragraph now starts as follows: "Finally, it is important to recognise the different sensitivities of the BQEs to pressure. **Because of this, we recommend that the most sensitive BQE to a supporting element is selected to set the threshold for that supporting element.** Thus ..." We agree with the comment about difficulties of communicating complex statistical approaches to non-technical audiences. However, it is also necessary to acknowledge the problems of simpler statistical approaches when faced with multiple interacting stressors. Our concern is that "less is more" will result in lenient standards that will not contribute to achieving WFD objectives.
Germany	A differing strength of stressor indication (also caused by different sensitivities of the BQE) by the different biological assessment systems of the MSs and between different biological methods is an major problem in this context (the used of multivariate approaches like Boosted Regression Trees). If most biological assessment systems are more or less indicative or non-indicative for different stressors why we should expect to get a realistic picture of the relative importance of different stressors by boostest regression tree analyses? Without taking into account the different strength of indications of the different biological assessment methods of MSs for different stressors, one cannot expect to get a valid statement regarding the relative importance of different stressors within a MS´s and also not across Europe.	To an extent, we agree with this comment. However, BRTs have only been used in this report to try to understand relationships amongst stressors not to set standards. See our response to an earlier comment on BRTs.

Germany	Synthesis- oxygen conditions Oxygen is not a pollutant	We've rephrased the first sentence of this point to read; "Low oxygen concentrations, BOD and ammonium-N often combine to form a "cocktail" of stressors (along with nutrients) associated with .."
France	Objectives Relationship between oxygen and BQE: This should be explained in more details : how is there an indirect relationship between BQEs and oxygen conditions? A lack of oxygen will entail quite directly the death of most aquatic communities, especially fishes. Regarding Fish Directive: this should be explained in some more details : the standards from the Freshwater Fish Directive do aim to protect aquatic communities, and especially fishes, which are the most sensitive to a lack of oxygen. Why are those standards not considered suitable?"	This is a fair point. We've simplified the sentence to read: "In some cases (e.g. oxygen conditions), there may be a combination of direct and indirect relationships, depending on the BQE and may also increase in significance as global warming raises water temperatures (Jane et al., 2021)." We have inserted the following sentence into the section on the Freshwater Fish Directive: "This was designed to "support fish life" and, as such, may be appropriate for WFD purposes. However, it is not clear whether this assumption has been widely tested during the WFD era." Sections 3.1.1.1 and 3.1.2.1 do show that many countries have set higher thresholds for DO, and also that current DO thresholds are not aligned with data from BQEs at sites in high and good status, so this is clearly an area where more work is needed.
France	Linking the standards to sensitive biological quality elements: As previously stated in the report, fish data is not currently available in the wise SoE database, but this BQE would be especially relevant for parameters such as oxygen or BOD5. Maybe similar comparisons could be made by each Member State with their own fish BQE data?	We agree with this statement. There are plans in EEA and the ETC-ICM to revitalise the dialogue with the MSs concerning their willingness to report Fish EQRs in the coming SoE data requests. Fish are indeed important in relation to oxygen conditions, which is also affected by climate change. In the meanwhile, it is a good suggestion that each MS could try to link fish EQR-data to oxygen or BOD data, to check their own boundaries.
France	Discussion-Issue 2 Which are the most important variables for explaining variation in ecological quality?For the results presented in this paragraph (Issue 2), it would be useful to clarify which data was used. Assuming the data is from different Member States, do you know if the sampling strategy is the same in all Member States? Could potential sampling effects affect the results?	Data were taken from different combinations of countries for each BQE to maximise the number of records that contained values for all of the variables used in the models. As this was an initial analysis no attempt was made to check the relative contributions of data by each country. For river phytobenthos 10 countries were used ("AT" "BE" "BG" "EE" "FI" "IT" "LU" "PL" "SE" "UK"); for macro-inverebtrates 12 countries were used ("AT" "BE" "BG" "CY" "FI" "IE" "IT" "LT" "LU" "RO" "SE" "UK"); for lake phytoplankton 6 countries were used ("BE" "IE" "IT" "LT" "RO" "SE"), for lake macrophytes 6 countries were used ("BE" "FI" "IE" "PL" "SE" "SI"). Sampling strategy: This is a useful comment. However, it is difficult to evaluate with the data that are available. Some preliminary explorations have

		been attempted, with the conclusion that sampling strategy is independent of datasets
France	Discussion- issue 3- Country specific and regional variation, should we expect consistent relationships between pressure variables and ecological quality? Physico-chemical typology, i.e. geology, might affect those relationship between pressure variables and ecological quality : this would be worth investigating. Indeed, Scandinavian countries are on the Baltic Shield (mostly granites) while northern Belgium lays mostly on sedimentary rocks. It would be interesting to include conductivity in the analysis, in order to assess the natural ionic charge and how it differs in various geological settings.	To some extent, variation due to physico-chemical typology are dealt with via the broad typology. Generally, we observed that "type" explains less variation than "country". Nonetheless, we recognise that this is a key point that needs to be considered further. The challenge is finding variables that are readily available within EEA datasets and which explain residual variability in datasets.
Italy	Experts think that it would be necessary to study the effect of the use of different combining rules (e.g. averages, one out all out, multimetric, etc.) of the supporting elements, for example applying the rules for all Member states to a common data set.	We agree that this is an important issue; however, it is better considered in discussions about classification than about the process of setting supporting element standards. We recognise in the Best Practice Guide that users need to consider how the degree of precaution are embedded in different approaches, and this interacts with national approaches to combination rules. From our perspective, these differences almost certainly contribute to some of the differences between countries for any given supporting element.
Netherlands	These are good reports and sound analysis. An important limitation is the lesser degree of comparability of the standards themselves: other parameters, other season, other aggregation and as consequence of that also poor comparability (as number of MSs) at level where comparisons should take place, the level of similar types. We can not help that at this moment but is an important limitation for the conclusions. This is even more valid for the CTWaters.	Thank you. Most of your points are already covered in paragraph 1 of 4.1. The following sentence has been added to the end of 1.2.5 (on interpretation of statistical effects of MS v type): "These analyses should be treated as broad indications of the extent to which variation amongst national standards is determined by factors other than the pressure in question."

	One question is about the statistical effect of MS, how should I interpret this? Because MS differ also in the occurrence of type this seems difficult to interpret for me.	
Netherlands	Transparency/Secci depth As final general conclusion is stated that MSs with lenient standards (<1m) should encourage to validate their standards against biological data. I think that we did very good work within the Central Baltic GIG especially for LCB1. We also concluded that LCB2 Secchi depth is not the best indicator. Such very shallow lakes suffer mostly from wind resuspension and can be very turbid although not by phytoplankton. Can you please make also reference to this good work? And reword this conclusion on validation with paying attention that some lake types are 'naturally' turbid (around 1m SD)?	We've added the following to the end of 3.2.2: "It is also important to recognise that some very shallow, unstratified lakes are naturally highly turbid (e.g. due to suspension of sediments by wind/wave action) and, as a result, Secchi depth may not be a suitable indicator of ecological status in that lake type."
Netherlands	pH: ''each national type while (13) countries (AT, BE, BG, CY, EE, ES, HU, IE, NL, NO, PL, PT, SI) present standards as a range''.-->For the NLs this is not a range but an optimum.	We think that this is already covered in 2.3.2.
Netherlands	Ammonia: The NLs has not a specific standard for ammonia as nutrient/eutrophication, but we have one for ammonium/ammonia as toxic substance, together with pH measurement. Isn't this included in the EEA data base? Please let us know if more information is needed.	See 2.3.1 and Table 2.1. There were not enough data to present "free ammonia" as a separate supporting element, and we have focussed on ammonium-N, which is both a nutrient and a precursor of free ammonia.
Netherlands	N and P In the analysis of Geoff MSs were also compared while correcting for the EU type. I liked this picture very much because that is the most proper way of comparing. It does not make sense to compare NO's standards to those of NLs because there are no or nearly no comparable types. Would it possible to redrawn this one?	Graphs showing national data split by type are in the appendix

Netherlands	Salinity Salinity is for us not purely an assessment criterion, at least for lakes. Salinity is also part of our typology, since naturally we have brackish lakes.	The following has been added to 3.8.1: "Salinity can also be a component of typology in countries where there are naturally brackish lakes (e.g. NL)."
Lithuania	1- Table 2.1: Check the table. Missing Lithuanian data 2- Table 2.2 Check the table. Missing Lithuanian data	Table 2.1 represents the data originally reported to and subsequently extracted from WISE, thus it is not appropriate to modify this table (Lithuanian data had not been reported to WISE at the time when the data were extracted). Table 2.2 has been updated
Lithuania	1- Results- Dissolved oxygen (rivers) Figure 3.7. Experts suggest specifying the Dissolved oxygen values to indicate the G/M status as a single values - 7.5 mg/l (RW-01, RW-04, RW-05) and 6.5 mg/l (RW - 04) 2- Results. Secchi depth (lakes) Figure 3.13 shows the range of transparency standard of H/M status.Experts suggest specifying the Secchi depth values to indicate the G/M status as a single values - 2 m (LW-03) and 1.3 m (LW-04) 3- Results. BOD rivers: Expert suggest supplementing the information of BOD adding the data of Lithuania. The G/M status value of BOD5 in Lithuania is 2.87 mg/l (the value calculated from BOD7 - 3.3 mg/l). This value (2.87 mg/l) is linked to RW-01, RW-04, RW-05 4- Results-Ammonia (rivers) Figure 3.25 shows the range of Ammonium-N standard of H/M status; expert suggest specifying the Ammonium values to indicate the G/M status as a single value - 0,2 mg/L (RW-01, RW-04, RW-05) 5- Results-Nitrate (Rivers) The figure 3.28 shows the range of Nitrate standard of H/M status. Expert suggest specifying the Nitrate values to indicate the G/M status as a single value - 2.3 mg/L (RW-01, RW-04, RW-05)	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated.

	6- Results TN (lakes) The figure 3.31 shows the range of total Nitrogen standard of H/M status. Experts suggest specifying the total nitrogen values to indicate the G/M status as a single value - 2.0 mg/L (LW-03, LW-04). 7- Results TN (rivers) The figure 3.34 shows the range of total Nitrogen standard of H/M status. Experts suggest specifying the total nitrogen values to indicate the G/M status as a single value - 3.0 mg/L (RW-01, RW-04, RW-05) 8- Results Orthophosphate (rivers) The figure 3.37 shows the range of orthophosphate standard of H/M status.Experts suggest specifying the orthophosphate values to indicate the G/M status as a single value - 90.0 ug/L (RW-01, RW-04, RW-05) 9- Results TP (lakes) The figure 3.40 shows the range of total phosphorus standard of H/M status, we suggest specifying the total phosphorus values to indicate the G/M status as a single values - 50 ug/L (LW-03) and 60 ug/L (LW-04). 10- Results TP (rivers) The figure 3.43 shows the range of total phosphorus standard of H/M status, we suggest specifying the total phosphorus values to indicate the G/M status as a single value - 140 ug/L (RW-01, RW-04, RW-05).	
Romania	1. ***Pg. 52, chapter 3.4.1, BOD_5 in rivers***: Romania (RO) is not nominated among countries that use a single value for each national type, although it was mentioned in the Draft from October 2019 (page 50);	This has now been corrected.

	2. ***Pg 57, chapter 3.5.1, Ammonium-N in rivers***: the same comment as above, (RO) is not nominated among countries that use a single value for each national type, although it was mentioned in the Draft from October 2019 (page 52).	
Sweden	Table 1.1. Mistake? SE is not included in the table.	SE data arrived too late for inclusion. We have now added a blank row to Table 1.1
Sweden	Results. Dissolved Oxygen (lakes) Quality standards for DO in lakes also have to address that brown stratified lakes may have low DO or even anoxic conditions in the hypolimnion during stratification under reference conditions. The macroinvertebrate community is then naturally controlled by DO concentration. The within type variation of reference hypolimnic DO can be large.	Thanks for this. The following comment has been added to 1.3.1: "We also note that humic lakes may have naturally low DO concentrations (sometimes even anoxic) in the hypolimnion during stratification, even at reference conditions. This emphasises the need for local knowledge to inform how standards are set and used."
Sweden	Results- Secchi depth (lakes) Sweden could deliver type wise ranges of reference values for secchi depth based on object specific modelling	We note the offer of additional data but given the resource constraints for this phase of the project we are unable to include these.
Sweden	Synthesis. Regarding variation in s between standards between countries: Comparisons between SE NO FI and DK shows that the differences in standards depend on different approaches and lack of cooperation rather than differences between the countries based on science. Skarbøvik, E., J. Aroviita, J. Fölster, A. L. Solheim, K. Kyllmar, K. Rankinen and B. Kronvang (2020). "Comparing nutrient reference concentrations in Nordic countries with focus on lowland rivers." Ambio 49(11): 1771-1783.Fölster, J., Ø. A. Garmo, P. Carlson, R. Johnson, G. Velle, K. Austnes, S. Hallstan, K. Holmgren, A. K. Schartau, F. Moldan and J. Aroviita (2021). "Acidified or not?A comparison of Nordic systems for classification of physicochemical acidification status and suggestions towards a harmonised system. SLU, Vatten och miljö: Rapport 2021:1."	Thanks. We've added these citations to 4.1

Sweden	Synthesis. Regarding Point 7. In this report nitrate is only discussed as a nutrient. In Sweden, there is a new standard for nitrate as a prioritised substance based on toxicity. This is is however controversial since the data supporting the standard was very poor. Is there a discussion in the ECOSTAT group or in other member´states on nitrate as a toxic substance?	Nitrate is a priority substance in groundwater, where the concentrations can be very high and dangerous to human health if used for drinking water supply. This can be further discussed in the working group on Chemicals and/or Groundwater. ECOSTAT does not include work on priority substances.
Sweden	The way forward. This work is limited to comparing G/M boundaries between countries and broad types mainly with statistical methods. There are then hopes that highlighting differences will stimulate the member states to voluntary adjusting the boundaries by following the tools developed by the ECOSTAT group. If this will not lead to more harmonized boundaries, it might be useful to look deeper into the reasons for these big difference including how the reference conditions are defined. For example, in the Nordic GIG, references were selected by a reference filter allowing <10% of agricultural land in the catchment. This does in practice mean that forest lakes and rivers were used as references for agricultural waters, although agriculture land is naturally more nutrient rich than forest land. Further, the REFCOND document stated that the reference state should relate to a pre-intensive agriculture practice. It would be interesting to compare how the reference state is defined for e.g. nutrients between in the agricultural landscape between the different member states.	Thank you for this comment. Reopening the issue of reference conditions is beyond the remit of this project and would need to be raised separately with the ECOSTAT secretariat.

2. Compilation of comments to PHC report – TRAC waters

MS	TRAC report	Expert/ESCOSTAT representative sending the comments
Cyprus	X	Gerald Dorflinger
Germany	X	Jens Arle
Italy	All	Aldo Marchetto
France	X	Sofia Vauclin
Netherlands	X	Marcel Van Berg
Lithuania	X	Diana Osadcaja / TRAC comments received later (pdf)
Romania	X	Carmen Hamchevici/Otilia Mihail
Slovenia	X	Natasa Dolinar
Ireland	X	Robert Wilkes – TRAC comments received later (pdf)

Country	Comment	Response
Cyprus	**TRAC waters:** Data mistake: Cyprus has not established any phys-chem boundaries yet, so the relevant reference in the Report is not correct. All boundaries included, concern only river water bodies, and not coastal ones. It seems that this mistake occurred while submitting the WFD reports of the previous cycle, which led to these conclusions that do not reflect the real condition. Due to the above, the results of the report regarding coastal waters do not reflect the real ecological status of Cyprus. Experts considers that the boundaries established from Greece about phys-chem, which refer to the same Intercalibration type with Cyprus (CW-Type_IIIE), would be appropriate for the assessment. *(Last update from April 2021):* "In Cyprus coastal waters, we monitor for WFD the following nutrient parameters: nitrates, orthophosphates, nitrites, ammonium, silicates and Dissolved inorganic nitrogen (DIN). We do not monitor TN and TP. As far as it concerns physicochemical parameters we monitor for WFD the	We have communicated by email to correct as many problems as possible in the dataset that underlies these analyses. All tables and figures have been updated, removing all values for Cyprus coastal waters as no boundaries for the parameters outlined are yet available.

	following ones: Dissolved oxygen, oxygen saturation, temperature, salinity, electrical conductivity, pH, oxidation-reduction potential. Concerning the <25% reference value that is reported in MSFD, this was defined primarily, until new threshold values are available. We have already contacted the authorities from Greece in order to begin an intercalibration exercise to define common threshold values for all common parameters that we consider appropriate, in the same way we had done in the past for chlorophyll-a and the other BQEs, that led to a nice result for the Eastern Mediterranean Subregion. We would be very much interested in doing a similar work for nutrients and physicochemical parameters and Greek colleagues informed us that they are, too. Unfortunately, this has not been implemented yet and we plan to do it soon. So, we inform you that we have to wait in order to proceed in the definition of reliable and realistic thresholds for the above parameters.”	
Germany	TRAC waters: Objectives: Regarding the paragraph on Fish Directive, German expert consider that the paragraph suggests that 6mg/l for oxygen is an old-fashioned and outdated standard but this is not true. The standard for oxygen is derived from physicologial considerations on the needs of benthic organisms (as correctly discussed in 4.2.1. Experrst suggest to delete this example here.	We have added the following stetement to this paragraph of Objectives: This standard despite derived from physiological considerations on the needs of some taxonomic groups, may not necessarily protect all organisms or life stages and might deserve further attention (but see discussion in section 4.2.1)
Germany	It needs to be acknowledged that the data used in this report reflect the status of reporting to WISE. Some EU Member States have set standards for physico-chemical supporting elements but they do not report these because these elements (ie. Secgcchi depth and Dissolved oxygen in Germany) are not considered in the classification of ecological status as long as the biological quality elements are still classified to be in moderate or worse status.	We had already stated in the Introduction (Objectives) that not all data in WISE was used in this report (see Table 2.1), but only the most commonly used and/or ecologically meaningful for this exercise. In any case, given this clarification by DE we added this additional sentence to the 1.2 Approach section: “In addition, some EU Member States may have set additional standards for physico-chemical supporting elements and not reported those elements to WISE; while these were not considered in

		the classification of ecological status as long as the biological quality elements are still classified to be in moderate or worse status. Nonetheless, where these standards were relevant and made available by MS (e.g. Secchi depth and Dissolved oxygen by Germany) they were also considered in this report."
Germany	Summary data for presentation As commented for freshwaters. Whether this is a meaningful approach remains to be discussed in Ecostat WG A. The resulting 50th percentiles (blue lines in the diagrams) are the product of the values of the MSs that have standards that measure the central tendency only. Other MSs that use other approaches for their standards are excluded. Furthermore the resulting 50th percentile is influenced by the number of different national types (MS with more national types contribute more values to the "overall 50 th percentile" than MSs with a low number of types. The blue lines (50 th percentiles) in the diagrams of the whole report are not representative because of these reasons. The conclusions based on comparisons of MS G/M Boundaries with these none representative "blue lines" are highly questionable.	The graphs treat all standards set as a central tendency in the same way and a further separation into those set using means and those which used medians would have been confusing. We recognise the issue that is being raised here but believe that the graphs in the report gain more from clarity of presentation than they lose through overlooking fine details of statistical distributions. We have also provided additional tables detailing MS differences in summary metrics and annual/seasonal measures use for each SE /water category, so that a more critical analysis is possible by the readers and MS while consulting this overview. The comment about the 50th percentile is valid. There is no perfect solution, but we will consider a more nuanced approach to calculating these percentiles in future reports. Once again, we emphasise that the "big picture" is unlikely to change radically as a result of this.
Germany	Secchi depth and oxygen standards were not reported to WISE, but German expert consider that they should be included in this work.	DE has sent these data (Secchi depth and oxygen standards in April 2021) and we have now included in the final version of the report, all graphs, tables and summary data has been updated in accordance.
Germany	Regarding Dissolved oxygen"two tailed effect": It could also be that the lower values are standards for stratified water bodies while the higher values are standards for non-stratified water bodies	Indeed, and we had cases where subtype specific reasons were raised to justify the standards presented as range, but often no additional details have been provided. We have now added this alternative explanation to the text considerations, in section 3.1 Oxygen: "It could also be that the lower range values are standards for stratified water bodies while the upper range values are standards for non-stratified water bodies. Countries reported differences regarding depth zone sampling (surface, bottom, or mix waters), but

		this information was usually not available to further understand the reasons for standards presented as ranges."
Germany	Results-Oxygen Statistical analyses: It is difficult to read all this statistical information. It should be better described how it can be interpreted, e.g. what does it mean if 89% of the variance is explained by country. I have made a suggestion in the text. Table 3.1 Analyses oft he variance for factorial model relating country and common typ. This type of analysis is completely meaningless to the reader here and in the wholre report. What should we "take home" from these results? It is meaningless whether the differences in G/M Boundaries for different Parameters in different broad types are statistically significant or not. As commented for freshwaters: It is not usefull to compare G/M Boundaries for dissolved oxygen between MSs that have quite different annual average temperatures / climatic conditions. Temperature is linked with the solubility of oxygen in the water Figure 3.1: Annual average temperatures differ between e.g. the Nordic MSs and the Mediterranean MSs but have direct influence on the solubility of oxygen in the water. At higher temperatures lower amount of dissolved oxygen is present at saturation (100%) than at lower tempeartures. The blue line in the diagram is therefore not representative to draw any interpretation about ambitions of MS`s or in order to compare or rate the G/M Boundaries between the MSs. Figure 3.5-Transitional waters. Comparison of TW Dissolved oxygen standards by country: Experts think that we are comparing pears and apples with this analysis. They suggest delete this graph for all elements.	Statistical analyses: We agree that with the available data, and given the lack of information for influencing environmental conditions as e.g. is the case of temperature influencing oxygen solubility in water, it is complicated to perform meaningful statistical analysis, especially when the data available per grouping factor such as geographic regions (GIG, marine region, etc..), which could provide more insight, is also very scarce. In this sense we have decided to drop these statistical analysis at this stage, for all PhCh SE. Graphical overviews: In any case, we still find it is valuable to have a graphical overview of the SE distribution across countries and the common IC types in all EU TRAC waters. We included other tables that may provide additional support to interpret the differences/similarities observed across standards reported by MS. DE data in the current report has been corrected following corrections mentioned by Germany.

	Results- Total Nitrogen Figure 3.22-Mistake. German values should all be black dots (single values) and no minima or maxima! In general the expert would prefer if the analysis is done per country and regional sea since otherwise it is suggested that countries have large ranges which just reflects different properties of the different regional seas that are NOT comparable. Results- Total Phosphorus German values should all be black dots (single values) and no minima or maxima! Values for the Baltic and North Sea should be distinguished e.g. by symbol or colour.	
Germany	HELCOM has only established nutrient boundaries for the open Baltic Sea basins. The nutrient boundaries used in coastal waters are provided by countries to HELCOM and follow national approaches so this conclusion is not justified and should be deleted.	Ok, we have removed/corrected accordimg to your comments, for accuracy, the following related statements: In Nitrogen discussion we removed sentence: "HELCOM has provided a good environmental status thresholds for sea sub-regions (HELCOM, 2015, 2017) but our comparison reveals that most TIN criteria 'present by Baltic countries for coastal waters exceeds HELCOM thresholds." In Phosphorus discussion we corrected the sentence: "Most of the Baltic countries provides TP threshold values in coastal waters, and these Good/Moderate boundaries are, in most of the cases, within the range values reported by MS to HELCOM (2015, 2017) for the good environmental status in the different Baltic sub-regions."
France	Introduction. Linking Trac data to IC type Table 1.1: Not sure that all French standards/thresholds for transitional waters correspond fully to European typologies: on the one hand, for lagoons France have 2 types (including one specific to France) and for estuaries, in the case of nutrients for example, we have ecotypes that have no link with European typologies.	Indeed, there are FR national types not matching common IC type. This was corrected and Table 1.1 updated in report.
France	Which physico chemical elements are used?	Unfortunately, we had to select those SE with minimum available data to allow some overview or comparison. At this stage it was not

	Expert suggest to analyses “other determinand for nutrient condition” as additional indicators.	possible to include “other determinand for nutrient condition” due to lack of information.
France	Figure 2.1: Mistake. Missing information on TDIN for transitional waters like estuaries (NEA 11).	TDIN is equivalent to the TIN in Fig. 2.1, a note was included in the Introduction to clarify that total inorganic nitrogen is equivalent to. dissolved inorganic nitrogen. We have checked Figure 2.1 to confirm and there is FR data in TW for this element.
France	Results-Dissolved oxygen (coastal waters) Experts clarifies that the value reported corresponds to quantile 10 of the bottom water in summer/autumn	Data corrected and Tables/figures updated.
France	Results-TDIN (CW and TW) Figures 3.18 and 3-20: For France and for european type CW-NEA1/26 and TW NEA 11, this is not really an annual average (AA-EQS), but an average on the monthly winter values over 6 years (November to February). Can the symbol (add winter) be corrected on the Figures (winter)?	Data corrected and Tables/figures updated.
France	Results-TDIN (TW) To include that France also uses salinity 33 standardization (not only SE), which means taking into account salinity/dilution gradient.	Data corrected and Tables/figures updated. Sentence corrected in section 3.4.2: “Three countries (BE, FR, SE) reported a set of G/M adjusted to salinity gradient.” You may also see more details in the Appendices.
France	Results-TN (TW) For poly- and euhalin lagoons in France, the 90th percentile is calculated on the 3 summer values (June-August) over 6 years	Data corrected and Tables/figures updated. In the Appendices you can consult details on season also.
France	Results-orthophosphate (TW) Note that for all nutrients (TN, TP, NO3 and PO4) in French lagoons, it is necessary to calculate 90th percentile from the monthly summer data (over 6 years).	Data corrected and Tables/figures updated. In the Appendices you can consult details on season also.
France	Discussion • Figures 4.4 and 4.5. Experts suggest to include the information that that the estuaries are monitored for nutrients in winter (November-February). • Figures 4.6, 4.8, 4.10: Experts suggest to include the information that that the coastal lagoons are monitored for nutrients in summer.	Although this comment refers to section 4.3 (Comparison of G/M boundaries between ECOSTAT 2014 questionnaires and WFD reporting) this aspect was better addressed in the updated section 3. We have included additional tables in all the results sections for each SE that allows to better scrutinize the measures and sampling details behind each summary metric. Also in the Appendices, you can now explore additional tables with this information.

		However, we did not always have this information for all countries reporting and we could not integrate it in a standardized way in the graphs. In section 4.3 the focus is on Countries differences within their own values updates through time and these aspects are not included in there, but provided in section 3 instead.
Italy	Experts think that it would be necessary to study the effect of the use of different combining rules (e.g. averages, one out all out, multimetric, etc.) of the supporting elements, for example applying the rules for all Member states to a common data set.	We agree that this is an important issue; however, it is better considered in discussions about classification than about the process of setting supporting element standards. We recognise in the Best Practice Guide that users need to consider how the degree of precaution are embedded in different approaches, and this interacts with national approaches to combination rules. From our perspective, these differences almost certainly contribute to some of the differences between countries for any given supporting element.
Netherlands	Coastal and Transitional waters: Only one question: how is dealt with salinity effects? We see a strong relationship between salinity and the concentration of nutrients. Also under reference conditions this would be the case. This advocates the use (also in comparison) a correction for salinity. I think that the types are too broad, at least at places with high riverine input.	We agree that salinity is an important factor to consider, however most countries do not refer this information. In the updated version of the report (or its appendices) we have now included some graphs illustrating the SE boundaries distribution along the salinity gradient whenever this information was provided. This was the case for some nutrients.
Romania	Introduction: Expert suggest replace “threshold concentrations” with “threshold values	We agree, not all SE treated in this overview are concentrations (e.g. Secchi depth). This has now been corrected as suggested.
Romania	Need clarification related to this statement: “we suspect, national conventions that predate the WFD and which may benefit from being revisited” There will be, in many cases, good reasons behind these choices, and the differing levels of precaution associated with particular approaches to aggregation may be reflected in the decision-making process. However, there are also, we suspect, national conventions that predate the WFD and which may benefit from being revisited.	To avoid ambiguity and confusion this statement was removed.

Romania	Results-Orthophospate (TW): Delete Romania	Data was updated according to MS request, and Orthophospate standards in TW deleted,; tables and graphs updated accordingly.
Slovenia	Coastal waters Small mistake on page 76, last sentence in the second paragraph of the Nitrogen section. Based on the results of the report, it should probably state Slovenia instead of Croatia, since Greece and Slovenia use similar threshold values for nitrogen, but not Croatia.	You are right. This was corrected in the text. Thanks!
Lithuania	Corrections to data and report (April 2021)	Values in database were corrected/updated according to comments received in April 2021 by MS. Tables and figures in the report were also updated accordingly. Salinity related graphs are also now available in cases where a meaningful number of countries has reported this information. See results of different SE or its correspondent appendices. Small typos in report were corrected according to pdf comments sent. Regarding TW correspondence to common IC types, the data revised in excel (2020) and tables sent (2021) seems to not correspond to comments to report received in April 2021, so these might need further correction in the future if the current updates are still not correct for TW National types T1, T2, T3.
Ireland	Corrections to data (April 2021)	Values in database were corrected/updated according to comments received in April 2021 by MS. Tables and figures in the report were also updated accordingly. Namely: • Match CW Nat Types to IC types – corrected in table. • TDIN is equivalent to the TIN in Fig. 2.1, a note was included in the Introduction to clarify that total inorganic nitrogen is equivalent to. dissolved inorganic nitrogen. • We agree that salinity ranges for each value are very relevant but most countries do not refer to it. In this revised version of the report, we included graphs where this information was available to illustrate its potential influence on boundaries

		set by MS. Please see new figures in report or corresponding appendices. • Likewise, we have included new tables (in report and Appendices) to allow explore this additional information where it has been reported by MS.
Ireland	IE do not use annual mean- we use winter or summer median	Corrections due: unfortunately, we have missed this correction, so IE values commented to be seasonal set at winter or summer have not yet been updated. Note to be updated in future work