We have taken all of the review comments into consideration. We would like to thank the reviewer for these comments which we believe have substantially improved the m/s. The methodological parts are now better described and the formulas are now given in a more correct fashion. Discussion about uncertainties have been expanded. Below we reply in detail on how and if we taken the individual comments into consideration.

**Reviewer #1**

Our response given below each comment

 1 General comments

The paper ‘Nitrogen surface water retention in the Baltic Sea drainage basin’ addresses relevant scientific questions within the scope of HESS. The topic is relevant both scientifically and supporting river basin management and control of N loading to the Baltic Sea. The methods are not very novel themselves, but estimation of N surface water retention in the whole of Baltic sea basin is. My proposal is major revision of the manuscript, taking into account all comments.

My major concern is almost total lack of uncertainty discussion. The authors present one number, 380 000 t of N as annual retention, but not any uncertainty estimates/ranges with different parameterizations by the MESAW model.

Discussion is somewhat short – the authors only mention that high retention in lakes is in accordance with earlier studies – but do not give proper credit to many published N retention studies in parts of the Baltic Sea catchment area, and compare their results to only those of Mörth et al. (2007). It is also misleading that in Intro, the authors refer mostly to **in-stream** retention studies, but in Discussion they point out that **in-lake** retention is of high importance.

ANSWER: We appreciate the comment that the estimation of nitrogen surface water retention in the Baltic Sea is novel. In fact we are not aware of any studies that have assessed the nitrogen retention in all the Baltic Sea drainage area besides the study by Mörth et al (2007).

Regarding uncertainty, we agree that this is a complex issue, and a quantitative assessment of the uncertainty associated with such complex mechanisms is, we believe, beyond the scope of this paper but offers potential for future work. Nonetheless, it is important to recognise and in a qualitative way discuss the uncertainty. We have thus included these uncertainty aspects at a variety of places throughout the revised m/s and have now included a particular uncertainty sub-section in the Results and Discussion.

We don’t agree that the references given in Introduction are biased towards those of references on instream retention and that lake retention is less emphasised. In fact most references in the Introduction is of general character including both instream and lake retention. In all cases, this obvious confusion is now better explained in the Introduction.

2 Specific comments

Title is good and abstract well written.

Section 2 Material and methods:

-the authors mention that retention is assumed to be the same for source categories P (point sources), dominated by inorganic load, and sources category S (total losses) which include varying shares of N, more in organic-N form. In these models, the assumptions are needed, but this assumption could be discussed in uncertainty discussion

ANSWER: We have modified the general formula 1 and removed R1, R2 and R3 and instead replaced it with R which was specific for this study. The formula 1 now reads as:



We agree that such differentiation of retention to the various source categories would have been ideal but is almost impossible to parametrise and would have required more data upstream the river mouths and also data on inorganic and organic nitrogen. Nonetheless this relevant issue has been better included in the discussion part

Section 3 Results and discussion:

-To make it more clear, the authors should mention also the estimated total gross N load, 950 000 t N annually. Also here, comparison to earlier estimates would be reasonable to have.

ANSWER: Thanks for the remark on the missing information about the estimated gross load from the model. It is now included in the very beginning of the Results section.

We are not aware of any other studies of gross emissions estimates for these 117 Baltic Sea basins besides a very old study conducted by the first-author 15 years ago (which was felt outdated to include).

-it is true that there is not apparent relationship between specific N load and share of wetland area, but from Fig 5b we can notice that load is always low in basins where wetland-% is >15%

ANSWER: The statistical analysis do not give any statistical significant parameters for wetlands (Table 1). It should be noted that the classification of wetlands is rather rough from the data source and given as joint expression of all wetlands ranging from marches to peatland bogs. We don’t have any possibility to include this is the analysis. In all cases for the reasons given in the paper it will have less importance for the overall objective to estimate the total N-retention in surface waters for the 117 basins. The reasons for the relatively low unit-area loads for the basins with >15% wetland area is due to the fact that they are all located in Finland or northern Sweden with low population densities and little agricultural area. This information has been included in the revised m/s.

-the term ‘Other’ is misleading, if these areas are practically all forests (are they?), the authors should include Fig 5d) of forests also into discussion

ANSWER: Indeed most of the land use category ‘Other’ is forest which was stated only once in the initially submitted m/s. We have included clarification of this in the revised m/s and in the Table headings.

We have included the following sentence on the missing comment on Figure 5d: There is a clear negative relationship between the unit-area loads of nitrogen and the share of ‘Other’ land (i.e. primarily forest’)

-the authors could also acknowledge PLC database by HELCOM which they use a lot, and to include reference /web-page. Which institutes provide data to that database?

ANSWER: The PLC-reports and the data source is already properly acknowledged in the Material and Methods section. We have added the web-site to make it even more clear ([www.helcom.fi](http://www.helcom.fi)). Below a list of the organizations providing data to PLC data base from the different countries and we believe it will become too exhaustive according to us to include it in the paper since it will include a lengthy addition of this form:

DCE -Danish Centre for Environment and Energy Aarhus University
Vejlsoevej 25
DK-8600 Silkeborg

Estonian Environment Agency
Mustamäe tee 33
EE-10616 Tallinn

Federal Environment Agency
Section II 2.2
Woerlitzer Platz 1
Dessau-Rosslau

Environmental Protection Agency
Environmental Status Assessment  Department
River Basin Management Division
Juozapavičiaus st. 9,
LT-09311 Vilnius

Institute of Meteorology and Water  Management,

National Research Institute
Jordana Str. 10/11
PL-40 056 Katowice

Saint-Petersburg Public Organization
"Ecology and Business"
Sabrikovskaya Str. 37, Office No. 307
Post Office Box 66
RU-197374 St. Petersburg

Department of Aquatic Sciences and Assessment
Swedish University of Agricultural  Sciences, SLU
P.O. Box 7050
SE-750 07 Uppsala

Finnish Environment Institute (SYKE) )

Mechelingatan 34a,

FI-00260 Helsingfors,

PB 140 Helsingfors

-the authors present very detailed results of lake and in-stream retention in Table A2, but do not discuss of the average percentages of these. How is the share between these estimates and how reliable/uncertain they are? For example for Neva river basin, retention in total surface water is estimated as 0.74, but lake+In-stream retention (0.91) seems not to be in accordance with the total?

ANSWER: Averaged over all basins, mean lake retention is 25% whereas the estimated instream retention is 5%. This information is included now.

Table A2 refer to the independent estimates of lake and in-stream retention respectively plus the total. For obvious reasons the independent percentages for lake and instream retention cannot be simply added (see methods). For example in Neva the instream and lake retention is 0.262 and 0.652, respectively. Certainly the combined retention is 0.74 according to the following simple calculation of total retention: 1- ((1-0.262)\*(1-0.652))=0.74.

We assume that the reviewer have anticipated an additive response which for obvious reasons is not true.

Given the confusion we have modified formulas 3 and 4 and replaced it with:

\*  *i = 1,2,...,n* (3)

|  |  |  |
| --- | --- | --- |
|  |  |  |

3 Technical corrections

-the last paragraph of Intro should be more concise and short, with no details on population and land use. Instead, there could be introducing parapgrah in 2 Materials and methods, describing the area

ANSWER: This text part is appropriately moved to Intro to Section 2.

-the estimates of annual N loads in Table A2 give an over-optimistic impression of the accuracy, e.g. Odra 70 289 195 kg N/yr !, I would propose to use tonnes N/yr

ANSWER: Unlike many other papers in similar fields, we in fact give all input data for the model as well as outputs for each single river basin so the analysis can be replicated if needed. We feel that this is in line with the recent trend to publish ‘rawdata’ in peer-review journals. We leave this question to the Editor-in-Chief to decide. We have no problem to make the indata and result files much more aggregated and without the decimal precision.

We have taken all of the review comments into consideration. We would like to thank the reviewer for these comments which we believe have substantially improved the m/s. The methodological parts are now better described and the formulas are now given in a more correct fashion. Discussion about uncertainties have been expanded. Below we reply in detail on how and if we taken the individual comments into consideration.

**Reviewer nr 2**

 The article is interesting to read, but seems to be lacking some information that is necessary for better understanding by the reader. I suggest that the authors describe certain parts more clearly and consider setting more subtitles to keep different parts of model description and results apart.

A comparison or discussion on pros and cons of the used model for retention in comparison to other possible models would be interesting.

ANSWERS: We have added more sub-section headings in the Results and Discussion section to increase the readability. It is now divided into

3.1 Parametrisation results

3.2 Major retention estimate results

3.3 Uncertainty aspects

We have expanded the discussion about pros and cons of the used model. A quantitative comparison with other models is outside the scope of this paper but we have included a qualitative discussion in the revised m/s (section 3.3). More precisely, the model used is an advanced regression model that goes beyond normal multiple regression analysis and can be seen as comparable with the SPARROW model (Schwarz, G.E., Smith, R.A., Alexander, R.B., and Gray, J.R., 2001).

Some questions to be clarified:

a) It is stated clearly which inputs are used, but the model description is confusing. Which parameters are estimated? Are all parameters areas specific, and if so do they vary a lot between areas? How is expert knowledge used in the fitting of the model. E.g. for equation 1 are there parameters estimated in all parts of this formuls (S, P, D and R) or are some of them observed or considered known. This information is given in the text later, should however be given right after formula 1 (e.g. page 10836 line 14 states what is assumed to be known, move this ahead).

ANSWER: The Model description has been substantially improved. All the formulas are now clearly given. Initially we described the general model given in Grimvall&Stålnacke (1996) but have in the revised version focused better on the adjustment made and parametric function used in this particular case study. We believe that this have increased the readability. In fact all the 4 given formulas have been changed. They now reads as:

 (1)

where *Li* is the load at outlet of basin i;

 *Si* is total losses from soil to water in basin *i*;

 *Pi* is the point source discharges (WWTP and industry) to waters in basin *i*;

 *Di* is the atmospheric deposition on surface waters in sub-basin *i*;

 *R* denote the retention for the source emissions *S, P* and *D*, respectively;

*n* is the number of basins, and

 *εI*is the statistical error term.

The total diffuse loss of N from soil to water, *Si*, in the *i*th sub-basin was assumed to be a function of the land cover (Eq. (2)):

 *Si = ( θ1a1i + θ2a2i + θ3a3i )* (2a)

where *a1i, a2i* and *a3i* in our study refer to the areas of three land cover classes, i.e. cultivated land, wetlands and other land (mainly forests), respectively.*θ1, θ2* and *θ3* are unknown emission coefficients for the three land use categories that are statistically estimated in MESAW jointly with the retention (see Eq. (3) below). The point source emissions, *Pi,* and atmospheric deposition on surface waters, *Di*, were assumed to be known (see Section 2.1).

Throughout the exploratory analysis we found that certain basins deviated from the relationship and in most cases also where geographically located near to each other. Thus we introduced a ‘grouping variable’ according to the following:

 *Si = ( θ1a1i + θ2a2i + θ3a3i ) \** ωj (2b)

where each group j consisted of 2 or more basins depending on the model run (see Table 1) and where ω is the unknown coefficient(s). The model was run with different combinations of basin sub-groups in order to obtain reasonable model coefficients and load estimates (i.e. little deviation between predicted and observed loads). The grouping of basins was based on prior knowledge of similarities between basins as well as geographic location. For example, the 10 smaller Danish sub-basins formed one group, as a residual analysis showed that these sub-basins deviated from the general relationships. In its practical meaning, we simply adjusted the ‘global’ diffuse emission coefficients to the local conditions (despite we don’t know the underlying causes). This can be justified since applying the same coefficient to such a large drainage basin (1 745 000 km2) seems less logic.

…… Irrespective of the exact retention mechanism, the parameterisation of the retention in the different basins was after several exploratory runs with alternative models done with the following empirical function (Eqs. (3) and (4)):

\*  *i = 1,2,...,n* (3)

where λ1 and λ2 denotes a non-negative parameter and *Ri* denote the retention in the *i*th basin. The empirical function were in our case derived from the conception that the removal of N takes place primarily in the surface waters (both instream and in lakes). The first part of the function reflects the instream retention whereas the second part reflects the retention in lakes and reservoirs.

Regarding the question if all the parameters are areas specific, and if so do they vary a lot between areas?

The final model include 9 estimated parameters (Model run #4 in table 1) and they don’t vary between the drainage basins besides the case with the grouping-variables (see answer under comment f) below). The diffuse emission parameters give the area-specific loads (i.e., source emissions). For example, Model run 4 for cultivated land gives a point estimate of 1073 kg km-2. Interestingly this is a value that normally could be monitored in small agricultural catchments in the Nordic/Baltic region (Stålnacke et al. 2014). We have included a better clarification of this in the revised m/s.

b) The total loss (S) is modelled from 3 land cover classes (cultivated, wetlands and other land). Do these 3 land cover classes add up to 100% of land cover? If so this should influence the estimation of the 3 parameters, since the variables will be linearly correlated. How is this handled? If there are land cover classes not in the model, this should be stated clearly.

ANSWER: Yes the 3 land cover classes adds up to 100% and are for sure inter-correlated. This will have less influence on the method applied although there is always a risk of multicollinearity of these kind of regression-type of models. It should be noted that the model inputs are areas of the land cover and not the percentages which will decrease the risk of multicollineariety. Experiences with the MESAW models as also given in the earlier quoted papers in different geographical areas (Liden et al; Vassiljev&Stålnacke, Vassilijev et al and Povilaitis et al) have not indicated any problem with possible interrelated explanatory variables..

In addition, parameter estimates displayed reasonable stability; little change occurred in the values of the most statistically significant model coefficients when additional variables were added in exploratory regressions.

c) Two formulas are given to compute/estimate retention. I the difference between them that one is used if there are lakes in the area, whereas the other one is used if there are no lakes? Or how do you choose between these for the different basins? Is lambda the same in these two models, i.e. if lambda a common estimate for both equations? State in the article. Hesse et al. ECOLOGICAL MODELLING Volume: 269 Pages: 70-85 made comparisons for different retention models. This might be interesting for you to comment in the article.

ANSWER: Both formulas for retention (Eq 3 and 4) is used in the simultaneous estimation of the source emission coffecients and retention coefficients. There are in fact 2 lamdas that is estimated. Formula 3 and 4 have been corrected accordingly

Given the confusion we have modified formulas 3 and 4 and replaced it with:

\*  *i = 1,2,...,n* (3)

A sentence that better explains this is included. The reference to Hesse et al have been included. Thanks for that reference.

d) The risk of overfitting/overparametrisation is mentioned and given as reason that retention parameters are the same for all source categories. Is this reasonable and can be motivated? How? How do you control for overfitting in this model, is it by only allowing a few parameters to vary or do you control it? Would any kind of cross-validation help to avoid overfitting?

ANSWER: We have the removed the sentence on ovefitting/overparametrisation. In total, 9 parameters were fitted on the 88 observations. Parameter estimates displayed reasonable stability; little change occurred in the values of the most statistically significant model coefficients when additional variables were added in exploratory regressions. Moreover, the diffuse source coeffcients (thetas) where all realistic in its value which is further explained in the revised m/s. We thus regard the issue with overfitting/overparametrisation as less likely.

e) In page 10837 line 9 you talk about the total N retention that is estimated. Does this regard fitting R\*Si+R\*Pi+ R\*Di, related to equation 1? When you do fitting on different groups, are parameter estimated individually for a group? If 10 danish subbasins form one group, how many parameters do you estimated from those, is it 4 (3 theta and 1 lambda) or more? Are estimates for thetas and lambda very different for the groups of basins? Parameter estimates should be given, at least as example.

ANSWER: We have now better explained how the total retention is estimated and how this is related to Eq1. The question on the grouping parameter/variable is explained under answer f) below. The parameter estimates is given in Table 1 and we have in addition included the thetas and lamda into the table heading for clarification and better references to the formulas given in Material and Methods

f) If groupings of basins is made due to geographical location or similarities, would not that suggest dependence/correlation between the basins and influence p-values (with the concept of statistical inference based on independent observations). The error term in (1) does not indicate that dependencies are taken into account. Can p-values be trusted?

ANSWER: This is a misunderstanding. The basins are not merged. Instead we during the modelling found that some basins deviated from the general relationship and most of these basins were in fact located geographically in the same geographical region. To the end, we identified 3 such ‘groups’ of basins (lower part of table 1). This will not by any means affect the independency criteria in this kind of statistical modelling. Instead we were with this ‘grouping’ able to differentiate eg the diffuse emission coefficients. For example, it is known that basins in Denmark and southern Sweden (due to more intensive agriculture) differ from the ones on northern Finland and Sweden. The procedure applied can be seen as introducing a dummy variable in normal multiple regression.

g) In the results unit-area specific loads are discussed. As the model is designed to predict N load rather than unit-area loads: was this expected? Could the model be adjusted if unit-area loads are interesting? Could this be a result of overfitting in the original model?

ANSWER: The model was fitted to river loads given in kg. We wanted to show-case the model results also as unit-area loads since this give higher credibility to the results and analysis. Principally, the model is generic and can also be applied with any dependent variable.

h) In figure 4 the relationship between estimated retention and total drainage area are given. In these figures it seems that drainage area has no influence on retention in %, whereas lake area (%) has a clear nonlinear relationship. How do these curves related to equations 3 and 4? Probably the equations and estimated parameter lambda are used to compute the estimated retention, i.e. the curves should reflect the relation in 3 and 4. Is this true? The line shown in the plot ‘retention and lake area’, why is it plotted there? How is it related to the model? Since this line does not fit well, does this indicate that the model does not fit well?

ANSWER: We agree that figure 4 can be confusing for the reader. The intention was to illustrate how the estimated retention (in %) is pair-wise correlated to the 2 main variables (lake are and drainage area) included in the retention expression. Apparently there is a strong curvelinear relationship between retention and the lake-share in a drainage basin and that there is a much weaker relationship between the retention and size of drainage basin. A further discussion about the interpretation of this is given in the revised m/s. We have also removed the fitted line in Figure 4 (left panel) since it is not connected to the parameter estimation at all.

i) Also the function fitted to specific load and lake area (%) is strange, why do you use this fitted line instead of an exponential/logarithmic relationsship or a square-root relationship. Where does the function come from? How is it motivated?

ANSWER:The figure 5 on area-specific N-loads vs lake area (%) is just given as an illustration on the relationships in the input data and just a support to the retention formula applied. It is given to the reader as an example. We have removed the fitted lines and the regression equations from the figure to avoid confusion.

Smaller notes

Relative differences are used to give equal weights to small and large basins. A motivation why this is a good choice in this context would be appreciated.

ANSWER: The model as given in formula 1 is based on loads at river mouths. In order to avoid that large basins (large basins will for most cases have more loads than small catchments) will have more effect in the parameter estimates we used the relative differences between observed and fitted loads. This is a standard procedure in many statistical analysis of this kind.

We have taken all of the review comments into consideration. We would like to thank the reviewer for these comments which we believe have substantially improved the m/s. The methodological parts are now better described and the formulas are now given in a more correct fashion. Discussion about uncertainties have been expanded. Below we reply in detail on how and if we taken the individual comments into consideration.

**Reviewer #1**

Our response given below each comment

 1 General comments

The paper ‘Nitrogen surface water retention in the Baltic Sea drainage basin’ addresses relevant scientific questions within the scope of HESS. The topic is relevant both scientifically and supporting river basin management and control of N loading to the Baltic Sea. The methods are not very novel themselves, but estimation of N surface water retention in the whole of Baltic sea basin is. My proposal is major revision of the manuscript, taking into account all comments.

My major concern is almost total lack of uncertainty discussion. The authors present one number, 380 000 t of N as annual retention, but not any uncertainty estimates/ranges with different parameterizations by the MESAW model.

Discussion is somewhat short – the authors only mention that high retention in lakes is in accordance with earlier studies – but do not give proper credit to many published N retention studies in parts of the Baltic Sea catchment area, and compare their results to only those of Mörth et al. (2007). It is also misleading that in Intro, the authors refer mostly to **in-stream** retention studies, but in Discussion they point out that **in-lake** retention is of high importance.

ANSWER: We appreciate the comment that the estimation of nitrogen surface water retention in the Baltic Sea is novel. In fact we are not aware of any studies that have assessed the nitrogen retention in all the Baltic Sea drainage area besides the study by Mörth et al (2007).

Regarding uncertainty, we agree that this is a complex issue, and a quantitative assessment of the uncertainty associated with such complex mechanisms is, we believe, beyond the scope of this paper but offers potential for future work. Nonetheless, it is important to recognise and in a qualitative way discuss the uncertainty. We have thus included these uncertainty aspects at a variety of places throughout the revised m/s and have now included a particular uncertainty sub-section in the Results and Discussion.

We don’t agree that the references given in Introduction are biased towards those of references on instream retention and that lake retention is less emphasised. In fact most references in the Introduction is of general character including both instream and lake retention. In all cases, this obvious confusion is now better explained in the Introduction.

2 Specific comments

Title is good and abstract well written.

Section 2 Material and methods:

-the authors mention that retention is assumed to be the same for source categories P (point sources), dominated by inorganic load, and sources category S (total losses) which include varying shares of N, more in organic-N form. In these models, the assumptions are needed, but this assumption could be discussed in uncertainty discussion

ANSWER: We have modified the general formula 1 and removed R1, R2 and R3 and instead replaced it with R which was specific for this study. The formula 1 now reads as:



We agree that such differentiation of retention to the various source categories would have been ideal but is almost impossible to parametrise and would have required more data upstream the river mouths and also data on inorganic and organic nitrogen. Nonetheless this relevant issue has been better included in the discussion part

Section 3 Results and discussion:

-To make it more clear, the authors should mention also the estimated total gross N load, 950 000 t N annually. Also here, comparison to earlier estimates would be reasonable to have.

ANSWER: Thanks for the remark on the missing information about the estimated gross load from the model. It is now included in the very beginning of the Results section.

We are not aware of any other studies of gross emissions estimates for these 117 Baltic Sea basins besides a very old study conducted by the first-author 15 years ago (which was felt outdated to include).

-it is true that there is not apparent relationship between specific N load and share of wetland area, but from Fig 5b we can notice that load is always low in basins where wetland-% is >15%

ANSWER: The statistical analysis do not give any statistical significant parameters for wetlands (Table 1). It should be noted that the classification of wetlands is rather rough from the data source and given as joint expression of all wetlands ranging from marches to peatland bogs. We don’t have any possibility to include this is the analysis. In all cases for the reasons given in the paper it will have less importance for the overall objective to estimate the total N-retention in surface waters for the 117 basins. The reasons for the relatively low unit-area loads for the basins with >15% wetland area is due to the fact that they are all located in Finland or northern Sweden with low population densities and little agricultural area. This information has been included in the revised m/s.

-the term ‘Other’ is misleading, if these areas are practically all forests (are they?), the authors should include Fig 5d) of forests also into discussion

ANSWER: Indeed most of the land use category ‘Other’ is forest which was stated only once in the initially submitted m/s. We have included clarification of this in the revised m/s and in the Table headings.

We have included the following sentence on the missing comment on Figure 5d: There is a clear negative relationship between the unit-area loads of nitrogen and the share of ‘Other’ land (i.e. primarily forest’)

-the authors could also acknowledge PLC database by HELCOM which they use a lot, and to include reference /web-page. Which institutes provide data to that database?

ANSWER: The PLC-reports and the data source is already properly acknowledged in the Material and Methods section. We have added the web-site to make it even more clear ([www.helcom.fi](http://www.helcom.fi)). Below a list of the organizations providing data to PLC data base from the different countries and we believe it will become too exhaustive according to us to include it in the paper since it will include a lengthy addition of this form:

DCE -Danish Centre for Environment and Energy Aarhus University
Vejlsoevej 25
DK-8600 Silkeborg

Estonian Environment Agency
Mustamäe tee 33
EE-10616 Tallinn

Federal Environment Agency
Section II 2.2
Woerlitzer Platz 1
Dessau-Rosslau

Environmental Protection Agency
Environmental Status Assessment  Department
River Basin Management Division
Juozapavičiaus st. 9,
LT-09311 Vilnius

Institute of Meteorology and Water  Management,

National Research Institute
Jordana Str. 10/11
PL-40 056 Katowice

Saint-Petersburg Public Organization
"Ecology and Business"
Sabrikovskaya Str. 37, Office No. 307
Post Office Box 66
RU-197374 St. Petersburg

Department of Aquatic Sciences and Assessment
Swedish University of Agricultural  Sciences, SLU
P.O. Box 7050
SE-750 07 Uppsala

Finnish Environment Institute (SYKE) )

Mechelingatan 34a,

FI-00260 Helsingfors,

PB 140 Helsingfors

-the authors present very detailed results of lake and in-stream retention in Table A2, but do not discuss of the average percentages of these. How is the share between these estimates and how reliable/uncertain they are? For example for Neva river basin, retention in total surface water is estimated as 0.74, but lake+In-stream retention (0.91) seems not to be in accordance with the total?

ANSWER: Averaged over all basins, mean lake retention is 25% whereas the estimated instream retention is 5%. This information is included now.

Table A2 refer to the independent estimates of lake and in-stream retention respectively plus the total. For obvious reasons the independent percentages for lake and instream retention cannot be simply added (see methods). For example in Neva the instream and lake retention is 0.262 and 0.652, respectively. Certainly the combined retention is 0.74 according to the following simple calculation of total retention: 1- ((1-0.262)\*(1-0.652))=0.74.

We assume that the reviewer have anticipated an additive response which for obvious reasons is not true.

Given the confusion we have modified formulas 3 and 4 and replaced it with:

\*  *i = 1,2,...,n* (3)

|  |  |  |
| --- | --- | --- |
|  |  |  |

3 Technical corrections

-the last paragraph of Intro should be more concise and short, with no details on population and land use. Instead, there could be introducing parapgrah in 2 Materials and methods, describing the area

ANSWER: This text part is appropriately moved to Intro to Section 2.

-the estimates of annual N loads in Table A2 give an over-optimistic impression of the accuracy, e.g. Odra 70 289 195 kg N/yr !, I would propose to use tonnes N/yr

ANSWER: Unlike many other papers in similar fields, we in fact give all input data for the model as well as outputs for each single river basin so the analysis can be replicated if needed. We feel that this is in line with the recent trend to publish ‘rawdata’ in peer-review journals. We leave this question to the Editor-in-Chief to decide. We have no problem to make the indata and result files much more aggregated and without the decimal precision.

We have taken all of the review comments into consideration. We would like to thank the reviewer for these comments which we believe have substantially improved the m/s. The methodological parts are now better described and the formulas are now given in a more correct fashion. Discussion about uncertainties have been expanded. Below we reply in detail on how and if we taken the individual comments into consideration.

**Reviewer nr 2**

 The article is interesting to read, but seems to be lacking some information that is necessary for better understanding by the reader. I suggest that the authors describe certain parts more clearly and consider setting more subtitles to keep different parts of model description and results apart.

A comparison or discussion on pros and cons of the used model for retention in comparison to other possible models would be interesting.

ANSWERS: We have added more sub-section headings in the Results and Discussion section to increase the readability. It is now divided into

3.1 Parametrisation results

3.2 Major retention estimate results

3.3 Uncertainty aspects

We have expanded the discussion about pros and cons of the used model. A quantitative comparison with other models is outside the scope of this paper but we have included a qualitative discussion in the revised m/s (section 3.3). More precisely, the model used is an advanced regression model that goes beyond normal multiple regression analysis and can be seen as comparable with the SPARROW model (Schwarz, G.E., Smith, R.A., Alexander, R.B., and Gray, J.R., 2001).

Some questions to be clarified:

a) It is stated clearly which inputs are used, but the model description is confusing. Which parameters are estimated? Are all parameters areas specific, and if so do they vary a lot between areas? How is expert knowledge used in the fitting of the model. E.g. for equation 1 are there parameters estimated in all parts of this formuls (S, P, D and R) or are some of them observed or considered known. This information is given in the text later, should however be given right after formula 1 (e.g. page 10836 line 14 states what is assumed to be known, move this ahead).

ANSWER: The Model description has been substantially improved. All the formulas are now clearly given. Initially we described the general model given in Grimvall&Stålnacke (1996) but have in the revised version focused better on the adjustment made and parametric function used in this particular case study. We believe that this have increased the readability. In fact all the 4 given formulas have been changed. They now reads as:

 (1)

where *Li* is the load at outlet of basin i;

 *Si* is total losses from soil to water in basin *i*;

 *Pi* is the point source discharges (WWTP and industry) to waters in basin *i*;

 *Di* is the atmospheric deposition on surface waters in sub-basin *i*;

 *R* denote the retention for the source emissions *S, P* and *D*, respectively;

*n* is the number of basins, and

 *εI*is the statistical error term.

The total diffuse loss of N from soil to water, *Si*, in the *i*th sub-basin was assumed to be a function of the land cover (Eq. (2)):

 *Si = ( θ1a1i + θ2a2i + θ3a3i )* (2a)

where *a1i, a2i* and *a3i* in our study refer to the areas of three land cover classes, i.e. cultivated land, wetlands and other land (mainly forests), respectively.*θ1, θ2* and *θ3* are unknown emission coefficients for the three land use categories that are statistically estimated in MESAW jointly with the retention (see Eq. (3) below). The point source emissions, *Pi,* and atmospheric deposition on surface waters, *Di*, were assumed to be known (see Section 2.1).

Throughout the exploratory analysis we found that certain basins deviated from the relationship and in most cases also where geographically located near to each other. Thus we introduced a ‘grouping variable’ according to the following:

 *Si = ( θ1a1i + θ2a2i + θ3a3i ) \** ωj (2b)

where each group j consisted of 2 or more basins depending on the model run (see Table 1) and where ω is the unknown coefficient(s). The model was run with different combinations of basin sub-groups in order to obtain reasonable model coefficients and load estimates (i.e. little deviation between predicted and observed loads). The grouping of basins was based on prior knowledge of similarities between basins as well as geographic location. For example, the 10 smaller Danish sub-basins formed one group, as a residual analysis showed that these sub-basins deviated from the general relationships. In its practical meaning, we simply adjusted the ‘global’ diffuse emission coefficients to the local conditions (despite we don’t know the underlying causes). This can be justified since applying the same coefficient to such a large drainage basin (1 745 000 km2) seems less logic.

…… Irrespective of the exact retention mechanism, the parameterisation of the retention in the different basins was after several exploratory runs with alternative models done with the following empirical function (Eqs. (3) and (4)):

\*  *i = 1,2,...,n* (3)

where λ1 and λ2 denotes a non-negative parameter and *Ri* denote the retention in the *i*th basin. The empirical function were in our case derived from the conception that the removal of N takes place primarily in the surface waters (both instream and in lakes). The first part of the function reflects the instream retention whereas the second part reflects the retention in lakes and reservoirs.

Regarding the question if all the parameters are areas specific, and if so do they vary a lot between areas?

The final model include 9 estimated parameters (Model run #4 in table 1) and they don’t vary between the drainage basins besides the case with the grouping-variables (see answer under comment f) below). The diffuse emission parameters give the area-specific loads (i.e., source emissions). For example, Model run 4 for cultivated land gives a point estimate of 1073 kg km-2. Interestingly this is a value that normally could be monitored in small agricultural catchments in the Nordic/Baltic region (Stålnacke et al. 2014). We have included a better clarification of this in the revised m/s.

b) The total loss (S) is modelled from 3 land cover classes (cultivated, wetlands and other land). Do these 3 land cover classes add up to 100% of land cover? If so this should influence the estimation of the 3 parameters, since the variables will be linearly correlated. How is this handled? If there are land cover classes not in the model, this should be stated clearly.

ANSWER: Yes the 3 land cover classes adds up to 100% and are for sure inter-correlated. This will have less influence on the method applied although there is always a risk of multicollinearity of these kind of regression-type of models. It should be noted that the model inputs are areas of the land cover and not the percentages which will decrease the risk of multicollineariety. Experiences with the MESAW models as also given in the earlier quoted papers in different geographical areas (Liden et al; Vassiljev&Stålnacke, Vassilijev et al and Povilaitis et al) have not indicated any problem with possible interrelated explanatory variables..

In addition, parameter estimates displayed reasonable stability; little change occurred in the values of the most statistically significant model coefficients when additional variables were added in exploratory regressions.

c) Two formulas are given to compute/estimate retention. I the difference between them that one is used if there are lakes in the area, whereas the other one is used if there are no lakes? Or how do you choose between these for the different basins? Is lambda the same in these two models, i.e. if lambda a common estimate for both equations? State in the article. Hesse et al. ECOLOGICAL MODELLING Volume: 269 Pages: 70-85 made comparisons for different retention models. This might be interesting for you to comment in the article.

ANSWER: Both formulas for retention (Eq 3 and 4) is used in the simultaneous estimation of the source emission coffecients and retention coefficients. There are in fact 2 lamdas that is estimated. Formula 3 and 4 have been corrected accordingly

Given the confusion we have modified formulas 3 and 4 and replaced it with:

\*  *i = 1,2,...,n* (3)

A sentence that better explains this is included. The reference to Hesse et al have been included. Thanks for that reference.

d) The risk of overfitting/overparametrisation is mentioned and given as reason that retention parameters are the same for all source categories. Is this reasonable and can be motivated? How? How do you control for overfitting in this model, is it by only allowing a few parameters to vary or do you control it? Would any kind of cross-validation help to avoid overfitting?

ANSWER: We have the removed the sentence on ovefitting/overparametrisation. In total, 9 parameters were fitted on the 88 observations. Parameter estimates displayed reasonable stability; little change occurred in the values of the most statistically significant model coefficients when additional variables were added in exploratory regressions. Moreover, the diffuse source coeffcients (thetas) where all realistic in its value which is further explained in the revised m/s. We thus regard the issue with overfitting/overparametrisation as less likely.

e) In page 10837 line 9 you talk about the total N retention that is estimated. Does this regard fitting R\*Si+R\*Pi+ R\*Di, related to equation 1? When you do fitting on different groups, are parameter estimated individually for a group? If 10 danish subbasins form one group, how many parameters do you estimated from those, is it 4 (3 theta and 1 lambda) or more? Are estimates for thetas and lambda very different for the groups of basins? Parameter estimates should be given, at least as example.

ANSWER: We have now better explained how the total retention is estimated and how this is related to Eq1. The question on the grouping parameter/variable is explained under answer f) below. The parameter estimates is given in Table 1 and we have in addition included the thetas and lamda into the table heading for clarification and better references to the formulas given in Material and Methods

f) If groupings of basins is made due to geographical location or similarities, would not that suggest dependence/correlation between the basins and influence p-values (with the concept of statistical inference based on independent observations). The error term in (1) does not indicate that dependencies are taken into account. Can p-values be trusted?

ANSWER: This is a misunderstanding. The basins are not merged. Instead we during the modelling found that some basins deviated from the general relationship and most of these basins were in fact located geographically in the same geographical region. To the end, we identified 3 such ‘groups’ of basins (lower part of table 1). This will not by any means affect the independency criteria in this kind of statistical modelling. Instead we were with this ‘grouping’ able to differentiate eg the diffuse emission coefficients. For example, it is known that basins in Denmark and southern Sweden (due to more intensive agriculture) differ from the ones on northern Finland and Sweden. The procedure applied can be seen as introducing a dummy variable in normal multiple regression.

g) In the results unit-area specific loads are discussed. As the model is designed to predict N load rather than unit-area loads: was this expected? Could the model be adjusted if unit-area loads are interesting? Could this be a result of overfitting in the original model?

ANSWER: The model was fitted to river loads given in kg. We wanted to show-case the model results also as unit-area loads since this give higher credibility to the results and analysis. Principally, the model is generic and can also be applied with any dependent variable.

h) In figure 4 the relationship between estimated retention and total drainage area are given. In these figures it seems that drainage area has no influence on retention in %, whereas lake area (%) has a clear nonlinear relationship. How do these curves related to equations 3 and 4? Probably the equations and estimated parameter lambda are used to compute the estimated retention, i.e. the curves should reflect the relation in 3 and 4. Is this true? The line shown in the plot ‘retention and lake area’, why is it plotted there? How is it related to the model? Since this line does not fit well, does this indicate that the model does not fit well?

ANSWER: We agree that figure 4 can be confusing for the reader. The intention was to illustrate how the estimated retention (in %) is pair-wise correlated to the 2 main variables (lake are and drainage area) included in the retention expression. Apparently there is a strong curvelinear relationship between retention and the lake-share in a drainage basin and that there is a much weaker relationship between the retention and size of drainage basin. A further discussion about the interpretation of this is given in the revised m/s. We have also removed the fitted line in Figure 4 (left panel) since it is not connected to the parameter estimation at all.

i) Also the function fitted to specific load and lake area (%) is strange, why do you use this fitted line instead of an exponential/logarithmic relationsship or a square-root relationship. Where does the function come from? How is it motivated?

ANSWER:The figure 5 on area-specific N-loads vs lake area (%) is just given as an illustration on the relationships in the input data and just a support to the retention formula applied. It is given to the reader as an example. We have removed the fitted lines and the regression equations from the figure to avoid confusion.

Smaller notes

Relative differences are used to give equal weights to small and large basins. A motivation why this is a good choice in this context would be appreciated.

ANSWER: The model as given in formula 1 is based on loads at river mouths. In order to avoid that large basins (large basins will for most cases have more loads than small catchments) will have more effect in the parameter estimates we used the relative differences between observed and fitted loads. This is a standard procedure in many statistical analysis of this kind.

**Nitrogen surface water retention in the Baltic Sea drainage basin**

**P. Stålnacke 1, A. Pengerud 1, A. Vassiljev 2, E. Smedberg 3, C.-M. Mörth3, H. E. Hägg3, C. Humborg3 and H. E. Andersen4**

[1]Bioforsk - Norwegian Institute for Agricultural and Environmental Research, Fr. A. Dahls vei 20, NO-1430 Ås, Norway. per.stalnacke@bioforsk.no, annelene.pengerud@bioforsk.no

[2] Tallinn University of Technology, Ehitajate tee 5, Tallinn 19086, Estonia. avasil@staff.ttu.ee

[3] Baltic Sea Centre, Stockholm University, 106 91 Stockholm, Sweden

erik.smedberg@su.se, magnus.morth@geo.su.se, hanna.hagg@stockholmresilience.su.se, christoph.humborg@itm.su.se

 [] Aarhus University, Department of Bioscience, Vejlsøvej 25, 8600 Denmark. hea@dmu.dk

Correspondence to: P. Stålnacke (per.stalnacke@bioforsk.no).

**Abstract**

In this paper, we estimate the surface water retention of nitrogen (N) in all the 117 drainage basins to the Baltic Sea with the use of a statistical model (MESAW) for source apportionment of riverine loads of pollutants. Our results show that the MESAW model was able to estimate the N load at the river mouth of 88 Baltic Sea rivers, for which we had observed data, with a sufficient degree of precision and accuracy. The estimated retention parameters were also statistically significant. Our results show that around 380 000 tons of N are annually retained in surface waters draining to the Baltic Sea. The total annual riverine load from the 117 basins to the Baltic Sea was estimated to 570 000 tons of N, giving a total surface water N retention of around 40%. In terms of absolute retention values, three major river basins account for 50% of the total retention in the 117 basins; i.e. around 104 000 tons of N is retained in Neva, 55 000 tons in Vistula and 32 000 tons in Oder. The largest retention was found in river basins with a high percentage of lakes as indicated by a strong relationship between N retention (%) and share of lake area in the river drainage areas. For example in Göta älv, we estimated a total N retention of 72%, whereof 67% of the retention occurred in the lakes of that drainage area (Lake Vänern primarily). The obtained results will hopefully enable the Helsinki Commission (HELCOM) to refine the nutrient load targets in the Baltic Sea Action Plan (BSAP), as well as to better identify cost-efficient measures to reduce nutrient loadings to the Baltic Sea.

**1 Introduction**

Expanding human activities have had a great impact on nutrient dynamics and nutrient export from watersheds (Hill and Bolgrien, 2011; Mayorga et al., 2010). Increased population densities, food production, sewage emissions and fossil fuel combustion are among the driving forces causing increased nutrient mobilisation and alterations to hydrological systems (Mayorga et al., 2010). Increased nutrient export from coastal watersheds has had severe impacts on the ecological functions and community composition of estuaries, with algal blooms, increased water turbidity, oxygen depletion, and severe fish deaths as the most prominent consequences (Kellogg et al., 2010; Mayorga et al., 2010; Hoffmann et al., 2009).

Several geomorphic, hydraulic and biological factors may interact to reduce nutrient export from watersheds (Wollheim et al., 2006). Hejzlar et al. (2009) define retention as the fraction of external nutrient inputs that is retained within watersheds, either in absolute values or relative to the input. For nitrogen (N), the term retention is widely used to describe the processes leading to a temporary immobilisation of reactive (non-N2) N by incorporation into biomass or sedimentation, or the permanent loss of reactive N by conversion into the non-reactive atmospheric form (N2) by denitrification (Billen et al., 2009). Results from mass-balance studies across a wide range of geographic scales indicate that watersheds could retain as much as 60-90% of total N inputs (Kellogg et al., 2010). Reduced N export can be achieved by increasing N retention in soils, sediments and biomass, reducing atmospheric and terrestrial N sources, and increasing in-stream N removal and retention processes (Hill and Bolgrien, 2011).

Water residence time is a major factor determining the retention of nutrients in watersheds (Hejzlar et al., 2009), while Hesse and co-workers emphasised the need for better understandding of terrestrial retention (i.e., in soils; Hesse et al., 2013). Nitrogen is primarily removed (or retained) from surface water by denitrification (i.e., the microbial production of N2 from fixed N), followed by processes such as sorption to sediment or organic matter, and biological uptake (Hejzlar et al., 2009). Watershed characteristics, such as hydrology and geomorphology, strongly control water residence time, and increased water residence time can enhance denitrification processes and thereby reduce N loads to coastal waters (Kellogg et al., 2010; Behrendt and Opitz, 2000). Total N inputs influence denitrification rates, whereas hydrology and geomorphology (or water residence time) influence the proportion of N inputs that are denitrified (Seitzinger et al., 2006). Certain areas within watersheds can be identified as sink areas with regard to N export, often being areas with a relatively long water residence time where biogeochemical processes can transform reactive N into organic N in biomass, or N gases via denitrification (Kellogg et al., 2009), or burial of N in sediments (Harrison et al., 2009). The mitigating effect of these sink areas could in some cases be negligible, especially in cases where such areas are bypassed by N-carrying water flows due to specific land management practices (e.g. tile drains or storm water overflows) (Kellogg et al., 2009). Denitrification processes are favoured in sediments and hypoxic or anoxic bottom waters, particularly in systems with abundant organic carbon (C) and nitrate (Harrison et al., 2009; Mulholland et al., 2008).

The question on how to quantify the retention of nutrients from source to river mouth remains one of the largest uncertainties in river basin management. Several authors (e.g. Mayorga et al., 2010; Seitzinger, 2008) emphasise the need for advances in methods and models for determining the impacts of human activities on nutrient inputs to coastal waters, and a better understanding of the processes leading to retention of N in watersheds. Seitzinger et al. (2002) argue that studies generally have focused on N removal in shorter sub-sections of rivers and emphasise the need for a river network approach if we are to quantify the retention of nutrients relative to total inputs. In later years, a number of models of different complexity have been developed for estimating surface water N retention (e.g. Billen et al., 2009; Grimvall and Stålnacke, 1996; Hejzlar et al., 2009; Hill and Bolgrien, 2011; Jung and Deng, 2011; Mayorga et al., 2010; Seitzinger et al., 2002).

In this paper, we estimate the surface water retention of N in the Baltic Sea drainage basin with the use of a statistical model for source apportionment of riverine loads of pollutants, the MESAW model (Grimvall and Stålnacke, 1996). Scientifically, estimation of retention is one of the largest challenges in river basin nutrient accounting (i.e. source apportionment and budget calculations), and so is also the case in the Baltic Sea drainage basin.

**2 Materials and methods**

ö

**2.1 MESAW input data**

Model input data included:

1. Land cover, including cultivated land, wetland, lake area, other land (mainly forest), and total drainage area (Corine Land Cover 2006 raster data)
2. Atmospheric N wet deposition (EMEP; http://emep.int/publ/helcom/2012/index.html)
3. Point source emissions, including emissions from waste water treatment plants (WWTPs) and industry (data from HYDE, EUROSTAT and OECD)
4. Observed annual riverine N load (kg N yr-1) as estimated from riverine N concentration and water discharge data for the time period 1994-2006 (PLC database by HELCOM ([www.helcom.fi](http://www.helcom.fi)) and data from Denmark from NERI)

For the estimation of WWTP emissions, we created a spatially distributed data set of people ‘connected’ or ‘not connected’ to WWTPs (primary, secondary and tertiary) within the Baltic Sea river basins. For this, we used spatially distributed population data and national level statistics on WWTP connection. Population numbers for the year 2005 divided into urban and rural population were obtained from the HYDE database (http://themasites.pbl.nl/en/themasites/hyde/). These data were redistributed into a 10x10 km grid. Percentages of population ‘connected’ and type of waste water treatment were compiled from EUROSTAT (European Commission) and the Organisation for Economic Co-operation and Development (OECD). For Russia, Belarus, Ukraine and Slovakia, only percentage of people ‘connected’ to any type of waste water treatment was available, so the distribution between primary, secondary and tertiary treatment was based on assumption and expert judgement. Based on these national statistics, the total number of ‘connected’ people in each country was calculated. The number of ‘connected’ people was then spatially distributed to the grid cells. The distribution was made based on the assumption that urban population and grid cells with higher population numbers would be more likely to have a municipal WWTP connection than rural and smaller populations. Applying this principle, the grid cells for each country were classified as ‘connected’ starting with urban populations in a descending order, and continuing with rural population in the same way until the number of ‘connected’ people reached the number specified by the national statistics. This procedure was carried out for all three treatment types; first tertiary, then secondary and last primary.. The number of people ‘not connected’ to any type of treatment plant was also calculated for each grid cell. Total N emission from WWTPs was then calculated for each grid cell based on the approach of Mörth et al. (2007).

**2.2 The MESAW model and model parameterisation**

MESAW is a statistical model for source apportionment of riverine loads of pollutants developed by Grimvall & Stålnacke (1996). This model-approach uses non-linear regression for simultaneous estimation of export coefficients to surface waters for the different specified land cover or soil categories and retention coefficients for pollutants in river basins. Examples of application of the MESAW model are given in Lidèn et al. (1999), Vassiljev and Stålnacke (2005), Vassiljev et al. (2008) and Povilaitis et al. (2012). To its character, MESAW, have many common features with the more well-known SPARROW model developed in the U.S. (Smith et al., 1997; Alexander et al., 2000).

The basic principles and major steps in the procedure included: (i) estimation of mean annual riverine N loads for a fixed time period (i.e., the years 1994-2006) at each of the 88 monitoring sites, (ii) derivation of statistics on land cover, lake area, point source emissions and atmospheric deposition (see Section 2.1) for each river basin, and (iii) use of a general non-linear regression expression with N loads at each river basin as the dependent/response variable and basin characteristics as covariates/explanatory variables. This gave the following generalised form of the model (Eq. (1)):

  *i = 1,2,...,n* (1)

where *Li* is the load at outlet of basin i;

 *Si* is total losses from soil to water in basin *i*;

 *Pi* is the point source discharges (WWTP and industry) to waters in basin *i*;

 *Di* is the atmospheric deposition on surface waters in sub-basin *i*;

 *R* denote the retention for the source emissions *S, P* and *D*, respectively;

*n* is the number of basins, and

 *εI*is the statistical error term.

The parameterisation of the model is flexible and study area specific depending on the data and expert knowledge. The model is fitted by minimising the sum of squares for the differences between observed and estimated loads. The model can be run based on absolute or relative values. If based on absolute values, the optimisation procedure finds the minimum sum of squares of the absolute differences between observed and estimated transport. This procedure implies that the influence of the different rivers/basins will be a function of size. If relative values are used, the optimisation procedure finds the minimum sum of squares of relative differences between observed and estimated transport. This procedure assumes that all rivers have the same weight in the optimisation routine. In this study, we used relative values in order to give equal weight to small and large river basins.

The total diffuse loss of N from soil to water, *Si*, in the *i*th sub-basin was assumed to be a function of the land cover (Eq. (2)):

 *Si = ( θ1a1i + θ2a2i + θ3a3i )* (2a)

where *a1i, a2i* and *a3i* in our study refer to the areas of three land cover classes, i.e. cultivated land, wetlands and other land (mainly forests), respectively.*θ1, θ2* and *θ3* are unknown emission coefficients for the three land use categories that are statistically estimated in MESAW jointly with the retention (see Eq. (3) below). The point source emissions, *Pi,* and atmospheric deposition on surface waters, *Di*, were assumed to be known (see Section 2.1).

Throughout the exploratory analysis we found that certain basins deviated from the relationship and in most cases also where geographically located near to each other. Thus we introduced a ‘grouping variable’ according to the following:

 *Si = ( θ1a1i + θ2a2i + θ3a3i ) \** ωj (2b)

where each group j consisted of 2 or more basins depending on the model run (see Table 1) and where ω is the unknown coefficient(s). The model was run with different combinations of basin sub-groups in order to obtain reasonable model coefficients and load estimates (i.e. little deviation between predicted and observed loads). The grouping of basins was based on prior knowledge of similarities between basins as well as geographic location. For example, the 10 smaller Danish sub-basins formed one group, as a residual analysis showed that these sub-basins deviated from the general relationships. In its practical meaning, we simply adjusted the ‘global’ diffuse emission coefficients to the local conditions (despite we don’t know the underlying causes). This can be justified since applying the same coefficient to such a large drainage basin (1 745 000 km2) seem less logic.

Retention was in our study used as a summarising expression for all hydrological and biogeochemical processes that may decrease or retard the transport of N, e.g. denitrification, sedimentation and biological uptake. Irrespective of the exact retention mechanism, the parameterisation of the retention in the different basins was after several exploratory runs with alternative models done with the following empirical function (Eq. (3)):

\*  *i = 1,2,...,n* (3)

where λ1 and λ2 denotes a non-negative parameter and *Ri* denote the retention in the *i*th basin. The empirical function were in our case derived from the conception that the removal of N takes place primarily in the surface waters (both instream and in lakes). The first part of the function reflects the instream retention whereas the second part reflects the retention in lakes and reservoirs. Our assumption was that the removal rate is proportional to drainage basin size and the ratio of the lake to the total drainage area. Subsequently this can be seen as an indirect expression of the water residence time in the river basin.

Moreover, for the sake of simplicity, we assumed that the retention is the same for source categories D, S and P. Finally, by combining the parametric expressions of losses from soil to waters and retention with the empirical data, the *θ1, θ2*, *θ3*, λ1λ2 and ω parameters were estimated simultaneously.

**3 Results and discussion**

3.1 Parametrisation results

As given in Section 2.2, the model was run exploratory in order to obtain reasonable parameter coefficients and load estimates (i.e. little deviation between predicted and observed loads). with 9 estimated parametersλ1 and λ2 (i.e., *θ1*) *θ2* (*θ3)* It should be noted that the classification of wetlands is rather rough from the data source and given as joint expression of all wetlands ranging from marches to peatland bogs. All the four grouping parameters ω1- ω4 were statistically significant.

Worthwhile to notice is that these diffuse losses parameters *θ1* - *θ3* (all are given in kg km-2 and thus can be interpreted as export or unit-area loss coefficients. Interestingly, our estimates corroborate well with the results of monitored losses from small catchments with relative uniform landuse. For example, the point estimate and standard error for cultivated land gave an estimate of 1073 kg km-2 and 109 kg km-2, respectively (Model run #4; Table 1). Stålnacke and co-workers compiled data from 35 small agricultural catchments in the Nordic and Baltic region (Stålnacke et al. 2014). They found that a majority of these catchments had a unit-area loss between 600-2500 kg km-2. In addition, our results showed that the nitrogen losses from agricultural land were almost four times higher than the corresponding losses from forested land (Table 4) which is found to be realistic and in line with other results (Lidèn et al., 1999; Vassiljev and Stålnacke, 2005, Vassiljev et al., 2008)

3.2 Major retention estimate results

The final model parameterisation using the 88 river basin data (i.e. Model run #4 in Table 1) was used to determine the surface water retention of N in all the 117 major river basins in the Baltic Sea drainage area. This included 78 river basins with observed N load (excluding the 10 smaller Danish sub-basins), and also an additional 39 unmonitored river basins.

The total annual riverine load from the 117 basins to the Baltic Sea was estimated to 570 000 tons of N compared to the estimated gross load of 950 000 tons of N (Figure 2). Thus, o, giving a total surface water N retention of around 40%. This is substantially higher than given by Mörth et al (2007) who reported a mean N retention of 15% in the Baltic Sea rivers. The spatial distribution of the relative surface water retention is shown in Fig. 3. Averaged over all basins, mean lake retention is 25% whereas the estimated in-stream retention is 5%. In terms of absolute retention values, three major river basins account for 50% of the total retention in the 117 basins; i.e. around 104 000 tons of N is retained in Neva, 55 000 tons in Vistula and 32 000 tons in Oder.

Most of the retention occurs in lakes, as indicated by a strong relationship between N retention (%) and share of lake area in the river drainage areas (up to 20% lake area; Fig. 4). In Göta älv, we estimated a total N retention of 72%, whereof 67% occurred in the lakes of that drainage area (Lake Vänern primarily). Other river basins with high retention were Kymijoki (70%), Motalaström (73%) and Neva (74%). All these basins are characterised by a high percentage of lakes. Low retention was estimated for lake-poor basins, e.g. Aurajoki (2%), Kasari (4%) and Kelia (3%). This is in accordance with earlier studies, where the highest N retention has been found in river basins with a large proportion of lakes. In a comparison of N retention in four selected watersheds in Europe, representing a wide range in climate, hydrology and nutrient loads, Hejzlar et al. (2009) found the highest retention values in the two watersheds with lakes as compared to the two other mostly or entirely lake-less watersheds. A global-scale analysis by Harrison et al. (2009) indicated that lakes and reservoirs are important sinks for N in watersheds, with small lakes (<50 km2) retaining about half of the global total. Despite the fact that reservoirs occupy only 6% of global lentic surface area, the reservoirs were estimated to retain about 33% of the total N retained by lentic systems.

3.3. Uncertainty aspects and outlook

It should be noted that there are considerable uncertainties related to estimates of nutrient loads and especially retention at the watershed scale. In a study comparing nutrient retention estimates by catchment-scale models of different complexity, Hejzlar et al. (2009) showed a large variation in nutrient retention values as estimated by the different models in four selected catchments in Europe. They further showed that retention values were directly proportional to nutrient sources within catchments, indicating a close relationship between uncertainties in quantification of diffuse nutrient sources and nutrient retention determination. They concluded that realistic modelling of nutrient export from large catchments is only possible with a certain level of measured data. However, modelling efforts that combine comprehensive datasets on population, land cover, water discharge and quality, etc., may serve as important tools for improved watershed management and for better identification of cost-efficient measures to reduce nutrient loading.In our study, the MESAW model was apparently able to estimate the N load at the river mouth of 88 Baltic Sea rivers for which we had observed data with a sufficient degree of accuracy (Fig. 1, upper panel; Table 1). However, when we show the obtained relationships using unit-area (specific) load, the model underestimates the load (Fig. 1, lower panel). Worth to notice is also that the 10 Danish sub-basins included (despite the effort with the grouping) deviate from the general relationship. These 10 smaller sub-basins have a high observed specific N load, which is not well predicted by the model.

Fig. 5a-d show the relationships between observed specific N load (kg N km-2) and share of various land cover categories and lake area in the 88 (78 for wetland) Baltic Sea basins with observed N loads. A high specific N load was generally found in river basins with a large share of cultivated land, as indicated by a strong positive relationship between specific N load (kg N km-2) and share of cultivated land (%; Fig. 5a). Opposite to cultivated land, specific N load was found to be negatively correlated with the share of ‘Other land’ (i.e., primarily forest; Fig 5d).

In their modeling of riverine N transport to the Baltic Sea, Mörth et al. (2007) found diffuse sources to contribute the most to the overall simulated riverine N loads. A review by Stålnacke et al. (2009) also emphasized the importance of diffuse sources (or share of cultivated land) in contributing to N loads in watersheds. HELCOM (2011) reports that 45-61% of the total waterborne inputs of N to the Baltic Sea are from diffuse sources. The importance of wetlands in determining N loads seems highly variable, with no apparent relationship between specific N load and share of wetland area in the river basins (Fig. 5b). This was less surprising since this land cover class included all kind of wetlands (from marshes to peatlands). The low specific load for drainages basin with a wetland coverage exceeding 15% are all located in middle/north Finland and also in the northern part of Sweden (Table A1). These basins are all characterised by low population density and low share of cultivated land. In a meta-analysis of the importance of wetlands for the removal of inorganic N and reduction of N export from watersheds, Jordan et al. (2011) found a large variation (0.25 to 100%) in N removal efficiency between individual wetlands. When grouped into different wetland classes, mean efficiency was highest for palustrine forested wetlands (63%) and lowest for estuarine emergent wetlands (33%).

 in our study(and the surface water area) apparently these explanatory variables are insLidèn et al., 1999; Vassiljev and Stålnacke, 2005; Vassiljev et al., 2008; Povilaitis et al., 2012the showed (Table 1)

 The predicted climate change is an additional factor that may significantly affect nutrient loads and retention in watersheds (Jeppesen et al., 2011). Changes in temperature and precipitation will most likely induce changes in agricultural land use, e.g. type of crops grown, rates and timing of fertiliser use etc., and thereby influence N cycling and export to coastal waters. However, given the uncertainties in predicting future climate and land use on a regional level, the predicted effects on nutrient budgets in watersheds remain highly uncertain (Jeppesen et al., 2011).

Despite these discerned uncertainties, it seems that the MESAW model seems to be a reliable tool for simultaneous estimation of sources and retention in a river basin. It was also evident that MESAW is flexible and can accommodate many functional relationships and explanatory variables. In addition, MESAW can be used to identify measurements that are outside the general patterns and relationships (i.e., outliers). The main advantages with the model are: (i) the simple structure of the model (ii) the simple input data (iii) all unknown parameters are derived from empirical data, (iv) that information from all water quality monitoring sites are used in an optimal way, and (v) v) it gives results on the base of all available measured data which is better than to apply emission coefficients received from literature; normally even extrapolated from other regions or up-scaled from small watersheds.

**4 Conclusions**

We claim that one of the largest scientific and management uncertainties are devoted to the question on how to quantify the retention from source to river mouth. In this study, we used the MESAW statistical model to estimate the surface water N retention in the 117 river basins draining to the Baltic Sea. The MESAW model was able to estimate the N load at the river mouth of 88 Baltic Sea rivers, for which we had observed data, with a sufficient degree of accuracy. The estimated retention parameters were also statistically significant. Our results show that around 380 000 tons of N are annually retained in surface waters draining to the Baltic Sea. The total annual riverine load from the 117 basins to the Baltic Sea was estimated to 570 000 tons of N, giving a total surface water N retention of around 40%. The largest retention was found in river basins with a high percentage of lakes.

The obtained results will hopefully enable the Helsinki Commission (HELCOM) to refine the nutrient load targets in the Baltic Sea Action Plan (BSAP), as well as to better identify cost-efficient measures to reduce nutrient loadings to the Baltic Sea.

**Acknowledgements**

This paper is a contribution from the RECOCA project, which has received funding from the BONUS + program funded jointly by European Community’s Seventh framework Programme (FP7/2007-2013, grant agreement 217246) and Baltic Sea national funding institutions. Additional funding was for the Norwegian authors provided under the research grant ‘AGRI-LOSS’ of the Norwegian Research Council. This work was for the Estonian c-author also supported by institutional research funding IUT (IUT19-17) of the Estonian Ministry of Education and Research.

**References**

Alexander, R.B., Smith, R.A. and Schwarz, G.E.: Effect of stream channel size on the delivery of nitrogen to the gulf of Mexico, Nature, 403, 758-761, 2000.

Behrendt, H., and Opitz, D.: Retention of nutrients in river systems: dependence on specific runoff and hydraulic load, Hydrobiologia, 410, 111-122, 2000.

Bergström, S., and Carlsson, B.: River runoff to the Baltic Sea - 1950-1990, Ambio, 23, 280-287, 1994.

Billen, G., Thieu, V., Garnier, J., and Silvestre, M.: Modelling the N cascade in regional watersheds: The case study of the Seine, Somme and Scheldt rivers, Agriculture Ecosystems & Environment, 133, 234-246, 2009.

Grimvall, A., and Stålnacke, P.: Statistical methods for source apportionment of riverine loads of pollutants, Environmetrics, 7, 201-213, 1996.

Harrison, J. A., Maranger, R. J., Alexander, R. B., Giblin, A. E., Jacinthe, P. A., Mayorga, E., Seitzinger, S. P., Sobota, D. J., and Wollheim, W. M.: The regional and global significance of nitrogen removal in lakes and reservoirs, Biogeochemistry, 93, 2009.

Hejzlar, J., Anthony, S., Arheimer, B., Behrendt, H., Bouraoui, F., Grizzetti, B., Groenendijk, P., Jeuken, M., Johnsson, H., Lo Porto, A., Kronvang, B., Panagopoulos, Y., Siderius, C., Silgram, M., Venohr, M., and Zaloudik, J.: Nitrogen and phosphorus retention in surface waters: an inter-comparison of predictions by catchment models of different complexity, Journal of Environmental Monitoring, 11, 2009.

HELCOM. The fifth Baltic Sea pollution load compilation (PLC-5). Balt. Sea Environ. Proc. No. 128. 2011.

Hesse, C., Krysanova, V., Vetter, T., and Reinhardt, J.: Comparison of several approaches representing terrestrial and in-stream nutrient retention and decomposistion in watershed modelling, Ecological Modelling, 269, 70-85, 2013.

Hill, B. H., and Bolgrien, D. W.: Nitrogen removal by streams and rivers of the Upper Mississippi River basin, Biogeochemistry, 102, 2011.

Hoffmann, C. C., Kjaergaard, C., Uusi-Kamppa, J., Hansen, H. C. B., and Kronvang, B.: Phosphorus Retention in Riparian Buffers: Review of Their Efficiency, Journal of Environmental Quality, 38, 1942-1955, 2009.

Jeppesen, E., Kronvang, B., Olesen, J. E., Audet, J., Sondergaard, M., Hoffmann, C. C., Andersen, H. E., Lauridsen, T. L., Liboriussen, L., Larsen, S. E., Beklioglu, M., Meerhoff, M., Ozen, A., and Ozkan, K.: Climate change effects on nitrogen loading from cultivated catchments in Europe: implications for nitrogen retention, ecological state of lakes and adaptation, Hydrobiologia, 663, 1-21, 2011.

Jordan, S. J., Stoffer, J., and Nestlerode, J. A.: Wetlands as Sinks for Reactive Nitrogen at Continental and Global Scales: A Meta-Analysis, Ecosystems, 14, 144-155, 2011.

Jung, H. S., and Deng, Z. Q.: Modeling of Nitrogen Retention in Amite River, Water Air and Soil Pollution, 215, 411-425, 2011.

Kellogg, D. Q., Gold, A. J., Cox, S., Addy, K., and August, P. V.: A geospatial approach for assessing denitrification sinks within lower-order catchments, Ecological Engineering, 36, 1596-1606, 2010.

Lidèn, R., Vassiljev, A., Loigu, E., Stålnacke, P. and Wittgren, H.B.: Nitrogen source apportionment using a physically and statistically based model, Ecological Modelling, 114, 235-250, 1999.

Mayorga, E., Seitzinger, S. P., Harrison, J. A., Dumont, E., Beusen, A. H. W., Bouwman, A. F., Fekete, B. M., Kroeze, C., and Van Drecht, G.: Global Nutrient Export from WaterSheds 2 (NEWS 2): Model development and implementation, Environmental Modelling & Software, 25, 837-853, 2010.

Mulholland, P. J., Helton, A. M., Poole, G. C., Hall, R. O., Hamilton, S. K., Peterson, B. J., Tank, J. L., Ashkenas, L. R., Cooper, L. W., Dahm, C. N., Dodds, W. K., Findlay, S. E. G., Gregory, S. V., Grimm, N. B., Johnson, S. L., McDowell, W. H., Meyer, J. L., Valett, H. M., Webster, J. R., Arango, C. P., Beaulieu, J. J., Bernot, M. J., Burgin, A. J., Crenshaw, C. L., Johnson, L. T., Niederlehner, B. R., O'Brien, J. M., Potter, J. D., Sheibley, R. W., Sobota, D. J., and Thomas, S. M.: Stream denitrification across biomes and its response to anthropogenic nitrate loading, Nature, 452, 202-205, 2008.

Mörth, C. M., Humborg, C., Eriksson, H., Danielsson, A., Medina, M. R., Lofgren, S., Swaney, D. P., and Rahm, L.: Modeling riverine nutrient transport to the Baltic Sea: A large-scale approach, Ambio, 36, 124-133, 2007.

Povilaitis, A., Stålnacke, P. and Vassiliev, A.: Nutrient retention and export to surface waters in Lithuanian and Estonian river basins, Hydrology Research, 43, 359-373, 2012.

Seitzinger, S., Harrison, J. A., Böhlke, J. K., Bouwman, A. F., Lowrance, R., Peterson, B., Tobias, C., and Drecht, G. V.: Denitrification across landscapes and waterscapes: A synthesis, Ecological Applications, 16, 2064-2090, 2006.

Seitzinger, S.: Nitrogen cycle - Out of reach, Nature, 452, 162-163, 2008.

Seitzinger, S. P., Styles, R. V., Boyer, E. W., Alexander, R. B., Billen, G., Howarth, R. W., Mayer, B., and Van Breemen, N.: Nitrogen retention in rivers: model development and application to watersheds in the northeastern USA, Biogeochemistry, 57, 199-237, 2002.

Smith, R.A., Schwarz, G.E. and Alexander, R.B.: Regional interpretation of water-quality monitoring data, Water Resources Research, 33, 2781-2798, 1997.

Stålnacke, P., Pengerud, A., Bechmann, M., Garnier, J., Humborg, C., and Novotny, V.: Nitrogen driving force and pressure relationships at contrasting scales: Implications for catchment management, International Journal of River Basin Management, 7, 221-232, 2009.

Stålnacke, P., Aakerøy, P.A., Blicher-Mathiesen, G., Iital, A., Jansons, V., Koskiaho, J., Kyllmar, K., Lagzdins, A., Pengerud, A., Povilaitis, A., Temporal trends in nitrogen concentrations and losses from agricultural catchments in the Nordic and Baltic countries, Agriculture, Ecosystems and Environment, 198, 94–103, 2014.

Vassiljev, A. and Stålnacke, P.: Source apportionment of nutrients in the Lake Peipsi drainage basin – experiences from a statistical model, Water Science and technology, 51, 309-317, 2005.

Vassiljev, A., Blinova, I. and Ennet, P.: Source apportionment of nutrients in Estonian rivers. Desalination, 226, 222-230, 2008.

Wollheim, W. M., Voosmarty, C. J., Peterson, B. J., Seitzinger, S. P., and Hopkinson, C. S.: Relationship between river size and nutrient removal, Geophysical Research Letters, 33, L06410, 2006.

**Figure captions**

Figure 1 Relationship between observed and predicted annual N load (kt N yr-1; upper panel) and specific observed and predicted N load (kg N yr-1 km-2; lower panel) in the 88 Baltic Sea basins with observed N load (lower panel).

Figure 2 Total estimated nitrogen (N) load (kt N yr-1) in the 117 basins of the Baltic Sea drainage area. Total retention is given as the difference between the estimated total load if no retention and the estimated total riverine net N load.

Figure 3 Relative total nitrogen (N) retention in the Baltic Sea drainage basins.

Figure 4 Relationship between estimated retention (%) and total drainage area (km2; upper panel) and share of lake area (% of total drainage area; lower panel) for 117 Baltic Sea basins.

Figure 5 Relationship between specific N load (kg N km-2) and share of (a) cultivated, (b) wetland, (c) lake and (d) other area (in % of total drainage area) in the 88 (78 for wetland) Baltic Sea basins with observed N load.

|  |
| --- |
|  |
|  |

Figure 1



Figure 2



Figure 3

|  |
| --- |
|  |
|  |

Figure 4

|  |  |
| --- | --- |
|  |  |
|  |  |

Figure 5

**Table legends**

Table 1 Results from the different MESAW model runs for estimation of total nitrogen (N) retention with different combinations of basin sub-groups. Results include estimated export coefficients (kg/km2) from different land use classes (i.e., cultivated, wetland and other), estimated retention coefficients (dimension-less) for lake area and total drainage area, and the coefficient of determination (R2) between observed and predicted annual loads. Standard error and t-ratio of the estimated coefficients are given for each model run.

Table 1

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model run** |  | **Diffuse emissions** | **Retention** | **R2 (observed vs. predicted)** |
|   |   |   | *Cultivated* *θ1* *(kg/km2)* | *Wetland* *θ2* *(kg/km2)* | *Other**θ3**(kg/km2)* | *Lake* *(dimension-less)* | *Instream* *(dimensionless)* |  |
| 1 | 88 monitored basins (1 group) | Est. coeff. | 1435 | 405 | 233 | 9 | 4E-03 | 0.94 |
|  |  | St. err. | 929 | 2527 | 443 | 16 | 4E-03 |  |
|  |  | t-ratio | 1.54 | 0.16 | 0.53 | 0.57 | 1.03 |  |
| 2 | 88 monitored basins (3 groups) | Est. coeff. | 1440 | 386 | 185 | 8 | 2E-03 | 0.98 |
|  |  | St. err. | 172 | 753 | 136 | 5 | 5E-04 |  |
|  |  | t-ratio | 8.38 | 0.51 | 1.36 | 1.78 | 3.60 |  |
| 3 | 88 monitored basins (4 groups) | Est. coeff. | 1137 | 208 | 220 | 11 | 8E-04 | 0.99 |
|  |  | St. err. | 115 | 668 | 126 | 5 | 3E-04 |  |
|  |  | t-ratio | 9.88 | 0.31 | 1.75 | 2.16 | 2.62 |  |
| 4 | 88 monitored basins (5 groups) | Est. coeff. | 1073 | 158 | 225 | 12 | 7E-04 | 0.99 |
|  |  | St. err. | 109 | 675 | 123 | 5 | 3E-04 |  |
|  |  | t-ratio | 9.85 | 0.23 | 1.83 | 2.23 | 2.27 |  |

*Estimated group ratios ± standard error for diffuse emissions coefficients:*

|  |  |  |
| --- | --- | --- |
| **Model run** | **Basin sub-group**  | **Ratio ± st. err.** |
| 2 | 2 Pregolia, Narva3 Daugava, Neva | 0.3± 0.22.2 ± 0.3 |
| 3 | 2 Pregolia, Narva3 Daugava, Neva 4 10 Danish+6 Swedish (south-west coast) basins | 0.4 ± 0.22.0 ± 0.32.0 ± 0.8 |
| 4 | 2 Pregolia, Narva3 Daugava, Neva 4 10 Danish+6 Swedish (south-west coast) basins5 27 Finnish basins | 0.4 ± 0.22.0 ± 0.22.1 ± 0.81.1 ± 0.2 |

**Appendix**

Table A1Input data to the MESAW model for estimation of total nitrogen (N) retention. Input data include land cover (cultivated, wetland, lake area, other and total drainage area; km2) and point source emissions (WWTP and industry; kg N yr-1). Observed annual loads are given with the retention results in Table S2.

Table A2Observed and predicted annual N loads (kg N yr-1) and total N retention as estimated by the MESAW model.

Table A1.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| River basin | ID |  | Land cover (km2) | Point sources (kg N yr-1) |
|  |  |  | *Cultivated* | *Wetland* | *Lake area* | *Other* | *Total area* | *WWTP* | *Industry* |
| Alterälven | 5 |  | 18 | 15 | 6 | 418 | 457 | 863 | 0 |
| Aurajoki | 231 |  | 388 | 20 | 0 | 448 | 856 | 49322 | 526 |
| Botorpströmmen | 98 |  | 120 | 1 | 94 | 772 | 986 | 0 | 0 |
| Dalälven | 25 |  | 1071 | 2158 | 1281 | 24128 | 28638 | 250825 | 393700 |
| Daugava | 62 |  | 18242 | 973 | 2480 | 62912 | 84608 | 2895199 | 83017 |
| Delångersån | 28 |  | 99 | 21 | 184 | 1665 | 1969 | 7258 | 0 |
| Emån | 97 |  | 559 | 23 | 272 | 3572 | 4427 | 73043 | 2200 |
| Eurajoki | 221 |  | 351 | 22 | 169 | 800 | 1342 | 0 | 24611 |
| Forsmarksån | 24 |  | 14 | 28 | 9 | 252 | 302 | 5 | 0 |
| Gauja | 61 |  | 3353 | 99 | 66 | 5432 | 8951 | 256499 | 0 |
| Gavleån | 26 |  | 169 | 44 | 165 | 2115 | 2494 | 85652 | 10600 |
| Gideälven | 341 |  | 50 | 277 | 103 | 3007 | 3437 | 37 | 0 |
| Göta älv | 151 |  | 6669 | 1484 | 9105 | 34206 | 51465 | 517002 | 882100 |
| Helge å | 91 |  | 1116 | 108 | 220 | 3237 | 4681 | 98586 | 0 |
| Iijoki | 14 |  | 193 | 1933 | 647 | 11369 | 14142 | 4400 | 0 |
| Indalsälven | 31 |  | 641 | 2162 | 1654 | 21382 | 25839 | 119903 | 35800 |
| Kalajoki | 173 |  | 771 | 234 | 127 | 3326 | 4457 | 17282 | 62 |
| Kalix älv | 8 |  | 92 | 3030 | 288 | 14286 | 17696 | 34519 | 55000 |
| Karvianjoki | 250 |  | 460 | 239 | 109 | 2673 | 3481 | 7550 | 0 |
| Kasari | 103 |  | 1085 | 211 | 2 | 1938 | 3236 | 34256 | 0 |
| Kelia | 47 |  | 335 | 40 | 1 | 336 | 712 | 29518 | 9300 |
| Kemijoki | 12 |  | 306 | 7633 | 1682 | 42892 | 52513 | 50135 | 125000 |
| Kiiminkijoki | 15 |  | 81 | 866 | 96 | 2850 | 3894 | 7666 | 0 |
| Kiskonjoki | 234 |  | 255 | 7 | 44 | 649 | 955 | 0 | 0 |
| Kokemäenjoki | 21 |  | 5180 | 402 | 2265 | 19281 | 27128 | 353484 | 201645 |
| Koskenkylänjoki | 404 |  | 338 | 5 | 35 | 571 | 950 | 0 | 114 |
| Kuivajoki | 132 |  | 25 | 355 | 30 | 942 | 1352 | 86 | 0 |
| Kymijoki | 41 |  | 2741 | 425 | 5971 | 27138 | 36275 | 189501 | 509272 |
| Kyrönjoki | 178 |  | 1442 | 320 | 46 | 3128 | 4936 | 45276 | 578 |
| Lagan | 143 |  | 903 | 275 | 598 | 4802 | 6579 | 104364 | 31000 |
| Lapuanjoki | 177 |  | 1052 | 206 | 83 | 2720 | 4060 | 12102 | 4650 |
| Lestijoki | 174 |  | 156 | 141 | 3 | 753 | 1053 | 0 | 0 |
| Lielupe | 63 |  | 10872 | 271 | 121 | 6549 | 17814 | 701774 | 66470 |
| Ljungan | 29 |  | 229 | 638 | 645 | 11092 | 12605 | 20964 | 0 |
| Ljungbyån | 96 |  | 140 | 4 | 8 | 854 | 1006 | 31242 | 0 |
| Ljusnan | 27 |  | 449 | 1645 | 705 | 17224 | 20024 | 57246 | 0 |
| Luleälv | 6 |  | 46 | 2016 | 1790 | 20702 | 24554 | 9045 | 0 |
| Lyckebyån | 95 |  | 36 | 8 | 33 | 721 | 797 | 6849 | 0 |
| Lögdeälven | 342 |  | 9 | 104 | 28 | 1361 | 1503 | 1600 | 0 |
| Motala ström | 99 |  | 3042 | 96 | 2937 | 8046 | 14121 | 341463 | 111400 |
| Mustijoki | 402 |  | 285 | 9 | 9 | 455 | 758 | 0 | 0 |
| Mörrumsån | 93 |  | 394 | 27 | 464 | 2490 | 3376 | 87147 | 2900 |
| Narva | 46 |  | 12437 | 1048 | 4789 | 39852 | 58126 | 1068418 | 132850 |
| Neman | 83 |  | 44359 | 554 | 1544 | 49469 | 95925 | 6206082 | 139420 |
| Neva | 42 |  | 7004 | 8126 | 45020 | 219436 | 279586 | 3522246 | 2586124 |
| Nissan | 145 |  | 254 | 100 | 179 | 2619 | 3152 | 62723 | 28000 |
| Norrström | 101 |  | 5457 | 302 | 2564 | 14754 | 23076 | 860051 | 379300 |
| Nyköpingsån | 100 |  | 952 | 32 | 579 | 2877 | 4440 | 61879 | 0 |
| Närpiönjoki | 202 |  | 241 | 70 | 5 | 706 | 1022 | 0 | 0 |
| Odra | 87 |  | 73524 | 225 | 1630 | 43559 | 118939 | 13758343 | 1133829 |
| Oulojki | 16 |  | 513 | 1828 | 2490 | 19411 | 24242 | 36077 | 50297 |
| Paimionjoki | 232 |  | 567 | 13 | 14 | 524 | 1118 | 0 | 0 |
| Perhonjoki | 175 |  | 313 | 327 | 57 | 1822 | 2519 | 0 | 15700 |
| PiteÄlv | 4 |  | 42 | 863 | 515 | 9732 | 11152 | 13057 | 0 |
| Porvoonjoki | 403 |  | 504 | 5 | 14 | 814 | 1337 | 38752 | 1135 |
| Pregolia | 84 |  | 9187 | 34 | 280 | 3919 | 13419 | 1276555 | 0 |
| Pyhäjoki | 172 |  | 440 | 204 | 179 | 2903 | 3727 | 2568 | 1412 |
| Pärnu | 601 |  | 2198 | 341 | 8 | 4053 | 6600 | 84972 | 0 |
| Rickleån | 1 |  | 49 | 52 | 104 | 1453 | 1658 | 824 | 0 |
| Rönneå | 142 |  | 661 | 21 | 67 | 1154 | 1903 | 46885 | 17100 |
| Råneälven | 71 |  | 24 | 994 | 70 | 3087 | 4175 | 3798 | 0 |
| Salaca | 602 |  | 1294 | 150 | 57 | 2007 | 3508 | 63498 | 0 |
| Siikajoki | 171 |  | 464 | 506 | 65 | 3074 | 4109 | 0 | 3629 |
| Simojoki | 131 |  | 47 | 597 | 148 | 2349 | 3141 | 501 | 0 |
| Skellefteälv | 2 |  | 98 | 1026 | 1152 | 9337 | 11613 | 23639 | 30500 |
| Torne älv | 10 |  | 264 | 6089 | 1405 | 32354 | 40112 | 52914 | 131000 |
| Töreälven | 72 |  | 3 | 70 | 14 | 417 | 505 | 719 | 0 |
| Ume älv | 35 |  | 252 | 2169 | 1318 | 23199 | 26939 | 38219 | 59100 |
| Uskelanjoki | 233 |  | 472 | 4 | 4 | 478 | 959 | 5763 | 26640 |
| Vantaanjoki | 401 |  | 541 | 12 | 52 | 1290 | 1895 | 39672 | 12032 |
| Venta | 80 |  | 6146 | 107 | 111 | 5328 | 11692 | 580890 | 0 |
| Vironjoki | 43 |  | 67 | 2 | 6 | 283 | 357 | 0 | 0 |
| Viskan | 149 |  | 365 | 14 | 136 | 1664 | 2178 | 44596 | 0 |
| Vistula | 85 |  | 124478 | 751 | 2268 | 66398 | 193894 | 20541873 | 547784 |
| Ähtävänjoki | 176 |  | 737 | 201 | 225 | 3155 | 4318 | 2349 | 14376 |
| Ätran | 147 |  | 558 | 50 | 196 | 2515 | 3320 | 6049 | 0 |
| Öreälven | 343 |  | 63 | 347 | 37 | 2600 | 3046 | 1801 | 0 |
| Ångermanälven | 33 |  | 398 | 2923 | 1921 | 26572 | 31815 | 36673 | 0 |
| Ry å |  |  | 214 |  | 2 | 69 | 285 | 23275 | 0 |
| Lindenborg å |  |  | 197 |  | 4 | 119 | 319 | 19624 | 0 |
| Skals å |  |  | 401 |  | 16 | 139 | 556 | 22700 | 0 |
| Karup å |  |  | 344 |  | 8 | 275 | 627 | 92696 | 0 |
| Gudenå |  |  | 1563 |  | 79 | 961 | 2603 | 404099 | 0 |
| Århus å |  |  | 183 |  | 10 | 130 | 324 | 134450 | 0 |
| Kolding å |  |  | 180 |  | 3 | 85 | 268 | 32126 | 0 |
| Odense å |  |  | 339 |  | 9 | 187 | 535 | 31007 | 0 |
| Ndr. Halleby å |  |  | 272 |  | 18 | 128 | 418 | 31204 | 0 |
| Suså |  |  | 478 |  | 24 | 254 | 756 | 107901 | 0 |
| Coast DE &Arkona Basin | 1011 |  | 1740 | 28 | 8 | 620 | 2395 | 74132 | 0 |
| Coast DE & Bornholm Basin | 1012 |  | 7858 | 81 | 191 | 2455 | 10585 | 336061 | 1535 |
| Coast DE & Fehmarn Belt | 1013 |  | 8041 | 52 | 280 | 2148 | 10522 | 377142 | 23830 |
| Coast DK &Arkona Basin | 2011 |  | 1109 | 14 | 3 | 500 | 1626 | 44132 | 55702 |
| Coast DK & Bornholm Basin | 2012 |  | 446 | 2 | 0 | 134 | 581 | 19962 | 9873 |
| Coast DK & Central Kattegat | 2018 |  | 9915 | 194 | 82 | 824 | 12459 | 71261 | 21316 |
| Coast DK & Fehmarn Belt | 2015 |  | 2471 | 14 | 62 | 1729 | 2961 | 228639 | 145739 |
| Coast DK & Northern Kattegat | 2017 |  | 376 | 16 | 13 | 463 | 629 | 78572 | 10873 |
| Coast DK &Samso Belt | 2014 |  | 7346 | 67 | 6 | 296 | 9204 | 0 | 2317 |
| Coast DK & Southern Kattegat | 2013 |  | 2141 | 27 | 0 | 237 | 3074 | 12299 | 183156 |
| Coast DK & The Sound | 2016 |  | 117 | 2 | 150 | 2199 | 422 | 370565 | 76197 |
| Coast EE & Baltic Proper | 3011 |  | 1102 | 201 | 40 | 3120 | 4463 | 38295 | 0 |
| Coast EE & Gulf of Finland | 3012 |  | 1953 | 181 | 16 | 3694 | 5843 | 90250 | 636400 |
| Coast EE & Gulf of Riga | 3013 |  | 1610 | 373 | 34 | 3375 | 5392 | 71976 | 0 |
| Coast FI & Baltic Proper | 4013 |  | 802 | 15 | 58 | 8112 | 3716 | 48827 | 1275161 |
| Coast FI &Bothnian Bay | 4011 |  | 1283 | 608 | 152 | 9272 | 10061 | 37028 | 393799 |
| Coast FI &Bothnian Sea | 4012 |  | 2255 | 165 | 292 | 2607 | 11844 | 5282 | 123771 |
| Coast FI & Gulf of Finland | 4014 |  | 1120 | 58 | 109 | 3999 | 5286 | 997 | 155412 |
| Coast LT & Baltic Proper | 5011 |  | 846 | 34 | 7 | 713 | 1599 | 23821 | 0 |
| Coast LV & Baltic Proper | 6011 |  | 2605 | 101 | 60 | 2491 | 5257 | 54410 | 0 |
| Coast LV & Gulf of Riga | 6012 |  | 1697 | 210 | 112 | 4052 | 6071 | 169910 | 0 |
| Coast North of Northern Kattegat | 9018 |  | 129 | 0 | 205 | 5857 | 464 | 178001 | 0 |
| Coast PL & Baltic Proper | 7012 |  | 7144 | 71 | 251 | 3313 | 10778 | 191663 | 14737 |
| Coast PL & Bornholm Basin | 7011 |  | 8207 | 64 | 10 | 2125 | 14333 | 67794 | 0 |
| Coast RU & Baltic Proper | 8011 |  | 3552 | 28 | 345 | 22109 | 5716 | 50496 | 334590 |
| Coast RU & Gulf of Finland | 8012 |  | 690 | 689 | 34 | 569 | 23832 | 50617 | 170000 |
| Coast SE &Arkona Basin | 9012 |  | 1182 | 0 | 2 | 154 | 1338 | 34653 | 0 |
| Coast SE & Baltic Proper | 9014 |  | 5022 | 66 | 315 | 4230 | 19925 | 54353 | 418800 |
| Coast SE & Bornholm Basin | 9013 |  | 1049 | 24 | 623 | 14215 | 5618 | 165765 | 223000 |
| Coast SE &Bothnian Bay | 9015 |  | 769 | 1195 | 821 | 16344 | 19129 | 105747 | 259100 |
| Coast SE &Bothnian Sea | 9016 |  | 1641 | 315 | 833 | 18550 | 21339 | 159982 | 1544000 |
| Coast SE & Central Kattegat | 9020 |  | 595 | 7 | 5 | 330 | 1878 | 9798 | 0 |
| Coast SE & Northern Kattegat | 9019 |  | 141 | 0 | 28 | 576 | 745 | 10765 | 7600 |
| Coast SE & Southern Kattegat | 9021 |  | 1054 | 26 | 80 | 1195 | 2234 | 21757 | 131000 |
| Coast SE & The Sound | 9011 |  | 2019 | 2 | 15 | 1138 | 2623 | 11479 | 11000 |
| Laihianjoki | 201 |  | 239 | 21 | 1 | 461 | 723 | 6010 | 0 |
| Isojoki | 205 |  | 176 | 80 | 3 | 895 | 1155 | 0 | 3000 |
| Sirppujoki | 222 |  | 143 | 7 | 3 | 270 | 424 | 0 | 0 |
| Iilolanjoki | 405 |  | 102 | 1 | 7 | 196 | 306 | 0 | 0 |

Table A2.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| River basin | ID |  | Annual N load (kg N yr-1) | Retention (kg N yr-1) | Relative retention |
|  |  |  | *Observed* | *Predicted* |  | *Total surface water* | *Lake* | *In-stream* |
| Alterälven | 5 |  | 96331 | 100515 | 17 290 | 0.15 | 0.134 | 0.014 |
| Aurajoki | 231 |  | 704385 | 600476 | 11 817 | 0.02 | 0.000 | 0.019 |
| Botorpströmmen | 98 |  | 173792 | 171655 | 197 934 | 0.54 | 0.526 | 0.021 |
| Dalälven | 25 |  | 5226692 | 4744728 | 3 296 123 | 0.41 | 0.343 | 0.102 |
| Daugava | 62 |  | 40351648 | 45292798 | 27 364 274 | 0.38 | 0.255 | 0.164 |
| Delångersån | 28 |  | 231231 | 255778 | 294 739 | 0.54 | 0.522 | 0.029 |
| Emån | 97 |  | 993846 | 953813 | 756 687 | 0.44 | 0.417 | 0.043 |
| Eurajoki | 221 |  | 639538 | 284749 | 434 449 | 0.60 | 0.594 | 0.024 |
| Forsmarksån | 24 |  | 90969 | 58633 | 20 485 | 0.26 | 0.250 | 0.012 |
| Gauja | 61 |  | 4467000 | 4443302 | 690 252 | 0.13 | 0.079 | 0.060 |
| Gavleån | 26 |  | 559846 | 456153 | 378 260 | 0.45 | 0.435 | 0.032 |
| Gideälven | 341 |  | 474231 | 568452 | 229 492 | 0.29 | 0.260 | 0.038 |
| Göta älv | 151 |  | 15496154 | 6222224 | 15 737 523 | 0.72 | 0.673 | 0.132 |
| Helge å | 91 |  | 2786308 | 2742915 | 1 697 732 | 0.38 | 0.354 | 0.044 |
| Iijoki | 14 |  | 2205385 | 2092250 | 1 372 973 | 0.40 | 0.348 | 0.074 |
| Indalsälven | 31 |  | 4321692 | 3260548 | 3 047 764 | 0.48 | 0.427 | 0.098 |
| Kalajoki | 173 |  | 2294615 | 1294900 | 507 684 | 0.28 | 0.249 | 0.043 |
| Kalix älv | 8 |  | 3505231 | 3021890 | 895 054 | 0.23 | 0.159 | 0.082 |
| Karvianjoki | 250 |  | 1398667 | 902551 | 378 474 | 0.30 | 0.267 | 0.038 |
| Kasari | 103 |  | 1949457 | 1593791 | 74 510 | 0.04 | 0.008 | 0.037 |
| Kelia | 47 |  | 831729 | 467939 | 12 609 | 0.03 | 0.009 | 0.018 |
| Kemijoki | 12 |  | 6372308 | 7897005 | 4 619 699 | 0.37 | 0.272 | 0.134 |
| Kiiminkijoki | 15 |  | 746000 | 719171 | 246 453 | 0.26 | 0.224 | 0.040 |
| Kiskonjoki | 234 |  | 351985 | 306363 | 173 663 | 0.36 | 0.349 | 0.020 |
| Kokemäenjoki | 21 |  | 9839231 | 5662122 | 6 747 574 | 0.54 | 0.493 | 0.100 |
| Koskenkylänjoki | 404 |  | 429846 | 376743 | 174 023 | 0.32 | 0.302 | 0.020 |
| Kuivajoki | 132 |  | 388538 | 252573 | 72 255 | 0.22 | 0.203 | 0.024 |
| Kymijoki | 41 |  | 5673077 | 3912199 | 8 968 328 | 0.70 | 0.657 | 0.114 |
| Kyrönjoki | 178 |  | 3274615 | 2201007 | 353 912 | 0.14 | 0.098 | 0.045 |
| Lagan | 143 |  | 2812308 | 2375686 | 2 785 262 | 0.54 | 0.515 | 0.052 |
| Lapuanjoki | 177 |  | 2052308 | 1525444 | 443 272 | 0.23 | 0.192 | 0.041 |
| Lestijoki | 174 |  | 448077 | 339396 | 20 572 | 0.06 | 0.037 | 0.021 |
| Lielupe | 63 |  | 13435786 | 11941659 | 2 100 218 | 0.15 | 0.073 | 0.082 |
| Ljungan | 29 |  | 1536538 | 1751972 | 1 256 416 | 0.42 | 0.374 | 0.070 |
| Ljungbyån | 96 |  | 241923 | 340462 | 40 219 | 0.11 | 0.087 | 0.021 |
| Ljusnan | 27 |  | 2822077 | 3149047 | 1 715 354 | 0.35 | 0.291 | 0.087 |
| Luleälv | 6 |  | 2920615 | 2572731 | 2 688 010 | 0.51 | 0.459 | 0.095 |
| Lyckebyån | 95 |  | 221462 | 158152 | 79 930 | 0.34 | 0.323 | 0.019 |
| Lögdeälven | 342 |  | 234077 | 272877 | 68 367 | 0.20 | 0.179 | 0.025 |
| Motala ström | 99 |  | 3017538 | 2067763 | 5 577 661 | 0.73 | 0.708 | 0.074 |
| Mustijoki | 402 |  | 620923 | 388273 | 59 428 | 0.13 | 0.117 | 0.018 |
| Mörrumsån | 93 |  | 834923 | 557809 | 951 339 | 0.63 | 0.616 | 0.038 |
| Narva | 46 |  | 5034077 | 5400364 | 6 902 138 | 0.56 | 0.490 | 0.140 |
| Neman | 83 |  | 44323731 | 46377160 | 20 173 375 | 0.30 | 0.158 | 0.172 |
| Neva | 42 |  | 44616846 | 36056404 | 104 559 254 | 0.74 | 0.652 | 0.262 |
| Nissan | 145 |  | 1368846 | 1231196 | 893 156 | 0.42 | 0.399 | 0.036 |
| Norrström | 101 |  | 3637692 | 4695747 | 7 182 353 | 0.60 | 0.564 | 0.093 |
| Nyköpingsån | 100 |  | 730077 | 792009 | 1 293 905 | 0.62 | 0.603 | 0.043 |
| Närpiönjoki | 202 |  | 657615 | 432931 | 32 255 | 0.07 | 0.049 | 0.021 |
| Odra | 87 |  | 70289195 | 73974593 | 31 717 905 | 0.30 | 0.138 | 0.188 |
| Oulojki | 16 |  | 2894615 | 2598002 | 3 708 470 | 0.59 | 0.545 | 0.095 |
| Paimionjoki | 232 |  | 900846 | 680805 | 114 622 | 0.14 | 0.125 | 0.022 |
| Perhonjoki | 175 |  | 815769 | 688291 | 209 547 | 0.23 | 0.208 | 0.033 |
| PiteÄlv | 4 |  | 1594231 | 1492285 | 966 100 | 0.39 | 0.350 | 0.066 |
| Porvoonjoki | 403 |  | 1303615 | 723513 | 107 528 | 0.13 | 0.108 | 0.024 |
| Pregolia | 84 |  | 4580143 | 4207286 | 1 429 755 | 0.25 | 0.195 | 0.072 |
| Pyhäjoki | 172 |  | 1127385 | 807208 | 504 218 | 0.38 | 0.359 | 0.039 |
| Pärnu | 601 |  | 3091070 | 3193671 | 219 537 | 0.06 | 0.013 | 0.052 |
| Rickleån | 1 |  | 283462 | 233266 | 181 333 | 0.44 | 0.422 | 0.027 |
| Rönneå | 142 |  | 2587846 | 1511841 | 682 779 | 0.31 | 0.291 | 0.029 |
| Råneälven | 71 |  | 540308 | 714315 | 177 607 | 0.20 | 0.164 | 0.042 |
| Salaca | 602 |  | 2287635 | 1585606 | 377 174 | 0.19 | 0.160 | 0.038 |
| Siikajoki | 171 |  | 1332615 | 1127084 | 266 866 | 0.19 | 0.157 | 0.041 |
| Simojoki | 131 |  | 748231 | 471133 | 286 563 | 0.38 | 0.355 | 0.036 |
| Skellefteälv | 2 |  | 1319385 | 1124731 | 1 475 798 | 0.57 | 0.536 | 0.068 |
| Torne älv | 10 |  | 5154615 | 5552103 | 3 319 728 | 0.37 | 0.290 | 0.119 |
| Töreälven | 72 |  | 89992 | 83532 | 28 567 | 0.25 | 0.244 | 0.015 |
| Ume älv | 35 |  | 3359846 | 3522076 | 2 619 449 | 0.43 | 0.363 | 0.099 |
| Uskelanjoki | 233 |  | 508182 | 652521 | 47 115 | 0.07 | 0.048 | 0.020 |
| Vantaanjoki | 401 |  | 1283000 | 753105 | 269 584 | 0.26 | 0.242 | 0.028 |
| Venta | 80 |  | 6649974 | 7118779 | 1 365 064 | 0.16 | 0.100 | 0.068 |
| Vironjoki | 43 |  | 213534 | 122758 | 26 503 | 0.18 | 0.167 | 0.013 |
| Viskan | 149 |  | 1568692 | 1016516 | 792 104 | 0.44 | 0.420 | 0.030 |
| Vistula | 85 |  | 112041104 | 116917897 | 55 292 179 | 0.32 | 0.120 | 0.229 |
| Ähtävänjoki | 176 |  | 419608 | 1049732 | 713 541 | 0.40 | 0.378 | 0.042 |
| Ätran | 147 |  | 2007769 | 1529797 | 1 155 539 | 0.43 | 0.408 | 0.037 |
| Öreälven | 343 |  | 491769 | 605790 | 110 818 | 0.15 | 0.123 | 0.036 |
| Ångermanälven | 33 |  | 4223154 | 3798849 | 3 450 419 | 0.48 | 0.413 | 0.107 |
| Ry å |  |  | 537807 | 495987 |  |  |  |  |
| Lindenborg å |  |  | 724156 | 458904 |  |  |  |  |
| Skals å |  |  | 683604 | 743902 |  |  |  |  |
| Karup å |  |  | 720306 | 866486 |  |  |  |  |
| Gudenå |  |  | 3075608 | 3177587 |  |  |  |  |
| Århus å |  |  | 442719 | 446084 |  |  |  |  |
| Kolding å |  |  | 651265 | 423277 |  |  |  |  |
| Odense å |  |  | 1071416 | 740493 |  |  |  |  |
| Ndr. Halleby å |  |  | 275250 | 473721 |  |  |  |  |
| Suså |  |  | 961097 | 952650 |  |  |  |  |
| Coast DE & Arkona Basin | 1011 |  |  | 1953735 | 139 436 | 0.07 | 0.036 | 0.032 |
| Coast DE & Bornholm Basin | 1012 |  |  | 7394183 | 2 176 860 | 0.23 | 0.174 | 0.065 |
| Coast DE & Fehmarn Belt | 1013 |  |  | 7094187 | 2 843 034 | 0.29 | 0.237 | 0.065 |
| Coast DK & Arkona Basin | 2011 |  |  | 1345482 | 61 852 | 0.04 | 0.018 | 0.026 |
| Coast DK & Bornholm Basin | 2012 |  |  | 529791 | 8 593 | 0.02 | 0.000 | 0.016 |
| Coast DK & Central Kattegat | 2018 |  |  | 9543967 | 1 505 255 | 0.14 | 0.071 | 0.070 |
| Coast DK & Fehmarn Belt | 2015 |  |  | 2715470 | 783 402 | 0.22 | 0.195 | 0.035 |
| Coast DK & Northern Kattegat | 2017 |  |  | 483611 | 127 542 | 0.21 | 0.195 | 0.017 |
| Coast DK & Samso Belt | 2014 |  |  | 7432683 | 537 534 | 0.07 | 0.007 | 0.061 |
| Coast DK & Southern Kattegat | 2013 |  |  | 2458534 | 91 878 | 0.04 | 0.000 | 0.036 |
| Coast DK & The Sound | 2016 |  |  | 232761 | 984 193 | 0.81 | 0.806 | 0.014 |
| Coast EE & Baltic Proper | 3011 |  |  | 1711747 | 261 982 | 0.13 | 0.094 | 0.043 |
| Coast EE & Gulf of Finland | 3012 |  |  | 3399885 | 289 177 | 0.08 | 0.031 | 0.049 |
| Coast EE & Gulf of Riga | 3013 |  |  | 2339545 | 295 661 | 0.11 | 0.068 | 0.047 |
| Coast FI & Baltic Proper | 4013 |  |  | 3285098 | 754 614 | 0.19 | 0.153 | 0.039 |
| Coast FI &Bothnian Bay | 4011 |  |  | 3211420 | 818 564 | 0.20 | 0.149 | 0.063 |
| Coast FI &Bothnian Sea | 4012 |  |  | 2382056 | 908 814 | 0.28 | 0.223 | 0.068 |
| Coast FI & Gulf of Finland | 4014 |  |  | 1783711 | 535 861 | 0.23 | 0.193 | 0.047 |
| Coast LT & Baltic Proper | 5011 |  |  | 1024349 | 78 227 | 0.07 | 0.046 | 0.026 |
| Coast LV & Baltic Proper | 6011 |  |  | 2919632 | 550 310 | 0.16 | 0.118 | 0.047 |
| Coast LV & Gulf of Riga | 6012 |  |  | 2347634 | 656 593 | 0.22 | 0.178 | 0.050 |
| Coast North of Northern Kattegat | 9018 |  |  | 291927 | 1 530 714 | 0.84 | 0.838 | 0.014 |
| Coast PL & Baltic Proper | 7012 |  |  | 6523722 | 2 349 513 | 0.26 | 0.213 | 0.065 |
| Coast PL & Bornholm Basin | 7011 |  |  | 8603413 | 770 891 | 0.08 | 0.008 | 0.075 |
| Coast RU & Baltic Proper | 8011 |  |  | 5280324 | 4 169 888 | 0.44 | 0.413 | 0.048 |
| Coast RU & Gulf of Finland | 8012 |  |  | 1082667 | 132 001 | 0.11 | 0.016 | 0.094 |
| Coast SE &Arkona Basin | 9012 |  |  | 1284175 | 55 321 | 0.04 | 0.018 | 0.024 |
| Coast SE & Baltic Proper | 9014 |  |  | 5408956 | 1 605 059 | 0.23 | 0.156 | 0.087 |
| Coast SE & Bornholm Basin | 9013 |  |  | 2193909 | 3 088 678 | 0.58 | 0.564 | 0.048 |
| Coast SE &Bothnian Bay | 9015 |  |  | 3191350 | 2 042 113 | 0.39 | 0.333 | 0.085 |
| Coast SE &Bothnian Sea | 9016 |  |  | 4986008 | 2 981 173 | 0.37 | 0.313 | 0.089 |
| Coast SE & Central Kattegat | 9020 |  |  | 687471 | 41 291 | 0.06 | 0.029 | 0.028 |
| Coast SE & Northern Kattegat | 9019 |  |  | 222276 | 101 656 | 0.31 | 0.301 | 0.018 |
| Coast SE & Southern Kattegat | 9021 |  |  | 1120775 | 520 464 | 0.32 | 0.295 | 0.031 |
| Coast SE & The Sound | 9011 |  |  | 2227338 | 232 256 | 0.09 | 0.063 | 0.033 |
| Laihianjoki | 201 |  |  | 356461 | 13 468 | 0.04 | 0.019 | 0.018 |
| Isojoki | 205 |  |  | 385234 | 21 891 | 0.05 | 0.032 | 0.022 |
| Sirppujoki | 222 |  |  | 195592 | 21 695 | 0.10 | 0.087 | 0.014 |
| Iilolanjoki | 405 |  |  | 124007 | 33 175 | 0.21 | 0.202 | 0.012 |