SOILS2SEA DELIVERABLE NO. 5.1

HYPE model for the Baltic Sea Basin

SOILS2SEA

Reducing nutrient loadings from agricultural soils to the Baltic Sea via groundwater and streams

This page has intentionally been left blank

SOILS2SEA DELIVERABLE NO. 5.1

HYPE model for the Baltic Sea Basin

May 2017

Authors: Alena Bartosova, Chantal Donnelly, Johan Strömqvist, René Capell, and Johanna Tengdelius-Brunell

SOILS2SEA

Reducing nutrient loadings from agricultural soils to the Baltic Sea via groundwater and streams

This report is a publicly accessible deliverable of the Soils2Sea project. The present work has been carried out within the project ‘Reducing nutrient loadings from agricultural soils to the Baltic Sea via groundwater and streams (BONUS Soils2Sea)’, which has received funding from BONUS, the joint Baltic Sea research and development programme (Art 185), funded jointly from the European Union’s Seventh Framework Programme for research, technological development and demonstration and from Innovation Fund Denmark, The Swedish Environmental Protection Agency (Naturvådsverket), The Polish National Centre for Research and Development, The German Ministry for Education and Research (Bundesministerium für Bildung und Forschung), and The Russian Foundation for Basic Researches (RFBR).

Bartosova A et al. HYPE model for the Baltic Sea Basin. BONUS Soils2Sea Deliverable 5.1. Swedish Meteorological and Hydrological Institute, Norrköping, May 2017, www.Soils2Sea.eu

Table of Contents

1. Background and objectives

The Baltic Sea Action Plan and the EU Water Framework Directive both require substantial additional reductions of nutrient loads (N and P) to the marine environment. The BONUS Soils2Sea project conducts research on a widely applicable concept for spatially differentiated regulation, exploiting the fact that the removal and retention of nutrients by biogeochemical processes or sedimentation in groundwater and surface water systems shows large spatial variations. By targeting measures towards areas where the local removal is low, spatially differentiated regulation can be much more cost-effective than the traditional uniform regulation.

To design and evaluate the effectiveness of spatially differentiated regulation requires improved knowledge on the nutrient transport and removal processes at local scale. Soils2Sea therefore conducts field studies with comprehensive data collection and modelling at four sites in Denmark, Sweden, Poland and Russia. Furthermore, Soils2Sea will conduct scenario analyses at the Baltic Sea Basin scale to assess how different regulatory measures as well as changes in land cover, agricultural practices and climate may affect the nutrient losses from the entire Baltic Sea Basin (BSB) to the Baltic Sea.

Evaluating the impacts of local scale spatially differentiated measures at a scale such as the 1.8 million km^2 Baltic Sea Basin poses a particular challenge. Multi-basin hydrological and nutrient models at this scale (e.g. Donnelly et al., 2013) are not able to simulate local scale spatially differentiated measures, because i) the models operate at a much coarser spatial resolution than the measures; ii) they often do not include local scale data but rather aggregated data which can vary in quality and resolution between countries; iii) they have a spatial resolution that is coarser than the heterogeneity of the physical system that controls the local variation in the nutrient reduction; and iv) they often have simplified process descriptions adequate for the input data complexity and model scale, but sometimes inadequate for simulating specific local scale measures such as field scale crop rotations,. Such measures can be simulated by comprehensive and data demanding local scale models (Hansen et al., 2014a); however, for computational and data access reasons these models are not operational at the Baltic Sea Basin scale. Therefore, other methods must be applied for upscaling the results from suitable local scale models to models operating at the Baltic Sea scale. Bronstert et al. (2007) provide one of the very few examples reported in literature of this type of upscaling based on dynamic combinations of small and large scale models.

The objective of the present deliverable report is to describe how the upscaling methodologies developed in Soils2Sea (Refsgaard et al, 2016) were implemented into E-HYPE by adjusting the model structure and process descriptions. Further, the revised model was calibrated to in-stream observation data as well as process data from other studies in the Baltic Sea Basin.

2. Methodology for Calibrating E-HYPE

Soils2Sea uses E-HYPE, a pan-European application (Donnelly et al. 2016, Hundecha 2016) of the Hydrological Predictions for the Environment (HYPE) code (Lindström et al.2010) as the modelling tool for the Baltic Sea Basin. Knowledge from local scale models developed in Soils2Sea is transferred into E-HYPE and will be used later in simulating the impacts of spatially differentiated measures at the Baltic Sea Basin scale

Section 2.1 gives an overview of the upscaling approach developed in D3.2 (Upscaling methodologies). E-HYPE calibration completes several steps in the overall upscaling approach. The full approach with a description on how the steps were or will be accomplished within Soils2Sea project is described here. Section 2.2 describes updates in E-HYPE model inputs that were carried out as a part of this project. Section 2.3 then describes the calibration process for E-HYPE.

2.1 Summary of Soils2Sea Upscaling Approach

This section gives an overview of the upscaling approach presented in D3.2 (Upscaling methodologies) and describes how the approach was implemented during E-HYPE development and which steps were completed. The general approach is similar for both groundwater and surface water processes:

- *STEP 1 – Compare concepts*. Check the consistency of the processes used in the local model and in E-HYPE. Identify possible needs for refined process representation or calibration of E- HYPE.
- *STEP 2 – Identify additional data requirements.* Assess whether additional data are required for E-HYPE at Baltic Sea Basin scale, for instance for new process descriptions, recalibration or evaluation of simulation results.
- *STEP 3 – Recalibration of* E-*HYPE*. Recalibrate E-HYPE if required.
- *STEP 4 – Upscaled* E-*HYPE parameters.* This is the core of the upscaling procedure. The local scale models are used to create relationships for how E-HYPE parameters should be modified to enable E-HYPE to simulate the effects of local scale processes in the scenario analyses.
- *STEP 5 – Use* E-*HYPE for Baltic Sea Basin simulations.*

The upscaling process is applied for groundwater processes and for surface water processes. Steps 1 and 2 were accomplished and reported in D3.2 (Upscaling methodologies). For groundwater processes, the following additional data and processes were identified and gathered for the E-HYPE development: presence of aquifers in deep soils, baseflow fraction for stream flows, nitrogen leaching from top soil layers to groundwater, and reduction of nitrogen in groundwater. There were no additional data needed for surface water processes, only a slight modification to HYPE model executable code that allowed changing local stream length, affecting the retention times in surface waters.

This report describes Step 3. Step 4 was documented in D3.2, where equations were developed to relate the effects of specific remediation measures to changes in HYPE model parameters or inputs. Step 4 as well as Step 5 will be fully implemented in D5.4 using the updated E-HYPE model.

2.2 E-HYPE model updates

E-HYPE v.3.1 was used as a starting point for the updates and subsequent recalibration. The set-up, calibration and validation of the water part of the E-HYPE v.3.1 are described in Hundecha et al. (2016). Before the recalibration, selected parts of the model were updated. The updates included: (a) point sources, (b) crop data, and (c) deep soils with active aquifers (implemented as new soil-land use classes to make differentiation of model parameters possible).

The first two updates were made to make model results comparable with nutrient modelling for the Baltic Sea Basin in other ongoing BONUS projects in addition to generally improving the quality of data used in the model. Point sources in E-HYPE were updated to match those in Urban Waste Water Treatment Directive (UWWTD) database (2016). The data were reviewed for consistency and completeness. Figure 1 shows countries where data were of sufficient quality and completeness and were incorporated in E-HYPE. Note that most countries in Baltic Sea Basin were included in the update. Crop data were updated to reflect more current agriculture practices. The crop information in the updated model is based on 2013 Eurostat data.

The third update is not an update of the input data but more of an adjustment of the model structure by adding more spatial hydrological units (soil-land use classes) that allowed us to simulate areas where deep soil aquifers play an important role in hydrology and in nitrogen reduction processes. Soils2Sea experts working at field and catchment scales and with good knowledge of local conditions in their countries agreed on the importance of the depth of a soil layer in the model for calculation of groundwater reduction for nitrogen. The deeper the soils, the larger the volume of water and the longer the residence time of water in the soil layer. The longer residence time means higher nitrogen reduction in the soil layer for the same denitrification rate.

Soil depths for the deepest layer in the original E-HYPE v.3.1, layer 3, ranged from 1 to 3.5 meters. To better represent nitrogen transformation and reduction processes in deeper soils, we extracted highly productive porous aquifers in Baltic Sea Basin from the European Hydrogeology map (BGR & UNESCO, 2014) (Figure 2). The E-HYPE model input files were then updated to include new soil and land use classes with deep soils, and here the soil depth was set to 10 meters. The sensitivity analyses tests revealed that increasing the soil depth beyond 10m had very little additional impact on simulated reduction of nitrogen in groundwater.

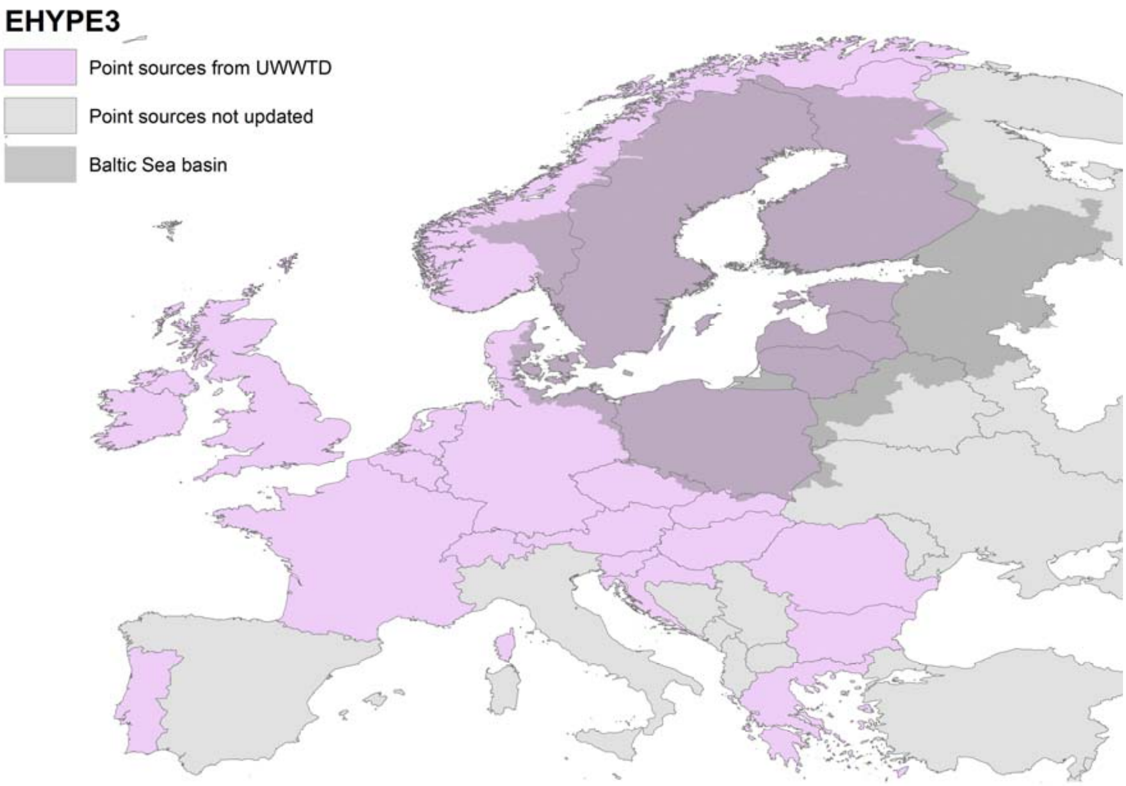

Figure 1. Countries where point source data were updated to match UWWTD.

Figure 2. Active aquifers in Baltic Sea Basin (BSB) defined as deep soils in E-HYPE. Extracted from BGR & UNESCO (2014)

2.3 Model calibration

To calibrate the model we used the stepwise, representative gauged basin approach described in Strömqvist et al. (2011) and Donnelly et al. (2016). We combined this with the upscaling methodology that was developed and tested in D3.2 (Upscaling Methodologies). In addition to evaluating model performance based on comparing simulated and observed concentrations at observation points, the E-HYPE model performance was also reviewed with respect to 3 process-based data for which maps could be provided: baseflow fraction, nitrogen leaching, and reduction of nitrogen in groundwater. When there was a discrepancy between the simulated and observed values, the relevant model processes were adjusted during a recalibration attempt through modifying selected model parameters (Table 1). Surface water retention was adjusted as needed in order to improve the match with the observed concentrations after calibrating for N leaching and N reduction in groundwater.

Table 1. HYPE parameters adjusted during re-calibration*

Parameter	Dependence	Description
Denitrlu	land use	denitrification in soil
Degradhn	land use	release of IN from slow-reacting N pool
Dissolhn	land use	release of ON from slow-reacting N pool
Minerfn	land use	release of IN from fast-reacting N pool
Dissolhp	land use	release of PP from slow-reacting P pool
Freuc	soil	adsorption to soil particles
Soilcoh	soil	soil resistance to erosion due to overland flow
Soilerod	soil	erosion caused by kinetic energy in rain

Note: * the reader is referred to HYPE documentation for full description of simulated processes and model parameters: http://www.smhi.net/hype/wiki/doku.php?id=start

2.3.1 Representative gauging basins (RGBs)

Representative gauging basins (RGBs) for each land use had to satisfy three important conditions: (a) sufficient monitoring data were available, (b) respective land use has a significant presence, and (c) point sources have no or minimal impact on water quality. However, there was a lack of sites for certain land uses where all three conditions were satisfied.

The land uses were thus divided into dominant land uses, which had a significant presence and could be calibrated directly, and secondary land uses, which were calibrated after the calibration of dominant land uses. Dominant land uses had a highly significant presence in calibration sites. Secondary land uses had a less significant presence in calibration sites but a significant presence in the area not covered by dominant land uses. This step-wise calibration enabled us to select a lower threshold for secondary land uses while preserving integrity of the calibration.

The dominant land use was considered having a significant presence if it comprised at least 70% of the contributing area. The secondary land use was considered having a significant presence if it comprised (a) at least 20% of the contributing area, and (b) at least 70% of the contributing area not covered by the dominant land uses.

Forest and agriculture were considered dominant and thus were calibrated directly. Pastures, open areas with vegetation, sealed urban areas, and wetlands were considered secondary land uses with less significant presence. Additional land uses (glaciers, water, and open areas without vegetation) were not recalibrated due to a lack of calibration sites. Calibrated land uses comprise 94% and 92% of E-HYPE and BSB areas, respectively (Table 2).

Water quality monitoring data were considered sufficient if (i) the monitoring period covered at least 5 years, (ii) there were at least 30 observations during the monitoring period, and (iii) at least 90% values were above the detection limit. Note that these are minimum requirements and most of the RGBs used in recalibration had significantly more water quality data than required. Ideally these RGBs should be spatially distributed across the modelled domain (Europe), although this wasn't always the case.

There were 89 sites selected for calibration (Figure 3). The calibration period was from January 1, 2000 to December 31, 2010 (11 years) with the model simulation starting from January 1, 1979. The initial period was considered a "warm up" period; it allows the model to achieve stable conditions for nutrient pools in various parts of the simulated system. The remaining sites were used to validate the model using the same time period.

Table 2. Land use composition in the full E-HYPE domain and in its Baltic Sea Basin area.

Land use	E-HYPE domain	Baltic Sea Basin
Agriculture	33%	22%
Forest	34%	50%
Glacier	0%	0%
Open w/o veg.	4%	1%
Open with veg.	17%	12%
Pasture	7%	4%
Urban	1%	0%
Water	2%	6%
Wetland	3%	4%

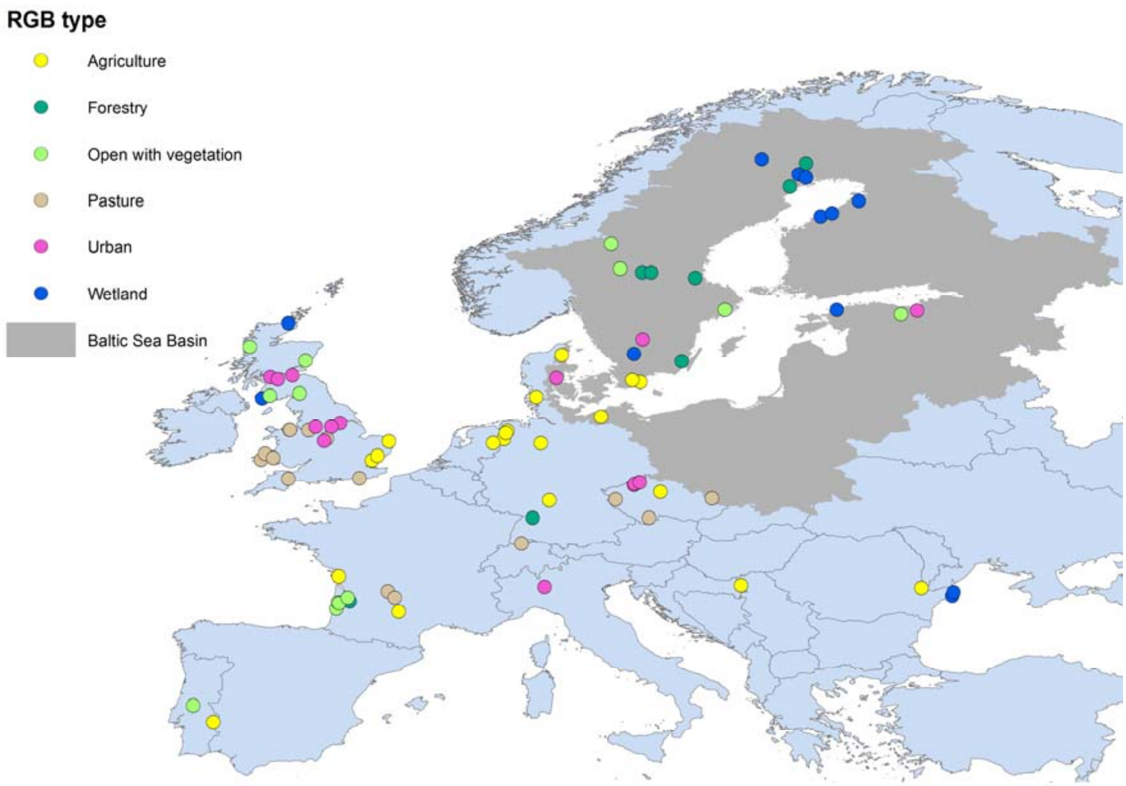

Figure 3. RGB sites used for E-HYPE re-calibration

The recalibration consisted of 3 steps. During the first step, representative gauging basins (RGBs) selected for dominant land uses were calibrated. During the second step, RGBs selected for the secondary land uses were calibrated. During the third step, sub-regional parameter for denitrification in soil layer 3 was introduced to address any regional patterns in residuals remaining after calibrating for the three process-based data sets.

2.3.2 Baseflow fraction

The first step in the methodology is to assure that the model simulates well the split between the flows entering the streams from shallow and deep soils. The baseflow fraction is estimated by applying the baseflow filter BFLOW (Arnold et al., 1995) to observed and simulated hydrographs. BFLOW uses three different passes where the hydrograph is separated in several steps. We used pass 3, giving the lowest baseflow fraction as this was identified by Hansen et al. (submitted) to be closest to the local model results.

The baseflow fraction (BF%) is calculated as the long-term total baseflow divided by the total discharge. The baseflow fraction is used as a soft criterion in the calibration, so that the calibration is deemed acceptable if the mean value of simulated baseflow is within +/-20% of the observed baseflow fraction. Note also that the E-HYPE model uses a distributed calibration so that model performance is not tuned locally to match observations, but rather regionally (over the whole continent) to best match as many observation points as possible. Baseflow fraction calculated for the observed flow time series is shown in Figure 4.

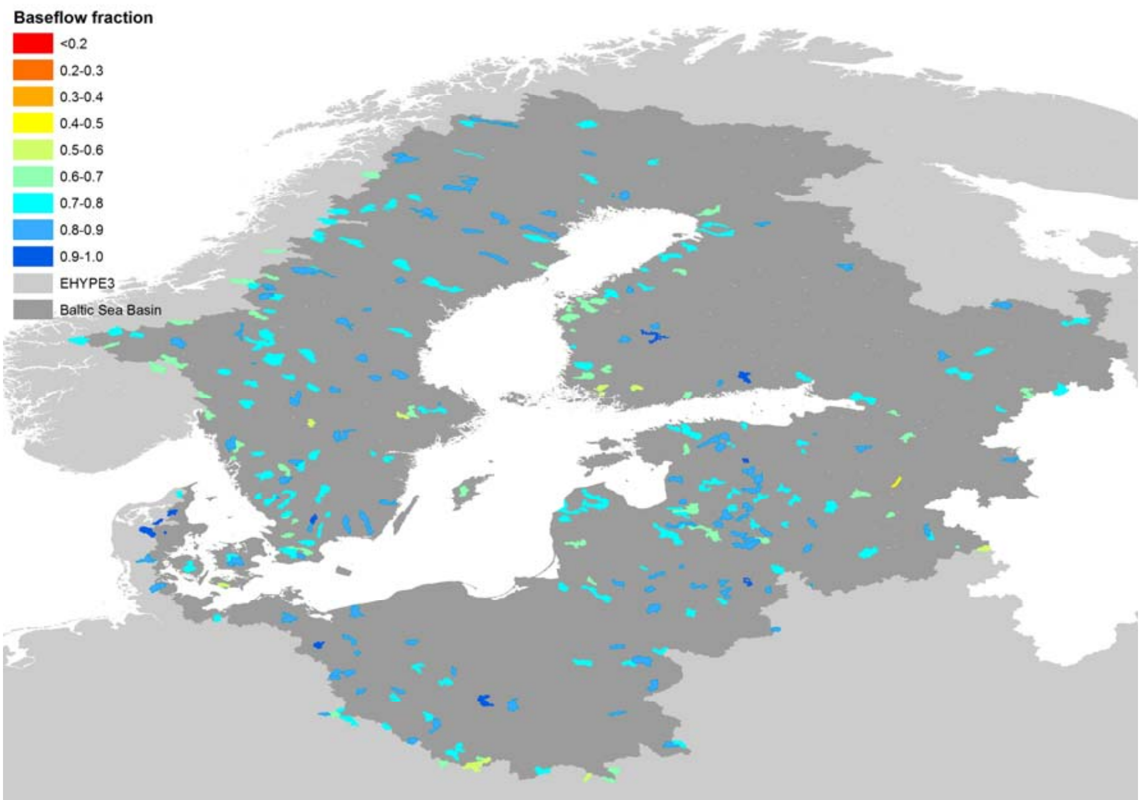

Figure 4. Baseflow fractions determined from observed flows in BSB

2.3.3 Nitrogen leaching

Constraining the N-leaching in E-HYPE requires a best possible estimate of N-leaching. Simulated values are used, as leaching observations are not systematically available. At the Baltic Sea Basin scale this is a 10 km grid scale map produced by Andersen et al. (2016). The grid values were matched to E-HYPE sub-basins and average values for each sub-basin were determined (Figure 5).

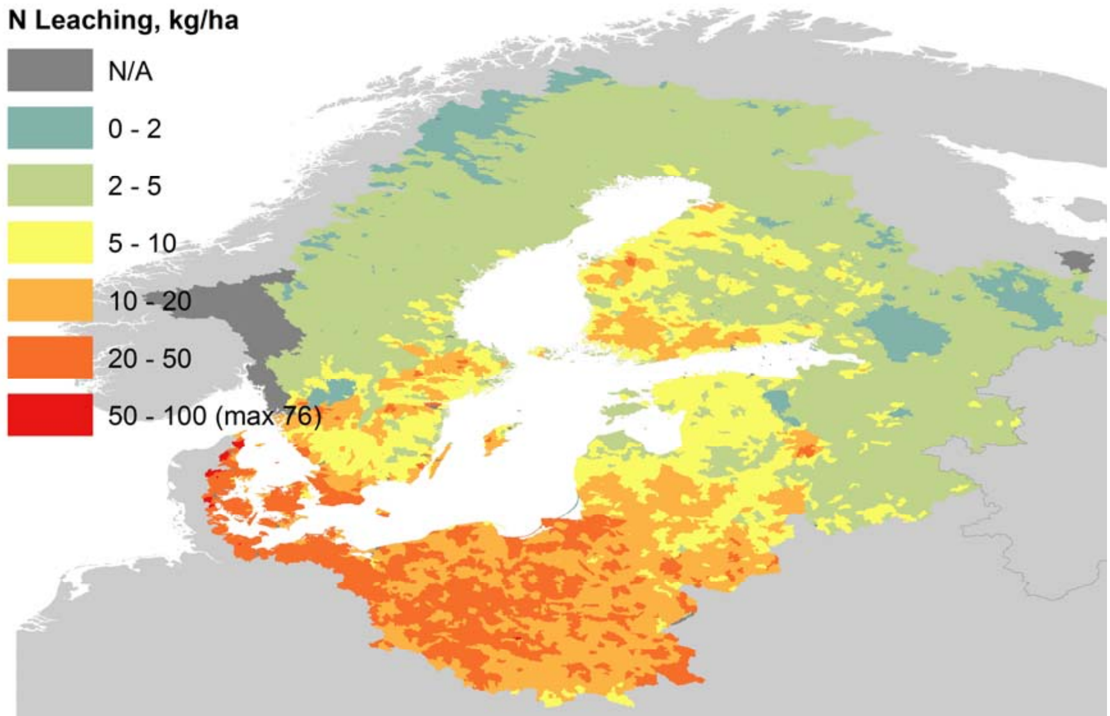

Figure 5.Estimated N leaching based on Andersen et al. (2016), aggregated to E-HYPE sub-basins

2.3.4 Reduction of nitrogen in groundwater

For constraining the groundwater N-reduction in E-HYPE we are utilising a new map with estimates on the groundwater reduction for the Baltic Sea Basin produced by Højberg et al. (2017). This map is based on the available national studies supplemented with expert judgement. For the Danish sub-basins the groundwater reduction map is based on the results from the Danish National N-model (Højberg et al., 2016). Figure 6 shows the percent reduction aggregated to E-HYPE subbasins where available. Data for Germany are not included; the estimate for Germany is a constant (54%) without any spatial variability needed to evaluate the model performance.

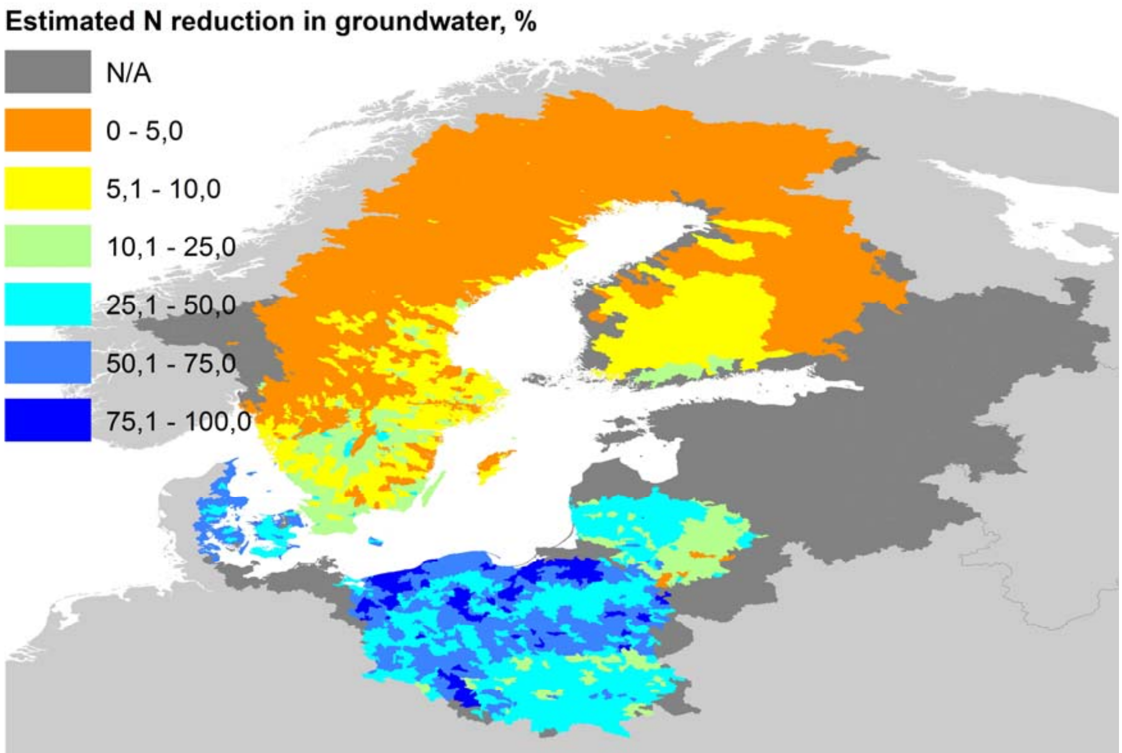

Figure 6. Estimated reduction of N in groundwater (Højberg et al., 2017) aggregated to E-HYPE subbasins. Average value for Germany (54%) is omitted.

2.3.5 Sub-regional calibration

The HYPE parameters can be grouped into different categories based on the scale at which they operate. Parameters that operate at the smallest scale (hydrologic response unit, or HRU) are generally soil type or land use dependent and their values are estimated based either on land use or soil type. There are also a group of parameters that are defined at a subbasin scale. These include general parameters which are assigned the same value through the model domain, and sub-regional parameters, which are region specific parameters.

HYPE offers the possibility to use model parameters on subbasin scale using so called sub-regional parameterization. This allows introducing variability of both general parameters and sub-regional parameters across the model domain based on spatial variability of physical controls of the processes the parameters describe aside from land use or soil variability. One has the choice to implement these estimators globally across the entire model domain or employ separate estimators for a number of different regions defined within the model domain. Any kind of grouping of subbasins can be employed and subbasins that constitute a given group need not be geographically contiguous.

The regionalization method used in the original E-HYPE v3.1.1 has been described in Hundecha et al. (2016). The regionalization of flow related parameters was retained, while the regionalization of water quality related parameters was removed prior to re-calibration. After the model was calibrated using the RGB approach, simulated reduction of nitrogen in groundwater was reviewed to identify any remaining spatial patterns. Regions identified in this manner were then coded to HYPE and used to calibrate sub-regional denitrification rate in soil layer 3. No other parameters were used on subbasins scale.

3. Results and Discussion

3.1 General discussion

There is a variety of approaches to calibration in large-scale modelling, but common to many large-scale models is that it is impossible to calibrate every single observation point in detail. The E-HYPE model regionalises parameters making them general, land use, soil-type, or region specific depending on the process represented by the parameterisation. Performance in single observations points is compromised to achieve the best possible performance across the model domain (Donnelly et al. 2016). The coarser spatial resolution also limits the number of monitoring sites that can be used for calibration and the overall resolution of the model input data. This means that for a single given catchment, large-scale models most likely use less calibration data and may be less fitted to the data that is available. On the other hand, the variation in performance can be used to estimate the uncertainty in simulating ungauged basins within the domain. Performance for a given catchment is generally considerably poorer than when compared to smaller scale models. Regarding nutrient calibration, there is also the potential for equifinality in nutrient reduction processes (Beven, 2006) because many combinations of parameter values with different splits between reduction in surface water and in groundwater can provide the same overall N-reduction.

At the E-HYPE scale, there is not always enough observation data to separate out whether reduction occurs in surface or groundwater. We use three process-based data sets to constrain the E-HYPE model calibration and followed a methodology developed in D3.2 (Upscaling Methodology) to define more clearly transport and transformation of nutrients on land, in soil, and in water. E-HYPE like most models is able to reach the same final calibration targets (water and nutrient fluxes at river gauging stations) via many different combinations of intermediate results (such as local scale flows, nutrient transport and reduction/retention). Our recalibration process in reality constrains E-HYPE to reproduce results at small scales that are comparable with those from detailed small scale models. In this way we expect that the equifinality level in E-HYPE will be reduced such that it to a greater extent will simulate the “right answers for the right reasons”. This improves the confidence in model predictions, when E-HYPE is used in scenario analyses to assess impacts of future changes in climate, land use and agricultural practice.

Although use of different models in the same study are not uncommon, only few other studies (e.g. Bronstert et al., 2007) have utilised this in an upscaling approach, where the local scale model is applied to train or develop a relationship for the large scale model. With the increased use of large scale models and the need to describe impacts of local scale interventions at the large scale, we believe that this approach hold a large potential for further development and wide application.

A critical assumption in this regard is that the calibration of E-HYPE made against data from small scale models will also be valid in other parts of the Baltic Sea Basin. The issue is not whether the large scale model simulations can match the small scale model in all aspects, but whether all necessary underlying conditions that determine the spatial variability in the processes are represented in the model structure and allow for thorough parameterization.

3.2 Baseflow fractions

The difference between the observed and simulated baseflow fractions is shown Figure 7. Comparison of the observed and simulated baseflow fraction shows a very close match between the two. The absolute difference is less than 20%, a threshold selected to indicate satisfactory calibration, for 95% of the gauging stations and less than 10% for 75% of the gauges in the Baltic Sea Basin (Figure 8). Performance across the whole E-HYPE domain was very similar. Only 16 out of 297 gages show absolute differences greater than 20% and only 3 gauges show absolute difference greater than 25% (Figure 7). No spatial pattern was identified among the differences. Based on these results, we decided that E-HYPE v3.1 simulated the observed baseflow fraction adequately and no further calibration of the flow was needed.

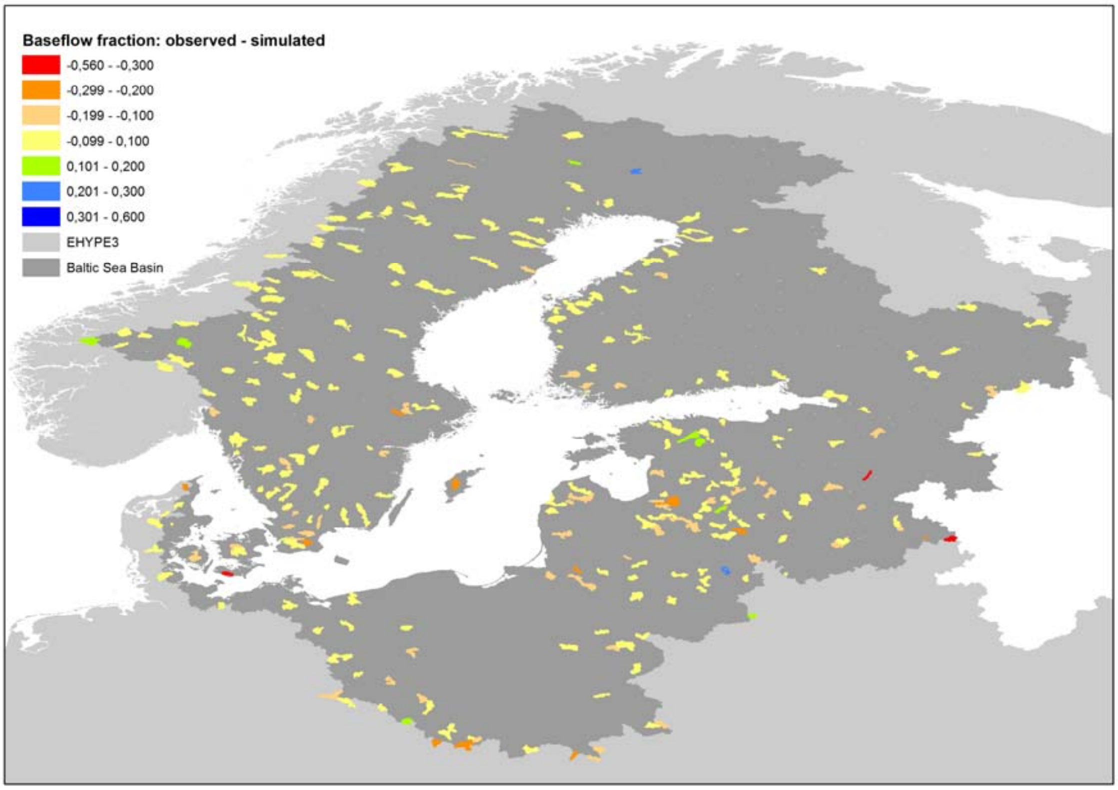

Figure 7. Absolute difference between the baseflow fractions determined from observed and simulated flows in Baltic Sea Basin.

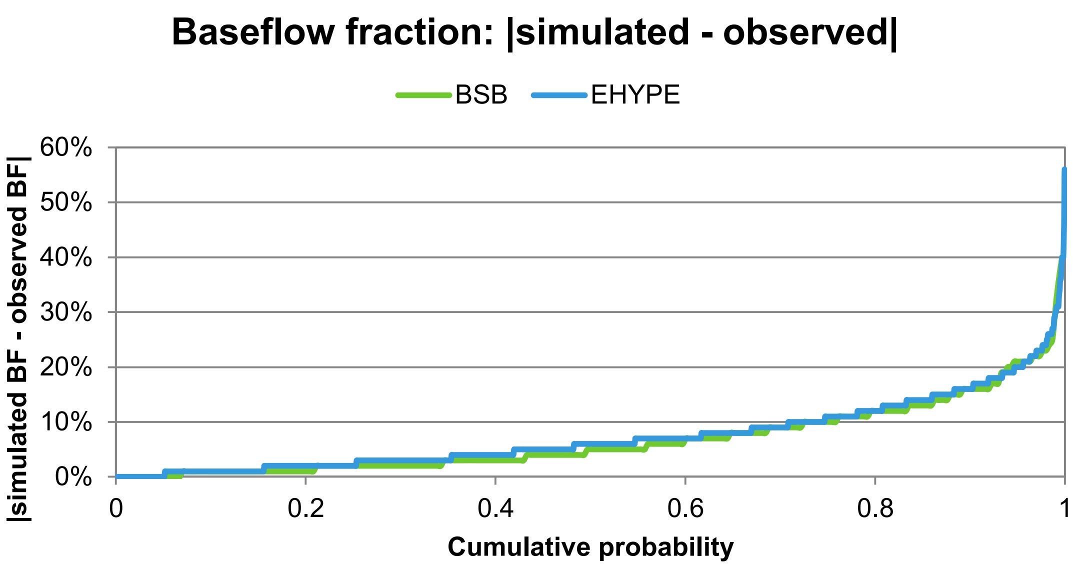

Figure 8. Cumulative probability of the difference between simulated and observed baseflow fraction in BSB and whole E-HYPE domain

3.3 Calibrating nitrogen processes

Both N processes, N leaching and reduction of N in groundwater, were calibrated at the same time. While at a smaller scale it is relatively easy to recalibrate the processes by altering model parameters to match observed values for one monitoring site, the calibration process becomes increasingly complex when it is applied at a large scale with many monitoring sites in many different watersheds. Often an improvement is made at some monitoring sites at the expense of others.

Figure 9 and Figure 10 show relationships between simulated (y-axes) and estimated by other studies (x-axes) values for (a) reduction of N in groundwater and (b) N leaching before and after calibration. Basic statistics showing the goodness of fit are also included. Calibration improved the match for N reduction in groundwater but not for N leaching. Note that both results are presented for runs without any sub-regional calibration present in E-HYPE 3.1 or the current version for water quality. The effect of sub-regional calibration is presented in Section 3.4.

Table 3. Calibrated parameter values for the original E-HYPE v.3.1.1 and the updated E-HYPE v.3.1.4.

Parameter	E-HYPE v.3.1.1.	E-HYPE v.3.1.4.
Denitrlu	0 - 0.03	0 - 0.02
Degradhn	0 - 0.000024	0 - 0.000024
Dissolhn	0 - 0.0000285	0 - 0.0000285
Minerfn	0 - 0.018	0 - 0.02788053
Dissolhp	0 - 0.0000039	0 - 0.0000039
Freuc	202.5 - 10400	150 - 10400
Soilcoh	0.15 - 10	0.011666667 - 10
Soilerod	0.0168 - 0.1638	0.0168 - 0.1338

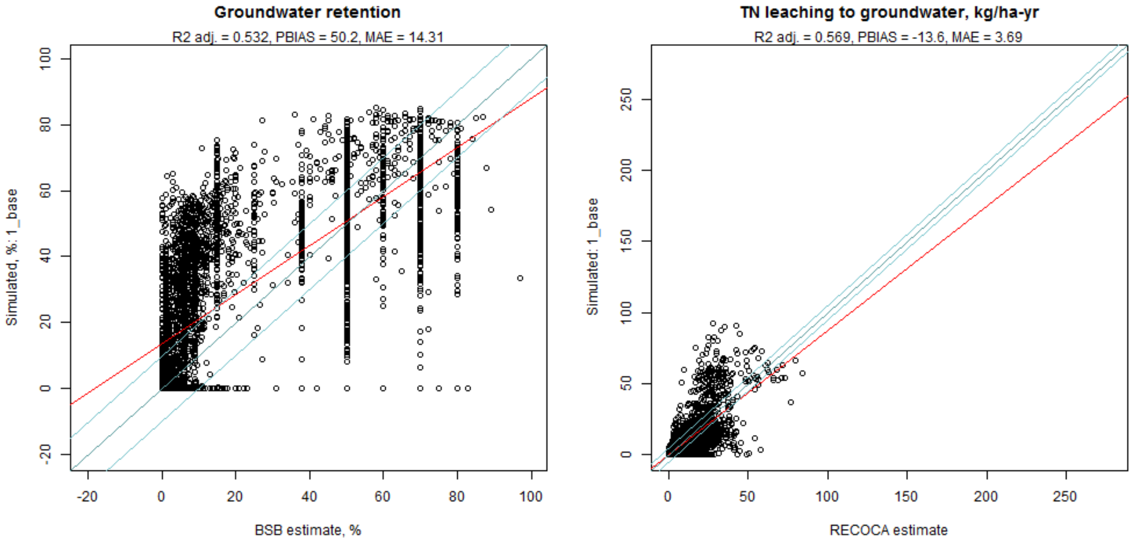

Figure 9. Comparison of observed (x-axes) and simulated (y-axes) nitrogen processes prior to calibration: (a) reduction of N in groundwater and (b) N leaching. BSB estimate refers to Højberg et al. (2017) and RECOCA estimate refers to Andersen et al. (2016).

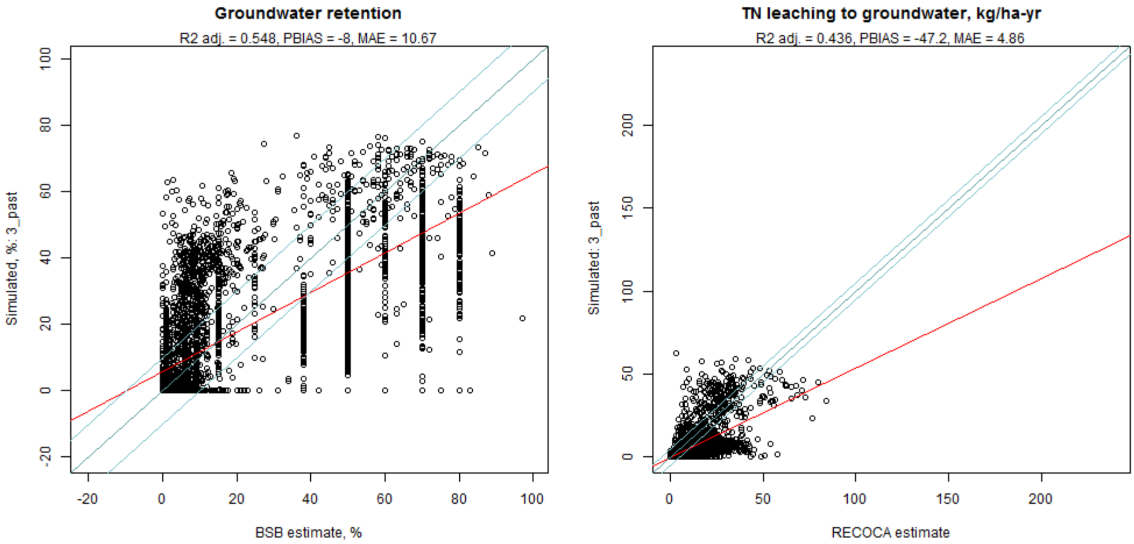

Figure 10. Comparison of observed (x-axes) and simulated (y-axes) nitrogen processes after calibration: (a) reduction of N in groundwater and (b) N leaching. BSB estimate refers to Højberg et al. (2017) and RECOCA estimate refers to Andersen et al. (2016).

3.3.1 Nitrogen leaching

The calibration of N leaching based on RGBs at a large scale did not achieve a significantly better match with observed concentration data in streams and lakes. There appears to be significant differences in watersheds; while individual calibration attempts resulted in better match with observed N leaching data for some RGBs, the overall performance was not improved and in fact, was slightly worse. Figure 11 shows N leaching simulated with the recalibrated E-HYPE model. The overall regional pattern is mostly matched, although the simulated values are not always within the same category as the values estimated by Andersen et al (2016). N leaching is lower in the northern part of the Baltic Sea Basin and increases in the southern part. The values simulated for Poland and east of Poland are noticeably lower than observed.

Figure 12 shows absolute differences between the N leaching simulated with recalibrated E-HYPE and the N leaching from Andersen et al (2016). The categories for display (0-2, 2-5, 5-10, 10-20, 20-50, and >50 kg/ha-year) were selected to differentiate more among the areas with the lowest N leaching. The N leaching is simulated within 5 kg/ha-year or 10 kg/ha-year of the values estimated by Andersen et al (2016) for about 70% and 84% HYPE subbasins, respectively (Figure 13). Leaching in most of the basins is slightly underestimated by 5-10 kg/ha-year. N leaching is underestimated by 10-20 kg/ha-year for Poland and countries northeast of Poland.

The following comparison combines the two categories with the lowest leaching into one (0-5 kg/ha-year). After combining the two lowest categories, about 60% of subbasins were assigned to the same category and 25% were assigned to the next category with majority of them showing underestimation by the model.

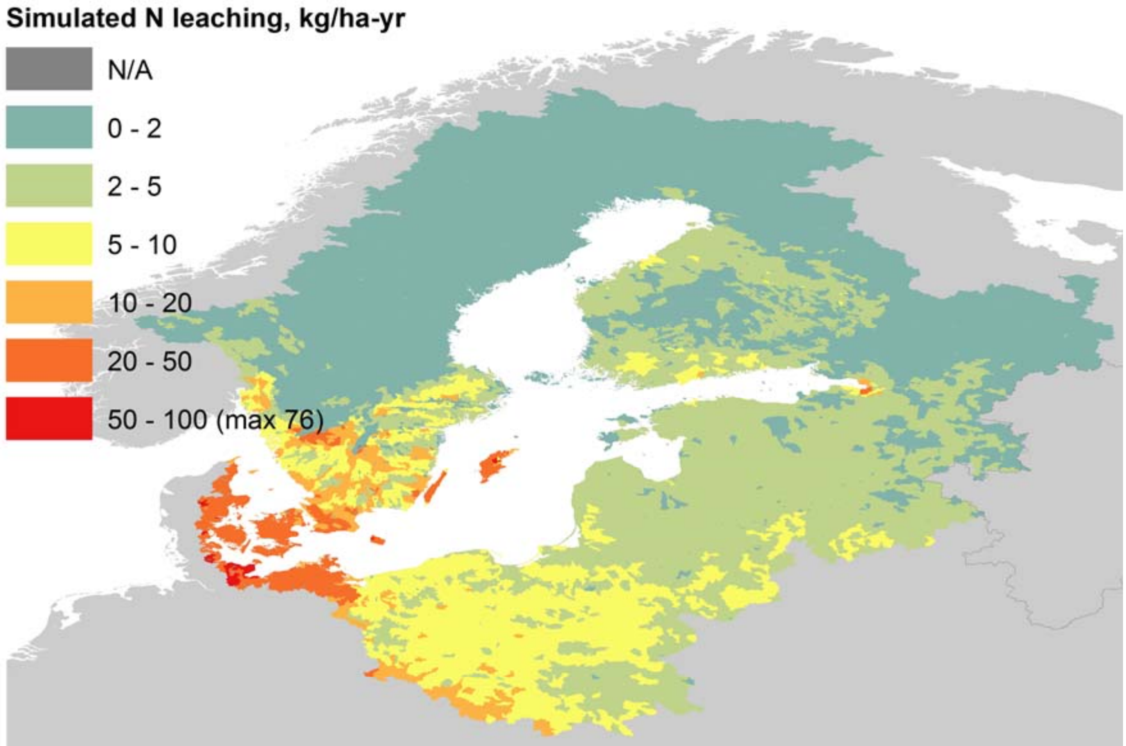

Figure 11. N leaching simulated with the re-calibrated E-HYPE

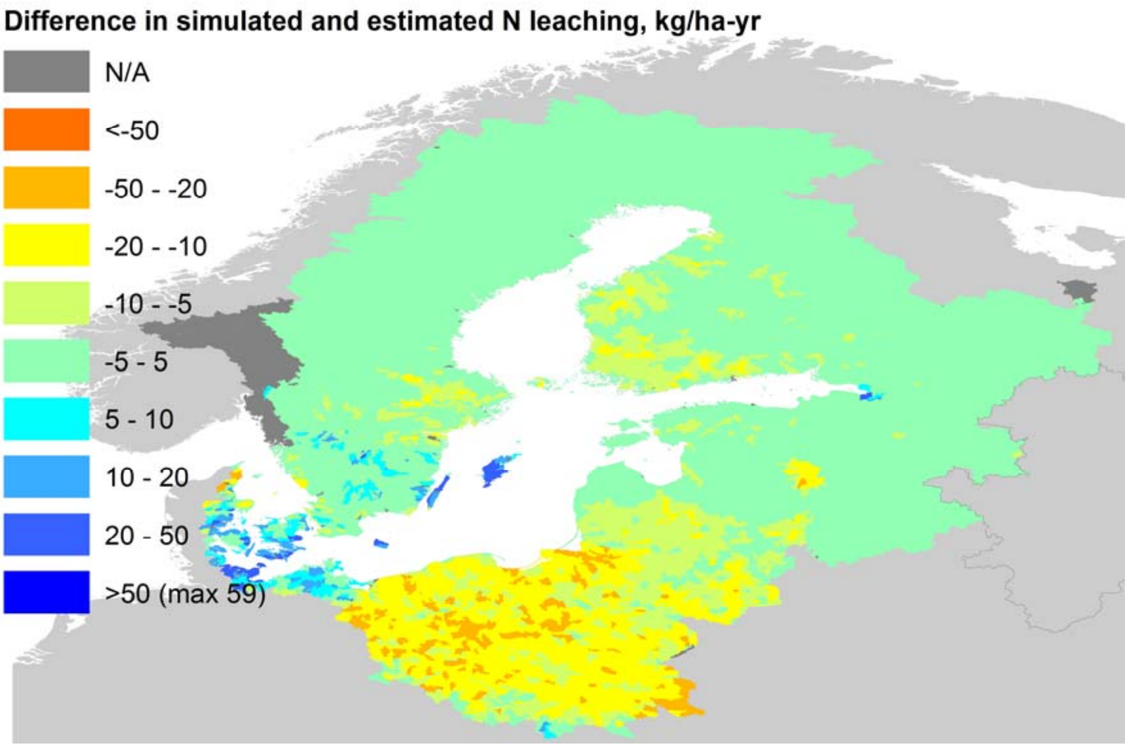

Figure 12. Difference in N leaching simulated with the recalibrated E-HYPE and N leaching based on Andersen et al. (2016).

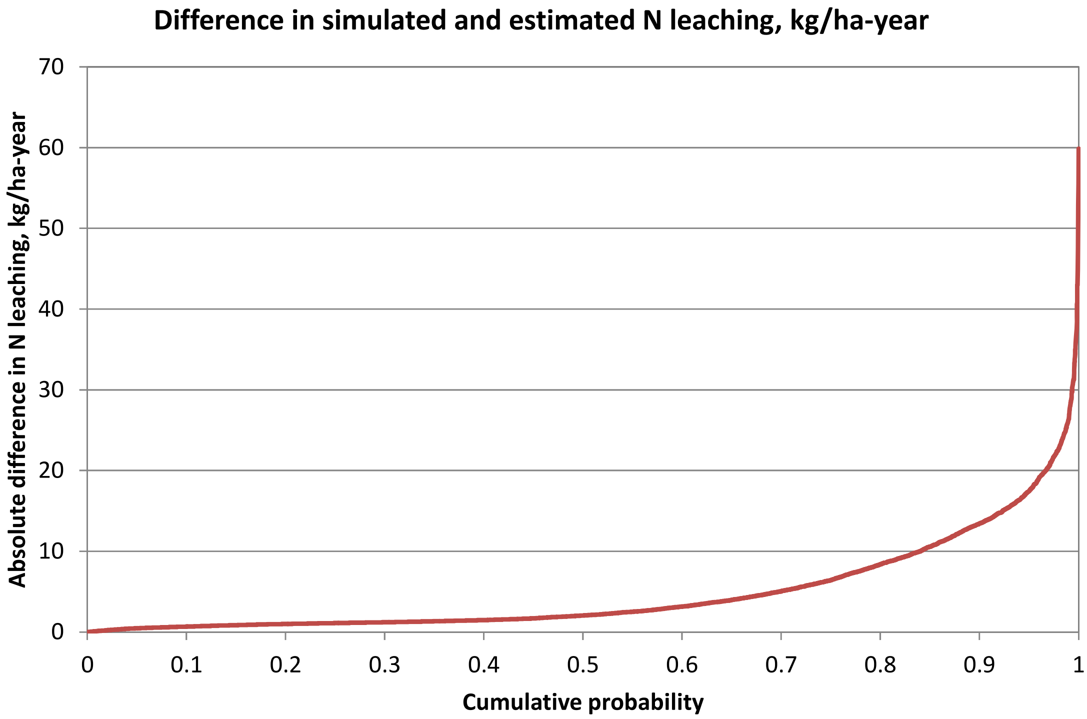

Figure 13. Cumulative probability for the difference in N leaching simulated with the recalibrated E-HYPE and N leaching based on Andersen et al (2016).

3.3.2 Reduction of nitrogen in groundwater

The recalibration of E-HYPE resulted in an overall improved simulation of N reduction in groundwater (Figure 14). The difference between the N reduction in groundwater simulated by the recalibrated E-HYPE prior to sub-regional calibration and estimated by Højberg et al. (2017) is shown in Figure 15. The northern part of the BSB is simulated well. However, some significant differences exist in southern Lithuania, Poland, southern Sweden and Denmark.

In particular, a large area in Poland was estimated by Højberg et al. (2017) to have a very low reduction of N in groundwater due to local hydrogeological conditions. E-HYPE model parameters typically vary with land use and soil during the calibration, aside from sub-regional parameters developed for flow during the original E-HYPE 3.1 calibration (Hundecha et al 2016). The variation of the parameter values with land use and soil means that patterns related to other characteristics (e.g., underlying hydrogeology) cannot be expressed in the model parameters directly. This shows the importance of using process-based data during calibration, especially for large-scale models. While the calibration may or may not result in a successful representation of all the patterns in the process-based data, it allows for a more thorough understanding of the model and better evaluation of model uncertainty.

The gridded estimate of N reduction by Højberg et al. (2017) shows categories of reduction for most of the countries rather than a continuous set of values. The data set was created by combining model estimates for some countries and expert judgement for other countries. When expert judgement was used, the reduction was estimated in discrete categories. The gridded estimate was aggregated to E-HYPE subbasin using median values. The goodness of fit statistics don’t necessarily capture how well the simulated continuous values compare to the observed discrete values. After recalibration, about 55% of sub-basins were simulated within the same general category (<5%, 5-25%, 25-50%, 50-75%, and 75-100%). In addition, about 37% sub-basins were simulated within the neighbouring category. The N reduction is simulated within 10 % or 20 % of the values estimated by Højberg et al. (2017) for about 72% and 86% HYPE subbasins, respectively (Figure 16).

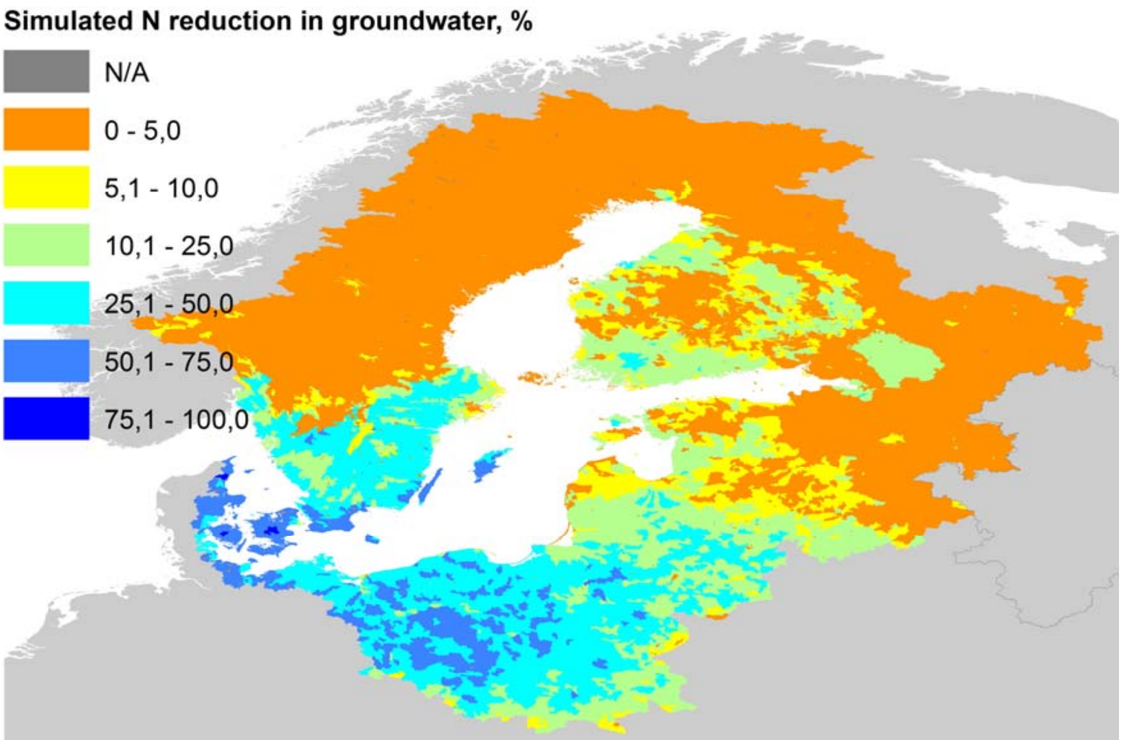

Figure 14. Reduction of N in groundwater simulated with the re-calibrated E-HYPE

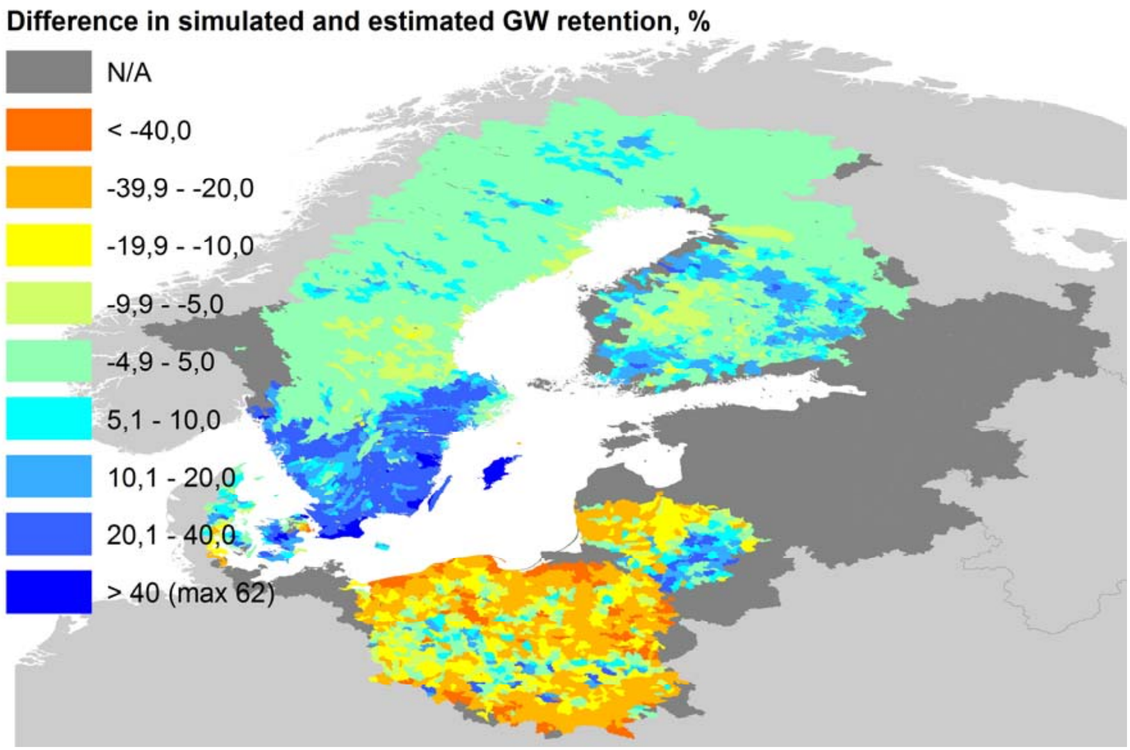

Figure 15. Difference in reduction of N in groundwater simulated with the recalibrated E-HYPE prior to sub-regional calibration and reported by Højberg et al. (2017)

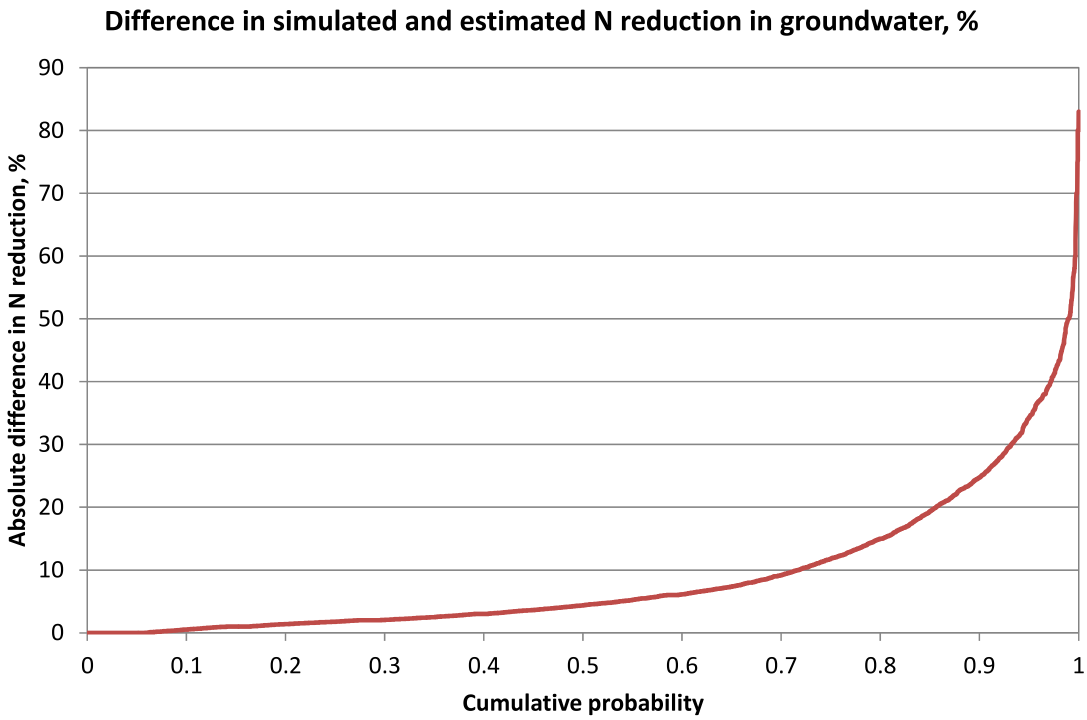

Figure 16. Cumulative probability for the difference in N reduction in groundwater simulated with the recalibrated E-HYPE and N reduction based on Højberg et al. (2017).

3.4 Sub-regional calibration

When comparing the simulated reduction with the reduction estimated by country experts (D3.2, Appendix B) and simulated N leaching with the N leaching by Andersen et al (2015), it became apparent that there were some regional patterns that E-HYPE was not capturing. The soil denitrification rate in E-HYPE 3.1 varies with land use but that cannot express the variability due to regional characteristics of deeper layers. The reduction of nitrogen in groundwater can vary significantly among different regions depending on the reduction potential, organic matter in the aquifer, or overall hydrogeological properties.

The HYPE model code was expanded to enable sub-region dependent calibration (rather than only land use dependent) of a denitrification coefficient that allows us to represent these regional differences. To preserve the integrity of the overall calibration of the soil and land use classes, the sub-regional parameters were only used in selected regions where differences in nitrogen reduction were identified in comparison with Højberg et al (2017). Three regions were selected, one in each of the following countries: Sweden, Lithuania, and Poland (Figure 17). The map of differences was reviewed together with a map of the values estimated by Højberg et al (2017) to select a region the model does not perform well and where a uniform denitrification may be expected at the same time.

Only denitrification rate in the deepest soil layer (layer 3) was calibrated for these 3 regions. Sub-regional denitrification rate is constant within each region but varies among the regions to best represent the reduction of nitrogen in groundwater in each region. Reduction of N in groundwater simulated by the recalibrated E-HYPE after the sub-regional calibration is shown in Figure 18. Figure 19 shows a significant improvement in the match between simulated and observed reduction of nitrogen in groundwater comparing to Figure 10. The improvement is very noticeable where the observed reduction of nitrogen is less than 30% while the higher nitrogen reduction values were not affected. There was no change to nitrogen leaching as expected.

The difference between the N reduction in groundwater simulated by the sub-regional recalibrated E-HYPE and estimated by Højberg et al. (2017) is shown in Figure 20. After sub-regional calibration, about 61% of sub-basins were simulated within the same general category (<5%, 5-25%, 25-50%, 50-75%, and 75-100%). In addition, about 33% of sub-basins were simulated within the neighbouring category. The number of sub-basins simulated outside the same or the neighbouring category decreased from 8% to 6%.

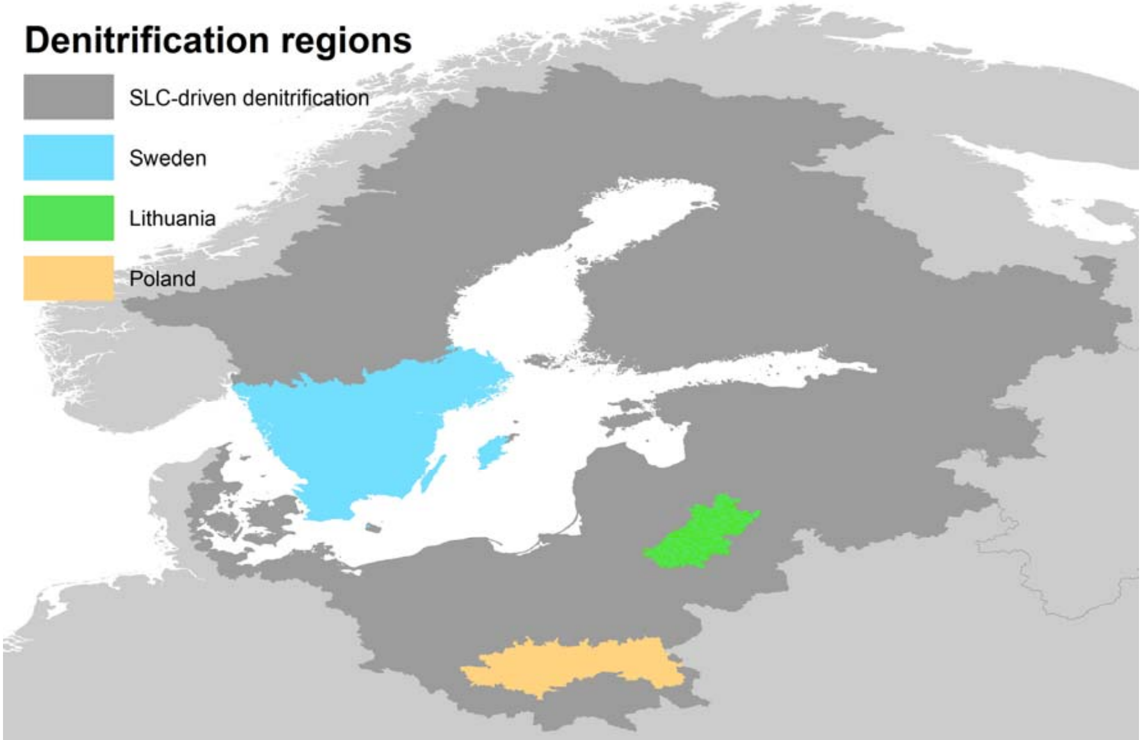

Figure 17. Selected areas for sub-region denitrification rates

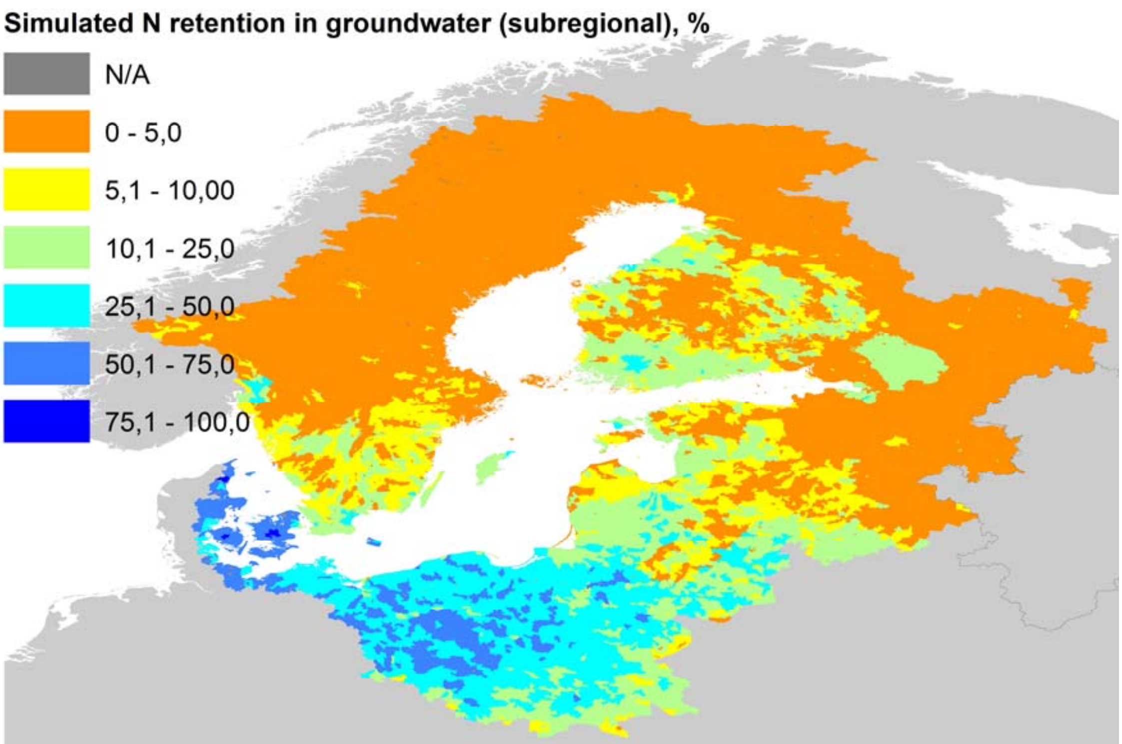

Figure 18. Reduction of N in groundwater simulated with the re-calibrated E-HYPE after sub-regional calibration

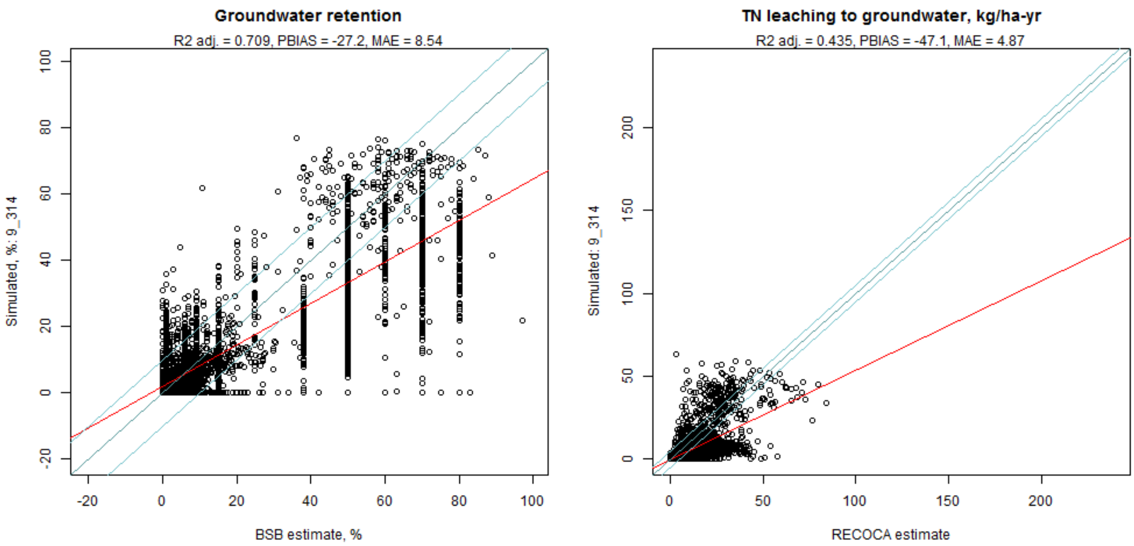

Figure 19. Comparison of observed (x-axes) and simulated (y-axes) nitrogen processes after sub-regional calibration: (a) reduction of N in groundwater and (b) N leaching

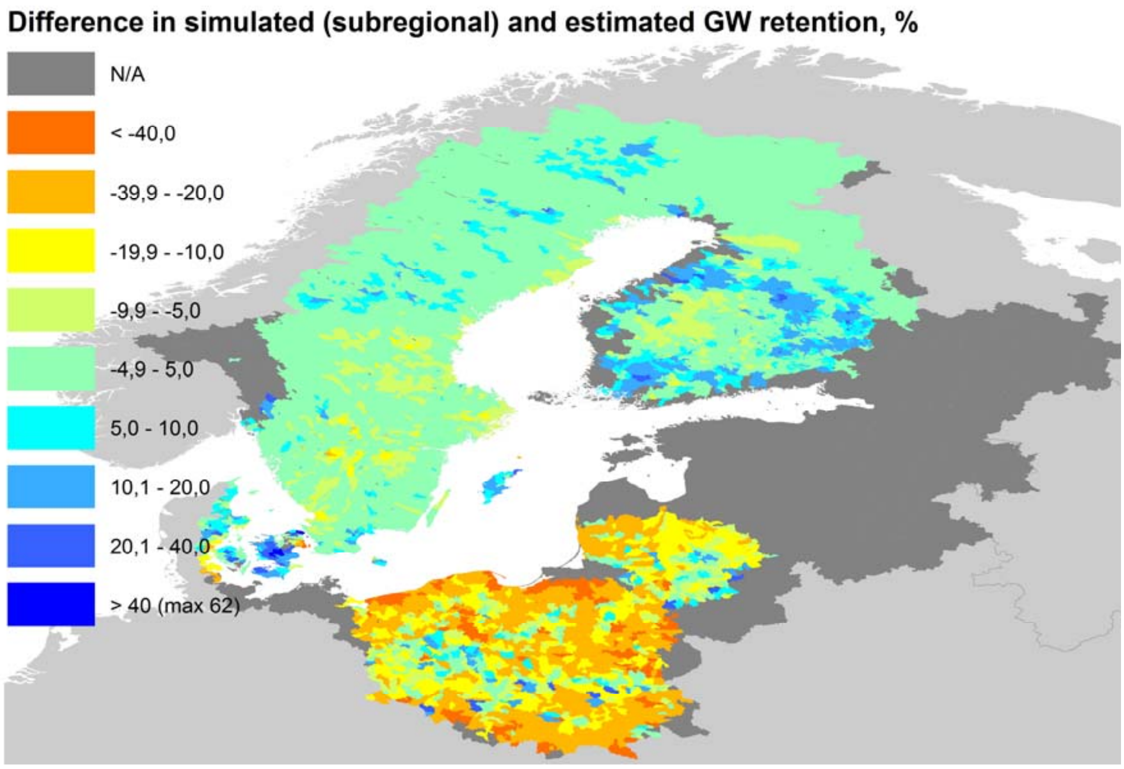

Figure 20. Difference in reduction of N in groundwater simulated with the recalibrated E-HYPE after sub-regional calibration and reported by Højberg et al., (2017)

Figure 21 shows average reduction of N in groundwater for the 3 denitrification regions as well as the values estimated by Højberg et al. (2017). All other areas are grouped together into a region called “other”. The simulated values are shown both after the recalibration and after the sub-regional calibration. The correlation between the observed and simulated values have significantly improved from 0,28 to 0,95 (Figure 22). The sub-regional calibration resulted in better representation of nitrogen processes in soils.

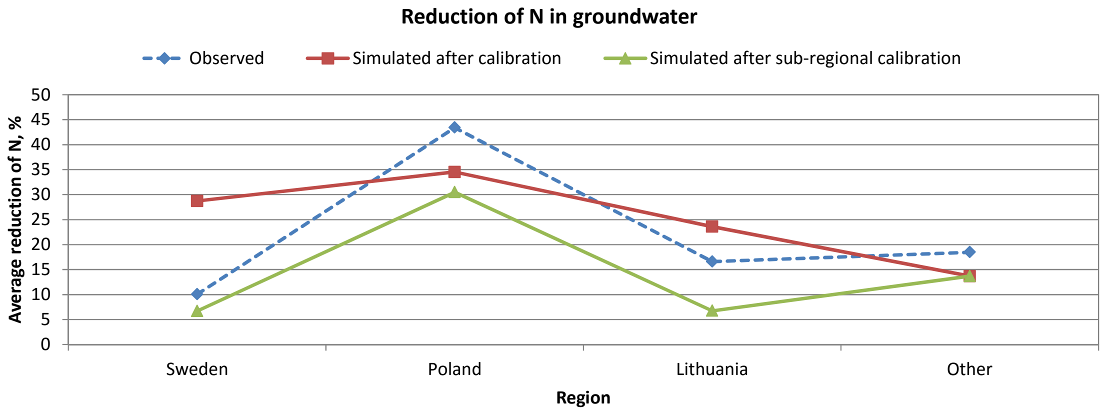

Figure 21. Average reduction of N in groundwater by denitrification regions

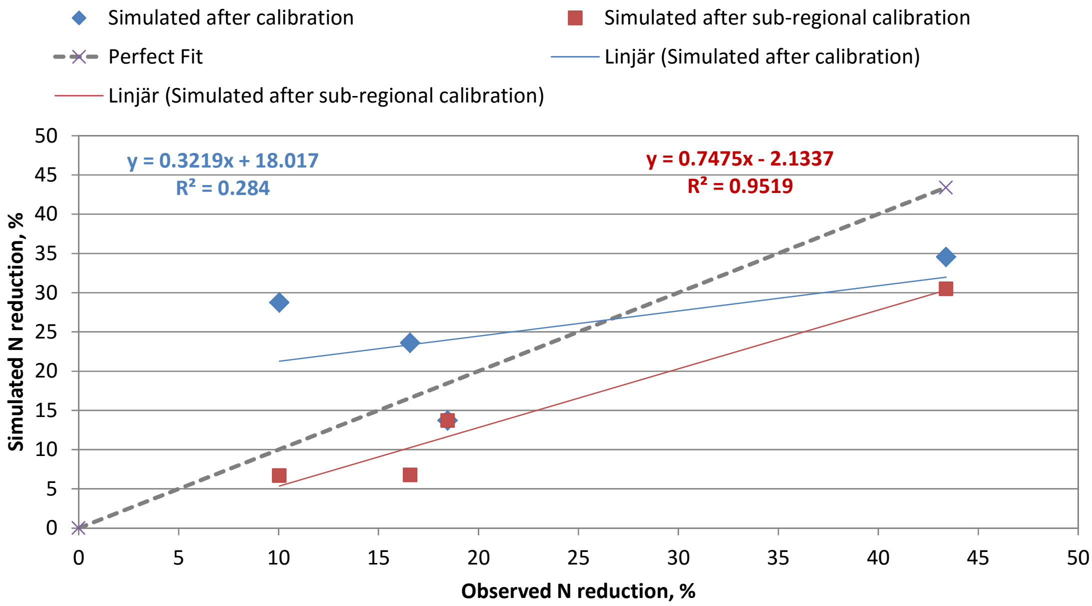

Figure 22 Comparison of model performance with respect to reduction of N in groundwater before (blue) and after sub-regional calibration (brown)

3.5 Model performance evaluation

The model update to E-HYPE v.3.1.4 resulted in an improved model behaviour with respect to groundwater and soil processes. The simulated internal model processes affecting nitrogen concentrations were now compared against data estimated by other studies or derived from observed data. The model was recalibrated with a focus on matching these internal processes better while keeping or improving the goodness-of-fit statistics for observed water quality data. While the recalibrated model matches the spatial patterns in the process-based data, the model performance with respect to goodness-of-fit statistics was not improved.

While the process description were improved, the overall change in the model performance was insignificant. Relative error (RE) in the recalibrated model varies among the monitoring sites, with 65% and 66% of sites having RE within 50% for TP and TN, respectively (Figure 23). The average RE is 42% and 14% and the median RE is 8% and -10% for TP and TN. Correlation coefficient (CC) in the recalibrated model varies with 7% and 30% of sites having CC greater than 0,5 for TP and TN, respectively (Figure 24). The average CC is 0,28 and 0,10 and the median CC is 0,08 and -0,28 for TP and TN. There is only a marginal difference between the model performance for the calibration sites and validations sites.

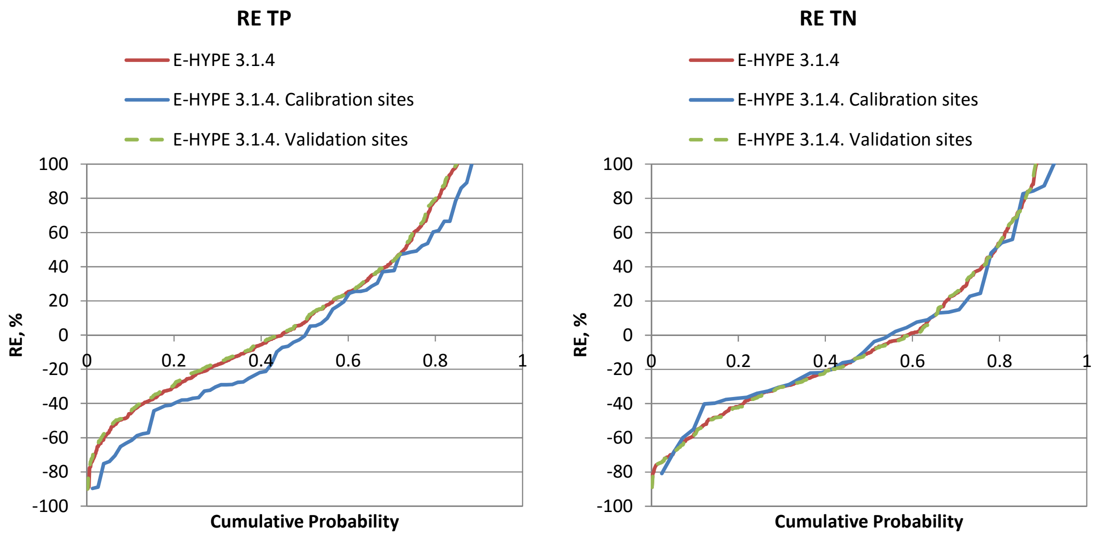

Figure 23. Distribution of relative error for TP and TN in recalibrated E-HYPE 3.1.4

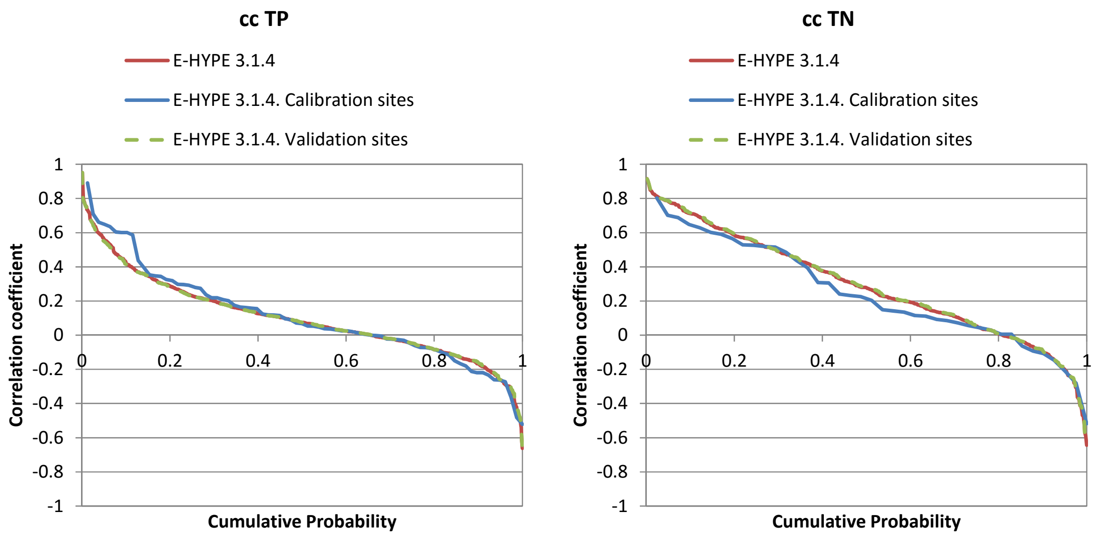

Figure 24. Distribution of correlation coefficient for TP and TN in recalibrated E-HYPE 3.1.4

TN and TP data were available from 523 and 845 sites across Europe, respectively. Of the selected 89 calibration sites, 41 had sufficient data for TN and 78 for TP. Note that TP data are available for about 60% more sites than TN data. Figure 25 and Figure 26 show spatial distribution of available water quality monitoring sites and their corresponding relative error for TN and TP, respectively. While the observed data were significantly expanded during the Soils2Sea project, large gaps still exists especially in southern Baltic countries.

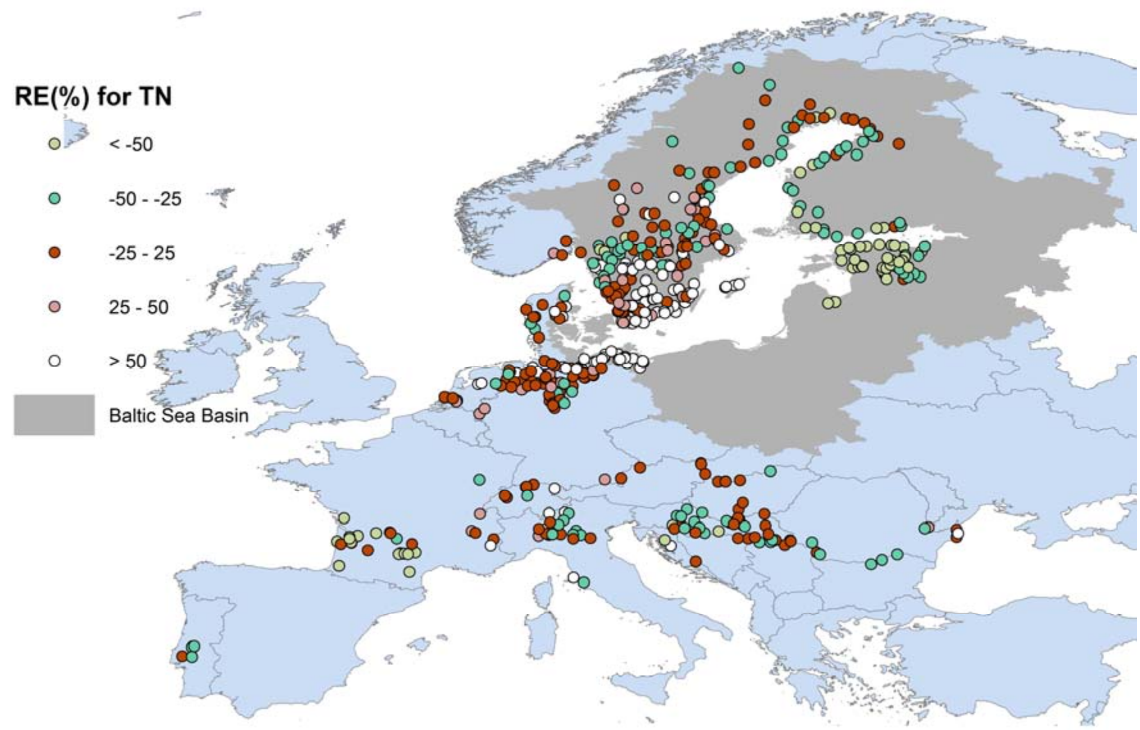

Figure 25. Relative error (%) for TN in E-HYPE 3.1.4

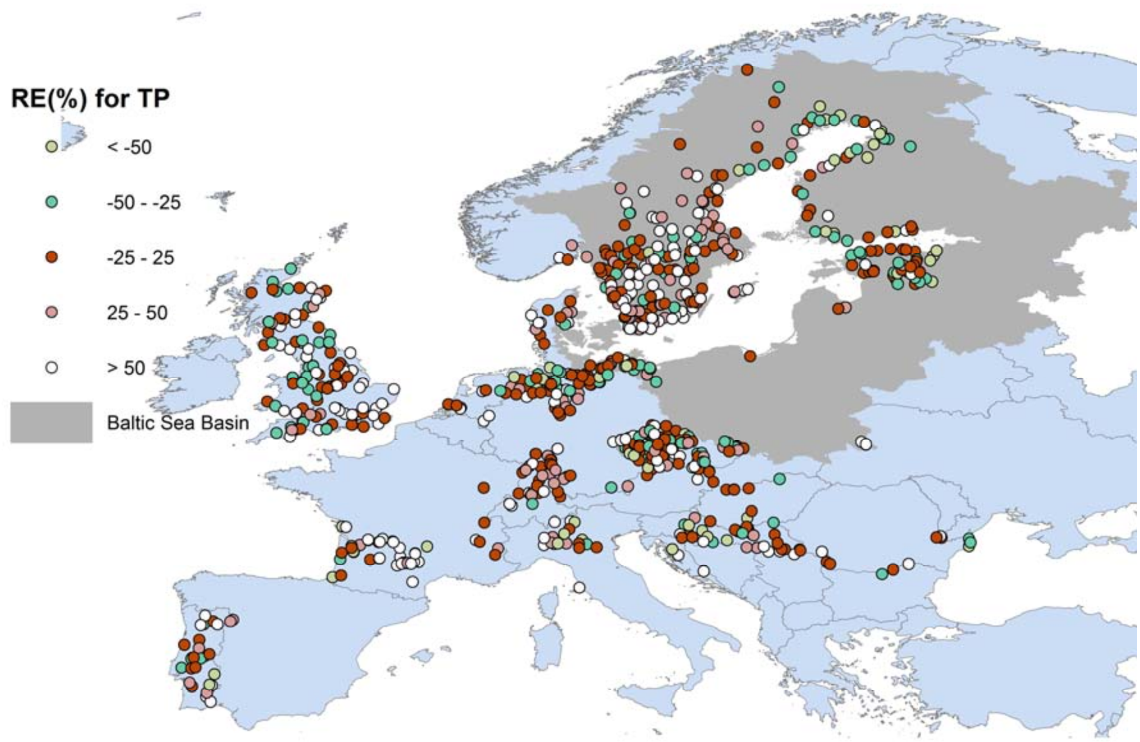

Figure 26. Relative error (%) for TP in E-HYPE 3.1.4

E-HYPE v.3.1.4 is ready to be used in simulating scenarios within the Soils2Sea project. The model performance is sufficient for such simulation, especially considering the improved match for nitrogen processes in soils. Simulating the soils processes correctly is very important for the next tasks in Soils2Sea project and should result in higher model reliability when evaluating impact of future conditions and various measures.

4. References

Arnold J.G., Allen P.M., Muttiah R., Bernhardt G., 1995. Automated Base-Flow Separation And Recession Analysis Techniques. Ground Water 33, 1010-1018, doi: 10.1111/j.1745-6584.1995.tb00046.x

Andersen HE, Blicher-Mathiesen G, Thodsen H, Andersen PM, Larsen SE, Stålnacke P, Humborg C, Mörth CM et al. (2016) Identifying hot spots of agricultural nitrogen loss within the Baltic Sea Drainage Basin. Water Air Soil Pollution, 227:38

BGR & UNESCO (eds.) (2014): International Hydrogeological Map of Europe 1 : 1,500,000 (IHME1500). Digital map data v1.1. Hannover/Paris .

Bronstert A, Bardoddy A, Bismuth C, Buiteveld H, Disse M, Engel H, Fritsch U, Hundecha Y et al. (2007) Multi-scale modelling of land-use change and river training effects on floods in the Rhine Basin. River Research and Applications, 23, 1102-1125.

Donnelly C, Arheimer B, Capell R, Dahné J, Strömqvist J (2013) Regional overview of nutrient load in Europe – challenges when using a large-scale model approach, E-HYPE. IAHS Publication 361, 49-58.

Donnelly C, Andersson JCM, Arheimer B (2015) Using flow signatures and catchment similarities to evaluate the E-HYPE multi-basin model across Europe. Hydrological Sciences Journal. http://dx.doi.org/10.1080/02626667.2015.1027710.

Hansen AL, Gunderman D, He X, Refsgaard JC.(2014a) Uncertainty assessment of spatially distributed nitrate reduction potential in groundwater using multiple geological realizations. Journal of Hydrology. 519, 225-237.

Hansen AL, Christensen BSB, Ernstsen V, He X, Refsgaard JC (2014b) A concept for estimating depth of the redox interface for catchment-scale nitrate modelling in a till area in Denmark. Hydrogeology Journal, 22(7), 1639-1655.

Hundecha Y, Arheimer B, Donnelly C , Pechlivanidis I (2016) A regional parameter estimation scheme for a pan-European multi-basin model. Journal of Hydrology: Regional Studies, 6, 90-111. http://dx.doi.org/10.1016/j.ejrh.2016.04.002

Højberg AL, Windolf J, Børgesen CD, Troldborg L, Tornbjerg H, Blicher-Mathiesen G, Kronvang B, Todsen H et al. (2015) En ny kvælstofmodel. Oplandsmodel til belastning og virkemidler. Metode rapport (A new nitrogen model. Catchment model for loads and measures. Methodology Report – In Danish). GEUS and Aarhus University. Available from http://www.geus.dk/DK/water-soil/water-cycle/Documents/national_kvaelstofmodel_metoderapport.pdf

Højberg A.L., Hansen A.L., Wachniew, P., Żurek A., Virtanen S., Arustiene J., Strömqvist J., Rankinen K., Refsgaard J.C. (2017) Review and assessment of nitrate reduction in groundwater in the Baltic Sea Basin. Journal of Hydrology: Regional Studies, 12, 50–68. http://doi.org/10.1016/j.ejrh.2017.04.001

Riml J, Wörman A (2011) Response functions for in-stream solute transport in river networks. Water Resources Research, 47, W06502.

Refsgaard JC, Jakobsen R, Hansen AL, Højberg AL, Żurek AJ, Różański K, Witczak S, Wachiniew P, Donnelly C, Capell R, Bartosova A, Strömqvist J, Wörman A, Morén I, Up-scaling methodologies. BONUS Soils2Sea Deliverable 3.2. Geological Survey of Denmark and Greenland, Copenhagen, November 2016, www.Soils2Sea.eu

The present work has been carried out within the project 'Reducing nutrient loadings from agricultural soils to the Baltic Sea via groundwater and streams (BONUS SOILS2SEA)', which has received funding from BONUS, the joint Baltic Sea research and development programme (Art 185), funded jointly by the EU and national funding agencies: Innovation Fund Denmark, The Swedish Environmental Protection Agency (Naturvådsverket), The Polish National Centre for Research and Development, The German Ministry for Education and Research (Bundesministerium für Bildung und Forschung), and The Russian Foundation for Basic Research (RFBR).