Reducing uncertainty in the calibration and validation of the INCA-N model by using soft data

J. Randall Etheridge, Ahti Lepistö, Kirsti Granlund, Katri Rankinen, François Birgand and Michael R. Burchell II

ABSTRACT

Process-based nutrient models are increasingly used to determine the impact of future changes in land use, agriculture production practices and climate on the quantity and timing of nutrients reaching surface waters. Calibration of catchment-scale models to observed conditions can be difficult due to parameter uncertainty and the heterogeneity of catchment processes. Soft data, i.e. knowledge of processes gained through experimentation, have been suggested as one method of reducing uncertainty and producing a more accurate model of the processes that occur in a catchment. In this work, the Integrated Catchment model for Nitrogen was calibrated and validated for the Yläneenjoki catchment in south-western Finland by incorporating soft data. The calibration for 2003–2008 produced an adequate model of the in-stream nitrate concentrations ($R^2 = 0.45$, NS = 0.42). However, model validation using data from 1997–2002 showed that the simulated in-stream nitrate concentrations were above the observed concentrations throughout the entire period ($R^2 = 0.34$, NS < 0). These results show that soft data can be used to constrain model parameters, resulting in a more accurate model of the catchment, but do not guarantee the best validation results as the simulated processes may not occur at the same time and rate as they did in the catchment.

Key words | catchment, modeling, nitrogen, soft data

J. Randall Etheridge (corresponding author)
François Birgand
Michael R. Burchell II
Department of Biological and Agricultural Engineering,
North Carolina State University,
Raleigh,
North Carolina,
USA
E-mail: *jretheri@ncsu.edu*

Ahti Lepistö
Kirsti Granlund
Katri Rankinen
Finnish Environment Institute,
Helsinki,
Finland

INTRODUCTION

Agricultural production has been identified as a major contributor of non-point source pollution in catchments throughout the world (e.g. Howarth *et al.* 2002; Räike *et al.* 2003; Cherry *et al.* 2008; HELCOM 2011). Although agricultural land use in Finland covers only 7% of the total land area, the losses of nutrients from agriculture are approximately 50% of the total nitrogen (N) loading (Vuorenmaa *et al.* 2002). The diffuse pollution from agriculture is concentrated in southern, south-western and western areas of Finland. The nutrient-rich waters leave the agricultural land and enter streams, lakes and estuaries leading to eutrophication (Vitousek *et al.* 1997) and potentially make the water supply unfit for consumption because of health risks (Ward *et al.* 1996; Townsend *et al.* 2003). The negative effects of nitrogen-rich waters on human health, biodiversity and climate change provide incentive to gain a better understanding of nitrogen processes and the effects of agriculture on the nitrogen cycle (Galloway *et al.* 2008).

Process-based models focused on hydrology and nutrient leaching are increasingly used to determine the impact of future changes in land use (agriculture, forestry, etc.) and climate on the nutrients reaching surface waters. Such models have been developed and used for assessing water quality issues at the small catchment scale (Lunn *et al.* 1996; Heng & Nikolaidis 1998). However, as decision tools for planners and managers, the use of these models is often limited due to high input data requirements, which prevent the calibration of the models for large river systems. At the catchment scale and in larger river-dominated basins, advanced process-based semi-distributed dynamic nutrient

doi: 10.2166/nh.2013.039

models such as the Integrated Catchment model for Nitrogen (INCA-N) (Whitehead et al. 1998a; Wade et al. 2002) can be applied over a wide range of spatial and temporal scales. INCA-N is a process-based model that uses a mass-balance approach to track mineral nitrogen within a catchment. It can integrate both point and non-point sources of nitrogen (Whitehead et al. 1998a, b; Wade et al. 2002). The model incorporates hydrology and different nitrogen processes such as mineralization and denitrification to simulate the mass of nitrogen in each part of the system.

A vital step in the process of modeling scenarios for planning and management purposes is to prove a model is capable of simulating what is currently occurring in a catchment through the calibration and validation process (Santhi et al. 2001; Jarvie et al. 2002; Granlund et al. 2004). In model calibration, the parameters are adjusted to more accurately simulate the observed results. The model parameters set during calibration are then applied to another period of time to assess the accuracy of the model in the validation phase (Refsgaard 1996).

A problem with this procedure is the possibility of obtaining a numerically correct result for the wrong reason (McIntyre et al. 2005; Kirchner 2006; Rankinen et al. 2006). The overestimation of nitrogen inputs to a system can be numerically compensated for by increasing nitrogen removal through plant harvest or denitrification, for instance. In this case, the modeled nutrient concentrations at the outlet of the catchment may provide a good fit with observed data, but the process rates within the model are not accurate. If the model does not accurately estimate what is occurring during normal conditions where there is observation data available for calibration, it is unlikely that the model will be reliable as the conditions move outside of those experienced currently (Kirchner 2006).

In water quality and hydrology modeling, soft data are knowledge about a catchment or process that is gained through experimentation, but cannot be compared directly to model output due to high uncertainty (Seibert & McDonnell 2002; Winsemius et al. 2009). Soft data do not provide an absolute number that can be used in calibration of the model such as in-stream nutrient concentrations or a continuous flow record, which are referred to as hard data. Hard data have an acceptable level of certainty and are used directly in support of model calibration (Winsemius et al. 2009). The uncertainty associated with soft data for use in process-based modeling is partially due to the process rates being measured at the field scale and the processes being simulated at the catchment scale (Seibert & McDonnell 2002; Wade et al. 2008). Another factor contributing to the uncertainty of experimental data is that experiments often provide a wide range of potential process rates based on a limited number of measurements. These two problems prevent the results of some field experiments from being used as hard data during model calibration, but the soft data can be used to specify a realistic parameter range to reduce model parameter uncertainty and provide a more realistic simulation of what is occurring in the catchment (Seibert & McDonnell 2002).

Some work has been done on the issue of parameter uncertainty in the INCA-N model (McIntyre et al. 2005; Wade et al. 2008; Rankinen et al. 2013). McIntyre et al. (2005) used a Monte Carlo analysis to show that the most sensitive model parameters had high uncertainty. They recommended the use of observed soil and groundwater concentrations to constrain model parameters, but did not have these data available for their simulations. The use of experimental data to reduce parameter uncertainty in the INCA-N model was recommended based on work with a virtual catchment by Raat et al. (2004). Rankinen et al. (2006) used the Generalized Likelihood Uncertainty Estimation (GLUE) methodology to determine the usefulness of soft data for automatic calibrations. Their work showed that the flow hydrograph and in-stream nutrient concentrations were not enough to adequately constrain the model parameters, but soft data could be used to reduce equifinality or the occurrence of multiple parameter sets producing the same simulation results. Despite these recommendations, the method and impact of using results from field experiments in manual calibrations and validations has not been fully explored for the INCA-N model. In previous studies, annual process rates from the literature have been used to adjust the nitrogen process parameters (e.g. Wade et al. 2006; Bärlund et al. 2009), but the steps to use this soft data in calibration are rarely shown.

The purpose of this paper is to show how soft data can be used to constrain model parameters in the manual calibration of the INCA-N model through a case study of the

Yläneenjoki catchment in Finland. The INCA-N model was calibrated for the years 2003–2008 and validated for the years 1997–2002. This study will focus on the use of published nitrogen process rates and an examination of the groundwater portion of the model to calibrate the in-stream nitrate (NO_3-N) and ammonium (NH_4-N) concentrations.

METHODS

Model description

INCA-N is a process-based model that uses a mass-balance approach to track mineral nitrogen in a watershed (Whitehead *et al.* 1998a; Wade *et al.* 2002). The model is semi-distributed and incorporates point sources, non-point source, hydrology, land-based nitrogen processes and in-stream nitrogen processes to simulate the daily flow, NO_3-N and NH_4-N concentrations in catchment streams. In this study, model version 1.11.10 was used.

The land-based portion of the model includes two zones: the groundwater zone and the soil zone. The model uses hydrologically effective rainfall (HER) as the input to the soil water. HER is defined as the portion of precipitation that reaches stream channels either through surface runoff or by groundwater discharge (Rankinen *et al.* 2002). The HER can be supplied for the whole catchment or for individual subcatchments within the model. The time it takes for rainfall to make it to the stream is driven by the base flow index (BFI), residence time constants and soil properties that lead to direct runoff. All of the HER that does not runoff infiltrates and passes through the soil zone. The BFI determines the portion of water that will pass through the groundwater zone after going through the soil zone. A higher BFI means a higher proportion of the water passes through the groundwater zone instead of going directly to the stream after passing through the soil zone. The residence time constants are defined separately for the groundwater and soil water zones.

Stream flow is modeled using a multiple reach approach that relates discharge Q to a mean flow velocity v based on the equation $v = aQ^b$. Ideally, the a and b constant flow parameters can be estimated from tracer experiments and/or based on channel properties, but can be calibrated based on the modeled hydrograph if the tracer experiments/channel properties are not available (Whitehead *et al.* 1998a).

The land-based nitrogen processes of mineralization, nitrification, plant uptake, denitrification and immobilization are simulated in the model. All of these process rates are temperature and moisture dependent. The process rates can be defined for up to six land use classes. Denitrification and nitrification are simulated in the in-stream portion of the model. The in-stream process rates are temperature dependent and can be altered between reaches.

INCA-N models transformation of nitrogen within the catchment, as well as leaching of inorganic nitrogen. It is assumed in the model that there is an infinite source of organic nitrogen that can be mineralized. Fertilizer applications (NO_3-N and NH_4-N), atmospheric deposition, point sources of nitrogen and biological nitrogen fixation are the other sources of nitrogen considered in the model. A mass-balance approach is taken to account for the transformations and movement of nitrogen through the watershed. The volume of water and mass of nitrogen is accounted for in each land use within each subcatchment. As water leaves one zone of the model and enters another zone, a mass of nitrogen is transferred between the modeled zones.

Study site

The River Yläneenjoki is one of two major rivers that discharge into Lake Pyhäjärvi (Figure 1). It is located in south-western Finland, which is a hotspot for agricultural nitrogen loading. Lake Pyhäjärvi eventually drains along the Eurajoki River to the Gulf of Bothnia to the west. The Yläneenjoki catchment has an area of 233 km^2 with 31% of the land in agricultural production. The agricultural production in the Yläneenjoki catchment is considered intensive for Finland with the primary products being cereals and animals (Lepistö *et al.* 2006). The soils in the river valley of the Yläneenjoki catchment are mainly clay and silt. Approximately 11% of the precipitation falls as snow and the long-term (1961–1990) average annual precipitation was 630 mm in the Yläneenjoki catchment (Hyvärinen 1999). The average discharge from the River Yläneenjoki is 2.1 $m^3 s^{-1}$ (Mattila *et al.* 2001) with the highest flows typically occurring during the spring snow melt and fall.

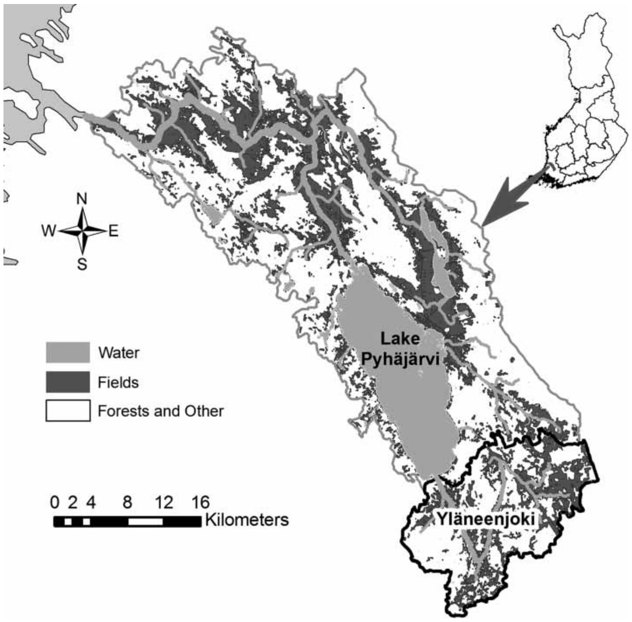

Figure 1 | The Eurajoki River catchment; the Yläneenjoki catchment is outlined in black.

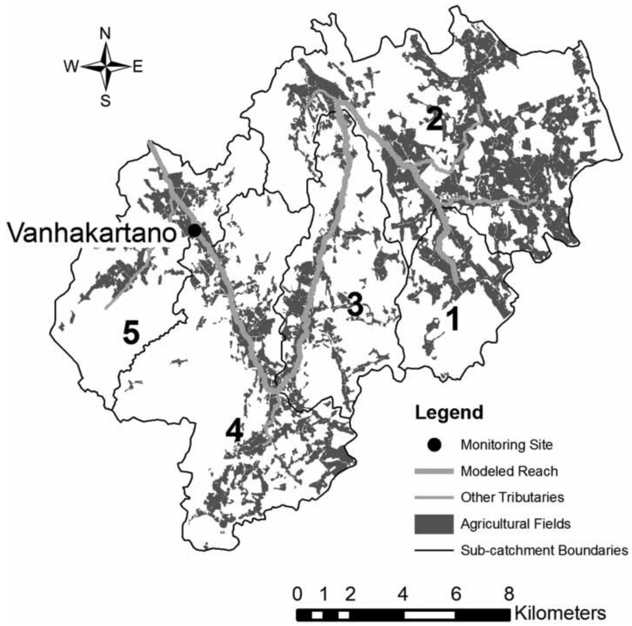

Figure 2 | The Yläneenjoki catchment showing the INCA-N subcatchments and the agricultural lands in dark grey.

Data collection

The daily discharge of the River Yläneenjoki has been monitored since the 1970s at the Vanhakartano measuring site (Figure 2), while nutrient concentrations have been monitored on a weekly to monthly basis. These data were available through the Environmental Information System (HERTTA) maintained by the Finnish environmental administration. The HERTTA data were supplemented by results from an automatic water quality station in the spring and the autumn of 2007 (Lepistö *et al.* 2008; Koskiaho *et al.* 2010). NO_3-N concentrations from the river were collected using sensor-based technology on an hourly interval during 27 March–27 April 27 and 4 October–20 November 2007. The measurements from the automatic water quality station provided a continuous record of observed data for two short periods of time, constituting only 4% of the calibration period. The daily average of these measurements was used in this work.

The model requires an input time series of HER, soil moisture deficit, air temperature and actual precipitation. These inputs were taken from the Watershed Forecast System (WSFS) of the Finnish Environment Institute (Vehviläinen & Huttunen 2010). For this study, the data collection went beyond the minimum time series inputs that was required to run the model and included soft data, i.e. any available information that could be used to improve the calibration of simulated nitrogen processes (Table 1). An example of soft data would be the nutrient process rates in agricultural soils determined through experimentation (Table 2). Scientific literature contains useful information about field experimentation, but collected data should also include reports of local agricultural practices and information on fertilization practices and crop yields, which are seldom available for Finnish catchments.

Model calibration: steps A–F

The INCA-N model has both a hydrologic and a nutrient component. Manual calibration of the INCA-N model begins with the hydrologic component because the movement of nitrogen is driven by water flow (Wade *et al.* 2002). Although the calibration process is iterative, the process generally follows a path similar to the path a drop of water would follow through the catchment, specifically by starting the calibration process in the land portion of the model then working to the stream portion of the model. The process used to complete the hydrologic calibration in this study was based on methods described in Rankinen *et al.* (2002) and Granlund

Table 1 | Source, extent and type of data that can be used to more accurately model catchment N processes using the INCA-N model

Data type	Source	Preferred extent
Fertilizer application rates and time of year	Producer interviews or published values	Each land use
Ranges of nitrogen process rates (e.g. denitrification, mineralization)	Published values or direct measurement	Each land use
Crop nitrogen uptake	Published values or direct measurement	Each land use
Crop growth rates	Published values or direct measurement	Each land use
Soil water NO_3-N concentration	Direct measurement or monitoring database	Each land use
Groundwater NO_3-N concentration	Direct measurement or monitoring database	Each land use or subcatchment
Soil water NH_4-N concentration	Direct measurement or monitoring database	Each land use
Groundwater NH_4-N concentration	Direct measurement or monitoring database	Each land use or subcatchment

Table 2 | Nutrient process rates used as soft data for model calibration

Land use	Process	Rate (kg N ha^{-1} a^{-1})	Source
Spring cereals	Nitrogen leaching	4–26	Salo & Turtola (2006)
Agriculture production	Nitrogen leaching	2–99	Vuorenmaa *et al.* (2002); Salo & Turtola (2006); Rankinen *et al.* (2007)
Forest	Nitrogen leaching	0.6–2.5	Vuorenmaa *et al.* (2002)
Spring cereals	Nitrogen uptake	40–112	Information Centre of the Ministry of Agriculture and Forestry (2007)
Winter cereals	Nitrogen uptake	30–116	
Grass	Nitrogen uptake	149–296	
Agriculture production	Mineralization	40–55	Rankinen *et al.* (2007)
Agriculture production	Denitrification	3–17	Svensson *et al.* (1991); Barton *et al.* (1999)
Forest	Denitrification	2	Barton *et al.* (1999)

et al. (2004). After satisfactory hydrologic calibration, nitrogen calibration is then conducted. As a rule, if the hydrologic parameters are adjusted, the nutrient calibration process should be restarted.

In this work, calibration of the nitrogen portion of the INCA-N model was completed for the years 2003–2008 with the use of soft data. A flow chart of the nitrogen calibration process using soft data is shown in Figure 3. This chart follows the example by Santhi *et al.* (2001) for the Soil and Water Assessment Tool (SWAT). An initial calibration was completed for the Yläneenjoki catchment by Lepistö *et al.* (2008). One problem noted by Lepistö *et al.* (2008) was that a fall application of manure was missed in their calibration. This was a likely cause of the simulated peak NO_3-N concentrations being lower than the observed concentrations during the fall of 2007 (Figure 4). Many animal producers in Finland apply manure to their fields in the autumn to empty their manure storage before winter.

To address this issue, a fall manure application was added at the beginning of this calibration process. The modeled fall manure application rates in the calibration varied between 24 and 50 kg N ha^{-1} depending upon the crop. The simulated annual application rates of fertilizers/manure did not exceed the regulation levels of the Nitrates Directive (EEC 1991) or the Finnish Agri-Environmental Programme, which promotes sustainability in agricultural practices (Rankinen *et al.* 2009). The addition of a new source of nitrogen required the calibration of the nitrogen portion of the model to start at step A in Figure 3.

The soft data were used in step B (Figure 3), where the simulated nitrogen process rates were compared to literature values. In this step of the calibration procedure, the quality of the calibration was judged primarily by the simulated process rates, but the fit of the in-stream nitrogen concentrations to the observed concentrations were still

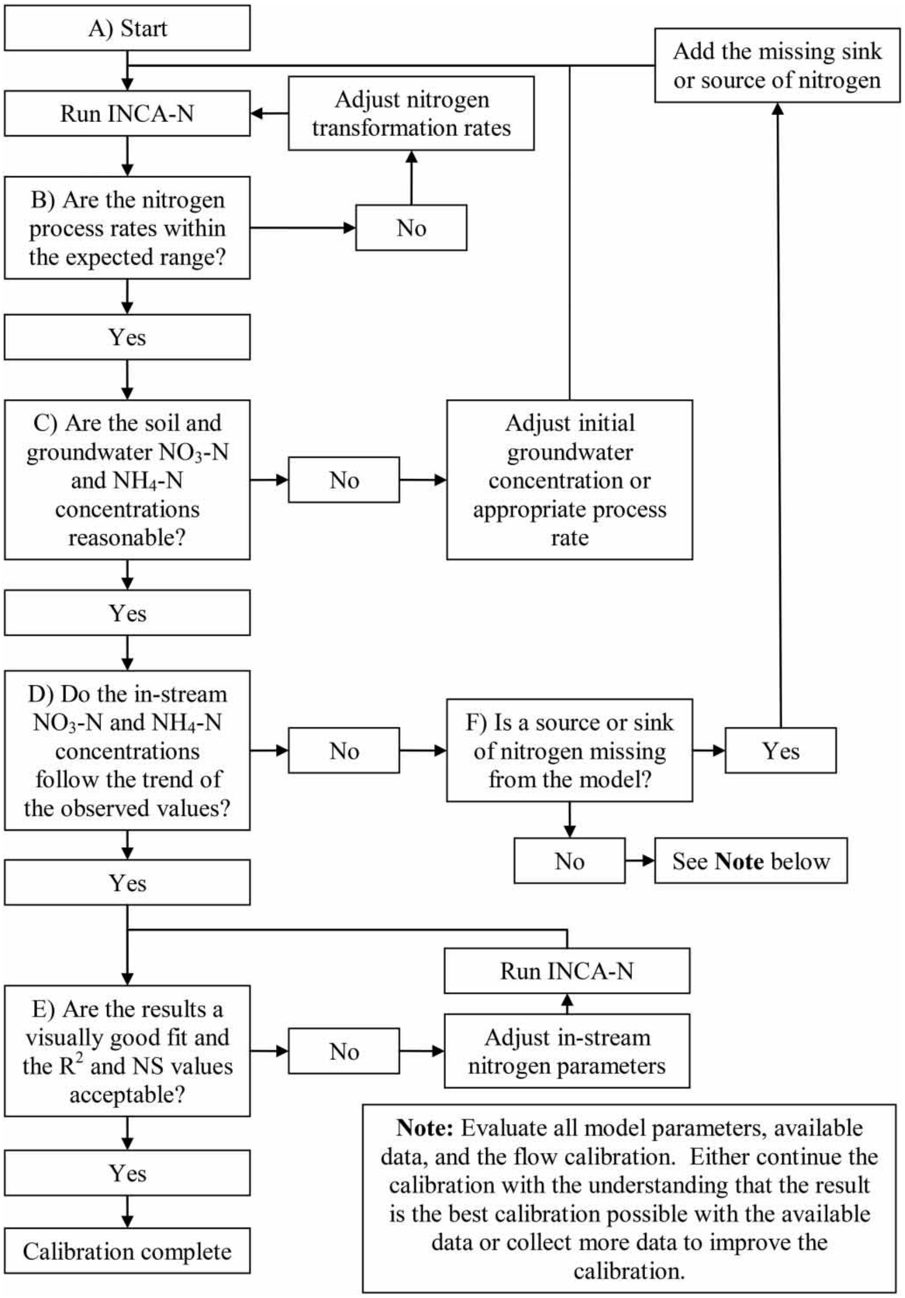

Figure 3 | Calibration procedure for the nitrogen portion of the INCA-N model using soft data.

considered as they show potential sources of error in the calibration. The processes include leaching, plant uptake, mineralization, nitrification, denitrification and fixation. The mass of nitrogen consumed or released by the individual processes were modeled for each land use. The model parameters were adjusted for each land use so that the simulated loads were in the range of values available from experimental data. Leaching was the only process that did not have an associated process rate that can be adjusted. The amount of leaching was based on the amount of flow through the soil and groundwater zones along with the mass of nitrogen stored in these zones. Elevated fertilizer inputs or mineralization rates are potential causes of leaching rates being above the expected range.

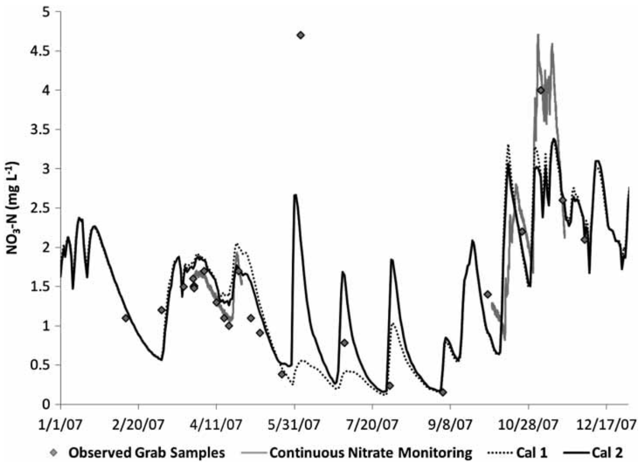

Figure 4 | INCA-N simulated NO_3-N concentrations from Calibration 1 and Calibration 2 compared to grab samples and continuous automatic monitoring for the Yläneenjoki catchment in 2007.

Step C (Figure 3) in the calibration process is closely linked to step B where the process rates were adjusted. The values reported in the load tables take into account both the soil and groundwater zones. Initial concentrations of NO_3-N and NH_4-N in both the groundwater and soil water zones are input values of INCA-N. The use of water samples collected in the catchment can provide a guide to the acceptable range of initial values, but may not provide the true concentration if the samples were collected at a time other than the beginning of the modeling period, or if there is heterogeneity of concentrations within each subcatchment. Denitrification and leaching are the only two nitrogen processes modeled by INCA-N in the groundwater zones, so the initial nutrient concentrations and rate of denitrification in the groundwater zone were adjusted to alter these process loads. Details about how these were adjusted in steps B and C are discussed in the results and discussion section.

After the process rates were constrained to the expected range and the nitrogen concentrations in the land portion of the model were deemed to be reasonable, the calibration proceeded to step D (Figure 3). In this step, the timing and relative magnitude of simulated peaks and drops in concentrations were examined to see if they matched the observed increases and decreases in concentrations. If the timing and magnitude of changes in concentrations did not match well, knowledge of the nitrogen processes and the study catchment were used to look for a potential source of error. The cause of errors in this case ranged from modeling the application of fertilizer at the wrong time to applying too little fertilizer for a certain crop. If a new source or sink was added to the model after a potential source of error was found, the nutrient calibration process was restarted at the beginning in step A. A change in timing or rate of fertilizer application has the potential to alter the annual amount of uptake, denitrification, nitrification and leaching.

Once the land-based portion of the calibration had all of the loads in reasonable ranges and the general dynamics of the simulated concentrations followed the observed nutrient concentrations, the calibration process continued to step E of Figure 3. Here the simulated nutrient concentrations were compared to the observed nutrient concentrations. Based on this evaluation, the in-stream processes of nitrification and denitrification were adjusted for each reach to improve the goodness of fit. The calibration was completed after the rates of the in-stream processes were adjusted so that the goodness of fit results were reasonable.

The model results were evaluated based on visual comparison to the observed data, the R^2 value and the Nash–Sutcliffe (NS) efficiency. An NS efficiency greater than zero indicates that the model output is better than using the mean of the observed data (Nash & Sutcliffe 1970).

Model validation

A model validation period was used to evaluate the model predictions. The INCA-N input parameters that were set during calibration were tested for a different set of dates that had an adequate set of observed nutrient concentrations and flow rates in the same catchment. Visual inspection, the R^2 value and the NS efficiency were used to evaluate the validation results as for the calibration procedure. The INCA-N model was validated for the Yläneenjoki catchment for the period 1997–2002.

RESULTS AND DISCUSSION

Calibration phase 1: preliminary calibration

The preliminary calibration (Cal 1) is shown in Figure 4 and was developed by adding a fall manure application to the calibration completed by Lepistö et al. (2008).

Comparing the discrete water quality samples to Cal 1 in 2007, there are observed NO_3-N concentrations above 3 mg L^{-1} that were not adequately simulated. The discrete samples provided information about the nutrient concentrations for only a brief period of time. The continuous NO_3-N data in the fall show that the NO_3-N concentration was above 3 mg L^{-1} for longer than 2 weeks. Visual inspection showed that simulated results do not adequately capture the long period of elevated NO_3-N concentration, despite the addition of a fall manure application to the model.

This initial model calibration (Cal 1) was checked against published values of annual nitrogen process rates to make sure they were within the range of published values (Step B). These published values were obtained from primary scientific literature, producer surveys or reports of local agricultural production practices. Examples of the published data used in this calibration are listed in Table 2. The results for Cal 1 show that the nitrogen uptake rates for spring and winter cereals (111–121 kg N ha^{-1} a^{-1}) were modeled at the upper end of the reported range. It is unlikely that the crop yields were at the upper end of the range continuously for 6 years. The modeled mineralization rates for both spring and winter cereals (71–101 kg N ha^{-1} a^{-1}) were also too high.

Calibration phase 2: including soft data

Using the information provided by these soft data, a second calibration loop was performed (Cal 2). Figure 4 shows the result of lowering the total nitrogen uptake and mineralization rates for spring and winter cereals to levels that were within the range of published values to produce Cal 2. The predicted NO_3-N concentrations during the fall of 2007 decreased slightly in Cal 2 due to the reduction in mineralization rate. The biggest change from Cal 1 to Cal 2 was with respect to the peak NO_3-N concentration values predicted for the summer. One observed concentration in early June was 4.7 mg L^{-1}. Cal 1 showed a peak NO_3-N concentration during this time of 0.5 mg L^{-1}, while Cal 2 had a peak NO_3-N concentration of 2.6 mg L^{-1}, showing an improvement of the calibration for peak summer NO_3-N concentrations. Simulated increases in NO_3-N concentration during the summer were due to the reduction in the plant nitrogen uptake rates guided by soft data. In this portion of the calibration procedure, the process rates required further examination after each model run. It was clearly evident from the process rates in Cal 2 that the simulated denitrification rates (26–35 kg N ha^{-1} a^{-1}) were too high in spring and winter cereals when compared to the values reported in Table 2.

Groundwater denitrification is included in the process rates calculated in INCA-N, so the groundwater NO_3-N concentrations were inspected as a part of the next calibration loop. This incorporates both step B and step C in the calibration process. Once NO_3-N reaches the deeper groundwater, a lack of carbon often reduces the potential for denitrification to occur. Some soils have an organic layer that provides the carbon needed for denitrification (Ambus & Lowrance 1991; Hill et al. 2004) and this has repeatedly been shown to be true in riparian zone soils (Gurwick et al. 2008; Messer et al. 2012). In the INCA-N model, the water and nitrogen from the groundwater zone is modeled as flowing directly into the stream. This does not allow for the modeling of a separate riparian or carbon-rich area where groundwater denitrification is likely to occur, so this NO_3-N removal is incorporated into the groundwater zone in the model calibration.

The groundwater NO_3-N concentrations for all of the land uses from Cal 2 are shown in Figure 5. The initial groundwater concentrations are set for each subcatchment instead of each land use. Figure 5 shows an increase of groundwater concentration over the 6-year calibration period for the three land uses where large amounts of fertilizer were applied. The constant increase shows that the groundwater denitrification

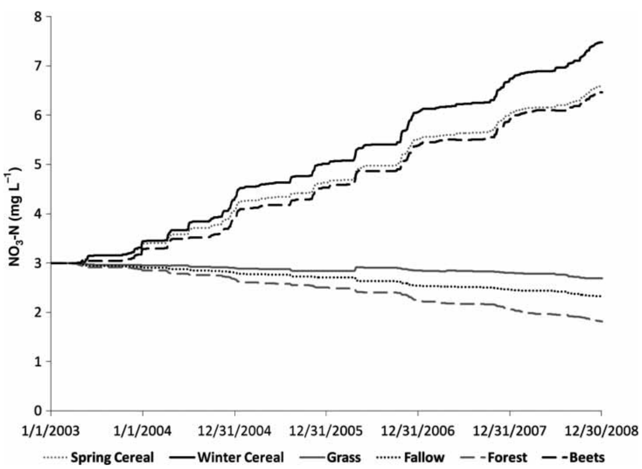

Figure 5 | INCA-N simulated groundwater NO_3-N concentrations from Calibration 2 for the Yläneenjoki catchment from 2003 through 2008.

rate was set too low. Due to the setup of the model, adjusting the groundwater processes at the subcatchment level is a process of balancing what is likely to occur for each land use.

As an example, consider the two different land uses of spring cereals and forests. If the initial NO_3-N concentration for the subcatchment is set too high, this will cause the forested area to have high NO_3-N leaching and a high rate of denitrification. If the NO_3-N concentration is set too low, the NO_3-N leaching from land planted in spring cereals will be too low and the denitrification rate in these soils will be lower than normal. There will also be a large increase in the spring cereal groundwater NO_3-N concentrations throughout the simulation period.

Calibration phase 3: denitrification and groundwater nitrate

The issue of denitrification rates being too high and the steady increase of groundwater NO_3-N concentrations in Cal 2 were addressed in Cal 3. One of the changes made in Cal 3 was adjusting the initial groundwater NO_3-N concentration and the groundwater denitrification rate to better balance the process loads expected in each land use. The resulting groundwater NO_3-N concentrations are shown in Figure 6. The initial NO_3-N concentration was adjusted from 3 to 1.75 mg L^{-1} to lower the simulated denitrification rates to reasonable values in the forested area. The higher initial NO_3-N concentration in Cal 2 caused a large amount of nitrogen to be removed through denitrification. With the adjustments in initial NO_3-N concentration and groundwater denitrification rates, the groundwater NO_3-N concentrations in the forest area approach a value near zero throughout the calibration period. The land uses that are heavily fertilized show a trend of increased groundwater concentrations, but the impact of denitrification can be seen in the seasonal dips in concentration that were not present in Cal 2 (Figure 5). It is clear that having to set the initial groundwater concentrations for each subcatchment has a large impact on the modeling results, especially in short-term modeling studies, as the groundwater concentrations change considerably during a modeling period of only 6 years (in this case $\pm$200%). Accurately simulating the amount of NO_3-N that flows from the groundwater to the stream is critical for correctly modeling stream water NO_3-N concentrations.

In addition to the changes made to the groundwater zone, the denitrification rates of spring and winter cereals (12–15 kg N ha^{-1} a^{-1}) were adjusted in the soil water zone in Cal 3 so that the process rates were within the range found in the literature. The change in model output from adjusting the land based denitrification rates is shown in Figure 7. The simulated peak NO_3-N concentrations increased from Cal 2 to Cal 3 from June 2007 through the end of the year. NO_3-N returns to base level concentrations (<0.5 mg L^{-1}) quicker in Cal 3 because there was less

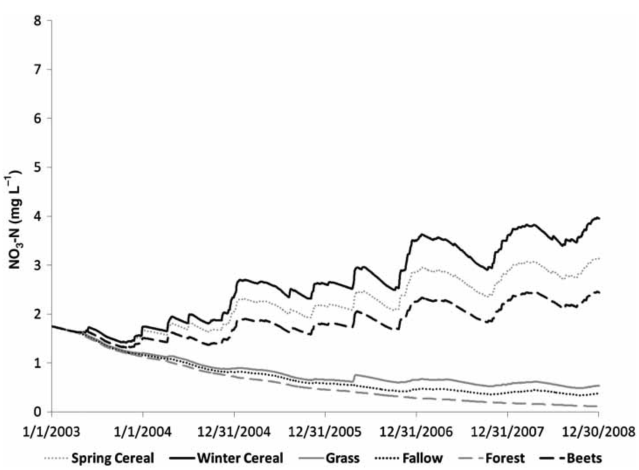

Figure 6 | INCA-N simulated groundwater NO_3-N concentrations from Calibration 3 for the Yläneenjoki catchment from 2003 through 2008. (*Note: y*-axis range of NO_3-N has been retained from Figure 5.)

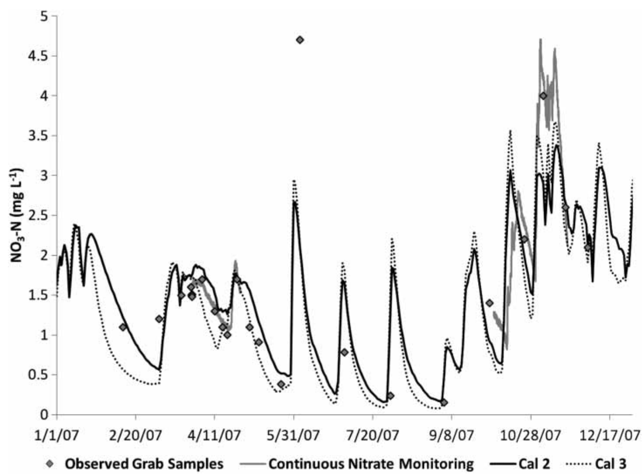

Figure 7 | INCA-N simulated NO_3-N concentrations from Calibration 2 and Calibration 3 compared to grab samples and continuous automatic monitoring for the Yläneenjoki catchment in 2007.

contribution from the groundwater sources as a result of the changes in this step of the calibration process. This is most evident in January and February 2007 where the NO_3-N concentration drops from over 2 mg L^{-1} to less than 0.6 mg L^{-1} in 22 days in Cal 3 instead of 35 days as it was in Cal 2.

Calibration phase 4: adjusting N process rates

The calibration process proceeded and adjustments were made to nitrogen process rates for all the different land uses to put them in the expected ranges, which produced Cal 4. These adjustments were made by following the same process (steps A–C) used for spring and winter cereals to produce Cal 2 and Cal 3. By visually inspecting the output, it was determined that the simulated NO_3-N concentration dynamics matched the changes in observed values reasonably well and step D was completed.

Final calibration

The calibration process continued to step E and the in-stream process rates were adjusted to produce the final calibration. The in-stream denitrification rate required little adjustment to provide the best fit, so the impact of adjusting in-stream process rates is better illustrated using NH_4-N. Figure 8 shows the reduction of in-stream NH_4-N concentrations from Cal 4 to the final calibration by increasing the in-stream nitrification rate.

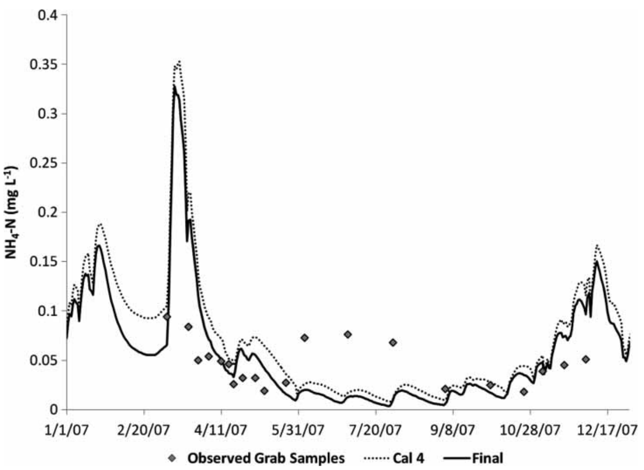

Figure 8 | INCA-N simulated NH_4-N concentrations from Calibration 4 and the final calibration compared to grab samples for the Yläneenjoki catchment in 2007.

The final NO_3-N calibration is shown in Figure 9 for 2007 and the final process parameters are in Table 3. The R^2 value for the entire calibration period for Cal 1 (0.35) increased to 0.45 for the final calibration. The NS efficiency increased from 0.33 to 0.42 from Cal 1 to the final

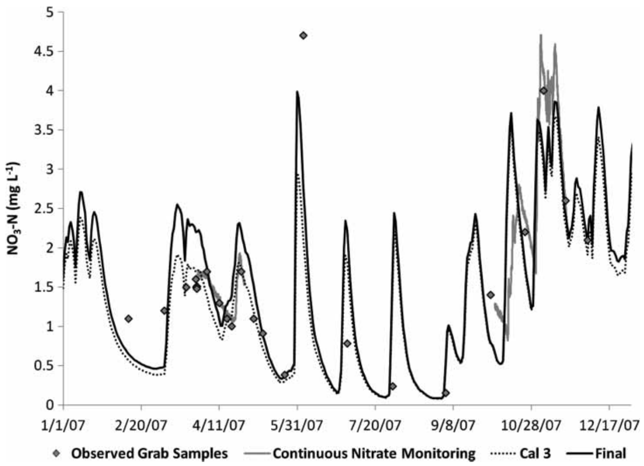

Figure 9 | INCA-N simulated NO_3-N concentrations from Calibration 3 and the final calibration compared to grab samples and continuous automatic monitoring for the Yläneenjoki catchment in 2007.

Table 3 | Nutrient process parameters in the final INCA-N calibration

Nutrient process	Land use	Parameter value
Soil water denitrification (m day^{-1})	Forest	0.0001
	Spring cereal	0.00025
	Winter cereal	0.0001
	Grass	0.0001
	Fallow	0.0001
	Beets	0.0003
Mineralization (kg N ha^{-1} day^{-1})	Forest	0.3
	Spring cereal	0.4
	Winter cereal	0.38
	Grass	0.41
	Fallow	0.3
	Beets	0.5
Nitrification (m day^{-1})	Forest	0.001
	Spring cereal	0.85
	Winter cereal	0.75
	Grass	0.95
	Fallow	0.6
	Beets	0.8
Groundwater denitrification (m day^{-1})	All	0.002
In-stream denitrification (day^{-1})	All	0.21
In-stream nitrification (day^{-1})	All	0.19

calibration. The major differences in the model output were the peak NO_3-N concentration values and how quickly the simulated NO_3-N concentrations decreased after a peak. The final calibration showed higher simulated NO_3-N concentrations during November 2007, but the concentrations were still lower than the observed values. The model output during April 2007 did not improve from Cal 1 to the final calibration. The Cal 1 output shows the NO_3-N concentrations between 0.1 and 0.2 mg L^{-1} above the observed data. The final calibration shows peak concentrations 0.4 mg L^{-1} above the observed peak concentrations, but during that period some concentrations are lower than the observed data. The final model simulation does show a higher peak NO_3-N concentration in June, when the observed concentration was 4.7 mg L^{-1}.

The results of the final calibration for 2003–2008 are shown in Figure 10. The R^2 and NS efficiency comparing the observed and simulated discharge were 0.69 and 0.67, respectively, for the whole calibration period. As stated earlier, the flow parameters were not altered during the calibration process described in this work. The timing of the observed discharge peaks was simulated well, but many of the observed peaks were higher than simulated flows. The variations between the model and the observed values could be caused by the inputs from WSFS not accurately capturing the spatial variability of precipitation in the catchment, or because the spatial heterogeneity of the catchment soils was not simulated in this semi-distributed model.

The results of the NO_3-N calibration for the entire period are shown in Figure 10 ($R^2 = 0.45$; NS $= 0.42$). The model simulated the low observed NO_3-N concentrations in summer well. In 2005, the modeled NO_3-N concentration does not drop to the low summer levels as quickly as the observed concentrations. The modeled maximum NO_3-N concentrations were also lower than the highest observed concentrations during many periods, primarily during the winter months. These missed peaks could have been caused by modeled denitrification rates that were too high in the winter.

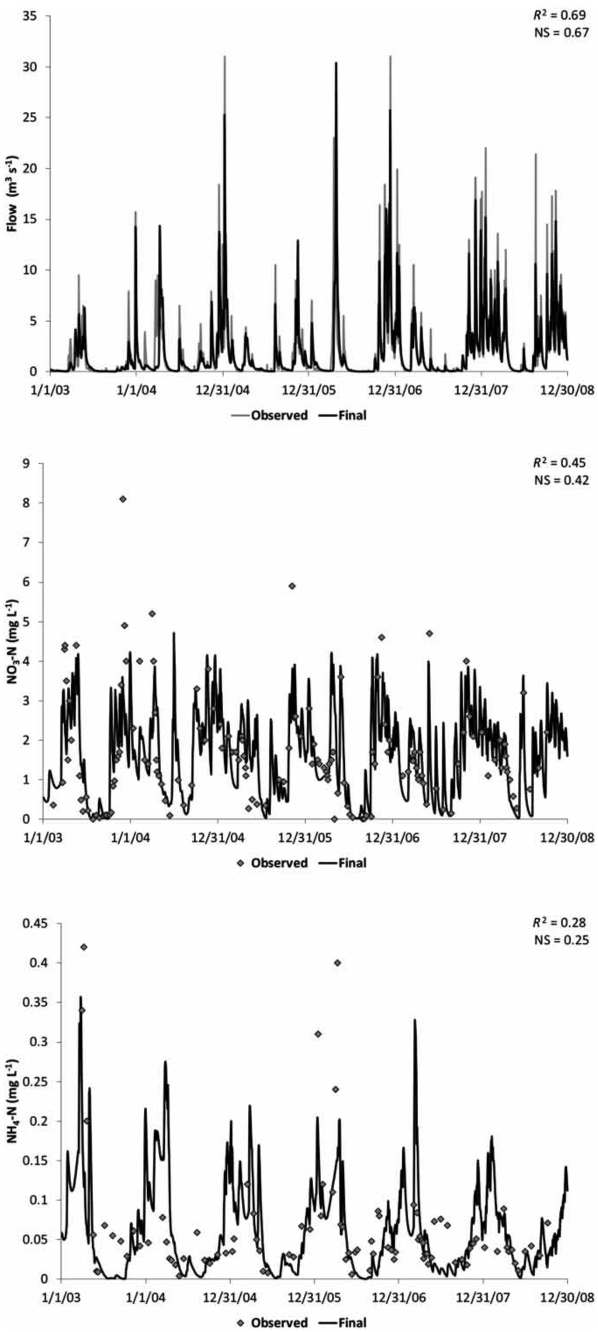

Figure 10 | Observed and simulated discharge, NO_3-N and NH_4-N concentrations at the outlet of the Yläneenjoki catchment as produced by the final calibration (2003–2008).

The shape of the simulated NH_4-N curve fit the observed data well, based on visual inspection. The R^2 value for NH_4-N was 0.28 and the NS efficiency was 0.25. The observed NH_4-N concentrations above 0.1 mg L^{-1} are simulated well except in late 2005 and early 2006. This could be caused by the simulated flow being above the observed flow and diluting the NH_4-N concentrations simulated during this time period. During the summer periods, the simulated NH_4-N concentrations dropped below 0.05 mg L^{-1} for long periods of time, while the observed concentrations are often above 0.05 mg L^{-1}. This is especially noticeable in 2003 and 2007. It is possible that a point source of NH_4-N that existed in the catchment was not simulated in the model. The high concentrations from a point source would not be diluted during the summer low-flow periods. The impact of a small point source may be more difficult to see during periods of higher flow.

Model validation

The model was validated for the years 1997–2002 and the model output is shown in Figure 11. The flow validation provided good results with $R^2 = 0.80$ and NS $= 0.79$. Although the goodness-of-fit statistics show good results, a visual inspection of the results show that most of the observed flow peaks were underestimated in the model. The timing of these peaks is modeled well, so the underestimation may have been caused by the *a* and *b* flow constants (responsible for relating discharge to mean flow velocity) being incorrect. It is also possible that the daily time series model inputs from WSFS could cause these errors. WSFS is used for flood prediction in Finland, so it is constantly being re-calibrated to most accurately model the most recent data. The inputs for this work were retrieved from WSFS in 2010, so the daily time series inputs for the validation period may not have been as accurate as they were for the calibration period.

The NO_3-N concentration results during the validation period were not as good as the flow results. The NO_3-N validation produced an R^2 of 0.34 and an NS efficiency below 0. These results and a visual inspection indicate that the timing of peak concentrations was simulated well, but that overall the NO_3-N concentrations were too high. The higher NO_3-N concentrations were most obvious during each spring. The overestimation occurred during the spring snow melt before any fertilizer was applied to the fields. Most of the nitrogen process loads were within the expected range of values, with the exception of the leaching loads. The amount of nitrogen lost to leaching was higher than expected. The rates of denitrification were lower than they were in the calibration period, but were still reasonable.

These results suggest that during the calibration period the simulated denitrification rates were high enough to prevent excess leaching from occurring. To further improve the model, the soil water NO_3-N concentrations and the minimum temperature at which nitrogen processes are allowed should be investigated further.

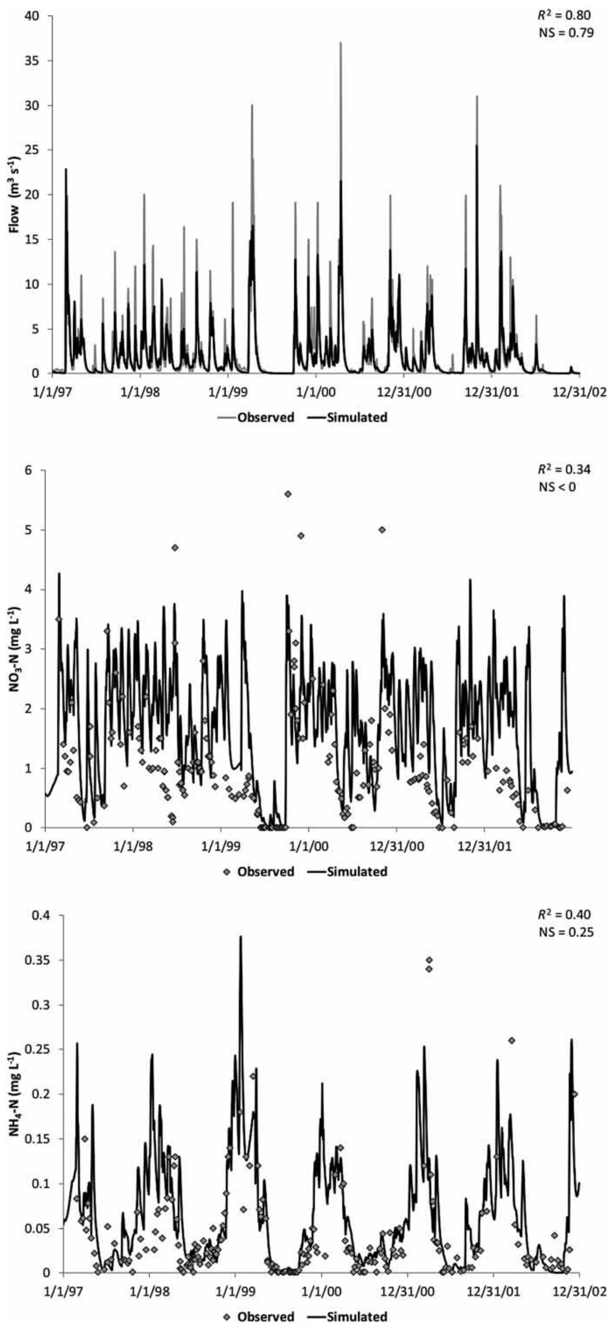

Figure 11 | Observed and simulated discharge, NO_3-N and NH_4-N concentrations at the outlet of the Yläneenjoki catchment from the validation period (1997–2002).

The model performed well during the validation period for simulating NH_4-N concentrations with an R^2 of 0.40 and an NS efficiency of 0.25. The timing of peaks in observed NH_4-N concentrations was simulated well. The modeled NH_4-N concentrations are above the observed concentrations from late 1997 to the summer of 1998. There are a few observed NH_4-N concentrations above 0.25 mg L^{-1} that are not simulated well near the end of the validation period. The simulated summer NH_4-N concentrations are a mix of overestimation and underestimation of observed values. The model does a better job of modeling the summer NH_4-N concentrations during the validation period than it does during the calibration period.

The two short periods of continuous water quality data provided a glimpse into the usefulness of continuous water quality monitoring for calibration of nutrient models. Collection of continuous records provides more insight into watershed hydrology and nutrient transformations than that provided by discrete sampling (Kirchner *et al.* 2000). The continuous data in this study showed that a significant export of NO_3-N was missed in the initial simulations.

Long-term continuous water quality monitoring may prove useful for reducing parameter uncertainty. In this study, however, continuous data were not available over a long enough period of time to test its impact on constraining the model. Raat *et al.* (2004) found that continuous data may not prove useful enough to warrant the effort and expense of its collection, based on simulations in a virtual catchment. The effort required to collect continuous data is being reduced through the introduction of new technology, so the collection of this high-frequency data should be encouraged in future modeling studies.

The structure of the INCA-N model makes it possible for two wrongs to result in a numerically correct answer. For instance, if mineralization rates in the model were set too high, higher simulated denitrification rates could remove excess nitrogen resulting in reasonable NO_3-N concentrations during the calibration period. In this case study, published literature values and other soft data were used to restrict nitrogen process rates to a reasonable range of values. Comparing the initial calibration to the final calibration shows an improvement in the goodness-of-fit statistics, but does not show a visual improvement in the results over the whole calibration period. A comparison of Cal 1 and the final calibration results during April 2007 does not show improvement in the simulation of in-stream NO_3-N concentrations. However, the nitrogen process rates in Cal 1 were not within the range of published values and were not as accurate in representing the processes actually occurring in the catchment as the final calibration. Using soft data to constrain process rates does not always lead to better goodness-of-fit results as it did in this case, but better goodness of fit should be sacrificed for a model that produces a more accurate representation of the system (Seibert & McDonnell 2002; Rankinen *et al.* 2006) and the potential that future simulations outside the calibration period may be more accurate.

Sensitivity analysis provides a method of determining what soft data are most important for model calibration. Rankinen *et al.* (2013) showed that, for simulations in an agriculturally dominated catchment in Finland, the INCA-N model was most sensitive to parameters that altered the nutrient process rates in the primary agricultural land use in the catchment. In the Yläneenjoki catchment, further research in quantifying the nitrogen process rates in spring cereals would likely be the most useful experimentation for further improving the model calibration. The sensitivity of the parameters change between watersheds with some of the variation being caused by differences in the land use and the relative importance of groundwater and surface water (McIntyre *et al.* 2005; Rankinen *et al.* 2006; Futter *et al.* 2009). To better guide modeling and data collection efforts, a sensitivity analysis should be conducted following initial calibration. The results of the sensitivity analysis could direct resources to areas that need more research or further constraints on the model parameters.

Restrictions on financial and time resources often prevent modelers from conducting experiments to measure the various nitrogen process rates that are simulated in the INCA-N model. Process rates for biogeochemical processes, such as mineralization, denitrification and leaching, are often available in a similar geographic region that can be used as soft data. The results of field experiments to measure biogeochemical process rates will generally have such high variability due to the heterogeneity of soils that using limited resources to measure these rates is not recommended if the sole purpose is for use in calibration of the INCA-N model.

With regards to biogeochemical processes, effort should be expended in finding the previous studies with the most similar climate, hydrologic regime and soil type that use a quality method of measuring the process rate. When resources for collection of soft data are limited, it is recommended that the modeler focus on gaining information on the local agricultural practices and crop yields. Variances in agricultural practices, such as using crop residues as animal feed instead of leaving them on the field surface, can have a large impact on nutrient loads in a catchment and can vary widely within a region (Lagzdins *et al.* 2012). Information on crop yields, the use of animal manure and fertilization rates is unlikely to be available in the scientific literature, but may be available from government publications. This local information that cannot be obtained from the scientific literature is where modelers should focus their resources for collection of soft data.

The results of the model validation show that the use of published process rates to assist in calibrating the model does not guarantee that the calibration is a completely accurate depiction of the processes occurring in the catchment. Despite the in-stream NO_3-N concentrations being modeled reasonably well in the calibration period, the simulated concentrations were above the observed concentrations in the validation period. The timing of major changes in the concentrations was simulated well, but the value of the simulated concentration was always too high. This result points to at least one nitrogen process being modeled incorrectly. It is possible that more information could be gained through validation of this calibration over a period of time that had different conditions in climate, land use or agricultural practices (Kirchner 2006). To further increase the accuracy of semi-distributed model in heterogeneous catchments, long-term calibrations are recommended (Wade *et al.* 2008). Long-term simulations will allow for the impact of changes in climate, agricultural practices and land use to be observed in the calibration phase of the model. These changes are not as drastic in short-term modeling periods and cannot be adjusted for in the calibration. The challenge to completing long-term simulations is the availability of long-term observed data to compare to the results of the model.

CONCLUSIONS

The results of this case study show that the accuracy of the INCA-N representation of a catchment and its internal processes can be improved through the use of soft data to constrain the nitrogen process loads to reasonable values. The modeler should also pay attention to model output other than the in-stream nitrogen concentrations to improve the accuracy of the model, as was illustrated through the examination of groundwater NO_3-N concentrations. A thorough examination of all model outputs and the use of soft data to constrain nitrogen process rates does not guarantee that the resulting calibration is a completely accurate depiction of the modeled catchment. Further work should be done to increase the confidence in model calibrations; modeling scenarios of climate, land use and management changes should proceed with the knowledge of potential sources of uncertainty within the INCA-N model.

ACKNOWLEDGEMENTS

The authors would like to thank the two anonymous reviewers for their helpful comments. This material is based upon work supported by the National Science Foundation under Grant No. DGE-0750733 and by the EU REFRESH project (FP7-ENV-2009-1/244121).

REFERENCES

Ambus, P. & Lowrance, R. 1991 Comparison of denitrification in two riparian soils. *Soil Science Society of America Journal* **55** (4), 994–997.

Bärlund, I., Rankinen, K., Järvinen, M., Huitu, E., Veijalainen, N. & Arvola, L. 2009 Three approaches to estimate inorganic nitrogen loading under varying climatic conditions from a headwater catchment in Finland. *Hydrology Research* **40** (2–3), 167–176.

Barton, L., McLay, C. D. A., Schipper, L. A. & Smith, C. T. 1999 Annual denitrification rates in agricultural and forest soils: a review. *Soil Research* **37** (6), 1073–1094.

Cherry, K. A., Shepherd, M., Withers, P. J. A. & Mooney, S. J. 2008 Assessing the effectiveness of actions to mitigate nutrient loss from agriculture: A review of methods. *Science of the Total Environment* **406** (1–2), 1–23.

EEC 1991 Council Directive 91/676/EEC of 12 December 1991 concerning the protection of waters against pollution caused by nitrates from agricultural sources. *Official Journal L* **375** (31), 1–8.

Futter, M., Helliwell, R., Hutchins, M., Aherne, J. & Whitehead, P. 2009 Modelling the effects of changing climate and nitrogen

deposition on nitrate dynamics in a Scottish mountain catchment. *Hydrology Research* **40** (2–3), 153–166.
Galloway, J. N., Townsend, A. R., Erisman, J. W., Bekunda, M., Cai, Z., Freney, J. R., Martinelli, L. A., Seitzinger, S. P. & Sutton, M. A. 2008 Transformation of the nitrogen cycle: Recent trends, questions, and potential solutions. *Science* **320** (5878), 889–892.
Granlund, K., Rankinen, K. & Lepistö, A. 2004 Testing the INCA model in a small agricultural catchment in southern Finland. *Hydrology and Earth System Sciences* **8** (4), 717–728.
Gurwick, N. P., Groffman, P. M., Yavitt, J. B., Gold, A. J., Blazejewski, G. & Stolt, M. 2008 Microbially available carbon in buried riparian soils in a glaciated landscape. *Soil Biology & Biochemistry* **40** (1), 85–96.
HELCOM 2011 The Fifth Baltic Sea Pollution Load Compilation (PLC-5). In *Baltic Sea Environment Proceedings No. 128. Helsinki Commission*. Helsinki.
Heng, H. H. & Nikolaidis, N. P. 1998 Modeling of nonpoint source pollution of nitrogen at the watershed scale. *Journal of American Water Resources Association* **34** (2), 359–374.
Hill, A. R., Vidon, P. G. & Langat, J. 2004 Denitrification potential in relation to lithology in five headwater riparian zones. *Journal of Environmental Quality* **33** (3), 911–919.
Howarth, R. W., Sharpley, A. & Walker, D. 2002 Sources of nutrient pollution to coastal waters in the United States: Implications for achieving coastal water quality goals. *Estuaries* **25** (4B), 656–676.
Hyvärinen, V. (ed.) 1999 *Hydrological Yearbook 1995. The Finnish Environment 280 ed.* Finnish Environment Institute, Helsinki.
Information Centre of the Ministry of Agriculture and Forestry 2007 *Yearbook of Farm Statistics 2006.* Information Centre of the Ministry of Agriculture and Forestry, Helsinki.
Jarvie, H. P., Wade, A. J., Butterfield, D., Whitehead, P. G., Tindall, C. I., Virtue, W. A., Dryburgh, W. & McGraw, A. 2002 Modelling nitrogen dynamics and distributions in the River Tweed, Scotland: an application of the INCA model. *Hydrology and Earth System Sciences* **6** (3), 433–453.
Kirchner, J. W. 2006 Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research* **42** (3), W03S04.
Kirchner, J. W., Feng, X. H. & Neal, C. 2000 Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature* **403** (6769), 524–527.
Koskiaho, J., Lepistö, A., Tattari, S. & Kirkkala, T. 2010 On-line measurements provide more accurate estimates of nutrient loading: a case of the Ylä neenjoki river basin, southwest Finland. *Water Science and Technology* **62** (1), 115–122.
Lagzdins, A., Jansons, V., Sudars, R. & Abramenko, K. 2012 Scale issues for assessment of nutrient leaching from agricultural land in Latvia. *Hydrology Research* **43** (4), 383–399.
Lepistö, A., Granlund, K., Kortelainen, P. & Räike, A. 2006 Nitrogen in river basins: Sources, retention in the surface waters and peatlands, and fluxes to estuaries in Finland. *Science of the Total Environment* **365** (1), 238–259.
Lepistö, A., Huttula, T., Bärlund, I., Granlund, K., Härmä, P., Kallio, K., Kiirikki, M., Kirkkala, T., Koponen, S., Koskiaho, J., Kotamäki, N., Lindfors, A., Malve, O., Pyhälahti, T., Tattari, S. & Törmä, M. 2008 New measurement technology, modelling and remote sensing in the Säkylän Pyhäjärvi area – Catchlake. *Reports of Finnish Environment Institute* **15**, 13–43.
Lunn, R., Adams, R., Mackay, R. & Dunn, S. 1996 Development and application of a nitrogen modelling system for large catchments. *Journal of Hydrology* **174** (3), 285–304.
Mattila, H., Kirkkala, T., Salomaa, E., Sarvala, J. & Haliseva-Soila, M. (eds) 2001 *Pyhäjärvi. Pyhäjärvi-instituutin julkaisuja 26 ed.* Pyhäjärven suojelurahasto.
McIntyre, N., Jackson, B., Wade, A. J., Butterfield, D. & Wheater, H. S. 2005 Sensitivity analysis of a catchment-scale nitrogen model. *Journal of Hydrology* **315** (1–4), 71–92.
Messer, T. L., Burchell II, M. R., Grabow, G. L. & Osmond, D. L. 2012 Groundwater nitrate reductions within upstream and downstream sections of a riparian buffer. *Ecological Engineering* **47**, 297–307.
Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models part I – a discussion of principles. *Journal of Hydrology* **10** (3), 282–290.
Raat, K. J., Vrugt, J. A., Bouten, W. & Tietema, A. 2004 Towards reduced uncertainty in catchment nitrogen modelling: quantifying the effect of field observation uncertainty on model calibration. *Hydrology and Earth System Sciences* **8** (4), 751–763.
Räike, A., Pietiläinen, O. P., Rekolainen, S., Kauppila, P., Pitkänen, H., Niemi, J., Raateland, A. & Vuorenmaa, J. 2003 Trends of phosphorus, nitrogen and chlorophyll *a* concentrations in Finnish rivers and lakes in 1975–2000. *Science of the Total Environment* **310** (1–3), 47–59.
Rankinen, K., Granlund, K., Futter, M. N., Butterfield, D., Wade, A. J., Skeffington, R., Arvola, L., Veijalainen, N., Huttunen, I. & Lepistö, A. 2013 Controls on inorganic nitrogen leaching from Finnish catchments assessed using a sensitivity and uncertainty analysis of the INCA-N model. *Boreal Environment Research* **18**. http://www.borenv.net/BER/pdfs/preprints/Rankinen.pdf.
Rankinen, K., Lepistö, A. & Granlund, K. 2002 Hydrological application of the INCA model with varying spatial resolution and nitrogen dynamics in a northern river basin. *Hydrology and Earth System Sciences* **6** (3), 339–350.
Rankinen, K., Karvonen, T. & Butterfield, D. 2006 An application of the GLUE methodology for estimating the parameters of the INCA-N model. *Science of the Total Environment* **365** (1–3), 123–139.
Rankinen, K., Salo, T., Granlund, K. & Rita, H. 2007 Simulated nitrogen leaching, nitrogen mass field balances and their correlation on four farms in south-western Finland during the period 2000–2005. *Agricultural and Food Science* **16**, 387–408.

Rankinen, K., Valpasvuo-Jaatinen, P., Karhunen, A., Kenttamies, K., Nenonen, S. & Bärlund, I. 2009 Simulated nitrogen leaching patterns and adaptation to climate change in two Finnish river basins with contrasting land use and climatic conditions. *Hydrology Research* **40** (2–3), 177–186.

Refsgaard, J. C. 1996 Chapter 2: Terminology, modelling protocol and classification of hydrological model codes. In: *Distributed Hydrological Modelling* (M. B. Abbott & J. C. Refsgaard, eds). Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 17–39.

Salo, T. & Turtola, E. 2006 Nitrogen balance as an indicator of nitrogen leaching in Finland. *Agriculture, Ecosystems & Environment* **113** (1–4), 98–107.

Santhi, C., Arnold, J. G., Williams, J. R., Dugas, W. A., Srinivasan, R. & Hauck, L. M. 2001 Validation of the SWAT model on a large river basin with point and nonpoint sources. *Journal of the American Water Resources Association* **37** (5), 1169–1188.

Seibert, J. & McDonnell, J. J. 2002 On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research* **38** (11), 1241.

Svensson, B. H., Klemedtsson, L., Simkins, S., Paustian, K. & Rosswall, T. 1991 Soil denitrification in three cropping systems characterized by differences in nitrogen and carbon supply. *Plant and Soil* **138** (2), 257–271.

Townsend, A. R., Howarth, R. W., Bazzaz, F. A., Booth, M. S., Cleveland, C. C., Collinge, S. K., Dobson, A. P., Epstein, P. R., Keeney, D. R., Mallin, M. A., Rogers, C. A., Wayne, P. & Wolfe, A. H. 2003 Human health effects of a changing global nitrogen cycle. *Frontiers in Ecology and the Environment* **1** (5), 240–246.

Vehviläinen, B. & Huttunen, M. 2010 *Hydrological forecasting and real time monitoring in Finland: the Watershed Simulation and Forecasting System (WSFS)*. Finnish Environment Institute. Available at: http://www.ymparisto.fi/download.asp?contentid=115457&lan=en.

Vitousek, P. M., Aber, J. D., Howarth, R. W., Likens, G. E., Matson, P. A., Schindler, D. W., Schlesinger, W. H. & Tilman, D. G. 1997 Human alteration of the global nitrogen cycle: Sources and consequences. *Ecological Applications* **7** (3), 737–750.

Vuorenmaa, J., Rekolainen, S., Lepistö, A., Kenttämies, K. & Kauppila, P. 2002 Losses of nitrogen and phosphorus from agricultural and forest areas in Finland during the 1980s and 1990s. *Environmental Monitoring & Assessment* **76** (2), 213–248.

Wade, A. J., Butterfield, D. & Whitehead, P. G. 2006 Towards an improved understanding of the nitrate dynamics in lowland, permeable river-systems: Applications of INCA-N. *Journal of Hydrology* **330** (1–2), 185–203.

Wade, A. J., Durand, P., Beaujouan, V., Wessel, W. W., Raat, K. J., Whitehead, P. G., Butterfield, D., Rankinen, K. & Lepistö, A. 2002 A nitrogen model for European catchments: INCA, new model structure and equations. *Hydrology and Earth System Sciences* **6** (3), 559–582.

Wade, A. J., Jackson, B. M. & Butterfield, D. 2008 Over-parameterised, uncertain 'mathematical marionettes' – How can we best use catchment water quality models? An example of an 80-year catchment-scale nutrient balance. *Science of the Total Environment* **400** (1–3), 52–74.

Ward, M. H., Mark, S. D., Cantor, K. P., Weisenburger, D. D., Correa-Villasenor, A. & Zahm, S. H. 1996 Drinking water nitrate and the risk of non-Hodgkin's lymphoma. *Epidemiology* **7** (5), 465–471.

Whitehead, P. G., Wilson, E. J. & Butterfield, D. 1998a A semi-distributed Integrated Nitrogen model for multiple source assessment in Catchments (INCA). Part I: model structure and process equations. *Science of the Total Environment* **210** (1–6), 547–558.

Whitehead, P. G., Wilson, E. J., Butterfield, D. & Seed, K. 1998b A semi-distributed integrated flow and nitrogen model for multiple source assessment in catchments (INCA). Part I: application to large river basins in south Wales and eastern England. *Science of the Total Environment* **210–211** (0), 559–583.

Winsemius, H., Schaefli, B., Montanari, A. & Savenije, H. 2009 On the calibration of hydrological models in ungauged basins: A framework for integrating hard and soft hydrological information. *Water Resources Research* **45** (12), W12422.

First received 13 February 2013; accepted in revised form 5 May 2013. Available online 25 June 2013