

# Application of the Bayesian calibration methodology for the parameter estimation in CoupModel

Y. Conrad and N. Fohrer

Department of Hydrology and Water Resources Management, Ecology Centre at the University of Kiel, Kiel, Germany

Received: 15 January 2009 – Revised: 13 March 2009 – Accepted: 28 April 2009 – Published: 10 August 2009

**Abstract.** This study provides results for the optimization strategy of highly parameterized models, especially with a high number of unknown input parameters and joint problems in terms of sufficient parameter space. Consequently, the uncertainty in model parameterization and measurements must be considered when highly variable nitrogen losses, e.g. N leaching, are to be predicted. The Bayesian calibration methodology was used to investigate the parameter uncertainty of the process-based CoupModel. Bayesian methods link prior probability distributions of input parameters to likelihood estimates of the simulation results by comparison with measured values. The uncertainty in the updated posterior parameters can be used to conduct an uncertainty analysis of the model output. A number of 24 model variables were optimized during 20 000 simulations to find the “optimum” value for each parameter. The likelihood was computed by comparing simulation results with observed values of 23 output variables including soil water contents, soil temperatures, groundwater level, soil mineral nitrogen, nitrate concentrations below the root zone, denitrification and harvested carbon from grassland plots in Northern Germany for the period 1997–2002. The posterior parameter space was sampled with the Markov Chain Monte Carlo approach to obtain plot-specific posterior parameter distributions for each system. Posterior distributions of the parameters narrowed down in the accepted runs, thus uncertainty decreased. Results from the single-plot optimization showed a plausible reproduction of soil temperatures, soil water contents and water tensions in different soil depths for both systems. The model performed better for these abiotic system properties compared to the results for harvested carbon and soil mineral nitrogen dynamics. The high variability in modeled nitrogen leaching showed that the soil nitrogen conditions are highly

uncertain associated with low modeling efficiencies. Simulated nitrate leaching was compared to more general, site-specific estimations, indicating a higher leaching during the seepage periods for both simulated grassland systems.

## 1 Introduction

The prediction of nutrient losses under agricultural land use is an important factor for the economic and ecologic evaluation of specific farming systems. The movement of nutrients and pollutants, which can be identical, namely nitrogen (N), but different in their load/concentration, has been concerned with various environmental effects (Lewis et al., 2003). Especially for intensive grassland systems, nitrogen leaching can produce unfavorable risks for surface water and near-surface groundwater. Various methods to quantify the complex interactions between components of the nitrogen balance have been developed, where extensive measurements would be too expensive and difficult (Jovanovic et al., 2008). General estimations of nitrogen losses are mostly based on simple mass balances at field or farm-gate scale. These N budgets measure or estimate the inputs and outputs of nutrients without detailed measurements of losses such as leaching, denitrification and volatilization (OECD, 2001). Such pure behavior imitation of “black box” approaches is necessary for the operational application of models at catchment scale, even if extrapolation purposes in space and time are limited (Casper, 2002).

Simulation models represent a modern alternative to observations. Greater understanding of nitrogen dynamics in soils at field scale can potentially improve agricultural practices regarding N use efficiency and minimizing pollution. Detailed process knowledge is necessary to evaluate model results. But point measurements or process observations are less useful for optimization and evaluation of models with increasing degree of model abstraction or higher spatial



Correspondence to: Y. Conrad  
(yconrad@hydrology-uni-kiel.de)

scale (Casper, 2002). Only robust parameterization techniques allow for plausible explanations of system behavior in physically based models. Due to complex transformations of soil N and carbon, many models express this complexity by a high number of input parameters, which implicates the need for a careful model optimization (Pappenberger and Beven, 2006). Until recently, calibration of highly parameterized models was performed by an intensive sensitivity analysis and fitting modeling results to measurements by “trial and error” procedure until observed values were reproduced well. Several procedures such as single, multiple or sequential parameter calibration procedures to optimize process-based models were tested, but the uncertainty in model input parameters and observations were not taken into account (Reinds et al., 2008). Consequently, only that parameter setting resulting in the best agreement between model and observations was accepted. But several combinations of input parameters may give the same model result, which makes it difficult to define a unique set of input parameters. Optimization methods that include uncertainties in model and measurements are thus to be preferred over “single best fit” methods without uncertainty assessment.

The Bayesian calibration methods have been used for optimization of forest ecosystems (Van Oijen et al., 2005; Klemetsson et al., 2008) or watershed models (Vrugt et al., 2006). Bayesian approaches include probability distributions of model input parameters, based on prior assumptions about their magnitude and uncertainty, combined with likelihood estimates of the model results by comparison with observations for model output variables. Consequently, parameter uncertainty can be quantified by this combined information. The updated parameter uncertainty can be used to analyze model output uncertainty. Reinds et al. (2008) defined the Bayesian calibration as a twofold extension of Maximum Likelihood estimation including prior information about input parameters and indentifying a single parameter vector with maximum probability and its uncertainty estimate. The Bayesian calibration technique was applied in this study to optimize the CoupModel (Jansson and Karlberg, 2004) on two different grassland systems in Northern Germany. Data on soil temperatures, soil water contents, groundwater (= saturation) level, nitrate ( $\text{NO}_3$ ) nitrogen concentration, soil nitrogen contents, denitrification and harvested carbon were available (Wachendorf et al., 2004; Herrmann, 2005; Lampe, 2005). Single-plot optimizations were carried out for 24 input parameters to investigate effect of parameter space and start value for each grassland system. The aim of this study was to test the applicability and usefulness of the Bayesian optimization technique for the parameterization of grassland systems. Obtained results could be helpful for parameterization of CoupModel in terms of the complex and uncertain soil N fluxes.

## 2 Materials and methods

### 2.1 Site description and measurements

Measured data were provided by the integrated project “The nitrogen project: A system approach to optimize nitrogen use efficiency on the dairy farm” located at the experimental farm “Karkendamm” in Northern Germany and carried out between 1997 and 2003 (Taube and Wachendorf, 2001). Multifactorial field experiments were conducted to investigate crop quality, soil nitrogen balance and groundwater quality. The climate at Karkendamm is maritime temperate with a mean annual temperature of  $8.6^\circ\text{C}$  and a mean precipitation of  $865\text{ mm year}^{-1}$  (Herrmann et al., 2005). Investigated grassland plots were dominated by perennial ryegrass (*Lolium perenne* L.) with up to four harvests per year. Two fertilization levels were considered: non-fertilized (N0) and highly fertilized (N300;  $300\text{ kg mineral N ha}^{-1}\text{ year}^{-1}$ ). Dominating soil types are Podzols (FAO, 2006) with low nutrient storage capacity and high leaching potential. The original soil profile was deep-ploughed in 1981 to improve the hydraulic conductivity leading to slanted soil layers between 0.27 and 0.80 m (Scholz, 1999). Sand contents varied between 85.7% and 93.0% up to a depth of one meter. Total organic carbon was highest in the upper soil profile with 3.8–5.6%, just as the total N with 0.21–0.27%. Available measurements are listed in Table 1 and were used to assess the CoupModel performance by comparing these observations with corresponding model results. In addition to measured soil water contents observed with TDR sensors and gravimetric technique, averaged soil water tensions in 30, 50 and 70 cm depth of the mown grassland were used to validate soil water conditions. Measurements were taken with puncture tensiometers with a rubber septum at the top during summer 1998 and 1999. Further information about data acquisition can be derived from Conrad and Fohrer (2009b). Observed soil solution concentrations can vary considerably, and measurement uncertainty is affected most likely by the spatial variability within the plot (Reinds et al., 2008). De Vries et al. (1999) reported a spatial variability of measured soil solution concentrations between 20–60% depending on depth and type of ion in forest soils from the Netherlands. In Reinds et al. (2008), an uncertainty of 30% was used for major ions at 40–60 cm depth for the Dutch forest plots. For low concentrations that are often measured for  $\text{NO}_3$  or aluminium, the measurement uncertainty is probably higher than 30%. Harmel et al. (2006) presented literature data about the uncertainty in sampling, preservation and analysis of solved nitrogen species in water samples. The error of solved  $\text{NO}_3$  nitrogen varied between –47% and –14%, when samples were refrigerated and analyzed within 54 h. The uncertainty in the laboratory analysis can range from  $\pm 75\%$  to 400% for  $\text{NO}_3$ , while the colorimetric technique showed error ranges from –4% to 9%. In this study, we assumed a relative error of  $\pm 20\%$  for the measured variables except for the  $\text{NO}_3$  nitrogen concentration in

**Table 1.** Available measurements at the Karkendamm site (Wachendorf et al., 2004), used for stochastic optimization.

Variable	Depth (m)	Measuring period	Numbers of samples
Soil temperature (°C)	0.05, 0.10, 0.15	1997–2002	1648
Soil water content (Vol.%)	0.10, 0.30, 0.40, 0.50, 0.60, 0.70, 0.80	1998–2002	20–198
Groundwater level (m)		1997–2002	146
NO <sub>3</sub> N concentration (mg N l <sup>-1</sup> )	0.60	1997–2002	995 (N0), 952 (N300)
Harvested carbon (g C m <sup>-2</sup> )	above ground	1997–1999	3 (N0), 0 (N300) <sup>a</sup>
Soil mineral N (Nmin) (g N m <sup>-2</sup> )	0–0.30, 0–0.60, 0–0.90	1999–2002	11–16
Soil NO <sub>3</sub> N (g N m <sup>-2</sup> )	0–0.30, 0–0.60, 0–0.90	1999–2002	11–17
Soil NH <sub>4</sub> N (g N m <sup>-2</sup> )	0–0.30, 0–0.60, 0–0.90	1997–2002	19–24
Denitrification (g N m <sup>-2</sup> )	non-fertilized plot	Apr–Jul 2001	32

<sup>a</sup> N0: additional estimates for 2000 and 2001; N300: estimated from harvested clover-grass (highly fertilized)

the non-fertilized grassland (relative error:  $\pm 50\%$ ). An absolute error was assumed for the soil temperatures with  $\pm 1^\circ\text{C}$  according to manufacturer's information on the temperature sensors and for the groundwater level with  $\pm 0.05$  m.

The NO<sub>3</sub> nitrogen leaching below the rooting zone (>60 cm) in the experimental data set from Karkendamm was not measured directly but calculated as a product of the averaged NO<sub>3</sub> nitrogen concentration and an estimated seepage water amount according to the climatic water balance equation (DVWK, 1996). This approach presented by Büchter (2003) is simple but best practice when water balance conditions can only be estimated and was called further model “Büchter”. In our study, results of CoupModel and model “Büchter” were compared regarding the calculations for seepage water amount and NO<sub>3</sub> nitrogen leaching.

## 2.2 CoupModel setup

CoupModel (Jansson and Karlberg, 2004) is an ecosystem process model used to calculate coupled heat, water, carbon and nitrogen fluxes under one-dimensional, unsaturated soil conditions. Several plant covers can be defined above the horizontal layered soil profile, where the Richard's equation is solved for water flow, and the Fourier's law of diffusion is used for heat fluxes. Lower boundary condition can be defined as free drainage or saturated. Potential transpiration is calculated from Penman's combination equation in the form given by Monteith (1965). The surface and aerodynamic resistance values are dependent on indices for the plant and leaf area, where these vegetation properties are calculated from the dynamic above ground biomass development. Compensatory water uptake by plant roots determines actual transpiration, where effects of soil temperature and salt on water uptake are ignored. Carbon and nitrogen dynamics are regulated by several plant and soil compartments linked by transfer and decay coefficients such as for biomass, litter or soil organic pools. Further information on nitrogen dynamics in soil and plant can be obtained in Jansson and Karlberg

(2004). The CoupModel has been applied since 2002 for roadsides (Lundmark, 2008), forests (Norman et al., 2008) and arable ecosystems (Karlberg et al., 2007; Zhang et al., 2007; Conrad and Fohrer, 2009a,b) at the plot scale. Major changes in the model structure associated with parameter uncertainty have happened since 2007, where the Bayesian (Klemetsson et al., 2008) and the GLUE approach (Lundmark, 2008) were introduced and tested. Following the results of sensitivity analysis in Conrad and Fohrer (2009a,b), the majority of input parameters was fixed at pre-defined values in this model setup. The remaining 24 parameters were selected for the stochastic optimization (Table 2). The first parameter in Table 2, i.e. *ThScaleLog(1)*, regulates soil heat flow from the uppermost soil layer. The other parameters are responsible for biotic system properties such as plant development, nutrient uptake, mineralization, nitrification and denitrification. The parameter space is defined by minimum (Min) and maximum (Max) values based on plausible reasoning. We have chosen mainly soil biotic parameters because they are highly uncertain due to small and site-specific data sets. These parameters are most sensitive on the NO<sub>3</sub> nitrogen leaching below the rooting zone.

## 2.3 Calibration method

The CoupModel was optimized using the Bayesian calibration applied by Van Oijen et al. (2005). The posterior probability distribution  $p(\theta/D)$  for the parameter vector  $\theta$  is derived by the likelihood function  $p(D/\theta)$  and the prior distribution  $p(\theta)$  of the parameter vector according to:

$$p(\theta/D) = c * p(D/\theta) * p(\theta) \quad (1)$$

Where the value of  $c(=1/p(D))$  is independent. The likelihood  $p(D/\theta)$  is computed assuming measurement errors are Gaussian and uncorrelated:

**Table 2.** Parameters selected for the stochastic optimization in CoupModel and their initial value and uncertainty ranges.

Property	Description	Start value	Ranges		Unit
			Min	Max	
<i>Soil thermal properties:</i>					
ThScaleLog (1), $x_{hf(1)}$	Scaling coefficient for the thermal conductivity in soil layer 1	0.5	-0.5	1	-
<i>Plant specific properties:</i>					
Specific LeafArea, $p_{l,sp}$	Leaf mass per unit leaf area	6	4	8	$\text{g C m}^{-2}$
RadEfficiency, $\varepsilon_L$	Radiation use efficiency	3.5	2	4	$\text{g d.w. MJ}^{-1}$
NUptFlexibilityDeg, $n_{Uptflex}$	Compensatory N uptake from soil	0.3	0.1	0.5	-
NUptMaxAvailFrac, $f_{NUpt}$	Fraction of mineral N for uptake	0.04	0.01	0.1	$d^{-1}$
<i>Decomposition and mineralization:</i>					
CN ratio microbes, $cn_m$	C:N ratio in microbes	10	9	11	-
Eff Litter1, $f_{e,l1}$	Efficiency of decay of litter 1	0.15	0.1	0.3	$d^{-1}$
Eff Litter2, $f_{e,l2}$	Efficiency of decay of litter 2	0.15	0.1	0.3	$d^{-1}$
Eff Humus, $f_{e,h}$	Efficiency of decay of humus	0.6	0.4	0.8	$d^{-1}$
HumFracLitter1, $f_{h,l1}$	Fraction of C and N from litter 1 to humus	0.2	0.1	0.4	$d^{-1}$
HumFracLitter2, $f_{h,l2}$	Fraction of C and N from litter 2 to humus	0.2	0.1	0.4	$d^{-1}$
Init H N Tot, $i_{h,N}$	Initial total N in humus	500	400	600	$\text{g N m}^{-2}$
Init L1 N Tot, $i_{l1,N}$	Initial total N in litter 1	5	4	6	$\text{g N m}^{-2}$
Init L2 N Tot, $i_{l2,N}$	Initial total N in litter 2	1	0.5	2	$\text{g N m}^{-2}$
RateCoefHumus, $k_h$	Coefficient for the decay of humus	5e-5	1e-5	1e-4	$d^{-1}$
RateCoefLitter1, $k_{l1}$	Coefficient for the decay of litter 1	0.1	0.01	0.5	$d^{-1}$
RateCoefLitter2, $k_{l2}$	Coefficient for the decay of litter 2	0.1	0.01	0.5	$d^{-1}$
RateCoefSurf L1, $l_{l1}$	Fraction of above ground residues to litter 1	1	0.05	1	$d^{-1}$
RateCoefSurf L2, $l_{l2}$	Fraction of above ground residues to litter 2	1	0.05	1	$d^{-1}$
<i>Nitrification process:</i>					
NitrateAmmRatio, $r_{nitr,amm}$	$\text{NO}_3:\text{NH}_4$ ration for nitrification	1	0.1	1.5	-
NitrificSpecificRate, $n_{rate}$	Specific nitrification rate	0.1	0.08	0.15	$d^{-1}$
<i>Denitrification process:</i>					
DenitDepth, $d_z$	Depth where the denitrification capacity ceases	-1.5	-1.6	-1.2	m
DenitNitrateHalfSat, $d_{NhalfSat}$	Effect of $\text{NO}_3$ concentration on denitrification	10	8	12	$\text{mg N l}^{-1}$
DenitPotentialRate, $d_{pot}$	Potential denitrification rate	0.1	0.05	0.2	$\text{g N m}^{-2} d^{-1}$

$$\log p(D/\theta) = \quad (2)$$

$$\sum_{i=1}^n \left( -0.5 \left( \frac{O_i - S_i}{M_i} \right)^2 - 0.5 \log(2\pi) - \log M_i \right)$$

Where the  $S_i$  are model results and  $O_i$  observations,  $n$  is the number of observations and  $M_i$  is the standard deviation or error of measured values. The logarithm of the data likelihood ( $\log p(D/\theta) = \text{LogLi}$ ) is used to avoid rounding errors when likelihood values are decreasing with increasing number of data points (Klemedtsson et al., 2008). We assume that structural errors are ignored and thus estimates of model output uncertainty show only the contribution from parameter uncertainty (Reinds et al., 2008). A numerical solution of Eq. (2) is often carried out in the form of a Markov Chain Monte Carlo (MCMC) approach with a large number of simulations. The Metropolis-Hastings random walk, a simple MCMC algorithm, was used to calculate posterior probabilities for an appropriate number of parameter combinations by randomly stepping through the parameter space. An adequate sampling during MCMC can be achieved by combination of the number of steps and the step size. If for a new

candidate point the product of the prior probability and the likelihood is higher than for the current point, the new point is accepted. The total LogLi is derived from the sum of all LogLi values of the validation variables in Table 1. This procedure results in a chain of points in the multi-dimensional parameter space, where the first 10% of the runs are ignored. The remaining chain includes all accepted parameter points which are used to calculate the posterior distribution of each parameter and correlation and covariance matrices.

In practical application, each grassland system was optimized separately in a “single-plot” calibration. The posterior distribution with its mean value and variability were derived for each optimized parameter depending on the selected output variables. A number of 20 000 simulation runs were carried out within the same parameter space for the tested systems according to Table 2.

### 3 Results

#### 3.1 Posterior parameter distributions

The Bayesian calibration provides the joint posterior distribution, which contains also correlations between parameters.

**Table 3.** Prior mean, posterior mean and coefficient of variation (CV) of the optimized parameters for both grassland plots at the Karkendamm site.

Property	Prior mean	Non-fertilized plot (N0)		Highly fertilized plot (N300)	
		Post mean	CV (%)	Post mean	CV (%)
<i>Soil thermal properties:</i>					
ThScaleLog 1), $x_{hf}(1)$	0.25	-0.50	0.9	0.66	1.5
<i>Plant specific properties:</i>					
Specific LeafArea, $pl_{,sp}$	6	6.8	0.6	4.0	0.6
RadEfficiency, $\varepsilon_L$	3	3.7	0.3	3.9	0.3
NUptFlexibilityDeg, $n_{Uptflex}$	0.3	0.32	1.0	0.38	2.0
NUptMaxAvailFrac, $f_{Nupt}$	0.055	0.006	1.4	0.02	3.5
<i>Decomposition and mineralization:</i>					
CN ratio microbes, $cn_m$	10	10.6	0.1	10.3	0.2
Eff Litter1, $f_{e,l1}$	0.20	0.23	1.1	0.10	2.7
Eff Litter2, $f_{e,l2}$	0.20	0.15	1.4	0.11	1.1
Eff Humus, $f_{e,h}$	0.60	0.55	1.0	0.78	1.1
HumFracLitter1, $f_{h,l1}$	0.25	0.19	0.9	0.21	1.9
HumFracLitter2, $f_{h,l2}$	0.25	0.23	1.5	0.20	3.1
Init H N Tot, $i_{h,N}$	500	505	0.4	481	0.7
Init L1 N Tot, $i_{l1,N}$	5	5.2	0.2	4.6	0.3
Init L2 N Tot, $i_{l2,N}$	1.25	0.63	0.1	0.59	3.3
RateCoefHumus, $k_h$	5.5e-5	5.9e-5	0.6	6.6e-5	1.3
RateCoefLitter1, $k_{l1}$	0.255	0.18	2.0	0.154	5.9
RateCoefLitter2, $k_{l2}$	0.255	0.303	2.4	0.119	3.1
RateCoefSurf L1, $l_{l1}$	0.525	0.85	1.0	0.920	0.6
RateCoefSurf L2, $l_{l2}$	0.525	0.59	3.0	0.930	1.0
<i>Nitrification process:</i>					
NitrateAmmRatio, $r_{nitr,amm}$	0.80	1.38	0.5	1.07	1.9
NitrificSpecificRate, $n_{rate}$	0.115	0.087	0.5	0.100	0.5
<i>Denitrification process:</i>					
DenitDepth, $d_z$	-1.40	-1.49	0.4	-1.54	0.4
DenitNitrateHalfSat, $d_{NhalfSat}$	10	9.8	0.2	10.7	0.2
DenitPotentialRate, $d_{pot}$	0.125	0.051	3.6	0.136	1.8

In this study, we focus on the marginal distributions expressed as posterior mean and coefficient of variation (CV) for the individual parameters. The variability or dispersion of a parameter is expressed as CV (= standard deviation divided by the mean), which is low in case of  $CV < 1$  (100%). Broad prior distributions shall narrow down with low CV values leading to reduced parameter uncertainty when measurements are conclusive. For most parameters, differences between the posterior mean occurred for both systems forced by the observations according to Eq. (2), even though same prior mean values were assumed (Table 3). Parameters with similar posterior mean values indicate a low sensitivity on individual parameters from the data behind for each system. This was found e.g. for the factor of the compensatory N uptake from different soil layers (*NUptFlexibilityDeg*), the CN ratio of the microbes (*CN ratio microbes*), the decay efficiency of litter pool 2 (*Eff Litter2*), the carbon (C) and N fractions transferred from litter pools 1 and 2 to the humus pool (*HumFracLitter1*, *HumFracLitter2*) and the initial N content in litter pool 2 (*Init L2 N Tot*). In this study, the CV var-

ied between 0.1 and 3.6% emphasizing a low variability of selected parameter values that were robust during optimization. The highest difference between the two grassland systems was found for the scaling factor of the thermal conductivity *ThScaleLog(1)*, but with minor influence on the model efficiency (RMSE,  $R^2$ ) for soil temperatures (Table 4). Remaining parameters differed in their posterior mean values between both systems, especially for plant specific parameters that are linked to the biomass production (*RadEfficiency*, *Specific LeafArea*), and for the denitrification (*DenitPotentialRate*). In general, the dispersion of the parameters (CV) was higher for the N300 than for the N0 plot.

### 3.2 Comparison with measurements

The total LogLi according to Eq. (2) was used as objective function to find the accepted simulations among 20 000 runs. Table 4 shows the resulting modeling efficiency of the 23 validation variables for both systems. The root mean squared error (RMSE) and the coefficient of determination ( $R^2$ ) were

**Table 4.**  $R^2$  and RMSE values for the comparison between simulated mean of the accepted runs and the observed values.

Efficiency measure	RMSE		$R^2$	
	N0	N300	N0	N300
System				
<i>Soil temperature (°C days)</i>				
0.05 m	1.64	1.46	0.93	0.94
0.10 m	1.38	1.26	0.95	0.95
0.15 m	1.25	1.11	0.95	0.96
<i>Soil water content (Vol.%)</i>				
0.10 m	6.45	5.52	0.60	0.59
0.30 m	6.73	5.05	0.34	0.51
0.40 m	3.04	3.83	0.85	0.91
0.50 m	10.66	4.77	0.35	0.55
0.60 m	6.85	2.23	0.95	0.94
0.70 m	12.6	7.58	0.41	0.57
0.80 m	7.56	5.18	0.43	0.72
<i>Soil water tension (hPa)<sup>a</sup></i>				
0.30 m	94	448	0.44	0.43
0.50 m	21	349	0.30	0.19
0.70 m	20	156	0.25	0.10
<i>Groundwater level (m)</i>				
Nearest point	0.15	0.21	0.41	0.39
<i>Nmin (g N m<sup>-2</sup>)</i>				
0–0.3 m	1.12	8.89	0.12	0.06
0–0.6 m	2.29	10.37	0.28	0.01
0–0.9 m	2.66	7.69	0.0005	0.03
<i>NO<sub>3</sub> N (g N m<sup>-2</sup>)</i>				
0–0.3 m	0.54	6.69	0.01	0.04
0–0.6 m	0.52	6.21	0.006	0.0005
0–0.9 m	0.54	8.45	0.03	0.14
<i>NH<sub>4</sub> N (g N m<sup>-2</sup>)</i>				
0–0.3 m	1.37	2.79	0.24	0.03
0–0.6 m	2.43	1.80	0.31	0.01
0–0.9 m	2.83	1.93	0.20	0.04
<i>NO<sub>3</sub> N concentration (mg N l<sup>-1</sup>)</i>				
Seepage period	3.37	22.1	0.02	0.14
<i>Denitrification (g N m<sup>-2</sup>)</i>				
(total)	0.0004	0.002	0.19	0.21
<i>Harvested carbon (g C m<sup>-2</sup>)</i>				
(cumulative)	71	730	0.27	0.008

<sup>a</sup> not optimized during the Bayesian calibration; posterior comparison to validate soil water conditions

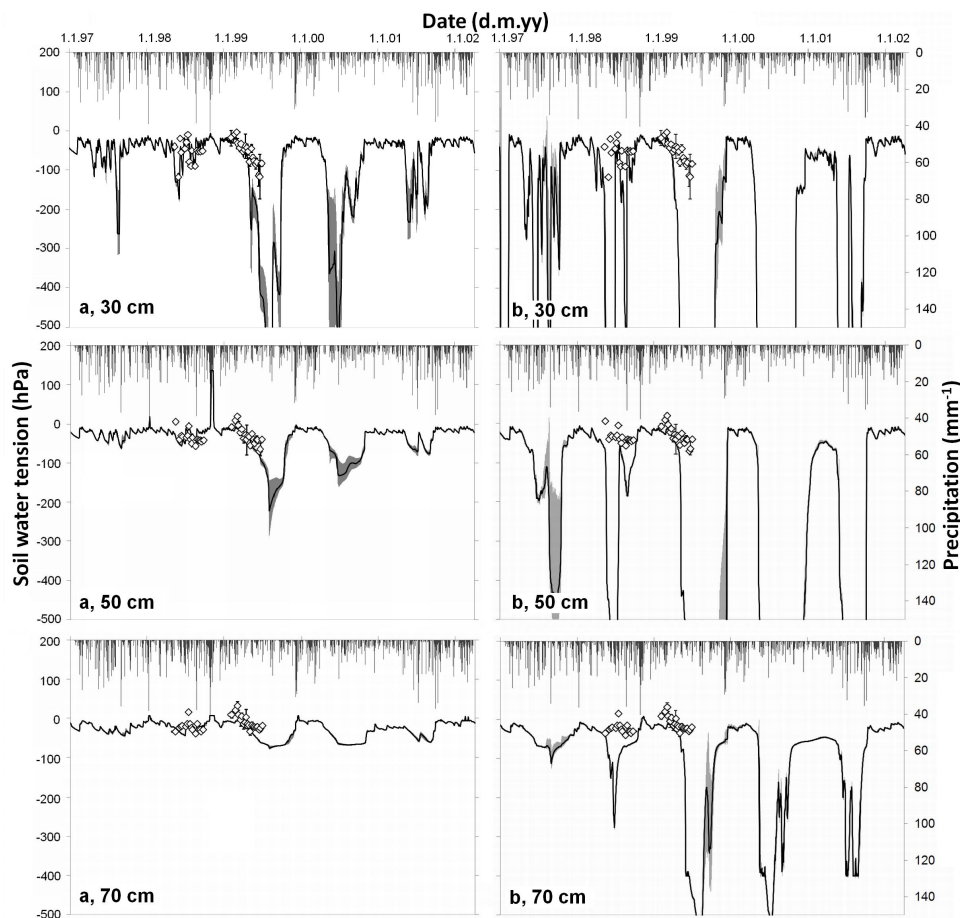
$R^2$ : coefficient of determination (1 or –1 [0; 1 or –1]; RMSE: root mean square error (0[–∞; +∞])

used to compare simulated mean values with the mean value of the measurements. Soil temperatures were simulated well for all depths with  $R^2 > 0.93$ . The RMSE varied between 1.11 and 1.64 with lowest deviations for the depth of 0.15 m in both systems, but with better values for N300 than N0. The agreement between modeled and observed soil water contents differed inside soil with  $R^2$  values from 0.43 (depth of 0.30 m) to 0.95 (depth of 0.60 m). The RMSE value was lowest in the depth of 0.40 m, and higher  $RMSE > 10$  were found for the N0 system in depths of 0.50 and 0.70 m. The groundwater level agreed satisfactorily with measurements from the nearest observation point with  $R^2$  values between 0.41 (N0) and 0.39 (N300). However, the model could not match the observed dynamic for both systems. But a lower RMSE value was achieved for the N0 than the N300 plot.

Soil water tensions, which represented mean values for the mown grassland plot during summer 1998 and 1999, were compared with model results for both fertilization levels in 30, 50 and 70 cm depth (Fig. 1). Model results were highly variable with soil water tensions between +146 hPa (saturated) and –2000 hPa ( $\approx$ pF 3.3 in 30 cm depth). Contrary to the model outcome, measured soil water tension was lowest with –117 hPa ( $\approx$ pF 2.1) at a depth of 30 cm, where also the best agreement between observation and model was achieved regarding  $R^2=0.44$  (N0) and 0.43 (N300) (see Table 4). RMSE values decreased with increasing depth, which indicated lower variability in the modeled results for deeper soil layers. Uncertainty was higher during summer than spring in both model and observation. Differences between CoupModel and field data were based on model limitations regarding slanted soil layers and the used water retention curve from the laboratory investigation. Fixed parameter values for each soil layer are possibly not representing heterogeneous field conditions with highly variable water contents within the field replicates (Karrasch, 2005). Secondly, measurements taken with a puncture tensiometer are biased, the technique must be installed carefully to prevent pressure changes in the tube during needle injection. The accuracy of the device is not as high as in systems with an in-site manometer or pressure sensor (Smith and Mullins, 2001). Systematic underestimation of actual tensions by up to 23 hPa was found for single-puncture tensiometer readings with a maximum error of approx. 10% (Greenwood and Daniel, 1996).

The comparison of daily simulated mean values for the soil mineral N with the measurements showed only low coefficients of determination  $R^2 < 0.31$  (Table 4) which did not suggest some evidence for an under- or overestimation. Soil mineral N contents between 0–0.90 m are shown in Fig. 2 indicating an overestimation of both the  $NH_4$  nitrogen in the N0 plot and the  $NO_3$  nitrogen content in the N300 system. In general, RMSE values were lower for the non-fertilized plot than for the N300 (Table 4) due to overestimations in CoupModel associated with the  $NO_3$  nitrogen content in the highly fertilized system. Modeled peaks in the right charts of Fig. 2 can identify the applied mineral fertilizer in the N300 plot. The few observations indicated only that the soil mineral content in spring was higher than in autumn. Deviations between modeled and measured results were possibly caused by an inadequate plant N uptake and soil nitrogen transformation (e.g. denitrification, see below) in CoupModel during summer.

Consequently, the total mineral N was also overestimated by CoupModel indicating much more uncertainty in the prediction of soil N dynamics of highly fertilized systems due to the input of mineral nitrogen. CoupModel considers mineral fertilizer input on the soil surface, where the fraction of  $NH_4$  nitrogen (the rest is  $NO_3$ ) and specific dissolution rate of the applied commercial fertilizer must be defined. In our study, fixed values for fractioning and dissolution rate were used in CoupModel, which was maybe not

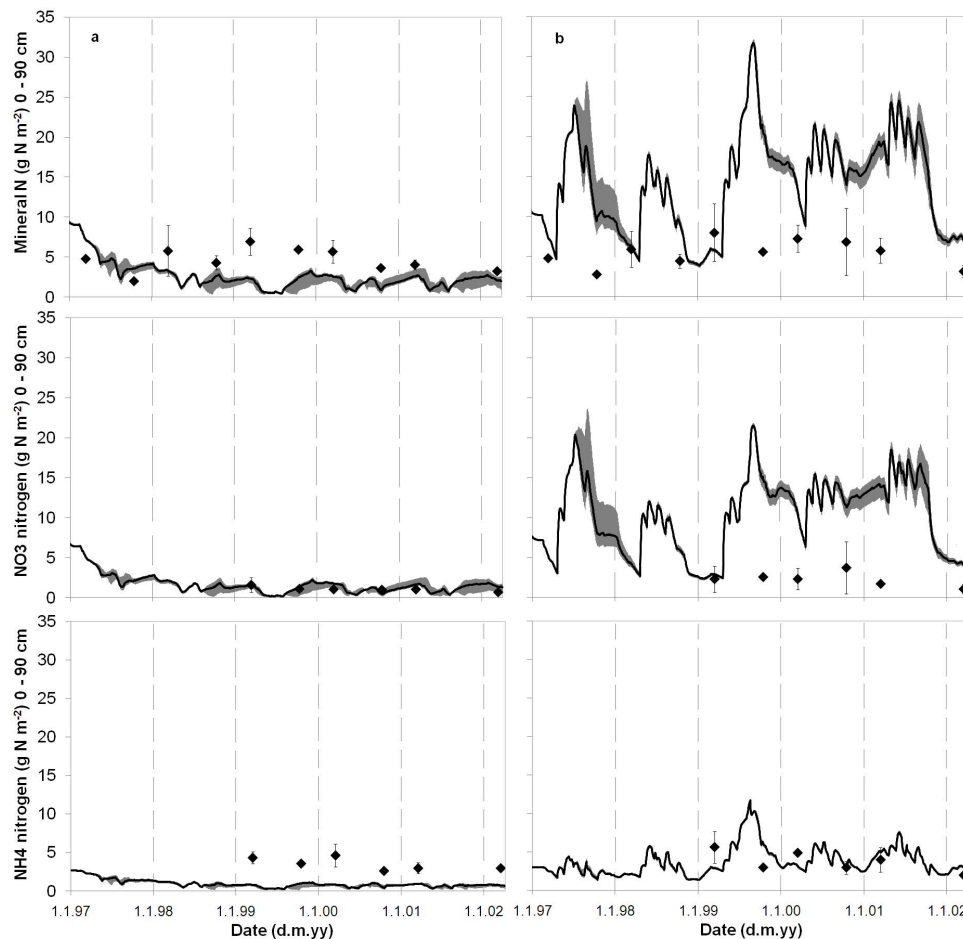


**Fig. 1.** Mean values of the soil water tensions (in hPa) in 30, 50 and 70 cm depth for the (a) non-fertilized and (b) highly fertilized mown grassland. Simulated results (solid line) are shown within the 5th and 95th percentiles and observations ( $\diamond$ ) of the mown grassland with their standard deviations.

representative for the applied  $\text{Ca}(\text{NH}_4\text{NO}_3)$  fertilizer. Further reason for the overestimation of modeled  $\text{NO}_3$ , which was not observed for  $\text{NH}_4$ , could be the uptake of mineral N by the grass vegetation that prefers ammonium (Maci et al., 2007). The denitrification process is also a possible sink for soil  $\text{NO}_3$  nitrogen. This gaseous nitrogen loss was very likely underestimated in the N300 system by CoupModel. The  $\text{NO}_3$  nitrogen content was reproduced better in the N0 plot than the  $\text{NH}_4$  nitrogen content, which was possibly caused by an underestimated nitrification rate. Reason for that can be found in the soil  $\text{NO}_3$  nitrogen concentrations that were seen as very uncertain with highly variable observations (Fig. 3). Concerning this, the simulated mean of the  $\text{NO}_3$  nitrogen concentration was forced to be within the standard deviation of the measurements during modeling. For the N0 system the simulated mean was only 94% ( $4.6 \text{ mg NO}_3 \text{ NI}^{-1}$ ) of the measured mean ( $4.9 \text{ mg NO}_3 \text{ NI}^{-1}$ ) indicating that too much  $\text{NH}_4$  nitrogen remained in the soil profile.

The  $R^2$  value of the averaged  $\text{NO}_3$  nitrogen concentrations was higher for the N300 (0.14) than for the N0 plot (0.02), which was maybe caused by less variable observations from day to day in the highly fertilized system. It can be discussed if the  $R^2$  value is an adequate performance measure for soil solution concentrations without considering data variability in model and measurements. The difference between modeled ( $24 \text{ mg NO}_3 \text{ NI}^{-1}$ ) and observed mean ( $11 \text{ mg NO}_3 \text{ NI}^{-1}$ ) of the N300 plot amounted to 125% demonstrating an overestimation of simulated soil  $\text{NO}_3$  nitrogen. The RMSE value was six times lower in the N0 plot, indicating also a good agreement for the non-fertilized system. One important process to reduce soil  $\text{NO}_3$  nitrogen in CoupModel is the denitrification process, which was compared with measurements from a temporary cutting period of a non-fertilized grassland plot (Fig. 4).

The RMSE from the N0 system was lower than for N300, while the  $R^2$  values were similar for this short period. In Fig. 3, the N300 plot shows a highly variable denitrification amount according to the input of mineral N and humid winter conditions. The denitrification decreased during summer



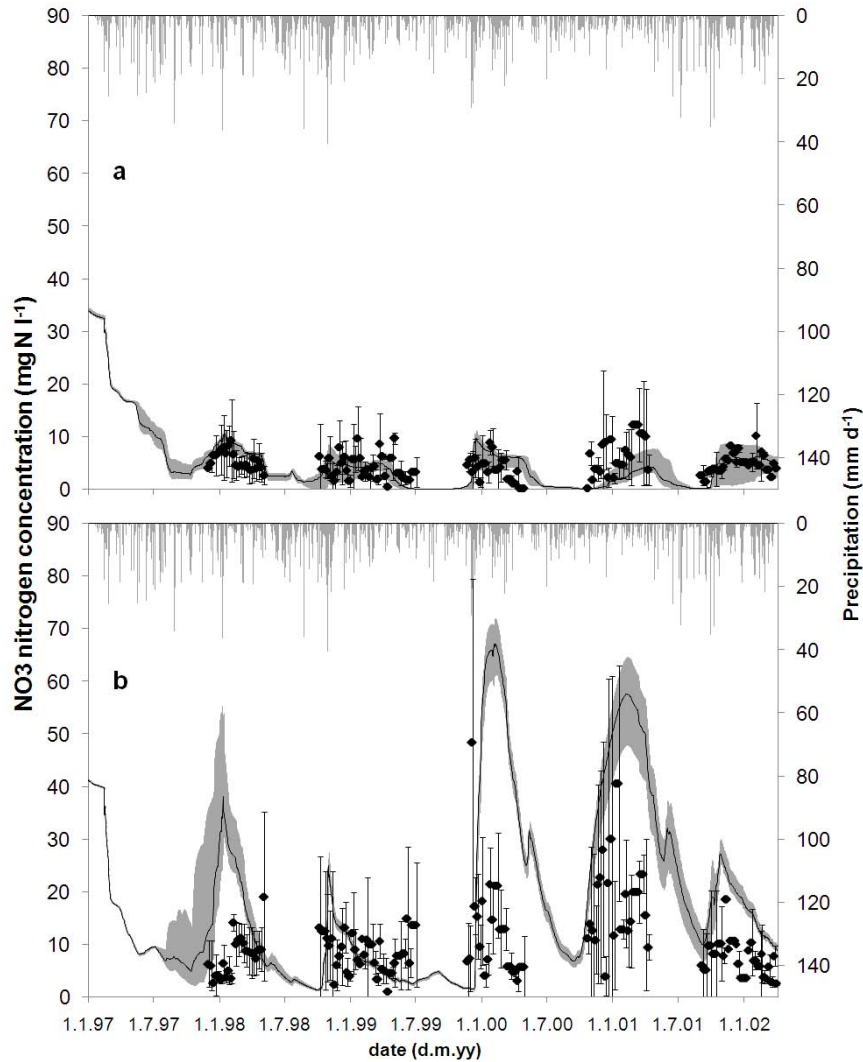
**Fig. 2.** Mean values of the soil total mineral, nitrate and ammonium nitrogen in the a) non-fertilized and b) highly fertilized grassland plots. Simulated results (solid line) are shown within the 5th and 95th percentiles and observations (◆) with their standard deviations.

due to lower soil water contents and significant N uptake for biomass production.

Modeled annual harvested C was compared with calculated values derived from observed pure grass biomass between 1997–1999 and estimations for 2000 and 2001 for the N0 plot (Ingwersen, 2002; see Table 1). The comparisons between harvested C for the N300 plot was based on estimations from harvested clover-grass biomass fertilized with  $300 \text{ kg N ha}^{-1} \text{ year}^{-1}$  (Trott, 2003). The agreement was better for the N0 plot ( $R^2=0.27$ ) than for the N300 plot ( $R^2=0.008$ ) indicating that the model was not able to reproduce realistic annual harvest amounts in both cases (Table 4). Biomass production was overestimated because the N uptake by plants was much less important than a good reproduction of measured soil N contents. The comparison with estimates from harvested clover-grass, which shows unstable harvest yields with increasing N fertilization due to the decline of clover (Herrmann et al., 2005), could be a second reason for the low agreement regarding the harvested C in the N300 system.

In Fig. 5, simulated  $\text{NO}_3$  nitrogen leaching below the rooting zone located between 0.60 and 0.65 m depth was compared to the model “Büchter” (Büchter, 2003) for the seepage periods. For the non-fertilized plot N0, CoupModel calculated an averaged leaching amount of  $18 \text{ kg NO}_3 \text{ N ha}^{-1}$  compared to  $10 \text{ kg NO}_3 \text{ N ha}^{-1}$  by the model “Büchter”. Averaged  $\text{NO}_3$  nitrogen leaching for the highly fertilized system in CoupModel added up to  $49 \text{ kg NO}_3 \text{ N ha}^{-1}$ , which was 63% higher than the leaching of  $30 \text{ kg NO}_3 \text{ N ha}^{-1}$  in the model “Büchter”. Major differences were found in the dynamic of the leaching, which showed more peaks in the CoupModel realizations and in the drainage water amount. CoupModel accounted a 65% higher drainage water amount for the N0 plot and 16% lower for the N300 system, compared to a standard value of 180 mm by the model “Büchter”. An additional reason for an underestimation by the model “Büchter” is the failure of the suction cups methodology to capture the whole  $\text{NO}_3$  nitrogen leachate in available soil pores. This sampling method for soil water solutions is widely used in unstructured soils, but it showed a tendency





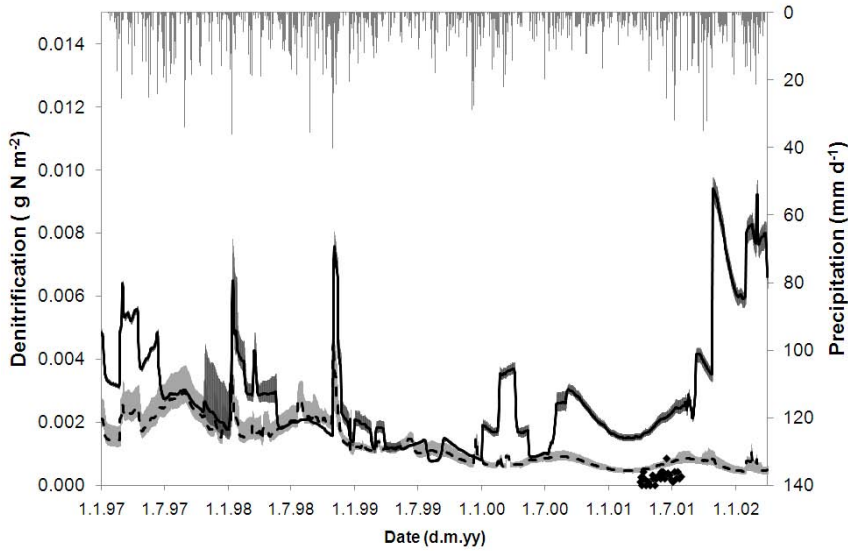
**Fig. 3.** Mean values of the soil  $\text{NO}_3$  nitrogen concentration in 60 cm depth for the (a) non-fertilized and (b) highly fertilized grassland plots. Simulated results (solid line) are shown within the 5th and 95th percentiles and observations ( $\blacklozenge$ ) with their standard deviations.

to underestimate N concentrations on an average of 8% compared to free drainage experiments (Erhart et al., 2007). Reasons for this deviation are possible anion absorption and the small cross-sectional area of the ceramic cups, which is maybe not representing spatial soil variability. An error of  $\pm 30\%$  or more should be expected in field studies, literature values must be handled with care, unless a representative number of ceramic suction cups are installed. Accordingly,  $\text{NO}_3$  nitrogen leaching has to be interpreted carefully, if obtained with this sampling method. Korsæth et al. (2003) modeled the  $\text{NO}_3$  leaching for grassland in Norway indicating that 5–23% of the N input can be leached out depending on soil, 98% of this amount was nitrate. In our study, where an additional atmospheric N deposition of  $20 \text{ kg N ha}^{-1} \text{ year}^{-1}$  was assumed, the  $\text{NO}_3$  nitrogen leach-

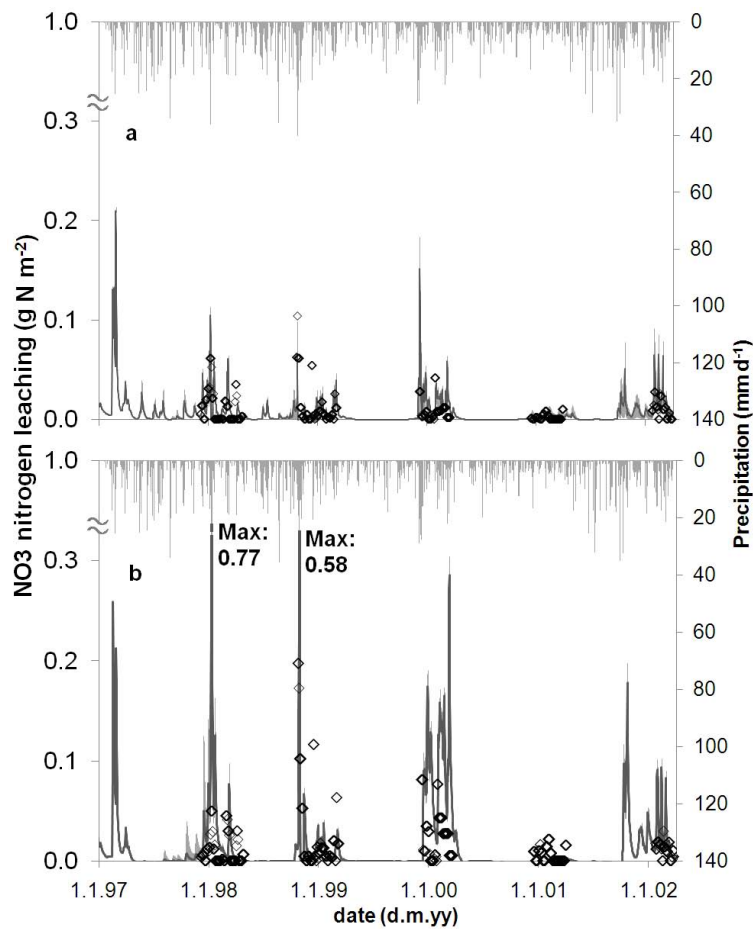
ing of the non-fertilized plot amounted to 90% of the N input and approx. 15% for the N300 system. Changes in the soil N storage were not considered in this simple N balance.

#### 4 Conclusions

The understanding of ecosystem processes involves an increasing model complexity with non-linear structural equations and useful methods for automated parameter estimation. Uncertainties in model parameterization and measurement have to be considered, but in practice rigorous uncertainty analysis is still rare (Stow et al., 2007). Model applications often fail to do an uncertainty assessment because many “competing methods” make it difficult to choose the most appropriate method and interpret the results (Pappenberger



**Fig. 4.** Total denitrification in the non-fertilized (dashed line) and highly fertilized (solid line) grassland plots. Simulated results are shown within the 5th and 95th percentile and observations (◆) of a non-fertilized system.



**Fig. 5.** NO<sub>3</sub> nitrogen leaching below the rooting zone (60–65 cm) in the (a) non-fertilized and (b) highly fertilized grassland plots. Coup-Model results (solid line) are shown within the 5th and 95th percentiles compared to the model “Büchter” (Büchter, 2003) (◇).

and Beven, 2006). Stochastic optimization can help to diminish the difficulties in terms of parameter estimation. In this paper, the applicability of the Bayesian calibration technique was demonstrated for the parameter optimization in the CoupModel. Efficient estimates of the most-likely parameter set and its underlying frequency distribution were provided during optimization runs. We performed two case studies demonstrating effects of multiple validation data on the parameter uncertainty. In spite of considerable differences between model and measurements, the prediction uncertainty associated with the parameter estimates is low, indicating that the main part of the uncertainty originates from the residuals between measurements and model predictions. Satisfying results were found for modeled abiotic properties, i.e. soil temperature, water content, water tension and groundwater level. On the whole, CoupModel results agreed better with observations for non-fertilized than highly fertilized conditions regarding soil N dynamics and harvested C. Uncertainty was highest for modeled soil NO<sub>3</sub> nitrogen concentrations in both systems, plausible results were also found for leached NO<sub>3</sub> nitrogen. One reason for the mismatch between model and measurements could be that simulated results were based on one soil profile, whereas observations were taken over a larger area ranging from few centimeters (e.g. installed equipment for soil temperature and water content) to several meters (e.g. mean values for harvested C and soil mineral N). Further work must be done to understand the limitation of this approach because of its subjective choice of the probability distribution and likelihood measure.

*Acknowledgements.* The study was supported by the German Environmental Foundation (DBU). The authors thank Per-Erik Jansson for the model support and both reviewers, Helge Bormann and one anonymous person, who helped to improve this article.

Edited by: B. Schmalz, K. Bieger, and N. Fohrer

Reviewed by: H. Bormann and another anonymous referee

## References

- Büchter, M.: Nitratauswaschungen unter Grünland und Silomais in Monokultur auf sandigen Böden Norddeutschlands. Ph.D. thesis, University of Kiel, Germany, Schriftenreihe des Instituts für Pflanzenbau und Pflanzenzüchtung – Grünland und Futterbau/Ökologischer Landbau (in German, English summary), 30, 1–116, 2003.
- Casper, M.: Die Identifikation hydrologischer Prozesse im Einzugsgebiet des Dürreychbaches (Nordschwarzwald). Mitteilungen des Instituts für Wasserwirtschaft und Kulturtechnik der Universität Karlsruhe (TH) (in German, English summary), Heft 210, 1–214, 2002.
- Conrad, Y. and Fohrer, N.: Modelling of nitrogen leaching under a complex winter wheat and red clover crop rotation in a drained agricultural field, *Phys. Chem. Earth, Parts A/B/C*, 34(8–9), 530–540, 2009.
- Conrad, Y. and Fohrer, N.: Modeling of nitrogen leaching in grassland systems with two different fertilization levels in Northern Germany, accepted in *J. Plant Nutr. Soil Sci.*, 2009b.
- De Vries, W., Reinds, G. J., Deelstra, H. D., Klap, J. M., and Vel, E. M.: Intensive Monitoring of Forest Ecosystems in Europe. Technical Report 1999. UN/ECE, EC, Forest Intensive Monitoring Coordinating Institute, 1–173, 1999.
- DVWK (ed.): Ermittlung der Verdunstung von Land- und Wasserflächen, DVWK-Merkblätter zur Wasserwirtschaft, 238, 1–134, 1996.
- Erhart, E., Feichtinger, F., and Hartl, W.: Nitrogen leaching losses under crops fertilized with biowaste compost compared with mineral fertilization, *J. Plant Nutr. Soil Sci.*, 170, 608–614, 2007.
- FAO: World reference base for soil resources 2006. 2nd edition. World Soil Resources Reports No. 103, Rome, Italy, 1–132, ISBN: 92-5-105511-4, 2006.
- Greenwood, K. L. and Daniel, H.: A double-puncture technique for improving the accuracy of puncture tensiometer measurements, *Aust. J. Soil Res.*, 34, 153–159, 1996.
- Harmel, R. D., Cooper, R. J., Slade, R. M., Haney, R. L., and Arnold, J. G.: Cumulative uncertainty in measured streamflow and water quality data for small watersheds, *Trans. ASABE*, 49(3), 689–701, 2006.
- Herrmann, A., Kelm, M., Kornher, A., and Taube, F.: Performance of grassland under different cutting regimes as affected by sward composition, nitrogen input, soil conditions and weather – a simulation study, *Europ. J. Agronomy*, 22, 141–158, 2005.
- Ingwersen, B.: Einfluss von Bewirtschaftungsmaßnahmen auf die Leistungsfähigkeit von leguminosenbasiertem Dauergrünland unter besonderer Berücksichtigung der Nährstoffbilanzierung. Ph.D. thesis, University of Kiel, Germany, Schriftenreihe des Instituts für Pflanzenbau und Pflanzenzüchtung – Grünland und Futterbau/Ökologischer Landbau (in German, English summary), 21, 1–207, 2002.
- Jansson, P.-E. and Karlberg, L.: Coupled heat and mass transfer model for soil-plant-atmosphere systems. Royal Institute of Technology, Department of Civil and Environmental Engineering, Stockholm, Sweden, 1–435, 2004.
- Jovanovic, N. Z., Hon, A., Israel, S., Le Maitre, D., Rusinga, F., Soltau, L., Tredoux, G., Fey, M. V., Rozanov, A., and van der Merve, N.: Nitrate Leaching from Soils cleared of Alien Vegetation, Report to the Water Research Commission, WRC Report No. K51696, 1–88, 2008.
- Karlberg, L., Gustafsson, D., and Jansson, P.-E.: Modeling carbon turnover in five terrestrial ecosystems in the boreal zone using multiple criteria of acceptance, *AMBIO: A Journal of the Human Environment*, 35(8), 448–458, 2007.
- Klemmedtsson, L., Jansson, P.-E., Gustafsson, D., Karlberg, L., Weslin, P., von Arnold, K., Ernfors, M., Langvall, O., and Lindroth, A.: Bayesian calibration method used to elucidate carbon turnover in forest on drained organic soil, *Biogeochem.*, 89(1), 61–79, 2008.
- Korsaeth, A., Bakken, L. R., and Riley, H.: Nitrogen dynamics of grass as affected by N input regimes, soil texture and climate: lysimeter measurements and simulations, *Nutrient Cycling in Agroecosystems*, 66, 181–199, 2003.
- Lampe, C.: Effect of nitrogen fertiliser and animal excrements on N<sub>2</sub>O emissions from permanent grassland using <sup>15</sup>N-labelling. Ph.D. thesis, University of Kiel, Germany, Schriftenreihe des

- Institut für Pflanzenbau und Pflanzenzüchtung – Grünland und Futterbau/Ökologischer Landbau, 37, 1–109, 2005.
- Lewis, D. R., McGechan, M. B., and McTaggart, I. P.: Simulating field-scale nitrogen management scenarios involving fertiliser and slurry applications, *Agr. Syst.*, 76, 159–180, 2003.
- Lundmark, A.: Monitoring Transport and Fate of De-icing Salt in the Roadside Environment – Modelling and Field Measurements. TRITA-LWR PHD 1038, PhD thesis, KTH, Stockholm, Sweden, 1–47, 2008.
- Lundmark, A. and Jansson, P.-E.: Estimating the Fate of De-icing Salt in a Roadside Environment by Combining Modelling and Field Observations. *Water Air Soil Poll.*, 195, 215–232, 2008.
- Maci, A., Rrogo, E., Kosegarten, H., and Mengel, K.: Nitrogen turnover in bare soil planted subsequently with grass as investigated by electro-ultrafiltration (EUF), *J. Plant Nutr. Soil Sci.*, 170, 81–86, 2007.
- Monteith, J. L.: Evaporation and Environment, in: *The State and Movement of Water in Living Organisms*, edited by: Fogg, G. E., 19th Symp. Soc. Exp. Biol., Cambridge: The Company of Biologists, 205–234, 1965.
- Norman, J., Jansson, P.-E., Farahbakhshazad, N., Butterbach-Bahl, K., Li, C., and Klemmedtsson, L.: Simulation of NO and N<sub>2</sub>O emissions from a spruce forest during a freeze/thaw event using an N-flux submodel from PnET-N-DNDC model integrated to CoupModel, *Ecol. Model.*, 216, 18–30, 2008.
- OECD: Environmental Indicators for Agriculture: Methods and Results. OECD, Paris, France, 2001.
- Pappenberger, F. and Beven, K. J.: Ignorance is Bliss: Or Seven Reasons Not to Use Uncertainty Analysis, *Water Resour. Res.*, 42(5), WO5302, doi:10.1029/2005WR004820, 2006.
- Reinds, G. J., Van Oijen, M., Heuvelink, B. M., and Kros, H.: Bayesian calibration of the VSD soil acidification model using European forest monitoring data, *Geoderma*, 145, 475–488, 2008.
- Scholz, D.: Bodenwasserhaushalt unterschiedlicher pflanzenbaulicher Produktionssysteme während der Vegetationsperiode von Böden des Versuchsgutes Karkendamm im Bereich der Schleswig-Holsteinischen Geest (in German), Diploma thesis, University of Hamburg, Germany, 1–99, 1999.
- Smith, K. A. and Mullins, C. E. (eds.): *Soil and Environmental Analysis. Physical Methods*, 2nd edn., Marcel Dekker, NY, USA, ISBN: 0824704142, 1–656, 2001.
- Stow, C. A., Reckhow, K. H., Qian, S. S., Lamon, E. C., Arhonditis, G. B., Borsuk, M. E., and Seo, D.: Approaches to Evaluate Water Quality Model Parameter Uncertainty for Adaptive TMDL Implementation. *Journal of the American Water Resources Association (JAWRA)*, 43(6), 1499–1507, 2007.
- Taube, F. and Wachendorf, M.: The Karkendamm project: A system approach to optimize nitrogen use efficiency on the dairy farm, in: *Grassland Farming – Balancing environmental and economic demands*, edited by: Søgaard, K., Ohlsson, C., Sehested, J., Hutchings, N. J., and Kristensen, T., *Grassland Science in Europe*, 5, 449–451, 2001.
- Trott, H.: Mittelfristige Auswirkungen einer variierten Bewirtschaftungsform und N-Intensität auf Leistungsparameter und die Stickstoffbilanz von Dauergrünland (in German, English summary), Ph.D. thesis, University of Kiel, Germany, Schriftenreihe des Instituts für Pflanzenbau und Pflanzenzüchtung – Grünland und Futterbau/Ökologischer Landbau, 28, 1–170, 2003.
- Van Oijen, M., Rougier, J., and Smith, R.: Bayesian calibration of process-based forest models: bridging the gap between models and data, *Tree Physiol.*, 25, 915–927, 2005.
- Vrugt, J. A., Gupta, H. V., Sorooshian, S., Wagener, T., and Bouten, W.: Application of stochastic parameter optimization to the Sacramento soil moisture accounting model, *J. Hydrol.*, 325(1–4), 288–307, 2006.
- Wachendorf, M., Büchter, M., Trott, H., and Taube, F.: Performance and environmental effects of forage production on sandy soils. II. Impact of defoliation system and nitrogen input on nitrate leaching losses, *Grass Forage Sci.*, 59, 56–68, 2004.
- Zhang, S., Lövdahl, L., Grip, H., Jansson, P.-E., and Tong, Y.: Modelling the effects of mulching and fallow cropping on water balance in the Chinese Loess Plateau. *Soil Tillage Res.*, 93, 283–298, 2007.