Smart Agricultural Technology 12 (2025) 101367

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Model for planning site-specific nitrogen fertilization of barley in northern conditions

Antti Halla [a,*], Liisa Pesonen [b], Jari Pohjola [a], Petri Linna [a], Tarmo Lipping [a]

[a] Tampere University, Pohjoisranta 11 A, FI-28100, Pori, Finland
[b] Natural Resources Institute Finland (Luke), FI-00790, Helsinki, Finland

ARTICLE INFO

Dataset link: https://doi.org/10.5281/zenodo.16735443

Keywords:
Grain yield
N-uptake efficiency
Sensitivity analysis
Model validation
Decision support

ABSTRACT

Nitrogen is essential for crop cultivation. Understanding the nitrogen cycle and the factors contributing to the availability of nitrogen for the crops enables estimating the need for additional fertilizer inputs during the growing season. Several computational models have been proposed for this purpose. We propose an alternative approach that predicts crop nitrogen uptake from estimated site-specific nitrogen uptake efficiency and soil moisture. We measured the model accuracy to predict crop yield and nitrogen content in barley, using data collected from two test sites during multiple growing seasons. The model achieved index of agreement of 0.97 for yield prediction and 0.93 for absolute nitrogen content. A global sensitivity analysis was further performed on the model to describe how its outputs can be attributed to the model inputs, showing the most and least influential variables in the model.

1. Introduction

Nitrogen (N) has a central role in crop cultivation as its availability to the crop is strongly linked to crop yield and quality. Additional N is applied as fertilizer input to the crops during the growing season, either once in the beginning of the growing season, or in multiple stages. N has also an important role when considering the environmental footprint of crop production. It has been estimated that only 30–40% of the N available in soil can be utilized by the crops while significant amount is leached to waterways or emitted to atmosphere [1].

Nitrogen cycles both in inorganic and organic forms in soil, plants, animals and the atmosphere. Fig. 1, adopted from Engel [2], illustrates the main processes related to the N dynamics. In crop cultivation, the main direct sources that replenish the plant available nitrogen in soil are fertilizer inputs and N mineralization process.

Mineralization N mineralization in soil is a consequence of microbial activity. When estimating the mineralization rate during the growing season, attention must be paid to the factors that affect the microbial activity in the soil [3]. The most important factors are the source of energy, the availability of the microbe essential nutrients and water as well as environmental conditions such as temperature, aeration, redox potential and acidity [4].

N fertilization Usually at a farm level, more nitrogen is taken out of the system as bio-products and losses than is returned by nutrient recycling, nitrogen fixation or precipitation from the atmosphere. In order to maintain high yields and economically meaningful production, nitrogen among other nutrients is added to production system as fertilizers. Artificial nitrogen fertilizers, containing inorganic ammonium, nitrate, or urea, are the most commonly used sources of additional nitrogen in spring cereal cultivation in Finland. The amounts used vary typically between 60–170 kg/ha, depending on crop, its variety, soil type, soil organic matter content and the crop residues from previous year in soil. Manure and sludge are typically used at those farms where domestic animal husbandry is part of production. Fertilizing effect of manure differs slightly from artificial fertilizer, because in manure or sludge nitrogen is both in organic and inorganic form. Inorganic ammonium and nitrate are immediately available to plants after application, while organic nitrogen and urea has to go through microbiological and chemical processes in order to be turned into the ammonium form. Therefore, manure or sludge has rather long-lasting fertilizing effect [5].

N uptake and yield formation The quantity and the quality of the grain yield respond strongly to nitrogen fertilization. Deficiency of nitrogen causes low yields and low grain protein content. Excessive nitrogen supply can also lead to problems through luxury nitrogen uptake, which

* Corresponding author.
E-mail address: antti.halla@tuni.fi (A. Halla).

https://doi.org/10.1016/j.atech.2025.101367
Received 4 June 2025; Received in revised form 22 August 2025; Accepted 23 August 2025

Available online 2 September 2025

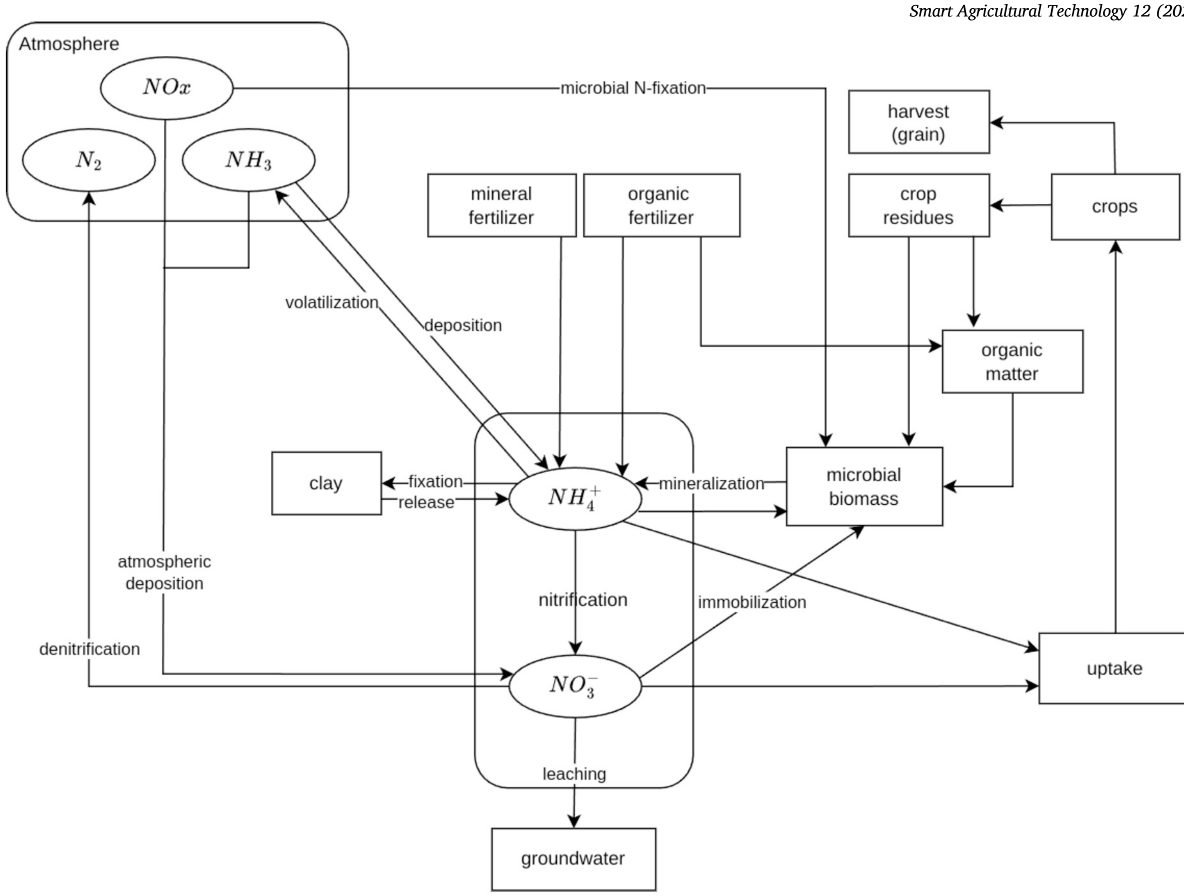

Fig. 1. Schematic of the nitrogen cycle. The main sources of nitrogen for crops are fertilizers and the mineralization process. Part of the soil N is lost to atmosphere and through leaching. (Adopted from Engel [2]).

may cause poor grain quality and yield losses due to lodging [6–8]. The nitrogen uptake by the plant is not only dependent on the available nitrogen in the soil but also on other growing factors, soil moisture content being the most important. Water is an essential medium in which nitrogen is carried from the soil to the plant and further to different parts of the plant [9]. Sufficient soil moisture conditions are also needed for the mineralization process [10].

1.1. Modeling N dynamics in soil

There have been numerous efforts to model the N-related features of crop growth. A simple model for practical management guidance called HERMES was introduced [11] and later evaluated [12] by Kersebaum. The model employs convection-dispersion equations for the transport of nitrate and first-order kinetics for the mineralization effect containing submodels for denitrification and plant growth. The model run starts at the harvest of the previous crop. Input data on weather conditions, soil properties and management are required. When evaluated using data from 5-year crop rotation cycle, the model simulated the soil water and nitrogen mineral content with index of agreement (IA) above 0.8 and 0.69, respectively. Above-ground biomass and N-uptake had IA of 0.93 and 0.71, respectively. More recently the HERMES model, appended with a drain flow component, was used to show that when using winter rye as a cover crop, nitrate loss to subsurface drainage was reduced [13].

A revised version of HERMES, called MONICA (MOdel for NItrogen and Carbon in Agroecosystems), was introduced by Nendel et al. [14] and evaluated for crop growth, soil moisture, and nitrogen dynamics. The main modification compared to the HERMES model was the introduction of model components related to carbon cycle in soil and plant. The model was calibrated using data from 8 sites over the period of 1992 to 2007. The performance of the calibrated model was evaluated based on experiments on 7 plots from 2 sites. While crop yield was estimated with high accuracy (IA close to 1), the model performance for N-uptake, crop N concentration and yield N concentration varied a lot with IA less than 0.5 in many cases. The authors noted that the model often underperformed for variables related to nitrogen.

Probably the most common model developed for simulating N dynamics is the Decision Support System for Agrotechnology Transfer model (DSSAT) [15] incorporating the *Fertilizer application* and *Soil N* modules. Cammarano et al. [16] used the DSSAT model to predict soil water content, soil mineral N, plant N, plant aboveground biomass, grain yield and grain N %. The 11-ha test field located in Scotland and growing barley, was divided into 4 productivity zones based on long-term yield maps. It was found that the goodness of fit between the predicted and observed grain yield was excellent (overall IA of 0.97), however, more modest results were obtained for plant N (IA of 0.73) and grain N % (IA 0.52). In addition to the plant growth related features they also modeled the mean marginal net return taking into account the premium paid by the industry if the grain N % is in optimal range. Another target of their optimization effort was minimizing the nitrate leaching.

1.2. Outline of the paper

In this paper we introduce a somewhat different approach for simulating N dynamics of crop growth. This is a process-based model that predicts crop nitrogen uptake from estimated site-specific nitrogen uptake efficiency and soil moisture. The primary purpose of the model is to support planning site-specific fertilization of cereals. The model considers two growing seasons – a reference year and a target year. By selecting different reference years, the model is capable of following the long-term changes in environmental conditions.

We first describe the model and provide the list of required inputs and parameters. We then evaluate the model using data from two test sites over various pairs of reference-target years. We also perform sensitivity analysis of the model to study the contribution of the various input factors to the simulation results. The sensitivity analysis results can be used to advise in required data accuracy when collecting data for practical use of the model.

2. Materials and methods

2.1. Model description

The model utilizes the N uptake efficiency of a homogeneous area within a field and the effect of temporally varying key factors to it (see Fig. 2). These are estimated based on the data from a reference season. The key temporally varying factor is the amount of available water at the site which means soil moisture in arable farming. Other factors influencing uptake efficiency would be weeds, pests and diseases, however, the closer study of these factors is framed out of this work. The test fields used in evaluating the model were observed and treated against weeds, pests and diseases. In the following the key components and parameters of the model are described in more detail.

N uptake efficiency N uptake efficiency (N_{eff}) is a quotient of the total N in the above-soil-surface plant parts ($N_{tot}\,[kg/ha]$) and total plant-available N in soil ($N_{soil}\,[kg/ha]$):

$$N_{eff} = N_{tot}/N_{soil}. \tag{1}$$

N_{tot} consists of the N content of the grain, straw and leaves. It can be determined by measuring the biomass and N concentration of the mentioned components. N_{soil} is determined as a sum of initial soluble N in the soil at the beginning of the growing period at seeding (N_{in}), N mineralized in the top soil during the growing season (N_m) and N applied as fertilizer (N_f):

$$N_{soil} = N_{in} + N_m + N_f. \tag{2}$$

N_{in} is measured from the soil samples taken from a soil layer 0–0.6 meters of depth. N_f is measured at application. N_m is calculated as a function of organic N content in the top soil (N_{org}), the depth of the top soil layer ($T_{soil}\,[cm]$), effective pH of the top soil (epH; see Equation (4)), volume weight of the top soil layer ($VW\,[kg/dm^3]$) and the microbial activity in the top soil (MA). Top soil is the visibly distinctive layer of soil that is mixed with organic matter such as plant residues. The factors 0.1 and 1 000 000 convert the result to kg/ha. The magnitude of the coefficient MA was validated using the net mineralization derived from the model construction data, i.e. the amount of mineralized nitrogen taken up by above-ground plant parts. Thus, the equation has to be corrected by the estimated amount of mineralized nitrogen uptake by roots. Hansson et al. [17] presented that the portion of the nitrogen uptake stored in roots is 21–28% of the total nitrogen uptake of barley plant. The average value 25% of this range is used in the model.

$$N_m = (T_{soil} * 0.1) * epH * VW * 1\,000\,000 * MA * (N_{org}/100)/0.75. \tag{3}$$

The typical response curves of soil microbial populations to the soil moisture content, temperature and pH are of similar parabolic shape [18]. These response curves have a top value in which microbial activity is at its maximum. Before and after this value the activity diminishes rapidly. The optimal value of pH for the most soil bacteria is 6.5–7. Below that value the activity diminishes rapidly, almost linearly. The slope is set to 0.45 within the pH range of interest in this context, namely the pH of arable soil, typically varying between 5 and 7.5. We therefore define the effective pH as:

$$epH = (1 - ((6.5 - pH) * e)), \tag{4}$$

where $e = 0.45$ if $pH < 6.5$, else 0. Grzebisz et al. [19] suggest that 1% of the organic N is mineralized during the growing season. Thus, MA gets the value 0.01 in the model.

Normalized site-specific N uptake efficiency The observed site-specific N uptake efficiency of the reference year (N_{eff}) is normalized with respect to the adequate top soil moisture content and growing density. This is done by adding respective coefficients ax and bx multiplied by the differences between the reference and observed soil moisture content ($diff_{SM_{nu}}\_1$) and growing density ($diff_{ED}\_1$), respectively. The soil moisture is an average value of weekly measurements within the relevant period in question (from seeding to ear emergence, from ear emergence to yellow ripening). Field capacity at 20 cm (SM_{fc}) is used as the reference value for soil moisture.

$$N_{effnorm}\_1 = N_{eff} + ax * diff_{SM_{nu}}\_1 + bx * diff_{ED}\_1, \tag{5}$$

where

$$diff_{ED}\_i = ED_{ref} - ED\_i, \quad i = 1, 2 \tag{6}$$

and

$$diff_{SM_{nu}}\_1 = SM_{fc} - SM_{nu}\_1. \tag{7}$$

Consistently adequate soil moisture content throughout the growing season is a precondition to maximum yield [20]. In our model implementation $ax = 0.005$, i.e., increase of 1 volume percent of soil moisture equals 0.5 percent increase in N uptake efficiency. This has support from the experimental results for the relationship between irrigation and nitrogen uptake by Tan et al. [21].

The coefficient $bx = 0.002$, i.e., increase of 1 head in the growing density per 0.25 m^2 equals 0.2 percent increase in N uptake efficiency and increase by 50 heads per 0.25 m^2 means N uptake increase by 10%. This assumption is valid up to the upper limit of 200 plants per 0.25 m^2. This is supported by empirical test data from Preiti et al. [22], from which derived values ranged from 0.15–0.32%.

Luxury N uptake Cereal plants like barley are able to uptake more nitrogen than they need, but above a certain limit the uptake efficiency decreases. The uptake above this level is called luxury N uptake. In the model, this decrease of uptake efficiency is 10%. The limit depends on the growth density: the more plants per area, the more the growth can uptake nitrogen per area. On the other hand, very dense growths are more prone for lodging compared to normal growths stands [23].

It was observed from the reference field data that the limit of total N_{soil} for barley variety used in the study not leading to lodging (N_{lux_limit}) was 180 kg/ha. When the growing density is lower than the optimal that maximizes the total yield at harvest, the limit of luxury N uptake lowers accordingly, and can be determined by multiplying the density difference by the coefficient bx.

Fertilization planning The obtained normalized site-specific N uptake efficiency for the reference year ($N_{effnorm}\_1$) is used when planning the site-specific N fertilization in the target year (N_f_2). The estimated site-specific soil moisture contents (SM_{nu}_2, SM_m_2) and growing density (ED_2) are used to define the total soluble N (N_{totsol}_2) and the N uptake efficiency (N_{eff}_2) in the target year:

$$N_{totsol}\_2 = (N_{soil}\_2) * (N_{effnorm}\_1 - (ax_2 * diff_{SM_{nu}}\_2) - (bx * diff_{ED}\_2)) \tag{8}$$

$$N_{eff}\_2 = N_{effnorm}\_1 - (ax_2 * diff_{SM_m}) - (bx * diff_{ED}\_2) \tag{9}$$

In Equations (8) and (9) the effect of the differences of the estimated soil moistures (SM_{nu}_2, SM_m_2) and growing density (ED_2) with respect to the values optimal for the N uptake efficiency, i.e., $diff_{SM_m}$, $diff_{SM_{nu}}$ and $diff_{ED}\_2$, is taken into account by subtracting the weighted differences, expressed as

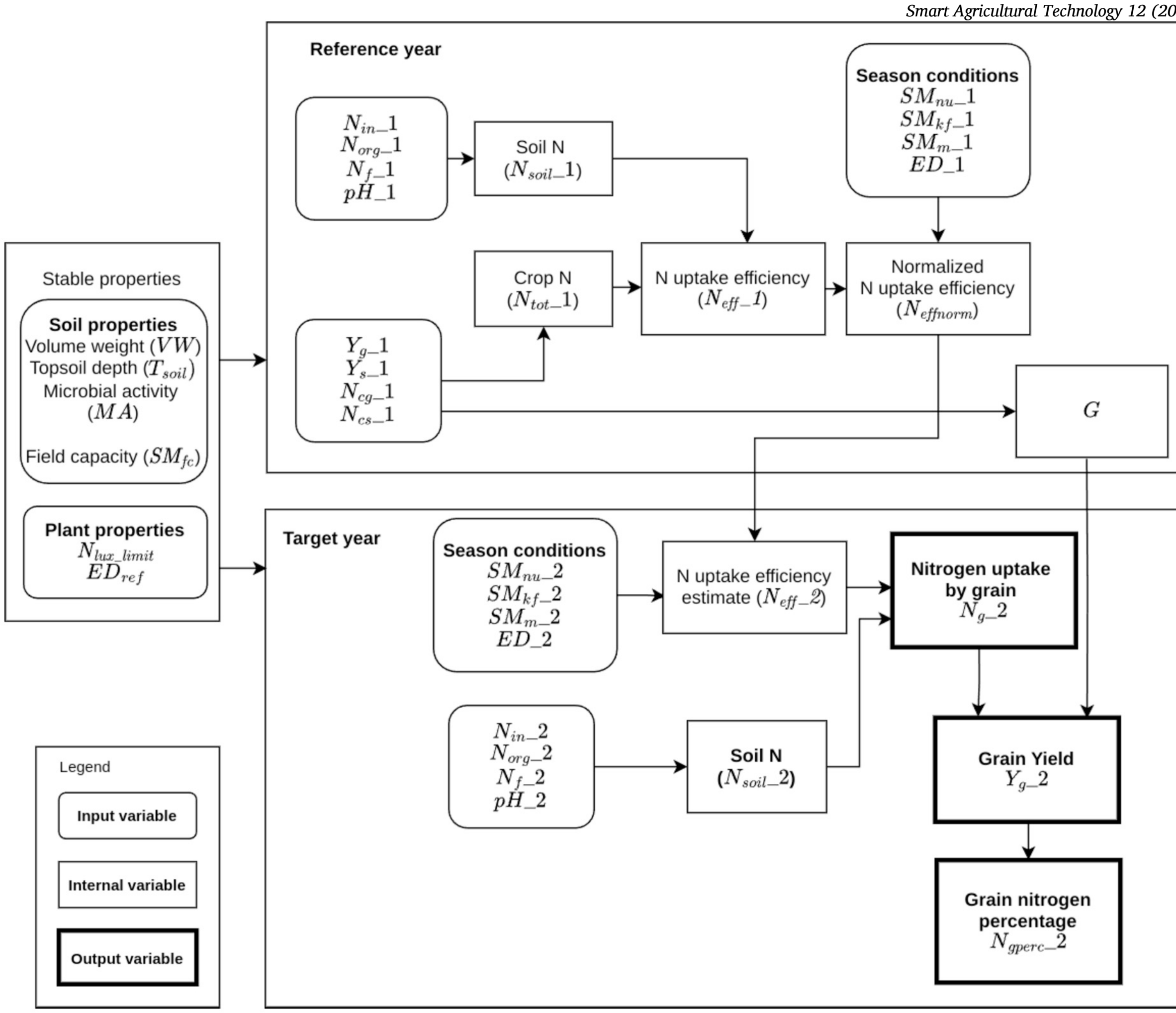

Fig. 2. Block diagram of the proposed model. The reference year measurements are done to estimate the site specific N uptake efficiency $N_{effnorm}$ and the gradient G, which are used for estimating the nitrogen uptake N_g_2 and grain yield Y_g_2 in the target year.

$$diff_{SM_m} = SM_{fc} - SM_m\_2 \tag{10}$$

$$diff_{SM_{nu}}\_2 = SM_{fc} - SM_{nu}\_2 \tag{11}$$

from the normalized site-specific N uptake efficiency ($N_{effnorm}\_1$).

The final step of the model is to calculate the outputs: the total N in the above-soil-surface plant parts (N_{tot}_2), the percentual N content in grain (N_{gperc}_2) and the grain yield Y_g_2 for the target year. The total N is obtained using the N uptake efficiency (N_{eff}_2), N mineralized in the top soil during the growing season (N_m) and the total soluble N (N_{totsol}_2):

$$N_{tot}\_2 = (N_{eff}\_2 * N_m) + N_{totsol}\_2 \tag{12}$$

The goal of the fertilization is to produce a certain amount of kernels with a certain quality. To estimate the amount of grain yield and its protein content, the amount of N in grain is calculated by multiplying the total N with the nitrogen harvest index (NHI). NHI is an indicator of N allocation efficiency into grain [24]. It describes the proportion of the total N uptake (above-ground parts) accumulated in the grain. Grain of six-row barley contains around 80% of taken total N [25]. NHI calculated from the Viikki field's data (see Table 4) supports this finding, and thus the NHI gets the value 0.8 in the model. The nitrogen in grain yield in the target year is then:

$$N_g\_2 = 0.8 * N_{tot}\_2 \tag{13}$$

We define a variable G to denote the ratio of the proportion of N in grain to the grain yield. An assumption is made, that when the other growing conditions are kept constant, this ratio determines the grain yield, given the N uptake level. Based on the reference year, G can be calculated as:

$$G = Nc_g\_1/Y_g\_1 = N_g\_1/Y_g\_1^2 \tag{14}$$

Note that Y_g_1 is squared as it appears also in the proportion of N in grain.

We can now calculate the dry matter grain yield in the target year assuming that the plant available water (soil moisture) at yield formation stage would be the same as in the first year by:

$$Y_g\_21 = \sqrt{\frac{N_g\_2}{G}} \tag{15}$$

If the soil moisture conditions differ, the amount of yield needs to be corrected with an amount of yield that is calculated by multiplying the moisture difference $diff_{SM_{kf}}$ with a coefficient DMW, which is set as 60. DMW describes the amount of change in produced dry matter per unit of soil moisture difference. That is, 1 volume-% change in water content of the 60 cm soil layer equals change of 60 kg/ha of yield. This reflects the results for barley water use efficiency reported by Singh and Kumar [26]. The relevant period for soil moisture measurement is from ear emergence to yellow ripening. Thus the target year grain yield can be expressed:

$$Y_g\_2 = Y_g\_21 + diff_{SM_{kf}} * DMW \tag{16}$$

Table 1
Description of input factors with their ranges considered in the sensitivity analysis.

Name	Unit	Min	Max	Description
Factors concerning the reference year				
N_{in}_1	[kg/ha]	10.000	55.000	Initial N content in soil, year 1
pH_1	-	5.000	7.000	pH of the topsoil layer, year 1
N_f_1	[kg/ha]	0.000	180.000	Fertilizer input, year 1
Yg_1	[kg/ha]	850.000	8000.000	Grain yield, year 1
Ys_1	[kg/ha]	750.000	5300.000	Straw yield, year 1
Nc_g_1	-	0.015	0.030	Relative N content in grain, year 1
Nc_s_1	-	0.003	0.010	Relative N content in straws, year 1
ED_1	[nr/0.25 m^2]	70.000	200.000	Ear density, year 1
N_{org}_1	[%]	0.050	0.500	Organic N content in topsoil, year 1
SM_{nu}_1	[%]	6.800	33.000	Soil moisture at Veg state at 20 cm, year 1
SM_{kf}_1	[%]	8.400	34.000	Soil moisture at Grfill stage at 60 cm, year 1
Factors concerning the target year				
N_{in}_2	[kg/ha]	10.000	55.000	Initial N content in soil, year 2
pH_2	-	5.000	7.000	pH of the topsoil layer, year 2
N_f_2	[kg/ha]	0.000	180.000	Fertilizer input, year 2
ED_2	[nr/0.25 m^2]	70.000	200.000	Ear density, year 2
N_{org}_2	[%]	0.050	0.500	Organic N content in topsoil, year 2
SM_{nu}_2	[%]	6.800	33.000	Soil moisture at Veg stage at 20 cm, year 2
SM_m_2	[%]	7.000	31.000	Soil moisture at Miner stage at 20 cm, year 2
SM_{kf}_2	[%]	8.400	34.000	Soil moisture at Grfill stage at 60 cm, year 2
Factors common for both years				
MA	[%]	0.005	0.015	Microbial activity
VW	[kg/dm^3]	0.650	1.400	Volume weight of the topsoil
T_{soil}	[cm]	15.000	40.000	Topsoil depth
SM_{fc}	[%]	15.000	35.000	Soil moisture at field capacity at 20 cm
ED_{ref}	[nr/0.25 m^2]	160.000	165.000	Normalized ear density
N_{lux_limit}	[kg/ha]	175.000	185.000	N luxury uptake threshold
NHI	-	0.750	0.850	Nitrogen Harvest Index
DMW	[kg/ha]	55.000	65.000	Dry matter water productivity

where

$$diff_{SM_{kf}} = SM_{kf}\_2 - SM_{kf}\_1 \quad (17)$$

The percentual N content in grain in the target year is then:

$$N_{gperc}\_2 = (N_g\_2/Y_g\_2) * 100 \quad (18)$$

The model variables are listed in Table 1, Table 2 and Table 3. The model was originally designed using Stella® 2.0 Systems Modeling Software (ISEE Systems) and ported into Python for technical verification and further development. The source code is published in Zenodo, https://doi.org/10.5281/zenodo.16735443.

2.2. Validation data

The proposed model was validated using data from two test sites: the Viikki site (from years 1993, 1994 and 1995) and the Vihti site (years 2005 and 2014). At the Viikki site the crop variety was barley Kalle while in Vihti two different varieties, barley Saana and barley Annabel were grown in years 2005 and 2014, respectively. Also, the experiments at the two sites and different years differed in the plot size: from Vihti, average data over 5 × 0.25 m^2 plots or a single 0.25 m^2 plot was available while in Vihti the plot size was 1 m^2. The data on the crop variety, pH, VW, N_{org} and fertilization are described in detail in Table 4.

2.3. Sensitivity analysis

We performed global sensitivity analysis (GSA), primarily to explore and describe the behavior of the model. GSA can be used to analyze how the variation in the model outputs can be attributed to the different model inputs. Secondarily, a set of model parameters were tested for their effect on model behavior. Both the inputs and the internal parameters that are allowed to vary within specified bounds are referred to as *input factors* in the sensitivity analysis.

Table 2
Intermediate variables.

Name	Unit	Description
N_{soil}_i	kg/ha	Nitrogen in soil, year *i*
N_{tot}_i	kg/ha	Total N in above-soil-surface plant parts, year *i*
N_m	kg/ha	N mineralized in the top soil
N_{eff}	-	N uptake efficiency

Table 3
Output variables.

Name	Unit	Description
N_g_2	kg/ha	Nitrogen in grain, year 2
N_{gperc}_2	%	Percentual N content in grain, year 2
Y_g_2	kg/ha	Grain yield, year 2

A comprehensive sensitivity analysis would require a more thorough treatment, such as the ones done for HERMES [27], MONICA [28] and DSSAT [29]. Instead, our aim here is to complement the model description and provide additional insight into its behavior with a brief yet systematic sensitivity study.

We chose a density-based GSA method PAWN, as it supports both screening of uninfluential variables and ranking of influential ones. It allows analysis of a given input-output sample and therefore supports filtering out samples producing invalid outputs, unlike methods that require the use of a specific sampling strategy. This method is also robust against skewed output distributions.

Our analysis assesses the influence of each input factor within the range of variability in similar settings. The range in our validation data set was used as a basis, which was expanded to cover additional expected variability. Furthermore, selected fixed internal model parameters were allowed to vary around their set values to verify the models' insensitivity to their exact value. A dummy parameter was included in the

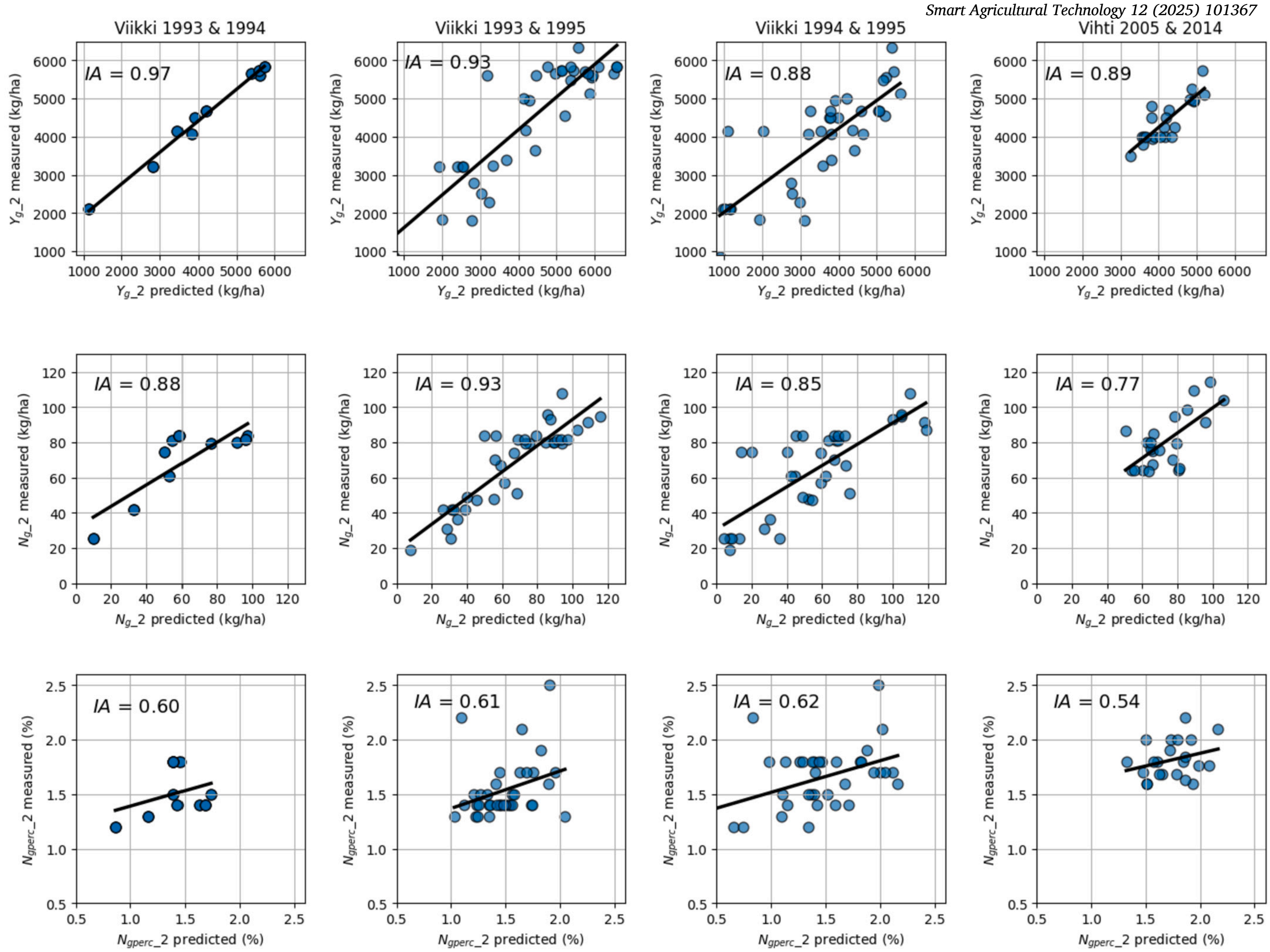

Fig. 3. Model validation results for the three output variables of dry matter grain yield (Y_g_2), N-uptake by grain (N_g_2) and N content in grain (N_{gperc}_2). For each pair of years the results of having measurements from the year 1 as a reference while predicting the outputs for the matching plots of the year 2 as well as using year 2 as a reference and predicting the outputs for year 1 are presented on the same graph.

input factors, which measures noise in the PAWN analysis and represents an expected index for a non-influential input, following Zadeh et al. [30]. The input factors and their ranges are described in Table 1. The input space was sampled using Latin hypercube sampling, where each input range is divided in N equally probable intervals for N samples. $N = 200,000$ was chosen for low noise, as measured by the dummy parameter. The model was run for each input sample, after which samples yielding an invalid output such as negative values were filtered out before the analysis. The sensitivity analysis was performed using the SALib software package ([31]).

PAWN compares the unconditional output distributions from the model runs where all inputs are allowed to vary simultaneously, with conditional output distributions where each input is fixed in turn to a specific value within their range. This difference is quantified by calculating the maximum absolute distance between the unconditional and conditional output distributions (Kolmogorov–Smirnov statistic, KS) at n fixed conditioning points of each input factor x_i. The sensitivity to a given input can vary in different regions of its range, and therefore the reported sensitivity index T_i for an input factor is then summarized by taking a statistic (the maximum in our case) of the KS statistics at the different conditioning points [32,33]:

$$T_i = \max_{x_i}[KS(x_i)], \tag{19}$$

where T_i takes a value between 0 and 1.

3. Results

3.1. Model validation

The model was run for each pair of years in the dataset, separately for both test sites. The results are shown in Fig. 3. Model performance is evaluated for three output variables: grain yield, grain N uptake and yield N % by calculating the index of agreement as:

$$IA = 1 - \frac{\sum_{i=1}^{n}(O_i - S_i)^2}{\sum_{i=1}^{n}(|O_i - \overline{O}| + |S_i - \overline{O}|)^2}, \tag{20}$$

where O_i are the observed values, S_i are the simulated values and $\overline{O}$ is the mean of the observed values.

The results indicate that the model performs better for predicting yield (IA 0.8–0.97) than for predicting N-uptake (IA 0.77–0.93). Yield N % appears to be more challenging to predict (IA 0.54–0.62) which can be anticipated since its value is derived from yield and N uptake, and the errors in their predictions are therefore combined.

Table 4
Description of the data used in model validation.

	Plot size	Crop	pH	VW [kg/dm3]	N_{org} %	Fertilization [kg/ha]	Other
Viikki 1993	Average of 5 x 0,25 m^2 plots	Six-row spring barley Kalle	6.3	1.411	5.7	70	Seeding with combined seed drill which places seed and fertilizers in different rows
			6.2	1.374	13.5	70	
			5.7	1.194	26.2	70	
			5.8	1.219	25.6	70	
			5.2	0.920	49.9	70	
Viikki 1994	Average of 5 x 0,25 m^2 plots	Six-row spring barley Kalle	6.0	1.411	5.7	40	Seeding with combined seed drill which places seed and fertilizers in different rows
			5.9	1.374	13.5	60	
			5.6	1.194	26.2	60	
			5.6	1.219	25.6	95	
			5.0	0.920	49.9	95	
Viikki 1995	Single 0,25 m^2 plot	Six-row spring barley Kalle	6.0	1.411	5.7	0	Seeding with combined seed drill which places seed and fertilizers in different rows
			6.0	1.411	5.7	60	
			6.0	1.411	5.7	120	
			6.0	1.411	5.7	180	
			6.2	1.374	13.5	0	
			6.2	1.374	13.5	60	
			6.2	1.374	13.5	120	
			5.5	1.194	26.2	0	
			5.5	1.194	26.2	60	
			5.5	1.194	26.2	120	
			5.8	1.219	25.6	0	
			5.8	1.219	25.6	60	
			5.8	1.219	25.6	120	
			5.0	0.920	49.9	0	
			5.0	0.920	49.9	60	
			5.0	0.920	49.9	120	
			5.7	0.920	49.9	30	
			5.0	0.920	49.9	30	
Vihti 2005	Single 1 m^2 plot	Two-row barley Saana	6.4	1.059	44	17	Seeding with VRA combined seed drill which places the seed and fertilizer in the same row
			6.7	0.863	31	30	
			6.9	0.675	32	20	
			6.3	0.675	31	20	
			6.3	1.139	41	20	
			7.0	0.959	17	37	
			6.8	0.703	17	37	
			6.6	0.705	20	37	
			6.4	0.697	30	24	
			6.4	0.724	32	19	
			6.4	0.902	32	20	
Vihti 2014	Single 1 m^2 plot	Two-row barley Annabel	6.1	1.059	44	40	Seeding with VRA combined seed drill which places the seed and fertilizer in the same row
			6.0	0.863	31	40	
			6.1	0.675	32	35	
			6.1	0.675	31	35	
			6.5	1.139	41	50	
			6.3	0.959	17	40	
			6.2	0.703	17	40	
			6.3	0.705	20	50	
			6.1	0.697	30	50	
			6.1	0.724	32	40	
			6.1	0.902	32	45	

3.2. Sensitivity analysis

The resulting PAWN sensitivity indices for the three main output variables, namely second year grain yield (Y_g_2), grain N content (N_{gperc}_2) and crop N uptake (N_{tot}_2), are shown in Fig. 4. A few factors dominate the model outputs while many seem uninfluential, including the fixed model internal parameters representing nitrogen harvest index (*NHI*), microbial activity (MA), luxury uptake limit (N_{lux_limit}) and water productivity (DMW).

4. Discussion

Validation The model appears to perform better on predicting yield and N uptake on the validation data, while N-content of the grain was more challenging to predict. This is comparable to the other more complex simulation models. HERMES model achieved IA of 0.93 and 0.71 for above-ground biomass and N-uptake, respectively [13] whereas MONICA achieved IA close to 1 for crop yield but less than 0.5 for crop N concentration and yield N concentration [14]. These are promising results, but more data on different crop types in more varied growing conditions would be required for making strong conclusions on generalizability of these predictions.

Sensitivity analysis The model behavior is dominated by a few factors while many seem rather uninfluential. This is common to environmental models with many parameters and inputs. The best predictor of the grain yield in the target year was grain yield in the reference year, followed by the fertilizer input, soil moisture at Grfill stage, relative N content in grain in the reference year, and ear density in the target year. The same parameters were also significant when predicting the grain N content, although in more uniform proportions. Grain N content was also affected by the fertilizer input in the reference year and the organic N content in top soil. The top three predictors of crop N uptake were fertilizer input, ear density, and grain yield in the reference year.

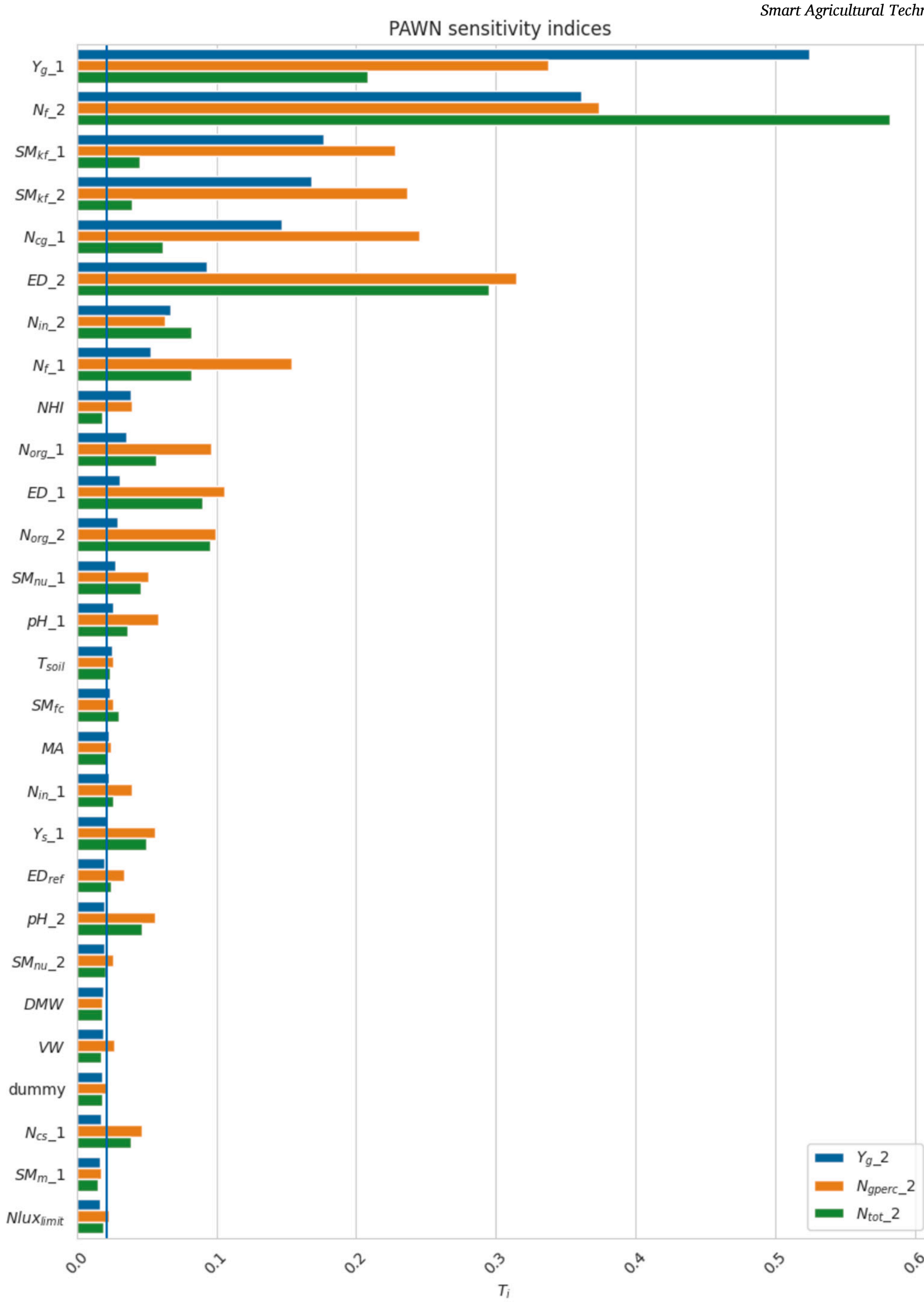

Fig. 4. PAWN sensitivity indices for the main target variables: second year grain yield (Y_g_2), grain N-content (N_{gperc}_2) and crop N-uptake (N_{tot}_2). The MA, VW and topsoil values were common for both years. A dummy parameter was included, which measures noise in the PAWN analysis and represents an expected index for a non-influential input. The results are sorted by decreasing influence to grain yield.

The vertical line in Fig. 4 indicates the sensitivity level of the dummy parameter. This level can be considered as a threshold under which the effect of the particular input factor has a negligible effect on the respective target variable. This is an important result when considering the real-life application of the model. In real-life use, collecting the data required by the model is often laborious and might prevent model adoption. The results of the sensitivity analysis indicate which input data are crucial and which inputs and parameters may be estimated from previous experience or adopted from the literature, for example.

In addition to PAWN, the Sobol and RDBFast sensitivity analysis methods were tested for comparison. They both showed similar results to PAWN. The main advantage of the PAWN method compared to Sobol was the possibility to provide an arbitrary input-output sample for analysis. Some critical viewpoints of the PAWN method have also been raised by Puy et al. [34] and Mora et al. [35].

Comparison to sensitivity analyses performed on other models can be indicative at best. The models and their parameters differ, but also the sensitivity analysis methods and target variables differ. In [29], the sensitivity analysis was done to DSSAT model using extended Fourier Amplitude Sensitivity Test (EFAST) for several variables. Of these variables, grain yield and grain protein content are the closest matches to our second year grain yield Y_g_2 and N-percentage in grain N_{gperc}_2 respectively. In their results, the most influential variables for grain yield were the conversion rate from photosynthetically active radiation to dry matter before the end of leaf growth (PARUE) and photosynthesis factor (SLPF). Our model, on the other hand, does not define such parameters, and it's yield output is most sensitive to the reference year yield along with nitrogen and soil moisture levels. This demonstrates that these models arrive at their results from different premises.

Fitness-for-purpose The results show promise for predicting the yield in given fertilization conditions, with a relatively simple theory-based computational model. In addition to its relative simplicity, this kind of two-stage (i.e., reference year – target year) model has the advantage of following the long-term changes in climate or other conditions affecting crop growth. By selecting different reference years, the farmer can adapt the use of the model and at the same time acquire new knowledge on the response in the productivity of the crop field to environmental conditions. The more the reference and target years differ from one another in growth conditions, the higher the anticipated prediction error. When the field observations required by the model are from several different growing seasons, different reference years could be used as starting points in the modeling when making risk assessments for fertilization, irrigation or seeding density plans.

As indicated by Fig. 3, the model is relatively robust with respect to the temporal difference between the reference and target years as well as the selection of test sites. In case of the Vihti site, there is a nine year difference between the reference and target years, however, the IA for grain yield is close to that of the Viikki site where the difference is one or two years. The IA of N uptake and N % in grain remains lower for the Vihti site though. Furthermore, the discrepancy between the model and the actual values could also be explained by the small size of the crop samples relative to the within-stand variation, even over a short distance in the cereal crop. Hannu Haapala's thesis [36] provides an example of this. Further validation of the model is required for better evaluation of the effect of the difference in environmental conditions of the reference and target years on the prediction performance of the model.

The Viikki 93 and 94 samples were collected from exactly the same location, the yield was averaged over five $0.25\,m^2$ samples. The samples from Viikki 95 were $1-5\,m$ off the 93–94 sample locations and there was only one $0.25\,m^2$ sample per observation. So that one $0.25\,m^2$ sample represents the entire site to varying degrees. In Vihti, the sample sites in different years were located in the same places with DGPS accuracy, but presumably did not coincide exactly in the same place. The sample size was larger, $1\,m^2$, but still may not represent the site sufficiently. In addition, the barley variety was different in Vihti in different years, which may also explain the difference in nitrogen uptake efficiency.

In addition to fertilization, the model can also be used to examine the effect of seeding density and irrigation to N uptake, yield amount and yield quality. The practicality of the required data acquisition remains an open question, while the sensitivity analysis hints at a possibility to simplify the model even further or to get approximate results even in the absence of all required measurements. The low computational requirements of this model also readily facilitate probabilistic treatment, such as providing uncertainty distributions for the model outputs using monte carlo simulations.

5. Conclusions

Choosing the nitrogen fertilization rate for field crops is a central optimization problem in open field farming. We presented a new process-based model for predicting nitrogen uptake in cereal crops. The model estimates the effect of given fertilizer input to yield using N uptake efficiency measured during a reference year. The model was validated for barley with data from two sites that were controlled for weeds and pests. A sensitivity analysis was performed to further describe the behavior of the model and identify the most and least influential inputs to the model.

Validation The validation results indicated that the model performs reasonably well for the yield and N uptake whereas the grain N % appears to be more challenging to predict.

Sensitivity analysis Sensitivity analysis identified influential and non-influential variables of the model. Knowing the influential variables can be used to direct data collection efforts where most precision is needed. Non-influential variables can more readily be estimated with less precision or potentially simplified out of the model with minimal loss of accuracy.

Significance/implications The presented model adds to the process-based approaches for estimating the effect of fertilizer input to the yield on a given field plot. The advantage of the model is the limited number of quantities to be measured, so it can be applied in practice on farms using smart farming technology. The model helps farmers to focus on the accuracy and quality of the measurements.

Limitations and further study The model was validated on barley on two separate sites. Data from seasons with different weather conditions and sites with varying soil types and management practices as well as other crop types and cultivars would further improve confidence in the results and their generalizability.

CRediT authorship contribution statement

Antti Halla: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Liisa Pesonen:** Writing – original draft, Resources, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jari Pohjola:** Writing – original draft, Visualization, Validation, Software, Investigation, Formal analysis, Data curation. **Petri Linna:** Writing – review & editing, Resources, Project administration, Funding acquisition. **Tarmo Lipping:** Writing – review & editing, Writing – original draft, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Source code and data for reproducing the results of this article are published in Zenodo, https://doi.org/10.5281/zenodo.16735443.

References

[1] N. Lehnert, H.T. Dong, J.B. Harland, A.P. Hunt, C.J. White, Reversing nitrogen fixation, Nat. Rev., Chem. 2 (10) (2018) 278–289, https://doi.org/10.1038/s41570-018-0041-7.

[2] T. Engel, Use of nitrogen simulation models for site-specific nitrogen fertilization, in: J. Stafford (Ed.), Proceedings of the First European Conference on Precision Agriculture, SCI, London, 1997, pp. 361–369.

[3] J.P. Schimel, J. Bennett, Nitrogen mineralization: challenges of a changing paradigm, Ecology 85 (3) (2004) 591–602, https://doi.org/10.1890/03-8002.

[4] G.E. Brust, Chapter 9 - management strategies for organic vegetable fertility, in: D. Biswas, S.A. Micallef (Eds.), Safety and Practice for Organic Food, Academic Press, 2019, pp. 193–212.

[5] W.C. Honeycutt, T.S. Griffin, Z. He, Manure nitrogen availability: dairy manure in northeast and Central U.S. soils, Biol. Agric. Hortic. 23 (2) (2005) 199–214, https://doi.org/10.1080/01448765.2005.9755320.

[6] M.J. Foulkes, G.A. Slafer, W.J. Davies, P.M. Berry, R. Sylvester-Bradley, P. Martre, D.F. Calderini, S. Griffiths, M.P. Reynolds, Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance, J. Exp. Bot. 62 (2) (2010) 469–486, https://doi.org/10.1093/jxb/erq300.

[7] R. Fischer, M. Stapper, Lodging effects on high-yielding crops of irrigated semidwarf wheat, Field Crops Res. 17 (3) (1987) 245–258, https://doi.org/10.1016/0378-4290(87)90038-4.

[8] M. Zhang, H. Wang, Y. Yi, J. Ding, M. Zhu, C. Li, W. Guo, C. Feng, X. Zhu, Effect of nitrogen levels and nitrogen ratios on lodging resistance and yield potential of winter wheat (Triticum aestivum L.), PLoS ONE 12 (11) (2017) e0187543, https://doi.org/10.1371/journal.pone.0187543.

[9] J.P.D.M.E. Abreu, I. Flores, F.M.G. De Abreu, M.V. Madeira, Nitrogen uptake in relation to water availability in wheat, Plant Soil 154 (1) (1993) 89–96, https://doi.org/10.1007/BF00011076.

[10] K.I. Paul, P.J. Polglase, A.M. O'Connell, J.C. Carlyle, P.J. Smethurst, P.K. Khanna, Defining the relation between soil water content and net nitrogen mineralization, Eur. J. Soil Sci. 54 (1) (2003) 39–48, https://doi.org/10.1046/j.1365-2389.2003.00502.x.

[11] K. Kersebaum, Application of a simple management model to simulate water and nitrogen dynamics, Ecol. Model. 81 (1) (1995) 145–156, https://doi.org/10.1016/0304-3800(94)00167-G.

[12] K.C. Kersebaum, Modelling nitrogen dynamics in soil–crop systems with HERMES, Nutr. Cycl. Agroecosyst. 77 (1) (2007) 39–52, https://doi.org/10.1007/s10705-006-9044-8.

[13] R. Malone, K. Kersebaum, T. Kaspar, L. Ma, D. Jaynes, K. Gillette, Winter rye as a cover crop reduces nitrate loss to subsurface drainage as simulated by HERMES, Agric. Water Manag. 184 (2017) 156–169, https://doi.org/10.1016/j.agwat.2017.01.016.

[14] C. Nendel, M. Berg, K. Kersebaum, W. Mirschel, X. Specka, M. Wegehenkel, K. Wenkel, R. Wieland, The MONICA model: testing predictability for crop growth, soil moisture and nitrogen dynamics, Ecol. Model. 222 (9) (2011) 1614–1625, https://doi.org/10.1016/j.ecolmodel.2011.02.018.

[15] G. Hoogenboom, C. Porter, K. Boote, K. Shelia, P. Wilkens, U. Singh, J. White, S. Asseng, J. Lizaso, L. Moreno, W. Pavan, R. Ogoshi, L. Hunt, G. Tsuji, J. Jones, The DSSAT crop modeling ecosystem, in: K. Boote (Ed.), Advances in Crop Modeling for a Sustainable Agriculture, Burleigh Dodds Science Publishing, Cambridge, United Kingdom, 2019, pp. 173–216.

[16] D. Cammarano, B. Basso, J. Holland, A. Gianinetti, M. Baronchelli, D. Ronga, Modeling spatial and temporal optimal N fertilizer rates to reduce nitrate leaching while improving grain yield and quality in malting barley, Comput. Electron. Agric. 182 (2021) 105997, https://doi.org/10.1016/j.compag.2021.105997.

[17] A.-C. Hansson, R. Pettersson, K. Paustian, Shoot and root production and nitrogen uptake in Barley, with and without nitrogen fertilization, J. Agron. Crop Sci. 158 (3) (1987) 163–171, https://doi.org/10.1111/j.1439-037X.1987.tb00258.x.

[18] R. Tate, Process control in soil, in: Soil Microbiology, John Wiley & Sons, Ltd., 2020, pp. 149–184, Ch. 5.

[19] W. Grzebisz, A. Niewiadomska, Nitrogen cycle in farming systems, Agronomy 14 (1) (2024) 89, https://doi.org/10.3390/agronomy14010089.

[20] A. Irmak, W.D. Batchelor, J.W. Jones, S. Irmak, J.O. Paz, H.W. Beck, M. Egeh, Relationship between plant available soil water and yield for explaining soybean yield variability, Appl. Eng. Agric. 18 (4) (2002) 471, https://doi.org/10.13031/2013.8748.

[21] Y. Tan, Q. Chai, G. Li, C. Zhao, A. Yu, Z. Fan, W. Yin, F. Hu, H. Fan, Q. Wang, Y. Guo, X. Tian, Improving wheat grain yield via promotion of water and nitrogen utilization in arid areas, Sci. Rep. 11 (1) (2021) 13821, https://doi.org/10.1038/s41598-021-92894-6.

[22] G. Preiti, A. Calvi, M. Romeo, G. Badagliacca, M. Bacchi, Seeding density and nitrogen fertilization effects on agronomic responses of some hybrid barley lines in a Mediterranean environment, Agronomy 11 (10) (2021) 1942, https://doi.org/10.3390/agronomy11101942.

[23] S. Dahiya, S. Kumar Harender, C. Chaudhary, Lodging: significance and preventive measures for increasing crop production, Int. J. Chem. Stud. 6 (1) (2018) 700–705.

[24] N.K. Fageria, Nitrogen harvest index and its association with crop yields, J. Plant Nutr. 37 (6) (2014) 795–810, https://doi.org/10.1080/01904167.2014.881855.

[25] S. Muurinen, J. Kleemola, P. Peltonen-Sainio, Accumulation and translocation of nitrogen in spring cereal cultivars differing in nitrogen use efficiency, Agron. J. 99 (2) (2007) 441–449, https://doi.org/10.2134/agronj2006.0107.

[26] K.P. Singh, V. Kumar, Water use and water-use efficiency of wheat and barley in relation to seeding dates, levels of irrigation and nitrogen fertilization, Agric. Water Manag. 3 (4) (1981) 305–316, https://doi.org/10.1016/0378-3774(81)90014-7.

[27] J.W.M. Pullens, K.C. Kersebaum, U. Böttcher, H. Kage, J.E. Olesen, Model sensitivity of simulated yield of winter oilseed rape to climate change scenarios in Europe, Eur. J. Agron. 129 (2021) 126341, https://doi.org/10.1016/j.eja.2021.126341.

[28] X. Specka, C. Nendel, R. Wieland, Analysing the parameter sensitivity of the agroecosystem model MONICA for different crops, Eur. J. Agron. 71 (2015) 73–87, https://doi.org/10.1016/j.eja.2015.08.004.

[29] Z.-h. Li, X.-l. Jin, H.-l. Liu, X.-g. Xu, J.-h. Wang, Global sensitivity analysis of wheat grain yield and quality and the related process variables from the DSSAT-CERES model based on the extended Fourier amplitude sensitivity test method, J. Integr. Agric. 18 (7) (2019) 1547–1561, https://doi.org/10.1016/S2095-3119(18)62046-5.

[30] F.K. Zadeh, J. Nossent, F. Sarrazin, F. Pianosi, A. van Griensven, T. Wagener, W. Bauwens, Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model, Environ. Model. Softw. 91 (2017) 210–222, https://doi.org/10.1016/j.envsoft.2017.02.001.

[31] J. Herman, W. Usher, SALib: an open-source python library for sensitivity analysis, J. Open Source Softw. 2 (9) (2017) 97, https://doi.org/10.21105/joss.00097.

[32] F. Pianosi, T. Wagener, A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environ. Model. Softw. 67 (2015) 1–11, https://doi.org/10.1016/j.envsoft.2015.01.004.

[33] F. Pianosi, T. Wagener, Distribution-based sensitivity analysis from a generic input-output sample, Environ. Model. Softw. 108 (2018) 197–207, https://doi.org/10.1016/j.envsoft.2018.07.019.

[34] A. Puy, S. Lo Piano, A. Saltelli, A sensitivity analysis of the PAWN sensitivity index, Environ. Model. Softw. 127 (2020) 104679, https://doi.org/10.1016/j.envsoft.2020.104679.

[35] E.B. Mora, J. Spelling, A.H. van der Weijde, Benchmarking the PAWN distribution-based method against the variance-based method in global sensitivity analysis: empirical results, Environ. Model. Softw. 122 (2019) 104556, https://doi.org/10.1016/j.envsoft.2019.104556.

[36] H.E.S. Haapala, Position dependent control (PDC) of plant production, Agric. Food Sci. 4 (3) (1995) 239–350, https://doi.org/10.23986/afsci.72612.