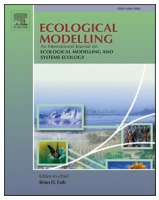
Ecological Modelling 483 (2023) 110433

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



NORNE, a process-based grass growth model accounting for within-field soil variation using remote sensing for in-season corrections

Anne-Grete Roer Hjelkrem *, Jakob Geipel, Anne Kjersti Bakken, Audun Korsaeth

Division of Food Production and Society, Norwegian Institute of Bioeconomy Research (NIBIO), 1431 Ås, Norway

ARTICLE INFO

Keywords:
Calibration
Drone
Dynamic model
Precision nitrogen fertilization
Yield

ABSTRACT

A process-based model was developed to predict dry matter yields and amounts of harvested nitrogen in conventionally cropped grassland fields, accounting for within-field variation by a node network design and utilizing remotely sensed information from a drone-borne system for increased accuracy. The model, named NORNE, was kept as simple as possible regarding required input variables, but with sufficient complexity to handle central processes and minimize prediction errors. The inputs comprised weather data, soil information, management data related to fertilization, and a visual estimate of clover proportion in the aboveground biomass. A sensitivity analysis was included to apportioning variation in dry matter yield outputs to variation in model parameter settings. Using default parameter values from the literature, the model was evaluated on data from a two-year study (2016–2017, 264 research plots in total each year) conducted at two locations in Norway (i.e. in South-East and in Central Norway) with contrasting climatic conditions and with internal variation in soil characteristics. The results showed that the model could estimate dry matter yields with a relatively high accuracy without any corrections based on remote sensing, compared with published results from comparable model studies. To further improve the results, the model was calibrated shortly before harvest, using predictions of above ground dry matter biomass obtained from a drone-borne remote sensing system. The only parameters which were hereby adjusted in the NORNE model were the starting values of nitrogen content in soil (first cut) and the plant available water capacity (second cut). The calibration based on the remotely sensed information improved the predictive performance of the model significantly. At first cut, the root mean square error (RMSE) of dry matter yield prediction was reduced by 20% to a mean value of 58 g m^{-2}, corresponding to a relative value (rRMSE) of 0.12. For the second cut, the RMSE decreased by 13% to 66 g m^{-2} (rRMSE: 0.18). The model was also evaluated in terms of the predictions of amounts of nitrogen in the harvested crop. Here, the calibration reduced the RMSE of the first cut by 38%, obtaining a mean RMSE value of 2.1 g N m^{-2} (rRMSE: 0.28). For the second cut, the RMSE reduction for simulated harvested N was 16%, corresponding to a mean RMSE value of 2.3 g N m^{-2} (rRMSE: 0.33). The large improvements in model accuracy for simulated dry matter and nitrogen yields obtained through calibration by utilizing remotely sensed information, indicate the importance of considering spatial variability when applying models under Nordic conditions, both for yield predictions and for decision support for nitrogen application.

1. Introduction

Forage based livestock production is a cornerstone in Norwegian agriculture both in terms of land use and economic value (Steinshamn et al., 2016). Its position, legitimacy and public support depend on its environmental performance, regarding resource use efficiency and emissions of greenhouse gases, acidifying and eutrophicating compounds. Tuning the application of nitrogen fertilizer to forage crop demand and production potential is regarded a key challenge in this context, both in time and space.

Precision agriculture is about matching resource application and agronomic practices with soil and crop requirements as they vary in space and time within a field (Whelan and McBratney 2000). The concept is usually associated with precision fertilization, that seeks to synchronize nutrient application with plant demand, accounting for spatial variation in order to maximize crop yields. High within field variation in nitrogen requirements is common because of highly varying crop status and soil characteristics, but still, fertilizers are commonly

* Corresponding author.
E-mail address: anne-grete.hjelkrem@nibio.no (A.-G.R. Hjelkrem).

https://doi.org/10.1016/j.ecolmodel.2023.110433
Received 6 December 2022; Received in revised form 8 May 2023; Accepted 31 May 2023
Available online 6 June 2023

applied at uniform rate over the entire field, according to the "average regime", in which yield expectations are assessed on field scale. To be able to implement precision fertilization to a forage crop in practice, it is crucial to develop a technological solution that enables predictions of required nitrogen inputs site-specifically within field (Stafford 2000; Auernhammer 2001).

By describing the underlying mechanisms in the system, process-based dynamic models may simulate the interactions between vegetation and environment. Such models are widely designed to mimic production and phenological development of forage grasses to better understand the underlying biological and physical processes, or to forecast yield and quality development for decision support in fodder production. Compared to empiric models, process-based models tend to be more readily adopted in grassland modeling because of their superior robustness, accuracy and flexibility of application (Hurtado-Uria et al., 2013). While process-based models rely on theory obtained from a range of experimental approaches, empirical models including machine learning models rely on observations. Machine learning can easily identify trends and patterns in data, but the method requires massive data sets to train generalizable models. Such models will predict accurately the behavior of a system for conditions similar to the training data, contrary to the process-based models that should work well in all the situations for which they are developed. Many of the sub-processes in crop growth are uncertain both in nature and their interactions, and as models often perform worse when unnecessary complexity is added, they should be kept as simple as possible. Under Nordic conditions, the process-based grass simulation models BASGRA (Höglind et al., 2016), BASGRA_N (Höglind et al., 2020) and CATIMO (Bonesmo and Bélanger 2002a, b), and the soil-crop model STICS (Brisson et al., 2003), simulate canopy growth as functions of weather conditions, soil characteristics and crop management factors. They are shown to work well under northern conditions (Korhonen et al., 2018), simulating yield development in pure stands of timothy at a field scale level. Neither of the models managed, however, to simulate yield response adequately to gradients in nitrogen fertilization rates (Korhonen et al., 2018). Timothy is commonly grown in mixtures with clover, but the mentioned simulation models have not been evaluated for that kind of stands because they do not account for clover nitrogen fixation.

Remotely sensed data may provide an instant and non-destructive estimate of grassland traits and, therefore, minimize costs and labor (Wachendorf et al., 2018). Recent studies demonstrated that the use of aircrafts and unmanned aerial vehicles (UAV, in the following referred to as drones) equipped with multi-/hyperspectral imagers can obtain good predictions of both grassland yield and quality through thorough radiometric calibration, multivariate statistics and machine learning (Capolupo et al., 2015; Geipel et al., 2021; Lussem et al., 2022; Pullanagari et al., 2018, 2016; Wijesingha et al., 2020).

Assuming that the remotely sensed information is more accurate at the time of observation than the corresponding prediction of the process-based growth model, the former may be used to update the latter. Few studies have emphasized the topic of improving the predictions of a process-based model by dynamic updates using remotely sensed information.

Various methods on how to couple remotely sensed data with crop models have been reviewed (Kasampalis et al., 2018; Jin et al., 2018; Moulin et al., 1998), but only the study by Clevers et al. (2002) have actually performed such a coupling. This study uses SPOT satellite data to successfully calibrate a wheat growth model during run-time. The study by Jin et al. (2018) focused on the crucial question on which method to select for the actual model update. Different options were compared and concluded that calibration was the best option.

The overall aim of this work was to develop a tool that enables farmers to time harvests and target nitrogen inputs in their forage production according to prevailing yield potential. For this purpose, the NORNE-model, a process-based model to predict daily dry matter yield and amount of nitrogen in yield according to nitrogen demand and availability, based on soil characteristics, sward clover proportion, local weather, and field management information was designed and tested. The model was designed to handle within field variation and to allow for in-season calibration of central model parameters, based on remotely sensed information and resulting estimates on above ground standing biomass. The model was kept relatively simple regarding required input variables and complexity in form of number of routines. To be able to cover within field variations, extra focus was placed on the water and nitrogen sub-models. Finally, a sensitivity analysis was performed to identify to what degree predicted dry matter yield was influenced by variation in settings of altogether 43 model parameters.

2. Material and methods

2.1. The core concept of NORNE

In order to develop a tool that enables farmers to time harvests and target nitrogen inputs in their forage production according to prevailing yield potential, a process-based model, the NORNE model, was developed to predict grass growth and amount of nitrogen in yield on a daily basis for a node network within the field (Fig. 1). To further improve prediction accuracy, the system allows for in-season and node-wise model calibration, using predictions of dry matter biomass from drone-borne remote sensing.

NORNE is an expansion of Grovfôrmodellen (Bakken 2016), which is an online decision support tool that simulates forage yield and quality development in species and species mixtures grown and harvested for silage production in Norway. In contrast to NORNE, Grovfôrmodellen has no routines that address and predict within field variation, and crop nitrogen content (and uptake) is expressed by a simple function of calculated standing crop biomass according to the critical nitrogen curve (Greenwood et al., 1990). Sub-optimal nitrogen supply according to crop demand is neither predicted nor allowed to retard crop growth rate.

The NORNE model requires daily input of weather data (air temperature, soil temperature, precipitation, global radiation, wind speed and relative humidity), soil type (sand, silty sand, sandy loam, loam, sandy silt, silty loam, clay loam, clay, silty clay loam and silt), clover percentage in the sward and some management data (timing and amount of nitrogen fertilization and timing of harvests). Field capacity, plant available water capacity and texture data are set according to Riley (2021) for the given soil type.

The model consists of three sub-models: (i) the plant growth and dry matter partitioning sub-model, (ii) the water sub-model and (iii) the nitrogen sub-model, that are further described in the sub-sections below.

2.1.1. The plant growth and dry matter partitioning sub-model

The NORNE model applies to daily (d) growth of grassland swards during the growing season in years of ley (not to the year of establishment). Growth start in spring was estimated to be the third day when a five-day floating average air temperature exceeded 5 °C, and the corresponding soil temperature (at 10 cm depth) exceeded 1 °C. Then, the potential daily growth (ΔDM_{pot}) was estimated as a sigmoid function of leaf area index (LAI) (Eq. (1)).

$$\Delta DM_{pot}(d) = \Delta DM_{max} \Big/ \left(1 + e^{a/(R_m - b \cdot LAI(d))}\right) \quad (1)$$

With ΔDM_{max} (maximum daily growth), a, R_m and b being fixed values set according to Grovfôrmodellen. LAI was defined as the one-sided green leaf area per ground area ($m^2\ m^{-2}$) and was calculated in line with Grovfôrmodellen. In line with Torssell and Kornher (1983), daily dry matter growth (ΔDM_{yield}) was further given by the potential growth limited through indices to account for effects of air temperature (TI), solar radiation (SI), age (AI) and water availability (WI) (Eq. (2)). The NORNE model also includes a module for nitrogen availability, and the potential growth is allowed to be limited by nitrogen supply according to the score of nitrogen index (NI) (Eq. (2)).

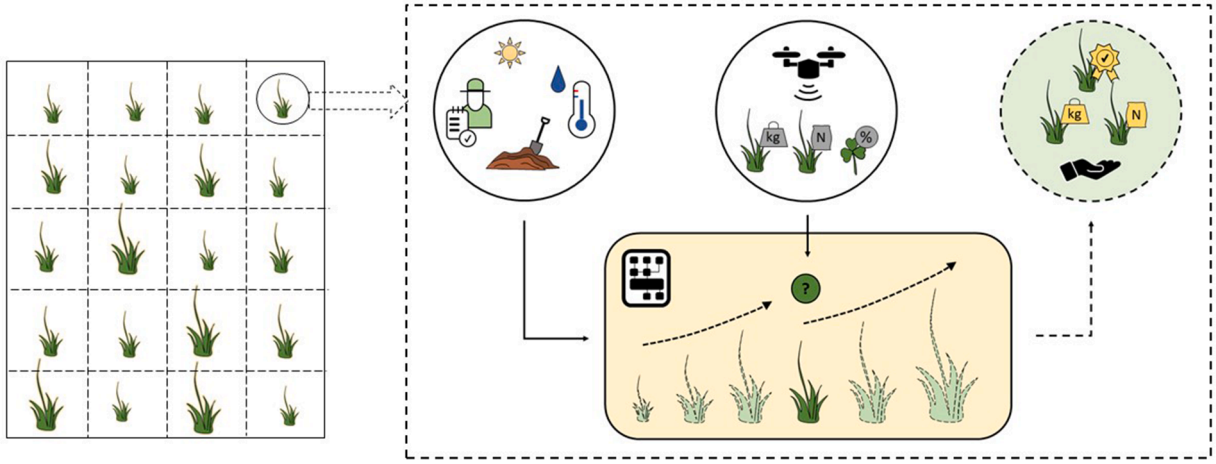


Fig. 1. Overview of the NORNE concept, which includes a process-based model to predict crop growth and amount of nitrogen in yield for a node network within field. To further improve prediction accuracy, the system allows for in-season model calibration of dry matter biomass predictions from drone-borne remote sensing.

$$\Delta DM_{yield}(d) = \Delta DM_{pot}(d) \cdot \min(TI(d), SI(d), WI(d)) \cdot NI(d) \cdot AI(d) \quad (2)$$

To form daily dry matter yield (g m^{-2}), a simple accumulation of the daily growth was performed. In addition to dry matter yield, the model estimates dry matter (Eq. (3)) of residues, stubble after harvest, roots in top soil layer (0 to 0.25 m below ground) and roots in sub soil level (0.25 to 0.5 m below ground).

$$DM_{tot}(d) = DM_{yield}(d) + DM_{stub}(d) + DM_{residues}(d) + DM_{roots,ts}(d) + DM_{roots,ss}(d) \quad (3)$$

Based on empirical data from field trials (unpublished), dry matter of residues ($DM_{residues}$) was set to be 2% of the dry matter yield. A standard stubble height of 0.07 m was assumed which is in agreement with the stubble height of the field experiments producing the data used in this study. The fraction of total above-ground dry matter ($DM_{yield} + DM_{stub} + DM_{residues}$) allocated to stubble was formulated as a declining function, adapted from the ENGNOR model (Baadshaug and Lantinga 2002), starting with allocating 85% of the total above-ground dry matter to stubble in the beginning of the growth period. After reaching an above-ground dry matter biomass of 460 g m^{-2}, no further allocation to the stubble was assumed. Root biomass was calculated according to an assumed root-shoot ratio of 1.15 (Fystro 1995). 90% of roots' dry matter was estimated to be in the top soil layer ($DM_{roots,ts}$), and the remaining 10% in the sub soil layer ($DM_{roots,ss}$) (Fystro 2001). At each harvest, 50% of the roots were assumed to die (Kertulis 2001).

2.1.2. Indices of aging, air temperature and solar radiation

The aging index, *AI*, retards daily growth rate until first cut according to the phenological stage of the grass crop. The stage is expressed as mean stage by count (MSC) (Moore et al., 1991), attaining values on a continuous scale from 1 to 4. *AI* is set to 1 until MSC exceeds 3, which corresponds to the stage when the inflorescence has emerged on 50% of the shoots. Afterwards, *AI* declines until MSC reaches 4, where it is set to 0. The development of MSC is predicted according to an equation adapted from Bonesmo (2004), with air temperature as the only driving variable. The original day length function was excluded because most first cuts in silage production systems are taken when the day length is longer than the threshold for retardation of phenological advancement. In regrowth after the first cut, there is no clear relationship between sward structure as expressed by MSC and aging, and both the MSC sub-model and *AI* are left out for this part of the season. Earlier versions of the functions leading to scores on *AI* were defined by Angus et al. (1980) and Torsell and Kornher (1983).

The temperature index, *TI*, was scored according to a bell-shaped function of air temperature, scaled between 0 and 1. The optimum was set at 17 °C, being the temperature corresponding to maximum growth (Torssell et al., 1982). Further, no growth was assumed at temperatures below 0 or above 33 °C (Torssell et al., 1982).

The solar radiation index, *SI*, was calculated according to an inverse exponentially growing function of global radiation. The function was scaled between 0 and 1 and was adapted from Torssell et al. (1982).

2.1.3. The water sub-model and the water index (WI)

The amount of water in top (W_{ts}) and sub soil (W_{ss}) layer was calculated daily in mm (l m^{-2}), following Eq. (4) and Eq. (5) (Torssell et al., 1982).

$$W_{ts}(d) = W_{ts}(d-1) + P(d) - AE_{ts}(d) - W_{percolation}(d) \quad (4)$$

$$W_{ss}(d) = W_{ss}(d-1) - AE_{ss}(d) + W_{percolation}(d) - W_{leaching}(d) \quad (5)$$

At growth start in spring, the soil moisture was assumed to be at field capacity in both soil layers, retrieved from soil type according to Riley (2021). Daily precipitation (*P*) was added to the water content in the top soil layer, and actual evapotranspiration (*AE*) subtracted from the water pool in both layers. When water content in the top soil layer exceeded field capacity, redundant water was assumed to percolate to the sub soil layer ($W_{percolation}$). Equally, when water content in the sub soil layer reached field capacity, redundant water was assumed to leach to the ground ($W_{leaching}$).

The potential evapotranspiration (*PE*) was estimated according to a function described in Riley and Berentsen (2009). This function has been developed and tested under Norwegian conditions, and requires inputs of weather data (global radiation, wind speed and vapor pressure deficit that is calculated from air temperature and relative humidity) and month number.

AE was calculated according to Torssell et al. (1982) (Eq. (6)). It was set to zero when water content was below the plant available water capacity.

$$AE(d) = PE(d) \cdot \frac{W_{ts}(d-1) + W_{ss}(d-1)}{FC_{ts} + FC_{ss}} \quad (6)$$

An index for water availability (*WI*) was constructed to retard simulated growth during periods with water deficit. It was defined as the ratio of actual to potential evapotranspiration and scaled between 0 and 1.

2.1.4. The nitrogen sub-model and the nitrogen index (NI)

Nitrogen content in soil was calculated for both soil layers and included the processes of fertilization, clover nitrogen fixation, nitrification, denitrification, percolation, leaching, mineralization,

humification and root death (Fig. 2). The model distinguished three pools of organic nitrogen: readily decomposable (litter 1), slowly decomposable (litter 2) and very slowly decomposable (humus), while the mineral nitrogen was split between ammonia (NH_4) and nitrate (NO_3) (Vold and Søreng 1997).

Mineral nitrogen application was added to the nitrogen content in soil, initially assumed to be 50% ammonia and 50% nitrate. Only the nitrogen content in form of nitrate was assumed to percolate or leach, and its amount was estimated from the nitrate concentration in the soil water and the amount of water that percolated or leached. Nitrification is the microbiological oxidation of ammonia to nitrate. The fraction of the ammonia nitrified to nitrate was calculated from soil temperature and soil water content in line with Sierra et al. (2003). Denitrification is the microbial process of reducing nitrate to gaseous forms of nitrogen. The calculation was based on the nitrate content in soil, soil temperature and water filled pore space in line with Heinen (2006). Decomposition of organic nitrogen compounds is the result of two contrasting processes, mineralization and humification. It was estimated separately for the two soil layers and for each of the organic components, litter 1, litter 2 and humus. Nitrogen mineralization was simulated as a first order decay, using rate constants from Vold and Søreng (1997) and Korsaeth et al. (2003) and soil climate rate that accounts for temperature, water and cultivar effect based on Andrén et al. (2004). Mineralization is the process by which organic nitrogen is converted to plant available inorganic forms and was estimated to account for 99% of the decomposition whereas the remaining 1% was regarded as humification. A fraction of the root dies at harvest, and the nitrogen content within this dry matter was retained to the soil. In line with the distribution of the root mass, 90% of this nitrogen was retained in the top soil layer and the remaining 10% in the sub soil layer. It was equally split between the organic components litter 1 and litter 2. Clover has the ability to fixate nitrogen from the air to the soil, and the estimation of the clover nitrogen fixation in the model was adapted from Lazzarotto et al. (2009) and depended on soil temperature, soil mineral nitrogen, dry matter of roots and clover proportion. All fixated nitrogen from clover was assumed directly taken up by the plant.

Nitrogen in mineral form was taken up by the plant from the soil on a daily basis, together with the daily clover nitrogen fixated. A maximum plant nitrogen concentration was set as a function of dry matter yield and used to limit the maximum daily nitrogen uptake. Further, the nitrogen uptake was partitioned between the different parts of the plant (Eq. (7)), which was done in accordance with Baadshaug and Lantinga (2002).

$$N_{tot}(d) = N_{yield}(d) + N_{stub}(d) + N_{residues}(d) + N_{roots,ts}(d) + N_{roots,ss}(d) \quad (7)$$

Nitrogen scarcity is another growth limiting factor, and the index for nitrogen availability (*NI*) was constructed to inhibit growth during periods with nitrogen deficit. *NI* was defined as the ratio of a plant's current nitrogen concentration to the plant's critical nitrogen concentration and scaled between 0 and 1. A plant's critical nitrogen concentration is defined as the minimum plant nitrogen concentration allowing for maximum growth rate (Greenwood et al., 1990). It declines with increased dry matter yield and was estimated in line with Grovfôrmodellen.

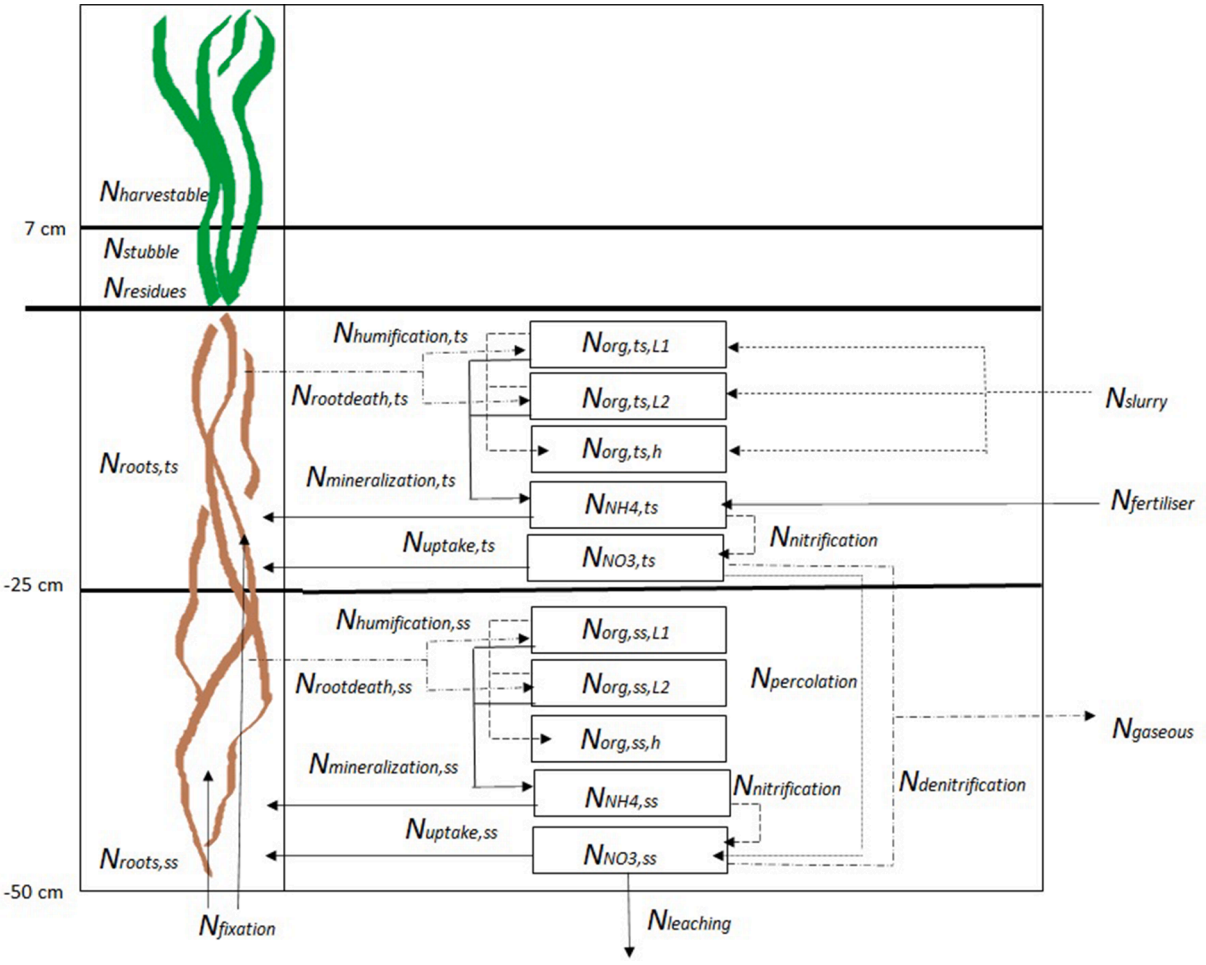


Fig. 2. Overview of the sub-processes included in the nitrogen sub-model, where nitrogen content in soil was calculated for both soil layers and included the processes of fertilization, clover nitrogen fixation, nitrification, denitrification, percolation, leaching, mineralization, humification and root death.

2.2. Field experimental data

Field experimental data from a previous study conducted at two locations in Norway (Geipel et al., 2021) were used for model calibration and evaluation. In South-East Norway, the field experiment was located at Apelsvoll (60.70 °N, 10.87 °E, 251 m above sea level) research station, with a humid continental climate. The soil was an imperfectly drained loam. In Central Norway, the field experiment was located at Kvithamar (63.49 °N, 10.87 °E, 25 m above sea level) research station, with a maritime climate. The soil was a poorly drained silty clay loam. At both locations, the experiment was laid out on soil with high variation in organic matter content and partly texture, using soil electric conductivity measurements as proxy variable (Fig. 3). This was done to mimic a typical situation where within field differentiation of fertilization rate is required.

The fields were established in 2015 with a mixture of timothy (*Phleum pratense* L.), meadow fescue (*Festuca pratensis*) and red clover (*Trifolium pratense* L.) and field trials were performed in the following two years, 2016 and 2017. The trials were laid out according to a split-plot design with five replicates in South-East and six replicates in Central Norway. Three different fertilization rates, low (13 g N m^{-2}), medium (20 g N m^{-2}) and high (27 g N m^{-2}) were applied to main plots, while time of harvest, early and normal first cut, was applied to sub-plots, totaling in 120 plots in South-East and 144 plots in Central Norway. The trials were managed with a three-cut regime, but data from the two first cuts were included in this study, only.

At each harvest day, a drone with a hyperspectral imager was used to capture canopy reflections of the grass legume mixture in the visual and near-infrared part of the solar spectrum right before cut and yield registration. Clover proportion was visually estimated in field in South-East. Then, and at both sites, the plants were cut at a stubble height of 0.07 m with an experimental plot harvester, followed by fresh and dry matter yield registration. Sub-samples of the dried yield were grinded and analysed by near-infrared reflectance spectroscopy, giving outputs for forage quality and proportion of clover (site in Central Norway, only).

After that, thorough radiometric and geometric calibration of the drone data was performed. The canopy reflections were transformed to average reflectance per plot and used as predictors to calibrate a powered partial least squares regression model to estimate dry matter yield. The validated prediction accuracy for a general dry matter yield model showed a root mean square error (RMSE) of 15.2% (55 g DM m^{-2}).

For consistency in the data and to exclude data with missing remote sensing estimates, only the field experimental data from South-East in 2016 and from Central Norway in 2017 was included in the current study. Moreover, due to high clover percentages observed in the main plots fertilized at low rates, these plots were omitted from the study as the NORNE model was not developed to handle such high percentages of clover that rarely occur in intensively managed grasslands in practice.

2.3. Weather data

Weather data with daily resolution were automatically recorded by weather stations provided by Agrometeorology Norway (https://lmt.nibio.no/). The data included air temperature (°C), soil temperature (°C), precipitation (mm), global radiation (MJ m^{-2}), wind speed (m s^{-2}) and relative humidity (%). Air temperature, wind speed and relative humidity were recorded at 2 m height, while soil temperature was recorded at 0.1, 0.2 and 0.5 m below ground. The measurements at 0.1 m below ground was used as the soil temperature in the top soil layer, while the mean temperature at 0.2 and 0.5 m below ground was used as the soil temperature in the sub soil layer. The distance between field trials and weather stations in this study was 350 m in South-East and 110 m in Central Norway.

2.4. Sensitivity analysis

A sensitivity analysis determines the parameters that are the key drivers of a model by investigating to what extent the variation in model output is influenced by variation in the model parameter settings (Saltelli et al., 2004). The screening method developed by Morris (Morris 1991) is adequate for complex models, like the NORNE model, where the number of parameters and also the computational cost, limit the possibility of numerical calculation. The method identifies the parameters being most influential to model output, in addition to which parameters have either a non-linear relationship with the output or interactions with other parameters. The elementary effects (EE) of model output are calculated by dividing the change in dry matter yield output from two consecutive model runs by the change in the input parameter (Eq. (8)).

$$EE_i(\theta^*) = \frac{y\left(\theta_1^*, \ldots, \theta_{i-1}^*, \theta_i^* + \Delta, \theta_{i+1}^*, \ldots, \theta_k^*\right) - y(\theta^*)}{\Delta} \quad (8)$$

Here, Δ is in the range of [1-(p-1), 1–1/(p-1)], p is the number of levels, $\boldsymbol{\theta}^*$ is any selected parameter vector mapped to the [0, 1] space and $\boldsymbol{\theta}$ is the selected parameter vector in the parameter space. The transformed point $\boldsymbol{\theta}$ from $\boldsymbol{\theta}^*+\boldsymbol{e}_i\Delta$ remains within the parameter space. For each index $i = 1,2,\ldots,k$, $\boldsymbol{e}i$ is a vector of zeros with a unit corresponding to its i'th component.

The finite distribution of elementary effects, denoted $EEi(\Theta)\sim Fi$, is constructed by r elementary effects that are sampled using an efficient design that constructs r trajectories of $k + 1$ points in the parameter space. Two sensitivity measures are calculated: (1) the mean value (μ), which evaluates the overall influence of the parameters on model output, and (2) the standard deviation (σ), which is used to detect parameters involved in interaction with other parameters or whose effect is nonlinear. To avoid the problem of effects of positive signs which occur when the model is non-monotonic, we will in this study use the mean of the absolute values (μ^*) (Campolongo et al., 2007).

For dynamic models that simulate daily outputs, the sensitivity of model parameters may change with time. It is consequently most

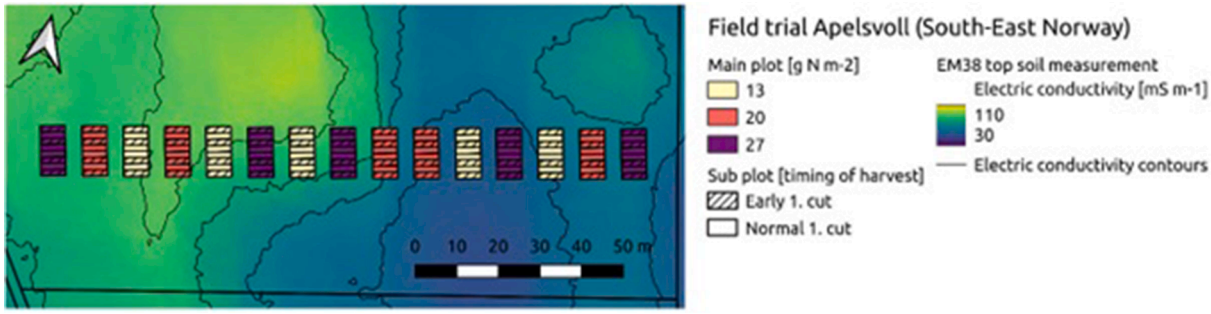


Fig. 3. Soil electric conductivity map showing high variation in terms of soil organic matter in the Apelsvoll field trial.

appropriate to consider the output over the whole time series (Lamboni et al., 2009), but the large number of responses that need to be evaluated makes this approach challenging for complex models. In this study, the total above ground yield at first and second cut was selected as the response.

The sensitivity analysis explored the space within the parameter boundaries, defined as uniform distributions with minimum and maximum defined as 50% and 150% of the default value. The sensitivity of a total of 43 model parameters were tested (Supporting Information, Table S1). For both cuts, the analysis was repeated for the different treatment combinations of fertilizer rate and timing of first cut. Thereafter, the mean of the four calculations was used as result. Further, the sensitivity of the model outputs to environmental conditions was included by introducing a parameter that was defined to vary between 1 and 4, and the selected value chose the location-year combination (South-East, Central, 2016, 2017).

2.5. In-season model calibration

The NORNE model is designed to handle within field variation and allows for in-season calibration of central model parameters to improve the prediction accuracy. Based on remotely sensed information and resulting estimates on above ground standing biomass (Geipel et al., 2021) calculated within 24 h before the cuts, within field, in-season update of the NORNE-model was performed by using the calibration procedure with the nonlinear least square algorithm. This optimization algorithm searches for the parameter set which would minimize the sum of squares of residual error.

The workflow for the in-season and node-wise model calibration is illustrated in Fig. 4. Here, two consecutive calibrations were performed. The parameters to be calibrated were selected due to the high within field soil variation that was deliberately included in the field experiment, by using fields with high variation in terms of soil organic matter and soil texture.

Firstly, the model was calibrated based on remotely sensed information from drone overflights within 24 h before first cut, to improve the model accuracy at first cut. Uncertainties regarding the starting value of nitrogen in soil were supposed to be the main source of discrepancy between simulated and actual observation at this early stage, and the starting values of mineral and organic nitrogen in soil, were selected for calibration. The organic nitrogen content was allowed to vary between 100 and 1000 g m^{-2}, while the mineral nitrogen content was allowed to vary between 0.5 and 3 g m^{-2}. A starting value of 2 g m^{-2} mineral nitrogen, initially consisting of 80% ammonium and 20% nitrate, and 540 g m^{-2} organic nitrogen was assumed for the entire field, and as the calibration was performed for a node-network within the field, node-specific initial values for soil nitrogen content were generated.

Secondly, the model was calibrated based on remotely sensed information from drone overflights within 24 h before second cut, to improve model accuracy at second cut. Water availability is important for optimal growth, and uncertainties regarding the soil texture determining, plant available water capacity was supposed to be the main source of discrepancy between simulated and actual observation at this later stage and was selected for calibration. The plant available water capacity in the top soil layer was allowed to vary within the boundaries of all soil texture types (Riley 2021), which were 35 and 95 mm. Further, the soil type containing that specific value for plant available water capacity in top soil layer was used and the corresponding field capacities and plant available water capacity in sub soil layer was used. Following the assumed soil type, starting value of 62.75 and 54.25 mm was used at South-East and 69.75 and 51.75 mm at Central Norway, for top and sub soil layer respectively. The results from the first calibration were incorporated and the calibration at second cut was performed for a node-network within the field, and node-specific field capacity levels were generated.

2.6. Model performance

Model performance was evaluated by four criteria to quantify the mismatch between simulated and observed quantity of dry matter yield and amount of nitrogen in plants. The four criteria were the absolute

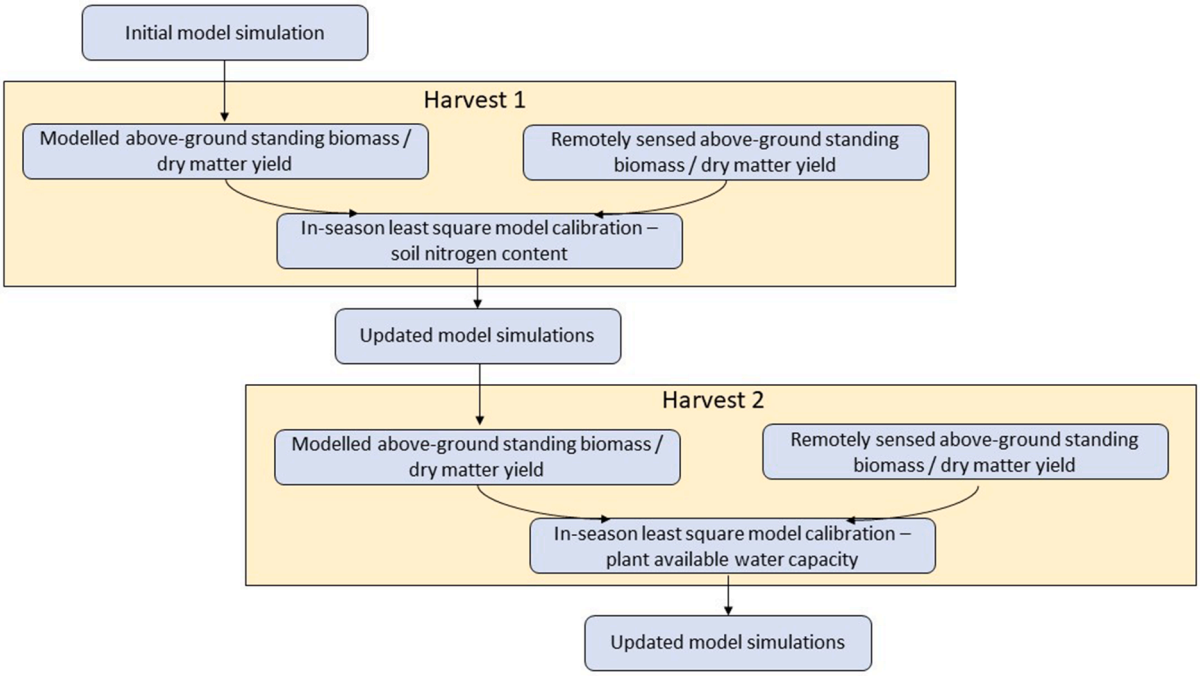


Fig. 4. Model calibration workflow to improve model predictions by remotely sensed in-season estimates of above-ground standing biomass around the time of the first and second cut. In-season model calibration is performed by an iterative non-linear least square matching approach, altering the initially assumed soil nitrogen content (cut 1) and field capacity (cut 2) in order to minimize the residuals in above-ground standing biomass on a node level.

error (AE), root mean square error (RMSE), the normalized (relative) RMSE (rRMSE) and the mean square error (MSE) as decomposed into its two components (squared bias and variance error) (Kobayashi and Salam 2000) (Eq. (9) - 12).

$$AE_i = S_i - O_i \tag{9}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2} \tag{10}$$

$$rRMSE = \frac{RMSE}{\overline{O}} \tag{11}$$

$$MSE = (\overline{S} - \overline{O})^2 + \frac{1}{n}\sum_{i=1}^{n}((S_i - \overline{S}) - (O_i - \overline{O}))^2 \tag{12}$$

Here, n is the number of observations, S_i is the simulated value, O_i is the observed value, $\boldsymbol{S}$ is the vector of model simulations, $\boldsymbol{O}$ is the vector of observations, $\overline{S}$ is the mean simulated value and $\overline{O}$ is the mean observed value. The first component of MSE refers to bias while the second component refers to variance error. Bias is the difference between the average simulation and observation and indicates error from erroneous assumptions in the model or an underfitted model. Variance error on the other hand, is a more random error that shows the variability of model simulation for a given data point or spread in the data. A high variance error may indicate an overfitted model. AE was only calculated separately for the sub-plots, while RMSE, rRMSE and MSE was calculated for each treatment.

3. Results

The sensitivity of 43 parameters to the dry matter yield output was assessed by a sensitivity analysis, looking at the dry matter yield outputs at first and second cut respectively (Fig. 5). The group of the 10 most sensitive parameters to dry matter yield at first cut (Fig. 5a) and the 11 most sensitive parameters to dry matter yield at second cut (Fig. 5b) were selected, both according to the main influential impact (mean value) and according to either a non-linear relationship with the output or interactions with other parameters (standard deviation). Within these two groups, four parameters were common: Θ_{24} - the optimum air temperature for grass growth, used to calculate the temperature index (TI), Θ_5 - the fraction of starting value of organic nitrogen in soil allocated to humus, Θ_6 – the fraction of starting value of organic nitrogen in soil allocated to litter 2 (among the fraction not allocated to humus) and Θ_3 - the environmental effect.

The overall two most sensitive parameters were Θ_{28} – the initial growth rate during spring growth (R_m) and Θ_{34} - the maximum daily growth during spring growth (DM_{max}) at first cut and correspondingly Θ_{29} – the initial growth rate during regrowth between first and second cut (R_m) and Θ_{35} – the maximum daily growth during regrowth between first and second cut (DM_{max}) at second cut. Additionally, the parameter Θ_{31} at first cut and Θ_{32} at second cut was in the group of the most sensitive parameters, also being a parameter value used to estimate potential daily dry matter yield during respectively spring growth and regrowth between first and second cut. Finally, at first cut, Θ_{15} - a temperature threshold value used to calculate LAI during first part of spring growth and Θ_{26} and Θ_{27}, being respectively the parameter determining the curvature of the radiation response curve and the insolation at light saturation of the stand used to calculate the solar index (SI). Further, at second cut, Θ_{14}, Θ_{16} and Θ_{17} which are all temperature threshold values used to calculate LAI and Θ_4 – the plant available water capacity at top soil level.

Dry matter yield and amount of nitrogen in yield at first and second cuts of a forage crop were simulated by the NORNE model using default parameter values, and the outputs were further compared to observed values using the field experimental data from South-East Norway in 2016 and from Central Norway in 2017. The error term RMSE was calculated individually for the different treatment combinations of fertilizer rate and timing of first cut at both locations. Further, the MSE was calculated and decomposed into its two components of squared bias and variance error and transferred into the RMSE of dry matter yield (Fig. 6) and amount of nitrogen in yield (Fig. 7).

For the first cut in South-East (Fig. 6a) and in Central Norway (Fig. 6c), an average RMSE of respectively 79.4 and 64.9 g m^{-2} was found for dry matter yield. For both locations, the RMSE was lowest for the early first cut treatments, while it generally increased with nitrogen fertilization rate (except for the early cut in Central Norway). The overall highest error term was detected for the treatment with normal first cut combined with high fertilization rate, both dominated by bias that contributed with respectively 63 and 78% of the error term for South-East and Central Norway. The error terms for both locations with early first cut combined with high fertilization rate were dominated by variance error (81 and 100%). The rRMSE was lowest for the early first

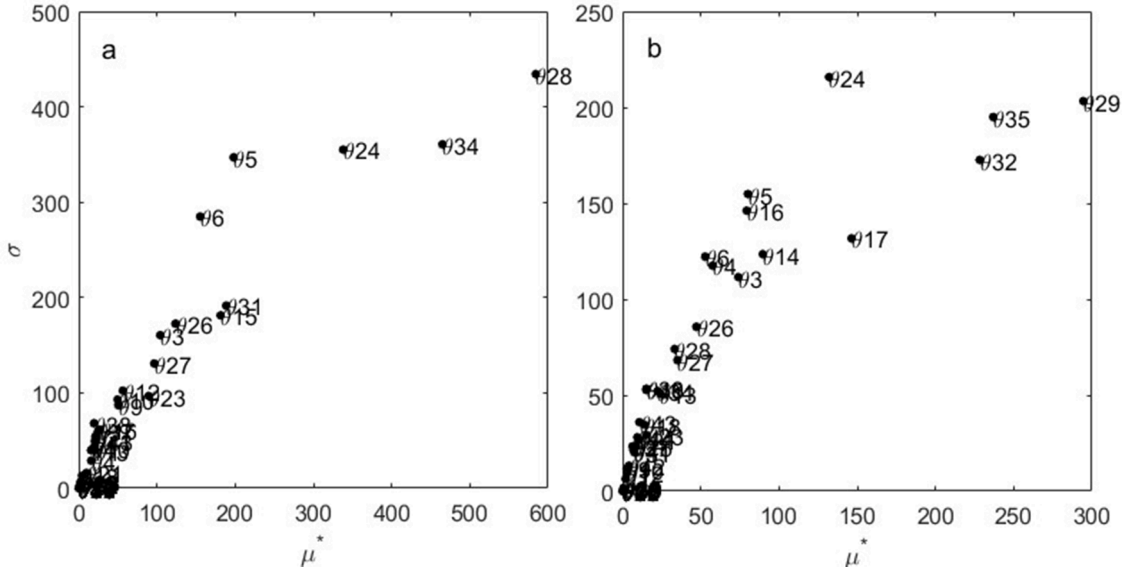


Fig. 5. Results from sensitivity analysis of the NORNE model, applying the Morris method for evaluation of the dynamic output at (a) 1st cut and (b) 2nd cut. The parameters are numerated, and a description of each is given in the Supporting Information.

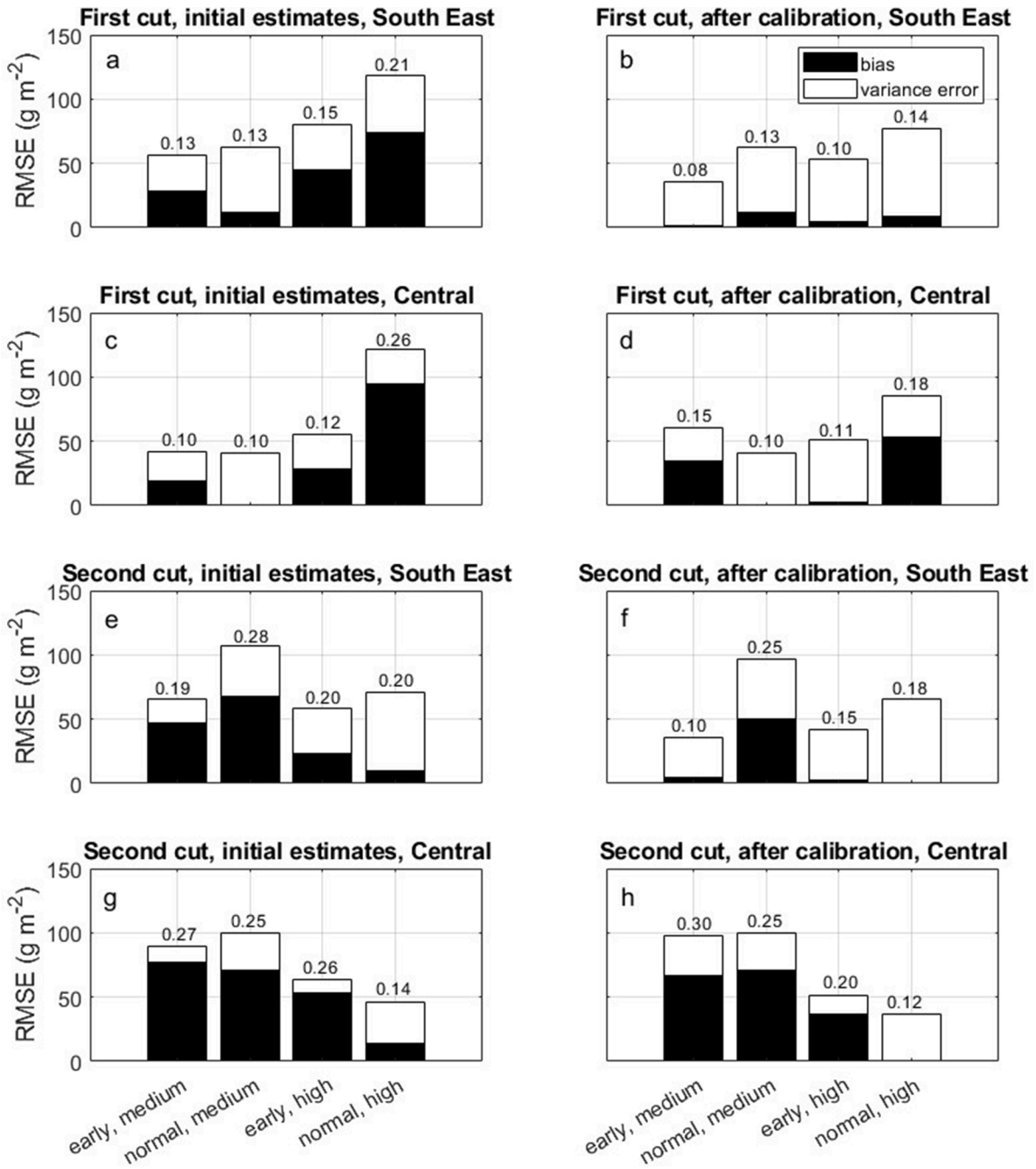


Fig. 6. The root mean square error (RMSE) of estimates obtained by the process-based model, decomposed according to mean square error (MSE) into bias (black) and variance error (white) for dry matter yield of first (a - d) and second (e - h) cuts, before (a, c, e, g) and after (b, d, f, h) calibration with remotely sensed information for South-East Norway 2016 (a, b, e and f) and Central Norway 2017 (c, d, g and h). The results comprise two different rates of nitrogen fertilization (medium and high) and two developmental stages at first harvest (early and normal). Relative RMSE (rRMSE) is given as a fraction above each bar.

cut, ranging between 0.10 and 0.13, while it ranged between 0.12 and 0.26 for the normal harvest.

For improved model fit at first cut, the initial values for organic and mineral nitrogen in soil were optimized through model calibration with the remotely sensed predictions of dry matter yield. This resulted in separate starting values through the sub-plots (nodes). The starting value for organic nitrogen in soil was allowed to vary between 100 and 1000 g N m^{-2}, while mineral nitrogen was allowed to vary between 0.5 and 3 g N m^{-2}. After calibration, the initial mineral nitrogen content varied between 0.5 and 3 g N m^{-2} in both South-East and Central Norway, while the initial organic nitrogen content varied between 100 and 672 g N m^{-2} in South-East and between 100 and 1000 g N m^{-2} in Central Norway. RMSE was re-calculated after calibration (Fig. 5b and d) showing an overall reduction, resulting in an average RMSE of respectively 57.0 and 59.5 g m^{-2} in South-East and Central Norway. For the treatments with early first cut combined with high fertilization rate, no changes in the initial values were obtained after calibration. For the treatment with normal first cut in both locations and the treatment with early first cut combined with medium fertilization rate in South-East Norway, improvements in the error term were found for dry matter yield at first cut, after calibration. The RMSE was reduced by 4 to 42 g m^{-2}, while rRMSE was reduced by up to 0.08. Overall, an 18% reduction in rRMSE was achieved. For the treatment in Central Norway with early first cut combined with medium fertilization rate, an increased error

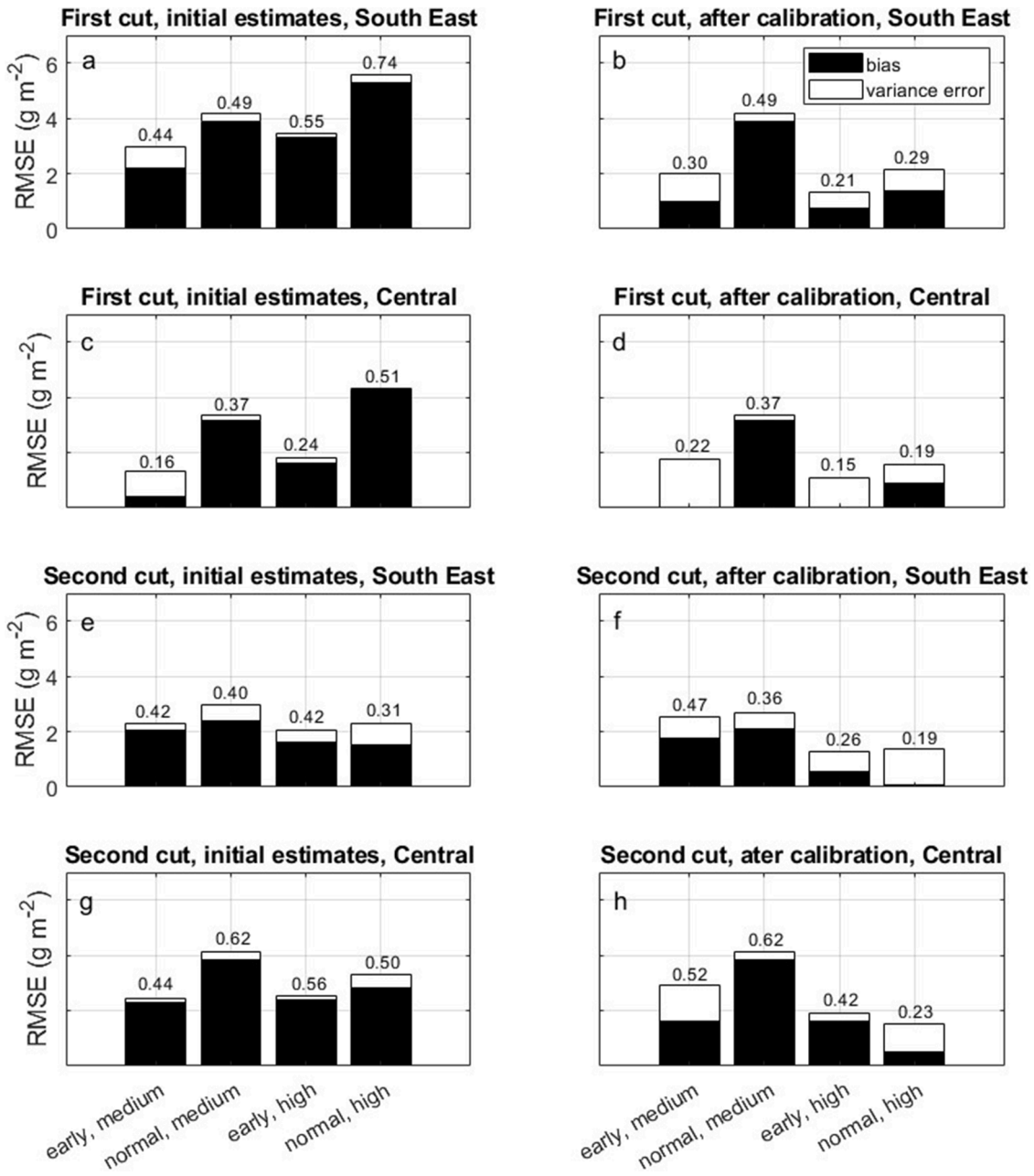


Fig. 7. The root mean square error (RMSE) of estimates obtained by the process-based model, decomposed according to mean square error (MSE) into bias (black) and variance error (white) for amount of nitrogen in dry matter yield of first (a - d) and second (e - h) cuts, before (a, c, e, g) and after (b, d, f, h) calibration with remotely sensed information for South-East Norway 2016 (a, b, e and f) and Central Norway 2017 (c, d, g and h). The results comprise two different rates of nitrogen fertilization (medium and high) and two developmental stages at first harvest (early and normal). Relative RMSE (rRMSE) is given as a fraction above each bar.

term was achieved after calibration due to an initial better fit for the NORNE-model compared to the remote sensing predictions. The error term was generally dominated by variance error.

For the second cut in South-East (Fig. 6e) and in Central Norway (Fig. 6g), an average RMSE of respectively 75.4 and 74.7 g m^{-2} was found for dry matter yield initially. For both locations, the RMSE was highest in the treatment with early first cut combined with high fertilization rate. The error term was generally dominated by bias, except the treatment with normal first cut in South-East Norway and the treatment with normal first cut combined with high fertilization rate in Central Norway where variance error accounted for most of the error term.

For improved model fit at second cut, the initial value for plant available water capacity in top soil layer was calibrated while the corresponding capacity in sub soil layer and the field capacity in both layers were set according to the soil type corresponding to the plant available water capacity identified by calibration. The parameters were allowed to vary within the boundary of all soil types. After calibration, the plant available water capacity in top soil varied between 56 and 92 l m^{-2} in South-East and 70 l m^{-2} in Central Norway. The model was re-run from growth start, using the new parameter values and RMSE at second cut was re-calculated (Fig. 6f and h). The RMSE varied between 35.4 and 96.6 g m^{-2} in South-East and between and 37.0 and 99.5 g m^{-2} in Central Norway. It was dominated by variance error, and the highest error term was detected for early first cut in Central Norway. The RMSE was thus reduced by 9.5 to 65.6 g m^{-2}, while RMSE% was reduced by up to 0.09. Overall, an 13% reduction in rRMSE was achieved. Only predictions of the treatments with early first cut combined with medium fertilization rate in Central Norway were not improved in the calibration process.

The AE was calculated individually for each sub-plot, with the distribution for the initial dry matter yield estimates at first (Fig. 8a) and second cut (Fig. 8c) given combined for all sub-plots that included the different treatment combinations of fertilizer rate and timing of first cut and both locations. A distribution shifted to the right was detected in both cases, with mean values of 41.3 and 57.2 g m^{-2} and standard deviations of 66.0 and 52.4 g m^{-2}, respectively for the first and second cut. The shift to the right, with positive mean values of AE indicate a general overestimation of dry matter yield by the NORNE model when compared to field observation data. The distribution of AE at first cut, after calibration (Fig. 7b) had a mean value of 6.4 g m^{-2} and standard deviation of 60.3 g m^{-2}. The reduced, but still positive mean value of AE indicates a general improvement in the model performance, but still with a minor overestimation of dry matter yield when compared to field data. Also the spread in fit between sub-plots was reduced. The distribution of AE at second cut, after calibration (Fig. 7d) showed a mean value of 41.6 g m^{-2} with a standard deviation of 58.7 g m^{-2}. The reduced, but still positive mean value of AE indicates a general improvement in the model performance, but still with a general overestimation of dry matter yield when compared to field data.

The RMSE for amount of nitrogen in yield was calculated and given in Fig. 6. For the first cut in South-East (Fig. 7a) and in Central Norway (Fig. 7c), an average RMSE of respectively 4.0 and 2.7 g N m^{-2} was found. The error term was dominated by bias (contributing with 74 to 98%), except for the treatment with early first cut combined with medium fertilization rate in Central Norway, where the bias only accounted for 30% of the error. For both locations, the RMSE was highest for the treatments with high fertilization rate. RMSE at first cut was recalculated after calibration (Fig. 7b and d). The error term was higher for the treatments with high fertilization rate compared to the treatments with medium fertilization rate. For the second cut in South-East (Fig. 7e) and in Central Norway (Fig. 7g), an average RMSE of respectively 2.4 and 3.1 g N m^{-2} was found for amount of nitrogen in yield initially. The error term was dominated by bias, contributing with 66 to 94% of the error term. After calibration, the RMSE was recalculated for South-East (Fig. 7f) and Central Norway (Fig. 7h). It was reduced at all treatments except the early first cut combined with medium fertilization rate in Central Norway. The variance generally dominated after calibration.

The AE was calculated individually for each sub-plot, with the distribution for the initial estimates of amount of nitrogen in dry matter yield at first (Fig. 9a) and second cut (Fig. 9c) given combined for all sub-plots that included the different treatment combinations of fertilizer rate and timing of first cut and both locations. A distribution shift to the right was detected in both cases, with mean values of 3.1 and 2.6 g N m^{-2} and standard deviations of 1.7 and 1.2 g N m^{-2}, respectively for the first and second cut. The large shift to the right, with positive mean values of the AE indicate a general high overestimation of amount of nitrogen in yield by the NORNE model when compared to field observation data. The distribution of AE at first cut after calibration (Fig. 9b) had a mean value of 1.5 g N m^{-2} and standard deviation of 1.8 g N m^{-2}. The reduced, but still positive mean value of AE indicates a general improvement in the model performance, but still with a minor overestimation of amount of nitrogen in yield when compared to field data. The distribution of AE at second cut, after calibration (Fig. 8d) had a mean value of 1.2 g N m^{-2} with a standard deviation of 1.9 g N m^{-2}. The reduced, but still positive mean value of AE indicates a general improvement in the model performance, but still with a general overestimation of amount of nitrogen in yield when compared to field data.

The simulated dry matter yield at harvest, before and after model calibration, is presented for each sub-plot within treatment and location in the supporting information section (Figures S1-S8). The observed yield and the yield predictions from remote sensing are also included in these figures. Among the altogether 176 sub-plots, six sub-plots from one site and treatment that differed in observed dry matter yield, were selected as an illustration of model performance (Fig. 10). For these six sub-plots, the model initially estimated a dry matter yield of 585 g m^{-2} at first cut and 324 g m^{-2} at second cut throughout the field. After calibration, sub-plot specific parameter values were applied, and the new simulation caught to a large extent the infield observed variation from 434 to 590 g m^{-2} in the first cut and from 252 to 308 g m^{-2} in the second cut (Fig. 10).

4. Discussion

When predicting dry matter yields with the NORNE model within

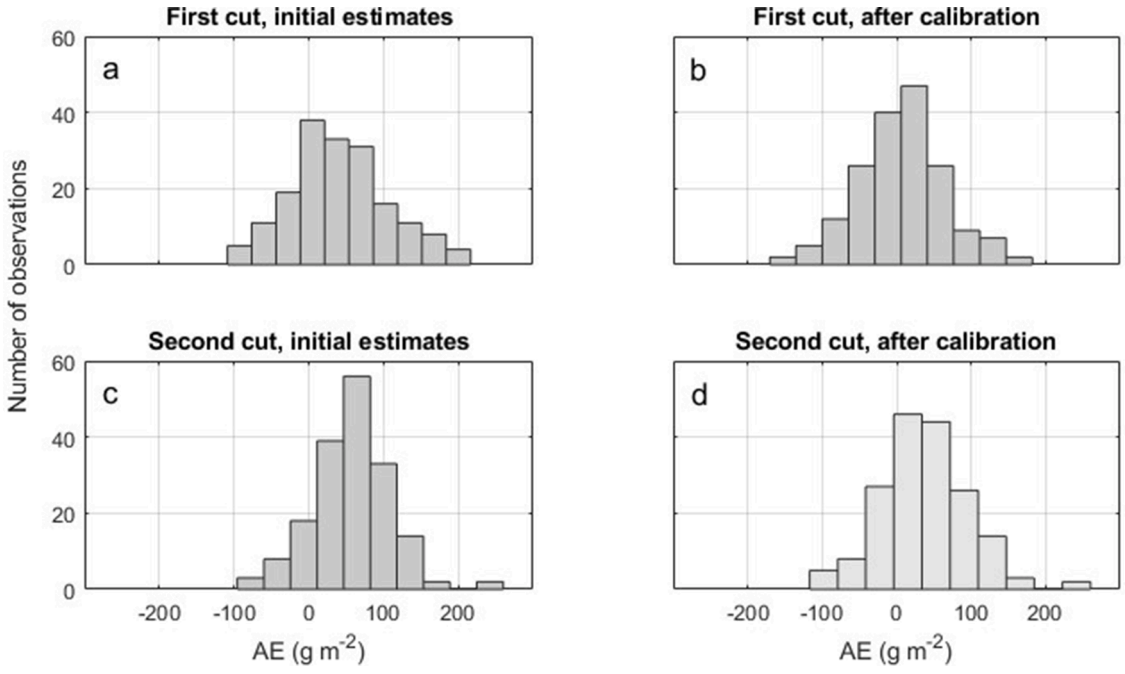


Fig. 8. The distribution of absolute error (AE) for dry matter yield predictions obtained by the process-based model of first (a – b) and second (c -d) cuts, before (a, c) and after (b, d) calibration with remotely sensed information for South-East Norway 2016 and Central Norway 2017 combined.

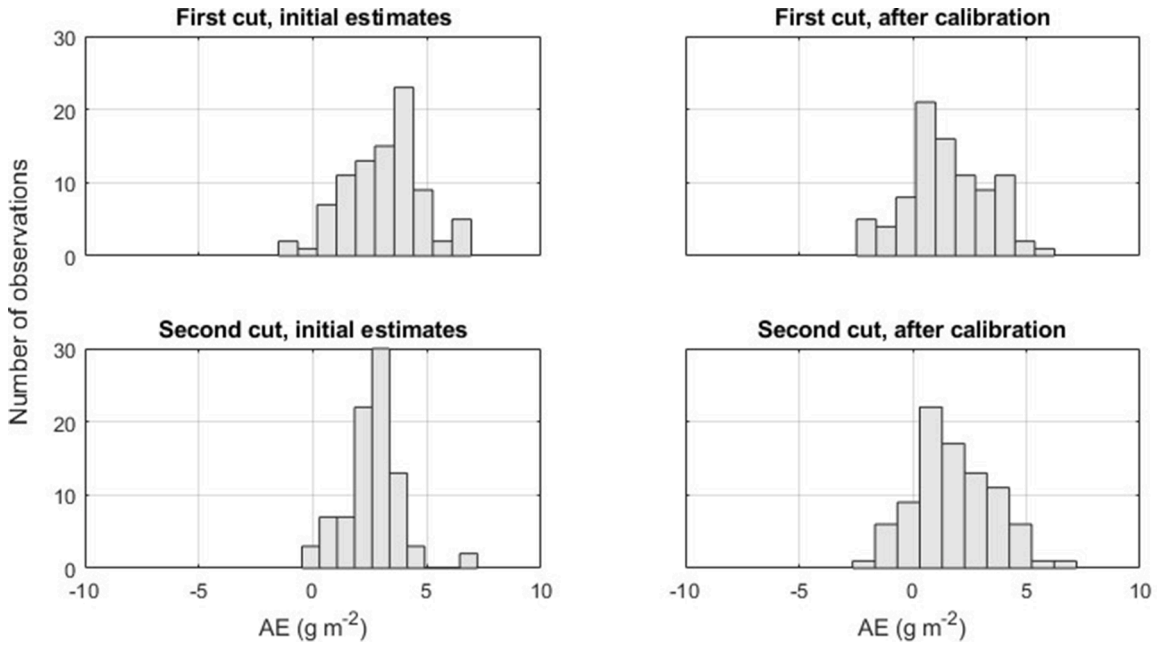


Fig. 9. The distribution of absolute error (AE) for amount of nitrogen in dry matter yield prediction obtained by the process-based model of first (a – b) and second (c -d) cuts, before (a, c) and after (b, d) calibration with remotely sensed information for South-East Norway 2016 and Central Norway 2017 combined.

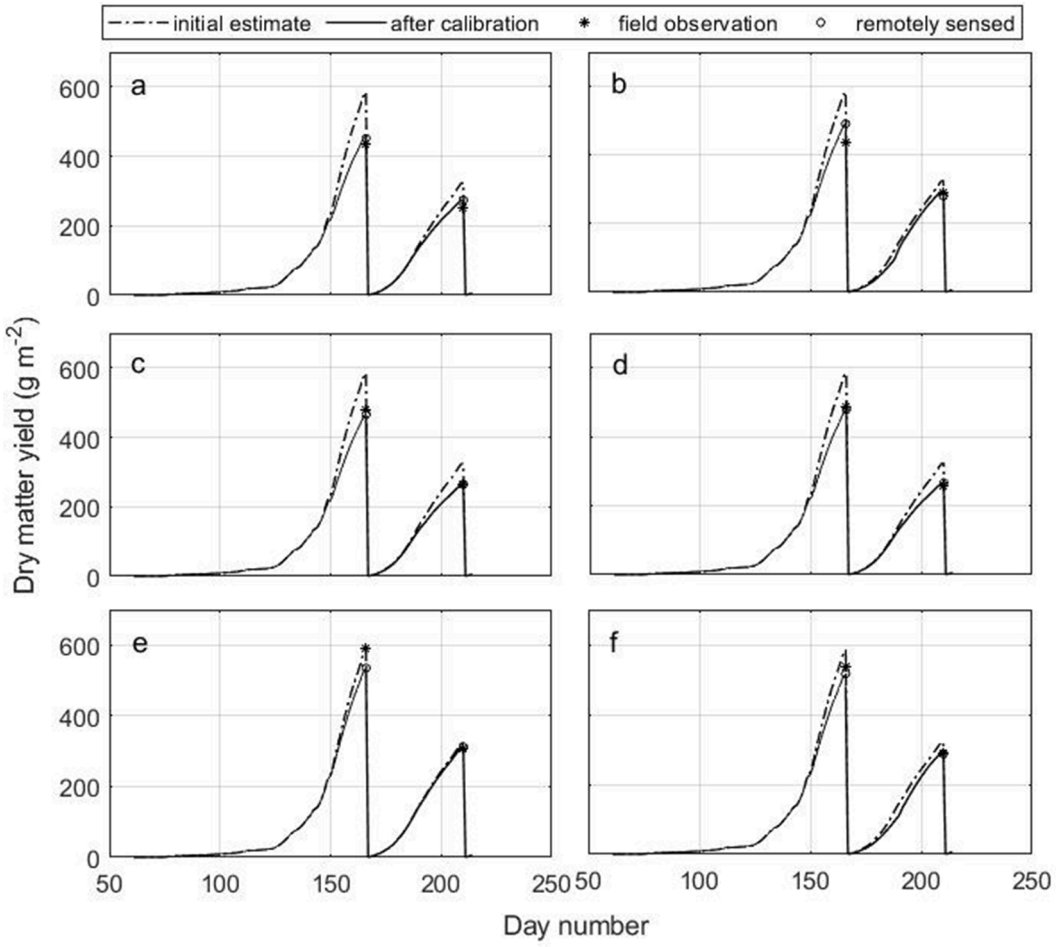


Fig. 10. Dry matter yields simulated with the process-based model before and after calibration with remotely sensed information, remotely sensed predictions and observed dry matter yield in first and second cuts of selected single nodes (sub-plots) in South-East Norway in 2016 for the treatment of normal first cut combined with medium fertilization rate.

two locations in Norway, at two rates of fertilization and for two different harvesting regimes, the model fitted the measured yields well. Overall, the model performed with an RMSE of 72 g m^{-2} at first cut and 75 g m^{-2} at second cut when using the default parameter values (i.e. without utilizing remotely sensed information). The rRMSE was 0.15 at first cut and 0.22 at second cut. These results where comparably better than those reported by Korhonen et al. (2018), who evaluated three growth models for timothy (BASGRA, CATIMO and STICS) under northern conditions, including data from Norway, Sweden, Finland and Canada. Our NORNE model achieved 0.30, 0.15 and 0.7 smaller errors (expressed as rRMSE) for first cuts and 0.08, 0.52 and 0.01 smaller errors for second cuts, respectively for BASGRA, CATIMO and STICS. The BASGRA model was later tested under Nordic conditions in a study by Höglind et al. (2020), allowing field specific calibration of some model parameters. BASGRA then performed with an rRMSE of 0.29 in Sweden and 0.27 in Norway for dry matter yield predictions, still worse than the rRMSE of yield predictions of NORNE using unchanged default parameter settings.

The amount of nitrogen in yield (g N harvested per m^2) was estimated, with less accuracy than dry matter yield, but still with a significant improvement after calibration to dry matter yield observations. Further improvement is to be expected if calibrations also were undertaken on basis of measurements of concentration of nitrogen in the yield. Höglind et al. (2020) included calculations and testing on crude protein content, with an rRMSE of maximum 0.30 for Swedish data and 0.18 for Norwegian. The NORNE model was not tested on crude protein, but for the amount of nitrogen in yield, it performed with a mean rRMSE of 0.30.

Nitrogen availability is crucial for optimal crop growth, and a comprehensive sub-model for soil-plant nitrogen dynamics was therefore developed within the NORNE model. Existing functions were adapted from literature and combined to describe the dynamic processes, with parameter values directly taken from literature. The inclusion of a separate nitrogen sub-model is a strength of the NORNE model compared to other crop growth models and may explain the better fit and its ability to catch the effects of different nitrogen application rates. It is also a necessity for further development of the model into a decision support tool for optimal plant nitrogen fertilization. The NORNE model managed to simulate dry matter yield responses to increased nitrogen fertilization rates. But still, using the initial parameter values, a better fit was achieved at first cut for the plots fertilized at medium rate compared to the fields fertilized at high rate. Contrary at second cut, the plots fertilized at high rate, achieved a better fit compared to the plots fertilized at medium rate. The study by Korhonen et al. (2018) showed that the STICS model, with a more detailed description of soil nitrogen mineralization and water balance compared to BASGRA and CATIMO, simulated dry matter yield response to increased nitrogen fertilization rates more accurately than the latter two models, which both underestimated the nitrogen fertilization response.

Within-field soil variation (e.g. varying distance to drainage networks) is typical in farmers' fields, and may cause spatial variability in dry matter yield. To mimic a typical situation, the field experiments were placed on soils with high internal variation in organic matter content and partly texture. Such non-linear effects can be accounted for using different methods (Li and Converino 2021; Runge et al., 2019), but a global sensitivity analysis was selected in this study to assess the sensitivity of 43 parameters in the NORNE model to dry matter yield outputs at first and second cut. The chosen screening method (Morris 1991) detected the parameters that were highly and weakly influential to changes in the dry matter yield output (linear relationship), additionally to detecting parameters with non-linear relationships to the dry matter yield output and the parameters interaction with other parameters. The results showed that parameters from the plant module were the most sensitive ones. Parameters related to the starting value of nitrogen in soil, optimal air temperature for crop growth and the environmental effect were important for the dry matter yield prediction at both first and second cut. Finally, the plant available water content was only among the most highly sensitive parameters to the model output at second cut.

Within 24 h before each cut, canopy reflection measurements were performed with a drone-borne hyperspectral imager. Based on these measurements, predictions of dry matter yield were calculated with a powered partial least square regression model at high spatial resolution, i.e. plot-level (Geipel et al., 2021). In contrast to the NORNE model that provides daily predictions over the entire growing season, the remote sensing model only applies to a very narrow time interval of one to two days around the day of measurement. Given a well-calibrated and robust remote sensing model, the remote sensing predictions were more accurate than the corresponding predictions of the NORNE model and assimilation of the outputs were used to improve the prediction accuracy of dry matter yields by the NORNE model. Nevertheless, the remote sensing model should be calibrated for a wider time-span of drone-borne measurements than just right before harvest to enable more suitable predictions for applications in farming practice. In line with drone-borne measurements, satellite data can be used for this purpose (Kasampalis et al., 2018) and in Zare et al. (2022), within-season assimilation of satellite observations was shown to improve the model fit of the PILOTE crop growth model used to forecast winter wheat.

Different assimilation methods, as calibration, forcing and updating are tested and discussed in the literature (Jin et al., 2018; Kasampalis et al., 2018). As forcing does not use any information in the crop model, but simply uses the remotely sensed data to replace the crop model simulation data, updating continuously adjusts model parameters in the crop model during the run time, while the calibration strategy adjusts selected parameters to achieve optimal consistency between the remote sensing data and the simulated state variables. Among the three methods, calibration is theoretically a better choice, but has a high computational cost due to the requirement of a lot of optimization iterations. In line with Jin et al. (2018), calibration was successfully used in this study. Calibration was first performed at first cut and then a re-calibration was performed at second cut later in the season.

The traditional calibration algorithm nonlinear least square was used in the assimilation step in this study. This optimization algorithm searches for the parameter set which would minimize the sum of squares of residual error. The more comprehensive Bayesian calibration (Berger, 1985) has the advantage that it, in addition to calibrating the parameter values, simultaneously quantifies the parameter uncertainty (Campbell, 2006). It is a flexible method, but due to a high computational cost when applied to complex models, the method was not used in this study. As the calibration process will be implemented in a decision support system, a faster routine is required.

Soil variation was deliberately included in the field experiment to mimic a typical situation where site specific fertilization is required. This was done by selecting a field with imperfectly drained soil with high variation in terms of soil organic matter and partly soil texture. Despite the variation, a uniform value was initially used for the entire field for both nitrogen in soil and soil texture. The initial values of both soil mineral and organic nitrogen were selected as the parameters to be calibrated at first cut. The rationale for this was that the amount of soil nitrogen is previously shown to be a sensitive model parameter (Hjelkrem et al., 2021), and that actual soil mineral and organic nitrogen contents are rarely known in practice, at least not at high resolution within a field.

The calibration of these initial values indicated high within field variation, and their adjustments improved the predictions of the NORNE model in all but one case. Overall, the model achieved an RMSE of 58 g m^{-2} for dry matter yield at first cut when the remotely sensed information was utilized. The rRMSE was 0.12, which corresponds to a reduction in the error term of 20% at first cut compared to the use of initial parameter values. Further development of remote sensing technology and more comprehensive data sets for empirical model calibration, will enable more accurate and reliable remote sensing predictions to adjust and optimize crop model parameters and improve crop model

simulation results at field and regional scales. In addition to improve model performance, the spatial information that is added to the process-based model from the incorporation of the remote sensing data allows the model to predict and advice site-specifically, which is required for precision agriculture.

In addition to the presence of nitrogen, water availability is important for optimal crop growth. Hurtado-Uria et al. (2013) evaluated the English crop growth model developed by Johnson and Thornley (1983) under Irish conditions and concluded that the inclusion of a water sub-model would improve the model's accuracy. A thorough sub-model for water was thus included in the NORNE model, highly depending on soil texture that was set uniform for the entire field despite the known within field variation. Plant available water capacity was given by the value of soil texture and was selected for calibration at second cut. The value of this parameter influenced the model output substantially, and, through calibration, improvements in the predictions of the NORNE model were achieved. Overall, the model performed with an RMSE of 66 g m^{-2} for dry matter yield at second cut when the remotely sensed information was utilized. The rRMSE was 0.18, which corresponds to a reduction in the error term of 19% at second cut compared to the use of initial parameter values. Thus, the special information added to the process-based model highly improved the model accuracy at second cut.

Using data from two geographically spread locations in Norway (South-East and Central), including two rates of nitrogen fertilization (medium and high) and two harvesting regimes, the NORNE model simulated dry matter yield and amount of nitrogen in yield with promising accuracy. Even though the model is process-based, it includes several empirically derived functions, and one could assume that the accuracy of the model would vary with climatic and other environmental conditions. Hence, model testing across a wide range of conditions is a key to strengthening the model. Therefore, a more comprehensive model validation is required, including a broader variation in crop composition and soil and weather conditions, in order to obtain a model of more generalized validity for northern conditions. Additionally, a routine to handle inputs and mineralization of organic nitrogen in animal manure and other organic fertilizers should be included in a future version, to increase the usability of the model for a wider range of agricultural practices.

With regard to the overall ambition, that NORNE will constitute the basis of a decision support system for timing of harvests and target nitrogen inputs to demand, routines for collection and interpretation of weather forecast with high spatial resolution need to be established. Presently, there are no forecasts for soil temperature and only two-day forecasts for global radiation, and proxies need to be developed to get the demanded input data. Further, the model must be implemented into a digital farm management service, along a web-based user interface for the integration of geo-referenced field, fertilization, remote sensing, and preferably soil data. Also, the use of satellite data in place of drone-borne sensor measurements should be evaluated in the future. As hyperspectral imagers are still rather costly, more affordable solutions are needed to enable a faster and more widespread implementation of remote sensing under today's practical conditions. Here, multispectral imagery from off-the-shelf drones (e.g. DJI P4MS) and publicly available satellite missions (e.g. ESA Copernicus Sentinel 2) represent promising alternatives, yet at the cost of a slight decrease in prediction accuracy.

5. Conclusions

The newly developed NORNE-model simulated crop dry matter yield and amount of nitrogen in yield with promising accuracy. The inclusion of in-season nodewise calibration of the parameters soil nitrogen content (i.e. mineral and organic nitrogen) at initiation of growth and soil field capacity improved the model performance significantly.

Overall, the model performed with an RMSE (rRMSE) of 58 g m^{-2} (0.12) for dry matter yield at first cut and 66 g m^{-2} (0.18) at second cut when the remotely sensed information was utilized, which corresponds to a reduction of respectively 20 and 16% for each cut when compared to the use of initial parameter values. For amount of nitrogen in yield, the improvements were even larger when utilizing remotely sensed information. Although the model was calibrated using remotely sensed predictions of dry matter yield only, a reduction in the estimated RMSE of 38% was achieved during first cut and 26% during second cut. Overall, the model performed with an RMSE (rRMSE) of 2.1 g N m^{-2} (0.28) during first cut and 2.3 g N m^{-2} (0.33) during second cut for simulated amount of nitrogen in yield when remotely sensed information on dry matter yield was utilized.

In future, a routine to handle inputs of animal manure and other organic fertilizers should be included, to increase the usability of the model further in agricultural practice. Also, the use of satellite data in place of drone-borne measurements should be evaluated. Finally, in a next step, a feedback loop for nitrogen demand will be added and the system will be included in a decision support system within a digital infrastructure to serve as a tool for farmers giving advice on farming practice.

CRediT authorship contribution statement

Anne-Grete Roer Hjelkrem: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing – original draft, Visualization. **Jakob Geipel:** Conceptualization, Methodology, Validation, Data curation, Writing – review & editing, Visualization. **Anne Kjersti Bakken:** Conceptualization, Methodology, Validation, Resources, Data curation, Writing – review & editing, Project administration, Funding acquisition. **Audun Korsaeth:** Conceptualization, Methodology, Validation, Resources, Data curation, Writing – review & editing, Funding acquisition.

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

The authors do not have permission to share data.

Acknowledgements

This study has been conducted as part of the project *Improved precision in forage crop management* and *Use of remote sensing for increased precision in forage production* (The Research Council of Norway, project number 280332 and 244251, respectively), funded by Forskningsmidlene for jordbruk og matindustri, Yara Norge, Yara Research centre Hanninghof Germany, Felleskjøpet Agri and Strand Unikorn.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2023.110433.

References

Andrén, O., Kätterer, T., Karlsson, T., 2004. ICBM regional model for estimations of dynamics of agricultural soil carbon pools. Nutr. Cycl. Agroecosyst. 70, 231–239.
Angus, J.F., Kornher, A., Torsell, B.W.R., 1980. A Systems Approach to Estimation of Swedish ley Production. Swedish University of Agricultural Sciences. Report 80.
Auernhammer, H., 2001. Precision farming – the environmental challenge. Comput. Electr. Agricult. 30, 31–43.
Baadshaug, O.H., Lantinga, E.A., 2002. ENGNOR, a Grassland Crop Growth Model For High Latitudes. Documentation. Agricultural University of Norway. Report no 2/2002.
Bakken, A.K., 2016. Grovfôrmodellen – verktøy eller leiketøy? BUSKAP 68 (3), 58–59.

Berger, O.J., 1985. Statistical Decision Therory and Bayesian analysis, New York.
Bonesmo, H., 2004. Phenological development in timothy and meadow fescue as related to daily mean air temperature and day kength. Grassland Sci. Europe 9, 799–891.
Bonesmo, H., Bélanger, G., 2002a. Timothy yield and nutrient value by CATIMO model. I. Growth and nitrogen. Agron. J. 94, 337–345.
Bonesmo, H., Bélanger, G., 2002b. Timothy yield and nutrient value by CATIMO model. II. Digestibility and fiber. Agron. J. 94, 345–350.
Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., 2003. An overview of the crop model STICS. Eur. J. Agron. 18, 309–332.
Campbell, K., 2006. Statistical calibration of computer simulations. Reliabil. Eng. Syst. Safety 91, 1358–1363.
Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Modell. Softw. 22, 1509–1518.
Capolupo, A., Kooistra, L., Berendonk, C., Boccia, L., Suomalainen, J., 2015. Estimating plant traits of Grasslands from UAV-acquired hyperspectral images: a comparison of statistical approaches. ISPRS Int. J. Geoinf. 4 (4), 2792–2820.
Clevers, J., Vonder, O., Jongschaap, R., Desprats, J-F., King, C., Prévot, L., Bruguier, N., 2002. Using SPOT data for calibrating a wheat growth model under mediterraniean conditions. Agronomie 22 (6), 687–694.
Fystro, G., 1995. In: Workshop Cost 814: Nitrogen supply and nitrogen fixation of crops for cool and wet climates. Holt, Norway.
Fystro, G., 2001. Tilpassa N-gjødsling til eng. Grønn forskning 5 (04).
Geipel, J., Bakken, A.K., Jørgensen, M., Korsaeth, A., 2021. Forage yield and quality estimation by means of UAV and hyperspectral imaging. Precis. Agricult. 22, 1437–1463.
Greenwood, D.J., Lemaire, G., Gosse, G., Cruz, P., Draycott, A., Neeteson, J.J., 1990. Decline in percentage N of C3 and C4 crops with increasing plant mass. Ann. Bot. 66, 425–436.
Heinen, M., 2006. Simplified denitrification models: overview and properties. Geoderma 133, 444–462.
Hjelkrem, A.G.R., Geipel, J., Bakken, A.K., Korsaeth, A., 2021. A novel dynamic model for estimating standing biomass and nitrogen content in grass crops harvested for silage production. Sensing – New Insights into Grassland Scienceand Practice. In: Astor, T., Dzene, I. (Eds.), Proceedings of the 21 st Symposium of the European Grassland Federation. Hosted by Universität Kassel, pp. 172–174. Germany. 17-19 May 2021, The Organising Committee of the 21 st Symposium of the European Grassland Federation, Universität Kassel, Grassland Science and Renewable Plant Resources, Witzenhausen, Germany.
Höglind, M., Cameron, D., Persson, T., Huang, X., Van Oijen, M., 2020. BASGRA_N: a model for grassland productivity, quality and greenhouse gas balance. Ecol. Modell. 417, 1–13.
Höglind, M., Van Oijen, M., Cameron, D., Persson, T., 2016. Process-based simulation of growth and overwintering of grassland using the BASGRA model. Ecol. Modell. 335, 1–15.
Hurtado-Uria, C., Hennessy, D., Shalloo, L., Schulte, R.P.O., Delaby, L., O'Connor, D., 2013. Evaluation of three grass growth models to predict grass growth in Ireland. J. Agricult. Sci. 151, 91–104.
Jin, X., Kumar, L., Li, Z., Feng, H., Xu, X., Yang, G., Wang, J., 2018. A review of data assimilation of remote sensing and crop models. Eur. J. Agron. 92, 141–152.
Johnson, I.R., Thornley, J.H.M., 1983. Vegetative crop growth model incorporating leaf area expansion and senescence, and applied to grass. Plant Cell Environ. 6, 721–729.
Kasampalis, D.A., Alexandridis, T.K., Deva, C., Challinor, A., Moshou, D., Zalidis, G., 2018. Contribution of remote sensing on crop models: *a review*. J. Imag. 4 (4), 1–19.
Kertulis, G.M., 2001. Effect of Nitrogen and Cutting Management On Root Growth and Productivity of a Kentucky Bluegrass (*Poa Pratensis* L.) and White Clover (*Trifolium Repens* L.) Pasture. West Virginia University.
Kobayashi, K., Salam, M.U., 2000. Comparing simulated and measured values using mean squared deviation and its components. Agron. J. 92, 345–352.
Korhonen, P., Palosuo, T., Persson, T., Höglind, M., Jégo, G., Van Oijen, M., Gustavsson, A.M., Bélanger, G., Virkajärvi, P., 2018. Modelling grass yields in northern climates – a comparison of three growth models for timothy. Field Crops Res. 224, 37–47.
Korsaeth, A., Bakken, L.R., Riley, H., 2003. Nitrogen dynamics of grass as affected by N inout regimes, soil texture and climate: lysimeter measurements and simulations. Nutr. Cycl. Agroecosyst. 66, 181–199.
Lamboni, M., Makowsk, D., Lehuger, S., Gabrielle, B., Monod, D., 2009. Multivariate global sensitivity analysis for dynamic crop models. Field Crops Res. 113, 312–320.
Lazzarotto, P., Calanca, P., Fuhrer, J., 2009. Dynamics of grass-clover mixtures – An analysis of the response to management with the PROductive GRASsland Simulator (PROGRASS). Ecol. Modell. 220 (5), 703–724.
Li, J., Convertino, M., 2021. Inferring ecosystem networks as information flows. Sci. Rep. 11, 7094.
Lussem, U., Bolten, A., Kleppert, I., Jasper, J., Gnyp, M.L., Schellberg, J., Bareth, G., 2022. Herbage mass, N concentration, and N uptake of temperate grasslands can adequately be estimated from UAV-based image data using machine learning. Remote Sens. (Basel) 14 (13).
Moore, K.J., Moser, L.E., Vogel, K.P., Waller, S.S., Johnson, B.E., Pedersen, J.F., 1991. Describing and quantifying growth stages of perennial forage grasses. Agron. J. 83, 1073–1077. https://doi.org/10.2134/agronj1991.00021962008300060027x.
Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33, 14.
Moulin, S., Bondeau, A., Delecolle, R., 1998. Combining agricultural crop models and satellite observations: from field to regional scale. Int. J. Remote Sens. 19 (6), 1021–1036.
Pullanagari, R.R., Kereszturi, G., Yule, I., 2016. Mapping of macro and micro nutrients of mixed pastures using airborne AisaFENIX hyperspectral imagery. ISPRS J. Photogr. Remote Sens. 117, 1–10.
Pullanagari, R.R., Kereszturi, G., Yule, I., 2018. Integrating airborne hyperspectral, topographic, and soil data for estimating pasture quality using recursive feature elimination with random forest regression. Remote Sens. (Basel) 10 (7), 1117.
Riley, H., 2021. Vanning til jord- og hagebruksvekster. En litteraturstudie av norske undersøkelser siden 1960. NIBIO Rapport 7 (160).
Riley, H., Berentsen, E., 2009. Estimation of water use for irrigation in Norwegian agriculture. Bioforsk Rapport. 4 (174).
Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., Sejdinovic, D., 2019. Detecting and quantifying causal associations in large nonlinear time series datasets. Sci. Adv. 5, 1–15.
Saltelli, A., Taranola, S., Campolongo, F., Ratto, M., 2004. Sensitivity Analysis in Practice, New York.
Sierra, J., Brisson, N., Ripoche, D., Noël, C., 2003. Application of the STICS crop model to predict nitrogen availability and nitrate transport in tropical acid soil cropped maize. Plant Soil 256, 333–345.
Stafford, J.V., 2000. Implementing precision agriculture in the 21st century. J. Agricult. Eng. Res. 76 (3), 267–275.
Steinshamn, H., Nesheim, L., Bakken, A.K., 2016. Grassland production in Norway. Grassland Sci. Europe 21, 15–25.
Torssell, B.W.R., Kornher, A., 1983. Validation of a yield prediction model for temporary grassland. Sedish J. Agricult. Res. 13, 125–135.
Torssell, B.W.R., Kornher, A., Svensson, A., 1982. Optimization of Parameters in a Yield Prediction Model For Temporary Grassland. Swedish University of Agricultural Sciences, Department of Plant Husbandry, Uppsala. Report 112.
Vold, A., Søreng, J., 1997. Optimization of dynamic plant nitrogen uptake, using apriori information of plant nitrogen content. Biometrical J. 39 (6), 707–718.
Wachendorf, M., Fricke, T., Möckel, T., 2018. Remote sensing as a tool to assess botanical composition, structure, quantity and quality of temperate grasslands. Grass Forage Sci. 73 (1), 1–14.
Whelan, B.M., McBratney, A.B., 2000. The "Null Hypothesis" of precision agriculture management. Precis. Agricult. 2, 265–279.
Wijesingha, J., Astor, T., Schulze-Brüninghoff, D., Wengert, M., Wachendorf, M., 2020. Predicting forage quality of Grasslands using UAV-borne imaging spectroscopy. Remote Sens. (Basel) 12 (1), 126.
Zare, H., Weber, T.K.D., Ingwersen, J., Nowak, W., Gayler, S., Streck, T., 2022. Combining crop modeling with remote sensing data using a particle filtering technique to produce real-time forecasts of winter wheat yields under uncertain boundary conditions. Remote Sens. 14, 1–26.