



## Grass modelling in data-limited areas by incorporating MODIS data products



Xiao Huang<sup>a,b</sup>, Gang Zhao<sup>c</sup>, Conrad Zorn<sup>d</sup>, Fulu Tao<sup>e</sup>, Shaoqiang Ni<sup>f</sup>, Wenyuan Zhang<sup>g</sup>, Tongbi Tu<sup>h,\*</sup>, Mats Höglind<sup>b</sup>

<sup>a</sup> State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan, China

<sup>b</sup> Norwegian Institute of Bioeconomy Research, Klepp Station, Norway

<sup>c</sup> Department of Global Ecology, Carnegie Institution for Science, Stanford, USA

<sup>d</sup> Department of Civil and Environmental Engineering, University of Auckland, Auckland, New Zealand

<sup>e</sup> Natural Resources Institute Finland (Luke), Helsinki, Finland

<sup>f</sup> Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China

<sup>g</sup> Department of Zoology, University of Oxford, Oxford, UK

<sup>h</sup> Center of Water Resources and Environment, Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), School of Civil Engineering, Sun Yat-Sen University, Guangzhou, China

### ARTICLE INFO

#### Keywords:

Process-based grass models  
Data-limited areas  
MODIS data products  
Bayesian calibration  
Ensemble Kalman filter  
BASGRA

### ABSTRACT

Process-based grass models (PBGMs) are widely used for predicting grass growth under potential climate change and different management practices. However, accurate predictions using PBGMs heavily rely on field observations for data assimilation. In data-limited areas, performing robust and reliable estimates of grass growth remains a challenge. In this paper, we incorporated satellite-based MODIS data products, including leaf area index, gross primary production and evapotranspiration, as an additional supplement to field observations. Popular data assimilation methods, including Bayesian calibration and the updating method ensemble Kalman filter, were applied to assimilate satellite derived information into the BASic GRAssland model (BASGRA). A range of different combinations of data assimilating methods and data availability were tested across four grassland sites in Norway, Finland and Canada to assess the corresponding accuracy and make recommendations regarding suitable approaches to incorporate MODIS data. The results demonstrated that optimizing the model parameters that are specific for grass species and cultivar should be targeted prior to updating model state variables. The MODIS derived data products were capable of constraining model's simulations on phenological development and biomass accumulation by parameter optimization with its performance exceeding model outputs driven by default parameters. By integrating even a small number of field measurements into the parameter calibration, the model's predictive accuracy was further improved - especially at sites with obvious biases in the input MODIS data. Overall, this comparative study has provided flexible solutions with the potential to strengthen the capacity of PBGMs for grass growth estimation in practical applications.

### 1. Introduction

Grassland is one of the largest ecosystems in the world (Suttie et al., 2005), occupying up to 40 % of the total terrestrial surface (Blair et al., 2014). In the high-latitude areas, grass-based forages are important sources for dairy and meat production, providing necessary nutrients (e.g. protein, fibre) to livestock (Dengler et al., 2020). However, the biomass productivity of grasslands within this region is instable under

climate change (Wiréhn, 2018). The low temperatures in winter (Cohen et al., 2012), as well as drought hazards during the growing season (Bakke et al., 2020), can significantly affect the survival and growth of perennial grasses (Höglind et al., 2010; Østrem et al., 2015). Ultimately, this can lead to substantial inter-annual variation in grass yields (Rende, 2019) and it is therefore of great importance to accurately estimate the grass growth dynamics to improve food security and adaptation to future climates (Höglind et al., 2013).

\* Corresponding author.

E-mail addresses: [damon19910125@gmail.com](mailto:damon19910125@gmail.com) (X. Huang), [gzhao@carnegiescience.edu](mailto:gzhao@carnegiescience.edu) (G. Zhao), [conrad.zorn@auckland.ac.nz](mailto:conrad.zorn@auckland.ac.nz) (C. Zorn), [fulu.tao@luke.fi](mailto:fulu.tao@luke.fi) (F. Tao), [nsq15@mails.tsinghua.edu.cn](mailto:nsq15@mails.tsinghua.edu.cn) (S. Ni), [wenyuan.zhang@zoo.ox.ac.uk](mailto:wenyuan.zhang@zoo.ox.ac.uk) (W. Zhang), [tutb@mail.sysu.edu.cn](mailto:tutb@mail.sysu.edu.cn) (T. Tu), [mats.hoglind@nibio.no](mailto:mats.hoglind@nibio.no) (M. Höglind).

Process-based grass models (PBGMs) have been developed to simulate grass growth processes under changing environmental conditions from field to regional scales. Such models, e.g. CENTURY (Parton, 1996), PaSim (Vuichard et al., 2007) and BASGRA (Höglin et al., 2016), are widely used to predict grass yields and advise optimal practices for grassland management (Graux et al., 2013; Höglin et al., 2013; Shah et al., 2020). For best performance and greater predictive accuracy of PBGMs, the assimilation of real observations (e.g. yield records) is advised to constrain the uncertainty of model parameters and structure. As PBGMs typically use a number of parameters to describe the empirical relationships between state variables (e.g. LAI, Leaf Area Index) and fluxes (e.g. photosynthetic rate), different parameter values are required to account for the diversity in the physiological properties of different grass cultivars (Van Oijen and Höglind, 2016). One common strategy to assimilate observations is to optimize key input parameters such that the model outputs and field observations agree within a desired level of accuracy (*calibration method*). A range of algorithms, including least squares algorithm (Marquardt, 1963), GLUE (Beven and Binley, 1992) and Bayesian inference (Box and Tiao, 2011), have been applied to optimize model parameters and evaluate the uncertainty of crop model simulations (Guérif and Duke, 1998; Varella et al., 2010; Hjelkrem et al., 2017; Huang et al., 2018).

In addition, with the ongoing development of satellite platforms and unmanned aerial vehicles (UAVs) for data capture, some state variables (e.g. LAI; SM, soil moisture) can be derived from the observed images with higher spatial-temporal resolution than field observations (Jin et al., 2018). These data products, e.g. MODIS-LAI (Knyazikhin, 1999) and AMSR-SM (Njoku et al., 2003), can be used to update model's state variables in real time to reduce the accumulative bias at each time step and thus improving the model's predictive skills (Ines et al., 2013; De Bernardis et al., 2016; Huang et al., 2016; Chen et al., 2018). Popular algorithms for such data assimilation approaches (*updating method*) include the ensemble Kalman filter (EnKF; Evensen, 1994) and particle filter (PF; Liu and Chen, 1998).

Remote sensing data however is seldom used in isolation, with field measurements also required. For example, measured LAI is needed to correct the satellite-based data products (Huang et al., 2015) or calibration of the biophysical process parameters is performed in advance by yield records (Ines et al., 2013; Zhang et al., 2020). However, it is not always possible to obtain enough field data for the data assimilation of PBGMs in practical applications, with many previous studies on grass modelling conducted under conditions with assumed "sufficient" field observations, irrespective of which assimilation strategy is applied (Wallach, 2011; Ben Touhami and Bellocchi, 2015; Höglind et al., 2020). In many cases, there are very few or even no observations in targeted study areas. Unlike cereal crops, grasslands are usually managed for pasture, silage and hay production, making it difficult to obtain such accurate records at different scales (as the grass is utilized at the farm and not sold as cash crop) for modelling and simulation. This is particularly an issue in some less developed regions (e.g. Africa), where there can be little field data available in terms of on-site measurements and/or regional scale statistics. How to perform robust and reliable predictions in data-limited regions still remains a challenge, as it obviously restricts the feasibility of PBGMs under different conditions.

Satellite-based data, including LAI, GPP (gross primary production) and ET (evapotranspiration), could be an easily accessible supplement in such data limited regions. However, to our knowledge, there have been no studies on grass modelling within data-limited areas yet, thus how to properly combine common data assimilation algorithms with satellite-based data for grass modelling purposes is unknown. Therefore, in this paper, two data assimilation strategies, including a calibration method and an updating method, are applied to assimilate the MODIS data products into the established BASGRA N model (Höglin et al., 2020) for grass growth simulation. Various assimilating schemes, including applying data assimilation algorithms independently or collaboratively (Ines et al., 2013; Zhao et al., 2013; Hjelkrem et al., 2017), transferring

optimal parameters from other sites (Patil and Stieglitz, 2015; van der Linden and Woo, 2003) and different data usage (Ines et al., 2013; Wang et al., 2013; Chen et al., 2018), are considered and compared given their popularity in previous studies. Applied to four grassland sites in Norway, Finland, and Canada, we conduct a range of simulations and compare the accuracy of different approaches. By comparing and analyzing the model simulations, our main objectives are therefore to: (i) explore the effectiveness of common data assimilation method in data-limited areas; (ii) illustrate both the strength and weakness of representative data assimilation methods in such modeling tasks; and (iii) recommend the most appropriate data assimilation strategies based on different scenarios of data availability.

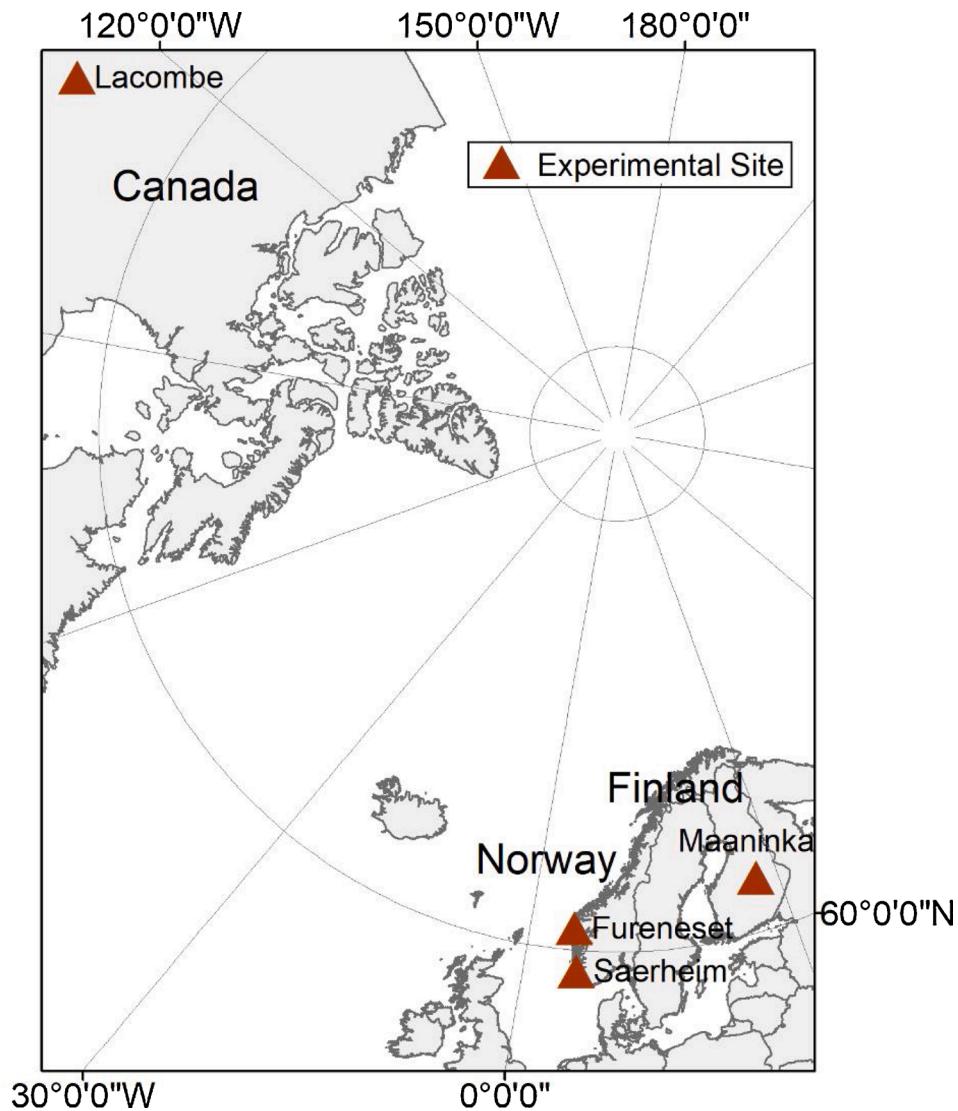
## 2. Materials and methods

### 2.1. Study area

Four experimental sites, including *Fureneset* and *Særheim* in Norway, *Maaninka* in Finland and *Lacombe* in Canada, were used for the modelling experiments (see Fig. 1). Timothy was the main grass species at all sites but the cultivar varied due to farmer's preference. At each site, several plot experiments (X-meter size) were conducted under different management practices surrounded by large-scale grasslands. Therefore, in this paper, the corresponding plots under management that are consistent with the surrounding large field were selected in representative of the average condition within the region. Detailed field observations for aboveground dry matter (DM), day matter yield (YIELD), specific leaf area (SLA), total tiller density (TITLTOT), carbon reserve (RES), frost tolerance (LT50) and LAI were available at these sites to calibrate model's parameters or validate the accuracy of model predictions. Daily climatic data measured from the automatic meteorology stations close to each site, as well as the records of management practices (e.g. fertilization, harvest date), were also obtained as inputs to drive model simulations. The selected sites in this study covered a wide range of climatic conditions, soil properties and management practices within the high-latitude regions that are critical to grass growth. The information about the four sites is provided in Table 1.

### 2.2. Satellite data products

The reprocessed MODIS version 6 LAI dataset from Land-Atmosphere Interaction Research Group at Sun Yat-sen University was used in this study (<http://globalchange.bnu.edu.cn/research/lai6#download>). This dataset reprocessed the raw data from MODIS MCD15A2H and MOD15A2H to fill the gaps in data continuity and consistency due to the cloud and seasonal snow cover (Yuan et al., 2011), with the final data generated using the Savitzky–Golay (SG) filter (Chen et al., 2004) to remove the noise in the raw data. For ET and GPP, we used the average of MODIS MOD16A2 and MYD16A2 as the MODIS-ET data (<https://modis.gsfc.nasa.gov/data/dataprod/mod16.php>), and the average of MODIS MOD17A2 and MYD17A2 as the MODIS-GPP data (<https://modis.gsfc.nasa.gov/data/dataprod/mod17.php>). The LAI, ET and GPP data products had an 8-day temporal frequency, with a 500-m spatial resolution for LAI and ET, and a 1-km spatial resolution for GPP data. No filtering processing was performed to ET and GPP datasets because their dynamics were greatly influenced by the short-term climatic conditions and we could not distinguish between noise and real fluctuations. At each experimental site, we used the average values of the cell where the coordinates in Table 1 were located and the surrounding 8 grid cells for LAI, GPP and ET. Given the small size of the plot experiments within the 1.5 km × 1.5 km region, plots under different management practices generally had a minor influence on the average values.



**Fig. 1.** Locations of the four studied sites.

**Table 1**  
Details of the selected study sites.

Site	Fureneset	Lacombe	Saerheim	Maaninka
Country	Norway	Canada	Norway	Finland
Coordinates	61.29 °N 5.04 °E	52.28 °N 113.44 °W	58.76 °N 5.65 °E	63.14 °N 27.32 °E
Annual precipitation (mm)	2280	429	1430	560
Mean daily temperature (°C)	7.7	3.5	7.8	4.2
Data Collection Period	2005–2006	2004–2005	2001–2002	2006–2007
Grass species	Timothy	Timothy	Timothy	Timothy
Grass cultivar	Grindstad	Climax	Grindstad	Tammisto II
Field observation Data	RES, DM, LAI, LT50, SLA, TITLTOT, YIELD	DM, LAI	RES, DM, LAI, SLA, TITLTOT	DM, LAI
Data source	Höglind et al. (2006)	Jing et al. (2012)	Höglind et al. (2010)	Virkajarvi et al. (2012)

### 2.3. BASGRA model

The BASGRA model is a PBGM that simulates the dynamics of leaves, roots, stems and tillers at daily time steps (Höglind et al., 2020). It has been well validated for its predictive accuracy under diverse environments (Hjelkrem et al., 2017; Woodward et al., 2020). The accumulation of dry matter in this model depends on the photosynthesis and remobilization of reserves. The photosynthetic module is based on the multiplication of intercepted radiation that is calculated using LAI and

the light use efficiency that is determined by the environmental factors including air temperature, CO<sub>2</sub> concentration and the Rubisco concentration in leaves. Meanwhile, detailed overwintering processes of frost and ice related stresses on tiller survival, e.g. cold hardening and dehardening, are implemented in this model, which affect the LAI and biomass accumulation in the subsequent growing season. The BASGRA model also simulates the water and nitrogen balance in soil environment and they in turn act as limiting factors for grass growth. LAI is a highly important state variable in BASGRA that links photosynthesis,

evapotranspiration and tiller development.

#### 2.4. Data assimilation methods

##### 2.4.1. Bayesian calibration

We applied Bayesian calibration (BC) as calibration method for data assimilation. The general framework followed is based on Bayes' theorem (Box and Tiao, 2011):

$$p(\boldsymbol{\theta}|Y) = \frac{p(Y|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})}{p(Y)} \propto p(Y|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}) \quad (1)$$

Where  $\boldsymbol{\theta}$  is the vector of calibrated parameters;  $Y$  is the matrix of observations (e.g. dry matter weight, LAI);  $p(\boldsymbol{\theta}|Y)$  is the posterior distribution of parameter given the observations used for calibration;  $p(\boldsymbol{\theta})$  is the prior probability of parameters;  $p(Y|\boldsymbol{\theta})$  is the likelihood of observations for any values of parameters;  $p(Y)$  is the marginal probability of observational data. The likelihood  $p(Y|\boldsymbol{\theta})$  can be calculated as:

$$p(Y|\boldsymbol{\theta}) = p(\epsilon_y = \tilde{Y} - Y) \sim f_{\epsilon_y}(\epsilon_y) \quad (2)$$

Where  $\tilde{Y}$  is the matrix of model outputs based on  $\boldsymbol{\theta}$ ;  $\epsilon_y$  is residual error between model observations and outputs;  $f_{\epsilon_y}$  is the probability density distribution of  $\epsilon_y$ .

In this study, we assumed Sivia's distribution for the prior marginal distribution of parameters. 25 parameters critical to grass modelling were selected for calibration with their descriptions, prior value ranges and default values shown in Table 2. Default values were generally determined as the average of minimum and maximum values. According to the direct effects of these parameters on targeted state variables, we categorized them into three classes (see Table 2): (i) Class A parameters were linked to biomass accumulation; (ii) Class B parameters were linked to leaf and LAI modelling; and (iii) Class C were linked to tiller development. Default values for other parameters are listed in Table S1. The probability of residual error was assumed to follow a Gaussian distribution. The Metropolis-Hastings algorithm was applied for the Markov chain Monte Carlo (MCMC) sampling, with the maximum chain length as 50,000 and the first 20 % for burn in. We used the model

outputs with the maximum posterior likelihood and the corresponding parameters in this paper.

##### 2.4.2. EnKF algorithm

The EnKF method is the Monte Carlo implementation of the original Kalman Filter for the ensemble forecasting of non-linear systems (Evensen, 2004). In this method, model state variables (e.g. LAI) are represented as ensembles following a Gaussian distribution and they are combined and updated with observational data in time as modelling proceeds. The fundamental step to update model states is defined as:

$$A^a = A + P_e H^T (H P_e H^T + R_e)^{-1} (D - HA) \quad (3)$$

where  $A^a$  and  $A$  are the analyzed (updated) and forecasted (modelled) matrices of ensemble states;  $P_e$  and  $R_e$  are the covariance matrices of forecasted ensemble and observation;  $H$  is the measurement operator and  $(D - HA)$  is the innovation vector. In this study, only one state variable LAI is assimilated into model. Therefore,  $A^a$  and  $A$  can be reduced to vectors,  $P_e$  and  $R_e$  can be reduced to variances. As model state is directly observed,  $H$  is an identity matrix and Eq. (3) can be simplified as:

$$A_i^a = A_i + \frac{P_e}{P_e + R_e} (D_i - A_i) \quad (4)$$

where  $A_i^a$  and  $A_i$  are the analyzed (updated) and forecasted (modelled) LAI for the  $i$ th ensemble member;  $P_e$  and  $R_e$  are the variances of modelled LAI and MODIS-LAI;  $D_i$  is the perturbed MODIS-LAI of the  $i$ th ensemble member.

In this study, the ensemble size was set as 200. As the simulating tasks were performed for data-limited area, the MODIS-LAI was not corrected using field measurements. We assumed a 25 % error in the MODIS LAI values after comparing the relative error between measured and MODIS-LAI data samples (see Fig. S1). At each time step, the perturbed LAI observation can be computed as:

$$D_i = D_{MODIS} \cdot f_e, f_e \sim U(0.75, 1.25) \quad (5)$$

where  $D_{MODIS}$  is the MODIS-LAI value and  $f_e$  is the random factor with a

**Table 2**  
Selected parameters used for model calibration.

Parameter	Unit	Description	Range	Default	Class
CLVI	$\text{gC}\cdot\text{m}^{-2}$	Initial value of leaf biomass	(0.0, 316.0)	31.6	A
LAI0	—	Initial value of LAI	(0.0, 10.0)	1.0	B
TILTOTTI	$\text{m}^{-2}$	Initial value of tiller density	(1000, 3000)	2000	C
CSTAVM	$\text{gC}\cdot\text{tiller}^{-1}$	Maximum size of elongating tillers	(0.1, 1.9)	1.0	A
DAYLB	$\text{d}\cdot\text{d}^{-1}$	Day length below which phenological stage is reset to zero	(0.0, 0.8)	0.4	B
DAYLP	$\text{d}\cdot\text{d}^{-1}$	Day length below which phenological development slows down	(0.3, 1.0)	0.65	B
FSLAMIN	—	Minimum SLA of new leaves as a fraction of maximum possible SLA	(0.00, 0.93)	0.47	B
K	$\text{m}^2\cdot\text{m}^{-2}\text{leaf}$	PAR extinction coefficient	(0.3, 0.9)	0.5	A
LAICR	$\text{m}^2\cdot\text{m}^{-2}\text{leaf}$	LAI above which shading induces leaf senescence	(1.9, 7.6)	3.8	B & C
LAIEFT	$\text{m}^2\cdot\text{m}^{-2}\text{leaf}$	Decrease in tillering with leaf area index	(0.1, 0.4)	0.2	C
LFWIDG	m	Leaf width on elongating tillers	(0.004, 0.017)	0.008	B
LFWIDV	m	Leaf width on non-elongating tillers	(0.0025, 0.0098)	0.0049	B
NELLVM	$\text{tiller}^{-1}$	Number of elongating leaves per non-elongating tiller	(1.0, 3.5)	2.1	B
RUBISC	$\text{g}\cdot\text{m}^{-2}\text{leaf}$	Rubisco content of upper leaves	(2.9, 11.6)	5.8	A
SHAPE	—	Area of a leaf relative to a rectangle of same length and width	(0.27, 1.00)	0.54	B
SLAMAX	$\text{m}^2\text{leaf gC}^{-1}$	Maximum SLA of new leaves	(0.03, 0.09)	0.06	B
TBASE	°C	Minimum value of effective temperature for leaf elongation	(1.8, 6.0)	3.6	B
RATEDMX	$^{\circ}\text{C}\cdot\text{d}^{-1}$	Maximum dehardening rate	(0.5, 2.5)	1.80	C
Hparam	$^{\circ}\text{C}^{-1}\cdot\text{d}^{-1}$	Hardening parameter	(0.005, 0.010)	0.0082	C
YG	$\text{gC}\cdot\text{g}^{-1}\text{C}$	Growth yield per unit expended carbohydrate	(0.65, 0.90)	0.84	A
DAYLG1G2	$\text{d}\cdot\text{d}^{-1}$	Minimum day length above which generative tillers can start elongating (by moving from TILG1 to TILG2).	(0.0, 1.0)	0.6	C
RGRTG1G2	$\text{tiller}\cdot\text{tiller}^{-1}\cdot\text{d}^{-1}$	Relative rate of TILG1 becoming TILG2	(0.0, 1.0)	0.9	C
NCSHMAX	$\text{gN/gC}$	Maximum N-C ratio of shoot	(0.02, 0.08)	0.04	A
TCNSHMOB	d	Time constant of shoot N remobilization	(1, 64)	8	A
TCNUPT	d	Time constant of soil mineral N uptake	(1, 64)	8	A

Class A: parameter related to biomass; Class B: parameter related to LAI; Class C: parameter related to tiller.

uniform distribution. With  $D_i$  obtained,  $R_e$  could be calculated afterwards following Burgers et al. (1998). We used the mean values of the total 200 ensemble predictions for the following analysis.

## 2.5. Design of numeric experiment

To simulate various options of data-availability, we assume the following scenarios:

- (i) **S0**: Detailed field observations at each site, as well as satellite-based data products, were available for data assimilation;
- (ii) **S1**: There were no field observations in the study area. However, data was available for use from other sites with similar grass cultivar;
- (iii) **S2**: There were no field observations in the study area. Only satellite-based data products were accessible for data assimilation;
- (iv) **S3**: There were very few field observations in the study area. DM is assumed to be the only information available (given it is the easiest to measure under normal conditions);

The different data assimilation strategies we simulate are:

- (i) **DA0**: Running the model simulation with the default parameter values listed in Tables 2 and Table S1 (in Supplementary Material);
- (ii) **DA1**: Running the model simulation with default parameter values and updating the model LAI with MODIS-LAI using the EnKF algorithm described in section 2.4.2;
- (iii) **DA2**: Calibrating the parameters in Table 2 using the BC method described in Section 2.4.1 and then running the model simulation with optimal parameter values;
- (iv) **DA3**: Running the model simulation with optimal parameter values obtained from DA2 and updating the model LAI with MODIS-LAI using the EnKF algorithm;

**Table 3** provides an overview of the scenario and data-assimilation strategy combinations we applied in this study. For parameter calibration DA2, (i) under scenarios **S1**, only the MODIS data at the corresponding site was used; (ii) under scenarios **S2**, the optimal parameters calibrated for another site using the field observations were used for the targeted site. For example, the optimal parameters at Fureneset, Lacombe and Saerheim obtained from DA2 were used to drive the

simulation at Maaninka separately; (iii) under scenario **S3**, both the field observation DM and MODIS data in the corresponding site were used; (iv) under scenario **S3**, to demonstrate if the modelling performance was caused by DM alone or DM and MODIS data together, we added Exp-S3-DA2-3 that only used field DM observations for model calibration. Unlike LAI, that was calculated directly from satellite signals, the GPP and ET products from MODIS platform were derived by simple models combining LAI and other information. As a result, in our simulated scenarios, we first used LAI alone (only LAI for parameter calibration) and then LAI + GPP + ET jointly (MODIS-LAI, GPP and ET all used for parameter calibration) for BC to compare the effectiveness of these two data combinations.

## 2.6. Evaluation

To evaluate the performance of model simulation under different scenarios and data assimilation strategies, we chose DM, LAI, TITLTOT and YIELD as the most important variables for grass growth and compared their simulated values with the field observations at each site. Two indicators, Root Mean Square Error (RMSE) and Normalized Root Mean Square Error (NRMSE), were utilized in this research to quantify the accuracy of model prediction.

## 3. Results

### 3.1. Model outputs under scenario S0 as baseline

The selected parameters in Table 2 were calibrated at each site

**Table 4**  
The relative deviation of optimal parameters from default values.

Unit: %	Fureneset	Lacombe	Saerheim	Maaninka	Standard variation <sup>a</sup>
Parameter (Class A - Biomass)	36.1	50.3	41.7	23.6	26.5
Parameter (Class B - LAI)	29.1	31.8	33.1	25.6	29.5
Parameter (Class C - Tiller)	12.8	27.5	21.3	28.4	19.1

<sup>a</sup> The standard variation among the four sites relative to the default value.

**Table 3**  
Combinations of scenarios and data assimilation methods used in this study.

Scenarios	Data assimilation	Experiment codes	Satellite data products	Field observations	Description
S0	DA0	Exp-S0-DA0	–	–	Using default parameter values
	DA1	Exp-S0-DA1	EnKF: LAI;	–	Exp-S0-DA0 + EnKF
	DA2	Exp-S0-DA2	–	BC: all the field observations at each corresponding site;	Using optimal parameter values
	DA3	Exp-S0-DA3	EnKF: LAI;	BC: all the field observations from other sites;	Exp-S0-DA2 + EnKF
S1	DA2	Exp-S1-DA2	–	BC: all the field observations from other sites;	Using optimal parameter values from other sites
	DA3	Exp-S1-DA3	EnKF: LAI;	BC: LAI;	Exp-S1-DA2 + EnKF
	DA2	Exp-S2-DA2-1	BC: LAI, GPP, ET;	BC: LAI;	Using optimal parameters calibrated by remote sensing data
	DA2	Exp-S2-DA2-2	BC: LAI;	–	Exp-S2-DA2-1 + EnKF
S2	DA3	Exp-S2-DA3-1	BC: LAI, GPP, ET; EnKF: LAI;	–	Exp-S2-DA2-2 + EnKF
		Exp-S2-DA3-2	BC: LAI; EnKF: LAI;	–	Using optimal parameters calibrated by both remote sensing data and DM
	DA2	Exp-S3-DA2-1	BC: LAI, GPP, ET;	BC: DM (at each site);	Exp-S3-DA2-1 + EnKF
		Exp-S3-DA2-2	BC: LAI;	–	Exp-S3-DA2-2 + EnKF
S3	DA3	Exp-S3-DA3-1	BC: LAI, GPP, ET; EnKF: LAI;	–	Using optimal parameter values only calibrated by DM
		Exp-S3-DA3-2	BC: LAI; EnKF: LAI;	–	–
	DA2	Exp-S3-DA2-3	–	–	–

individually under Exp-S0-DA2. We calculated the relative deviation of the optimal parameters from default values. The averages of each parameter class (A, B, and C) and the standard variations across the four sites are presented in Table 4. The results showed that site-specific optimal parameters were required in our study regions to ensure the accuracy of grass growth simulations, among which parameters in Class A showed the highest deviation in most sites. With regards to the variations between the four sites, at least 19 % difference in parameters was found despite Timothy being dominant at all sites. The highest variations were shown for Class A and Class B which showed the diversity of grass's response in different natural environments.

In Fig. 2, the model outputs for DM, LAI, TITLTOT and YIELD in daily steps at the four experiment sites under scenario S0 are displayed and compared with the field measurements. The results demonstrated significant diversity among different modelling schemes. LAI simulations using the EnKF method (Exp-S0-DA1 and Exp-S0-DA3) had obvious fluctuation in short term as MODIS-LAI was assimilated into the model, while it tended to be smoother for the other three variables. Meanwhile, we see the RMSE and NRMSE values for each site in Fig. 3, where simulations from Exp-S0-DA2 using optimal parameters had the highest accuracy for all the four variables at all sites. The additional application of EnKF using the same parameters (Exp-S0-DA3) did not improve the predictive accuracy as expected given MODIS-LAI is not bias-corrected. Similarly, model predictions using default parameter values (Exp-S0-DAO and Exp-S0-DA1) showed much lower accuracy. The effects of applying EnKF differed a lot for different variables at different sites. For example, it greatly improved the simulations of LAI at Fureneset and Saerheim, but decreased the estimation accuracy of DM at Maaninka. The overall performance ("Average" in Fig. 3) was similar between Exp-S0-DAO and Exp-S0-DA1 with the exception of the TITLTOT parameter. As a result, we adopt the results from Exp-S0-DAO and Exp-S0-DA2 in the following analysis as the baseline to assess the inaccuracies and accuracies of the remaining scenarios, respectively.

### 3.2. Model outputs under scenario S1

The model performance under scenario S1 using transferred optimal parameters is illustrated in Fig. 4. For example, at Fureneset, the optimal parameters in the other three sites obtained individually from Exp-S0-DA2 were used to independently drive model predictions and provide a range of estimates. In Fig. 4a-d, the results from Exp-S1-DA2

demonstrated clearly for all sites, that the worst model predictions (upper boundary) could be even less accurate than simply adopting default parameters in the model (Exp-S0-DAO). On the contrary, the best model predictions (lower boundary) showed a similar predictive accuracy as Exp-S0-DAO in most cases. Meanwhile, the results in Fig. 4e-h from Exp-S1-DA3 showed similar model performance with Exp-S1-DA2. The only exception was that the worst performing simulations of LAI were greatly improved for the Fureneset site (Fig. 4e) as MODIS-LAI data that was assimilated using the EnKF method was generally accurate during the sampling period.

The average RMSE and NRMSE of the four sites are given in Table 5. Generally, we see the high uncertain interval and instability of cross use of optimal parameters for grass modelling when the cultivars are different, as users cannot typically test which optimal parameter settings will provide a better performance in advance of the simulation.

### 3.3. Model outputs under S2

The results of model predictions under scenario S2 that use MODIS data to calibrate model parameters are shown in Fig. 5. At Fureneset (Fig. 5a&e), those simulations using only MODIS-LAI for parameter optimization (Exp-S2-DA2-2 and Exp-S2-DA3-2) led to high biases in simulation of DM and YIELD. However, the model performance of Exp-S2-DA2-1 and Exp-S2-DA3-1 was seen to greatly improve by incorporating MODIS-GPP to constrain the model prediction on biomass accumulation. The similar benefit of supplementing GPP into data assimilation on DM was found at the Lacombe site. At Saerheim, assimilating MODIS data only improved the model's performance for LAI but the simulations of TITLTOT were even worse than the initial Exp-S0-DAO scenario. Meanwhile, at Manninka, the accuracy of results from the four schemes were only slightly improved under scenario S2 compared to those under S1 and were still worse than that of Exp-S0-DAO. At this site, we observed significant deviations between MODIS-LAI and measured LAI (see Fig. S2), which likely influenced both the effectiveness of parameter optimization and the accuracy of MODIS-GPP and MODIS-ET input data. Moreover, in the four experiments under scenario S2, the difference in accuracy between data assimilation DA2 (Fig. 5a-d) and DA3 (Fig. 5e-h) was negligible. This was mainly because the MODIS-LAI dataset was used to calibrate model parameters and therefore the LAI simulations expectedly matched the MODIS-LAI data well. Additional assimilation of MODIS-LAI by EnKF did not change the

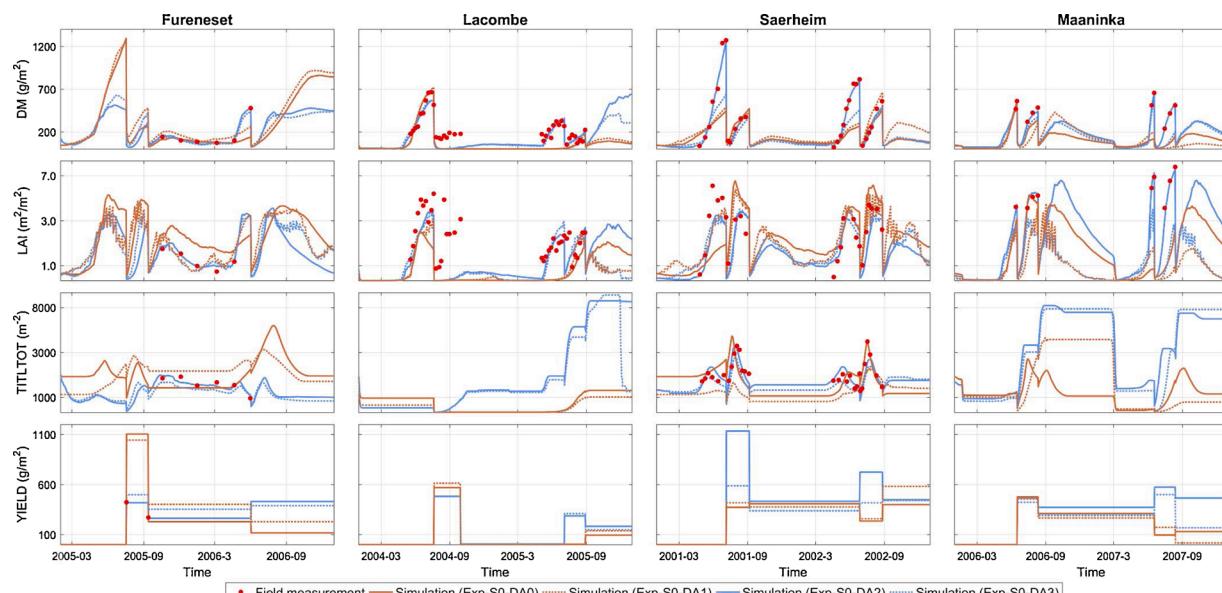


Fig. 2. Field measurements and daily model outputs under scenario S0.

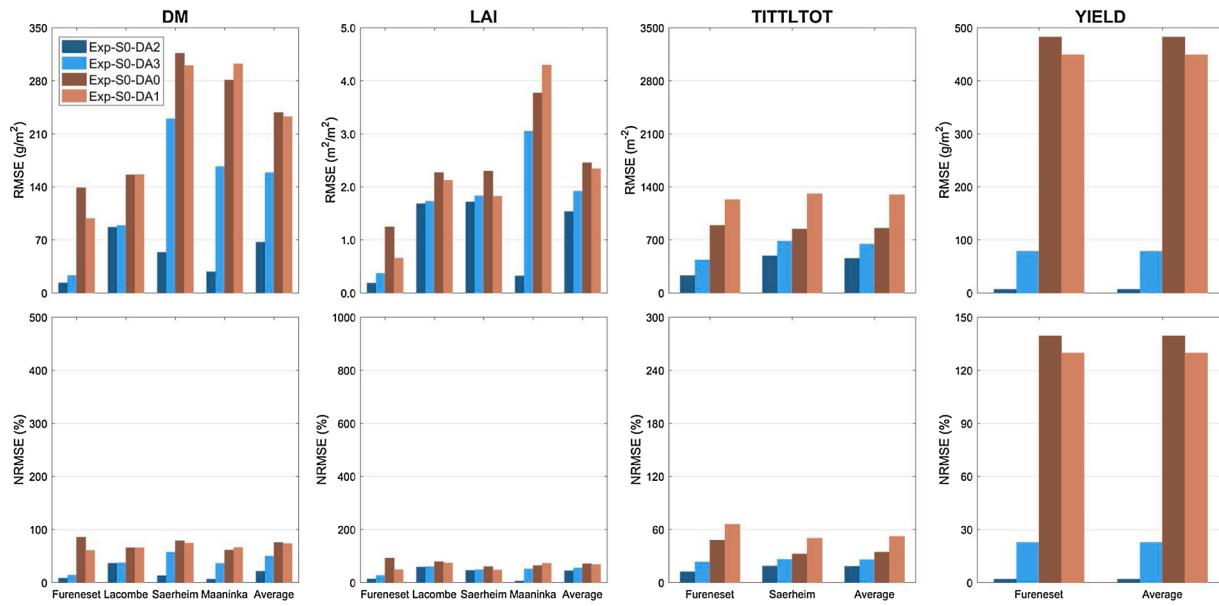


Fig. 3. Model performance under scenario S0.

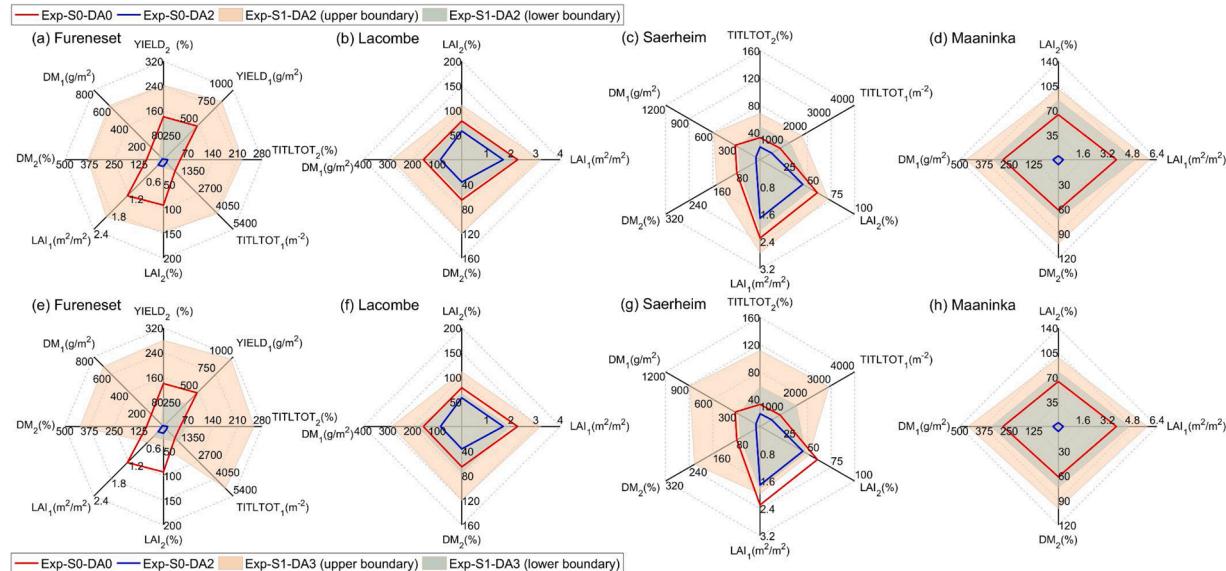


Fig. 4. Model Performance under scenario S1. (a)-(d): using data assimilation method DA2 (without EnKF), (e)-(h): using data assimilation method DA3 (with EnKF); Subscript 1: RMSE, subscript 2: NRMSE.

model's state variables considerably (see Fig. S3).

#### 3.4. Model outputs under scenario S3

Based on the S2 scenarios, we further incorporated field measurements of DM into parameter calibration, with these results shown under scenario S3 in Fig. 6. Firstly, under Exp-S3-DA2-3, the accuracy of DM simulations was guaranteed as measured DM was used for parameter calibration. However, with regards to other 3 variables, the corresponding accuracy decayed similar to or even worse than Exp-S0-DAO. It proved that calibrating model only with few field observations might lead to the overfit on a certain variable while sacrificing the robustness of model predictions on other variables, which would be risky and unreliable in changing environment. By comparison, the experiments integrating DM and MODIS data showed more balanced improvement on all variables. At Furenset (Fig. 6a&e), the predictive accuracy was

improved considerably compared with S2 and the results of Exp-S3-DA2-1 and Exp-S3-DA3-1 show accuracy close to Exp-S0-DA2. At Lacombe and Saerheim, the results were improved over the S2 scenarios by assimilating additional DM information. Predictive accuracy on most variables was similar to Exp-S0-DA2 while the performance of predicting TITLTOT at Saerheim was significantly enhanced compared with S2. At Manninka, with the assistance of measured DM, model simulations were better constrained, and the model performance exceeded the Exp-S0-DAO simulated results and got closer to Exp-S0-DA2 than S2 scenario. This demonstrated that even a few field measurements could effectively improve the robustness of parameter estimation in regions where the MODIS data was not accurate enough to be relied upon in isolation. Finally, the application of EnKF did not significantly change the final results under S3 for the same reason as explained in section 3.3 above.

**Table 5**

The overall performance of model predictions under scenarios S0 and S1.

Indicator	Experiment	DM	LAI	TITLTOT	YIELD
RMSE	Exp-S0-DAO	238	2.4	852	482
	Exp-S0-DA2	67	1.5	454	6
	Exp-S1-DA2	[261, 412]	[2.8, 3.1]	[1201, 2368]	[471, 835]
	Exp-S1-DA3	[270, 569]	[2.5, 3.0]	[1521, 3311]	[323, 969]
	Exp-S0-DA0	75.3	71.0	34.3	139.4
NRMSE	Exp-S0-DA2	21.2	44.4	18.3	1.9
	Exp-S1-DA2	[82.5, 130.3]	[81.4, 90.1]	[48.3, 95.3]	[136.1, 241.5]
	Exp-S1-DA3	[85.4, 180.0]	[73.0, 88.3]	[61.2, 133.2]	[93.5, 280.0]
	Exp-S0-DA0				

### 3.5. Comparison between different scenarios

The NRMSE for scenarios S1, S2 and S3 at all sites across DM, LAI, TITLTOT and YIELD parameters are shown in Fig. 7. In general, the results under scenario S3 had greater accuracy (as indicated by the darker colors) than the other two scenarios. This illustrated that in data-limited areas, grass growth estimation could be improved if even a small number of field measurements were available to be assimilated into the model. Under scenario S2, the model simulations totally relied on the MODIS data products. The predictive accuracy for DM was improved at Fureneset and Lacombe by adding MODIS-GPP into the parameter calibration. The results also showed that in data-limited areas without any field measurements, using MODIS data products was feasible in some cases but has lower accuracy than under scenario S3 (in which some field data is available). Under scenario S1, the optimal parameters from other sites were used for grass modelling, and the corresponding results varied a lot depending on which site is used. The best model performance (minimum value) under S1 was not as good as that of S3, but was comparable with or even better than that of S2. However, the worst performance (maximum value) under S1 was not acceptable for grass growth estimation.

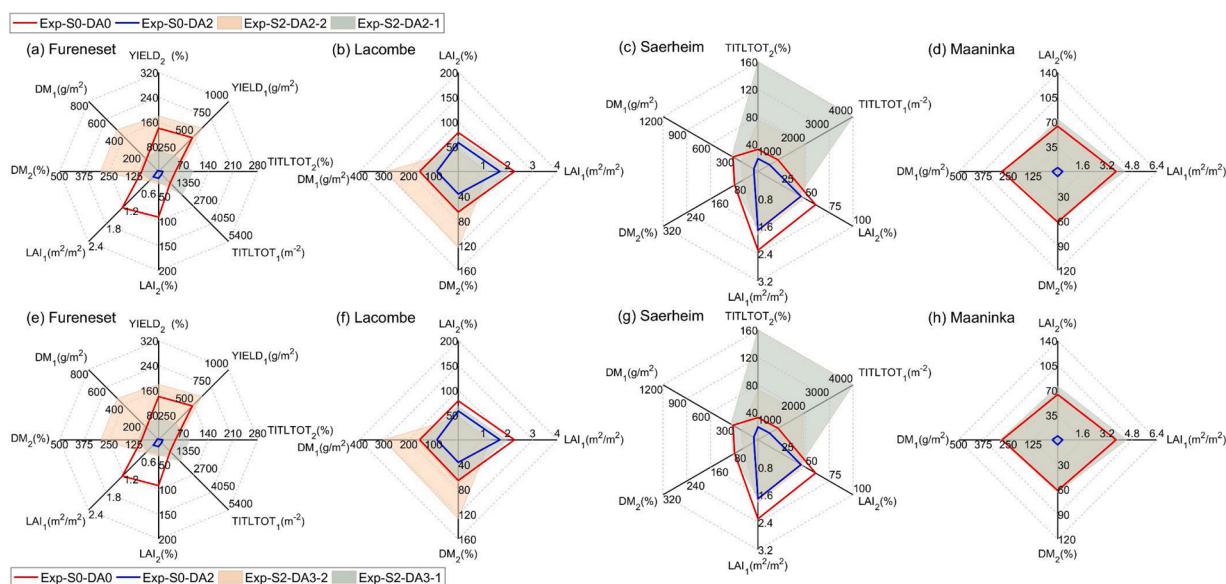
In addition, the relative deviations of parameters under S2 and S3 from optimal parameters of Exp-S0-DA2 are shown in Fig. 8. In terms of

parameters in Class A, smaller deviations were observed for scenario S3 compared with S2 at almost all site except Lacombe. Therefore, assimilating measured DM into BASGRA model had the most straightforward effect on processes on biomass accumulation. Under S2, additional assimilation of GPP was also helpful to constrain parameter estimation by comparing Exp-S2-DA2-1 with Exp-S2-DA2-2. However, when MODIS-GPP and measured DM were used together (Exp-S3-DA2-1), the mismatch between the satellite-based and measured data seemed to lower the effects in comparison with Exp-S3-DA2-2. Similar patterns applied to parameters in Class B at Fureneset, Lacombe and Saerheim. At Maaninka, as the MODIS-LAI were obviously biased from measurement, parameter estimation was constantly improved as more and more information was assimilated. Moreover, parameters in Class C showed the minimum deviation and different data usage for parameter optimization seemed to have an insignificant effect on results.

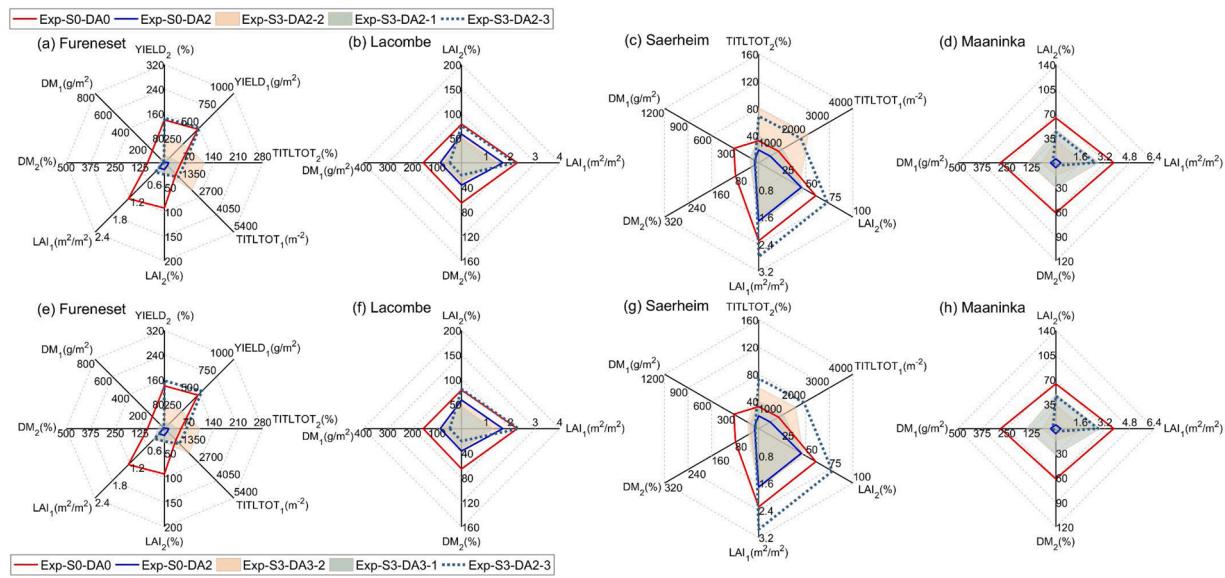
## 4. Discussion

### 4.1. Model transferability

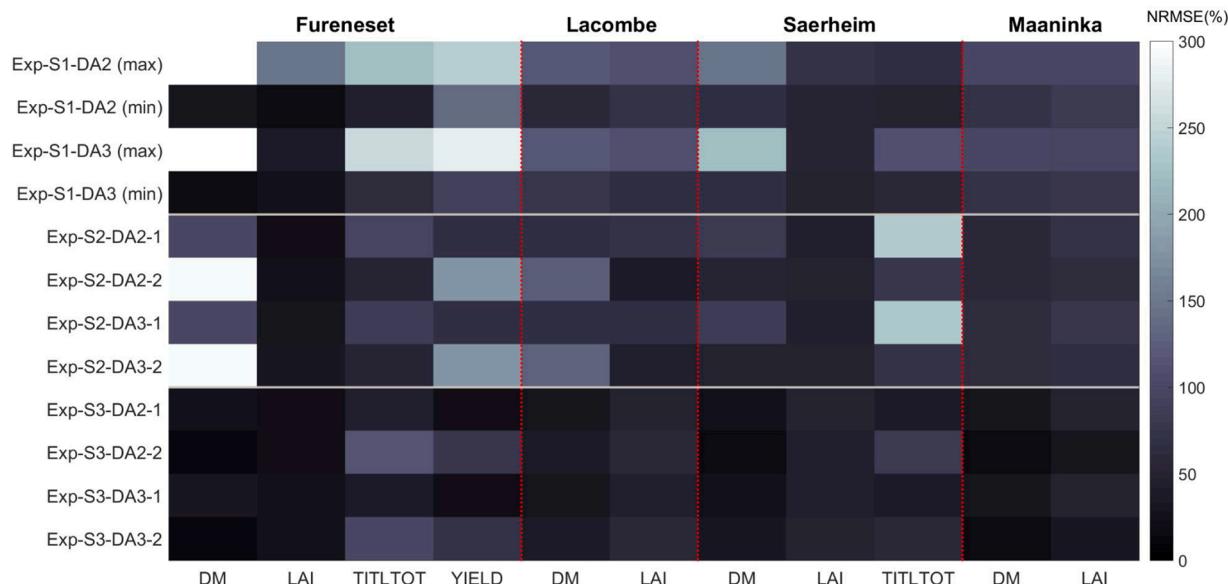
The results under scenario S1 clearly showed that parameters that were optimal to a certain site were not necessarily transferable to other sites. Given the grass species was consistent across all sites, we could have theoretically expected the optimal parameters to show similar predictive accuracies if the grass model could generally capture the key processes for the same grass type. However, the performance of parameter cross transfer demonstrated high uncertainty for the modelling of the targeted variables because of the different cultivars. It showed the grass parameters in BASGRA model were more cultivar-specific than species-specific, which implied that parameterization work, as well as model development, should further step into cultivar level. The physiological processes in grass or crop models could be classified into 2 type: (i) species-sensitive ones, and (ii) cultivar-sensitive ones. By this, users are able to decide which parameters are transferable in data-limited areas depending on the data availability of other sites. Meanwhile, for regional or global scale simulations, current applications usually used the same parameterization schemes for the same grass/crop type without considering the cultivar effects. More observations, e.g. the remote sensing data products used in this study, could be utilized to improve the representativeness of parameterization for large-scale modelling.



**Fig. 5.** Model Performance under scenario S2. (a)-(d): using data assimilation method DA2 (without EnKF), (e)-(h): using data assimilation method DA3 (with EnKF); Subscript 1: RMSE, subscript 2: NRMSE.



**Fig. 6.** Model Performance under scenario S3. (a)-(d): using data assimilation method DA2 (without EnKF), (e)-(h): using data assimilation method DA3 (with EnKF); Subscript 1: RMSE, subscript 2: NRMSE.



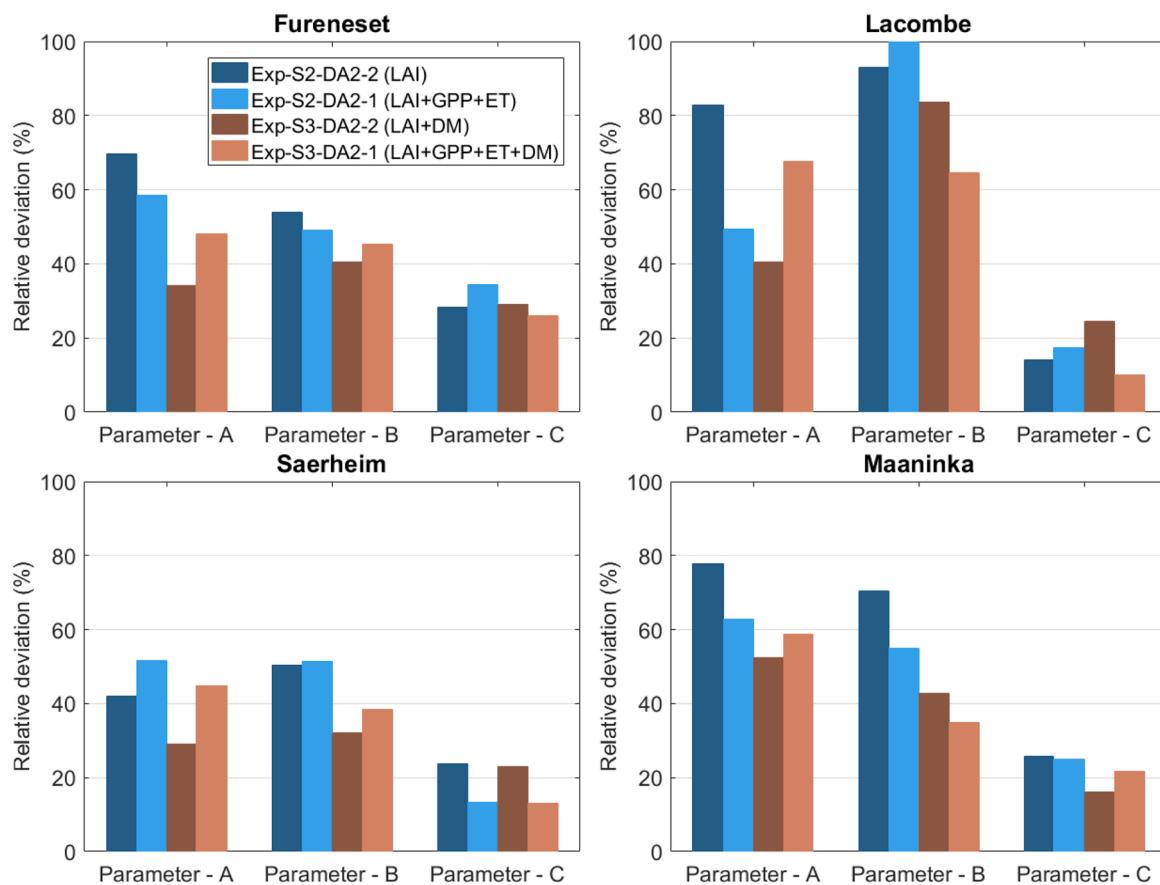
**Fig. 7.** The comparison of predictive accuracy between different scenarios across different model parameters.

#### 4.2. Data assimilation methods

Compared with other studies (Ines et al., 2013; Tewes et al., 2020; Zhao et al., 2013), we found the updating method using the EnKF algorithm did not lead to improvements in predictions for this study. On the one hand, it was clear that the effectiveness of the updating methods greatly depended on the robust estimation of model parameters (calibration method). As shown in the results of Exp-S0-DA1 and Exp-S1-DA3, model simulations driven by inaccurate parameterization could have even been worsened by assimilating MODIS-LAI. In comparison, under scenario S2 and S3, the predictive accuracy was enhanced with calibrated parameters. Therefore, ensuring the robustness and confidence in the model parameters should be the foundation for grass growth modelling. It should be noted that in high-latitude areas like our study area, MODIS observations were often associated with high uncertainties due to the frequent cloudy weather. In previous studies, postprocessing of remote sensing data was usually performed using field

observations to reduce the biases of raw data. However, under the assumption and objective of this study to simulate data-limited conditions, it was not necessary to conduct the correction for the MODIS-LAI values. The inaccurate remote sensing observations used here were also likely to lower the effectiveness of updating method.

This study also showed that different calibration methods and updating methods differed in their ways to influence grass growth simulations. Calibration methods could optimize model parameters and these parameters therefore adjusted more state variables throughout the whole simulating period. For example, under scenario S2 and S3, the MODIS-LAI used for parameter optimization not only directly modified the model simulation on LAI, but also considerably changed the simulations of grass leaf biomass and winter survival by calibrating the relevant parameters. In contrast, by using updating methods, model state variables were replaced at each time step, while model parameters were kept unchanged. To the best of our knowledge, in almost all studies using the EnKF updating method for crop modelling, LAI was the only



**Fig. 8.** The relative deviation of optimal parameters using MODIS data products from the optimal parameters using field observations. Class A: biomass accumulation parameters, Class B: LAI parameters, Class C: tiller development parameters.

variable used related to crop growth. However, none of the computing scheme was available to reverse any of the other model state variables (e.g. leaf carbon pool) when LAI was updated, although in a mechanistic sense, the variables were connected. For example, in Fig. 2 the fluctuation of LAI was very evident. But its influence on DM, TITLTOT and YIELD modelling was flattened in the simulations as the curves were quite smooth. We believe that the performance of updating methods could be boosted by integrating more detailed reversing schemes, transferring changes in LAI to one or more state variables instead of simply updating LAI in the simulation process.

#### 4.3. MODIS data products

In data-limited areas, we usually do not have enough field measurements to correct the MODIS data products. However, the results under scenario S2 and S3 demonstrated raw MODIS data was still applicable for grass modelling. Compared with field sampling, which were usually low in frequency, MODIS-LAI generally captured the spatial pattern of grass growth process well although its magnitude might be uncertain. As a result, such information was still helpful to optimize model parameters if it was not greatly biased from the real condition (e.g. Maaninka). Besides LAI, which was the most common remote sensing data for crop modelling, the present study also showed that MODIS-GPP and MODIS-ET were applicable for data assimilation to avoid an obvious discrepancy from reality. More importantly, although there were inaccuracies in the model simulations by only using MODIS data, supplementing few field measurements seemed to be very helpful to correct the bias in MODIS data and it therefore could more effectively constrain model's trajectory. On the other hand, the performance of simulating tasks using these remote sensing data products highly

depended on the accuracy of this data itself. The accuracy of such plant-related data products should be further improved for data-limited areas.

#### 4.4. Uncertainty and limitation

In this paper, the BASGRA model was used for grass simulation. As the modeling approaches and complexity differed largely between grass models, their responses to data assimilation methods and data products were expected to be diverse. Thus, the results from the numerical experiments were limited to the performance of the BASGRA model. More PBGMs could be used in the future to provide a more comprehensive vision on this topic.

Moreover, we used MODIS data products to represent the potentially available remote sensing information. This was because only the MODIS platform had large-scale data available in the long term. Other satellite platforms, e.g. Sentinel, are expected to provide more accurate data products with higher resolution in the future. In this research, we used the traditional version of BC and EnKF to test the performance of data assimilation in data-limited areas. However, many modified data assimilation methods based on BC and EnKF have been developed to improve the accuracy of crop modelling. These methods should be tested in future research focusing on data-limited areas.

#### 5. Conclusion

In this study, we conducted a number of modelling experiments with different combinations of data availability and data assimilation methods for grass growth estimation in data-limited areas. We concluded that, when using the BASGRA model, the MODIS- LAI, GPP and ET data could effectively constrain a model's simulations of

phenological development and biomass accumulation if they were not greatly biased from real conditions. Although the prediction accuracy may not be as good as when the model was entirely calibrated by field observations from the region where it was used, simulations with such data may improve the prediction accuracy substantially compared with model outputs driven by default parameters. Adding a few field measurements into data assimilation with satellite data products can further improve the predictive skills. Model performance increased in accuracy with less biased MODIS data, while it also improved significantly even if MODIS data was not accurate at one study site (Maaninka). We demonstrate the importance of a calibration method for parameter optimization for robust grass modelling before other data assimilation methods are applied. To improve the usefulness of the updating methods, the quality of input remote sensing data products should be targeted and additional reversing schemes should be explored to update other state variables that can also be used to assess model performance in data-limited areas.

## 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.

## Acknowledgments

X.H. acknowledges support from the "Visiting Researcher Fund Program" (2019SWG01) of State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University. Thanks go to the two anonymous reviewers and editor for their constructive comments on the paper. We also thank Dr. Panu Korhonen and Dr. Gilles Bélanger for providing the data used in this study.

## Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.fcr.2021.108250>.

## References

- Bakke, S.J., Ionita, M., Tallaksen, L.M., 2020. The 2018 northern European hydrological drought and its drivers in a historical perspective. *Hydrol. Earth Syst. Sci. Discuss.* 2020, 1–44. <https://doi.org/10.5194/hess-2020-239>.
- Ben Touhami, H., Bellocchi, G., 2015. Bayesian calibration of the Pasture Simulation model (PaSim) to simulate European grasslands under water stress. *Ecol. Inform.* 30, 356–364. <https://doi.org/10.1016/j.ecoinf.2015.09.009>.
- Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrol. Process.* 6 (3), 279–298. <https://doi.org/10.1002/hyp.3360060305>.
- Blair, J., Nippert, J., Briggs, J., 2014. Grassland ecology. In: Monson, R.K. (Ed.), *Ecology and the Environment*. Springer, New York, New York, NY, pp. 389–423. [https://doi.org/10.1007/978-1-4614-7501-9\\_14](https://doi.org/10.1007/978-1-4614-7501-9_14).
- Box, G.E., Tiao, G.C., 2011. *Bayesian Inference in Statistical Analysis*. John Wiley & Sons.
- Burgers, G., Jan Van Leeuwen, P., Evensen, G., 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Weather Rev.* 126 (6), 1719–1724. [https://doi.org/10.1175/1520-0493\(1998\)126<1719:ASITEK>2.0.CO;2](https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2).
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sens. Environ.* 91 (3), 332–344. <https://doi.org/10.1016/j.rse.2004.03.014>.
- Chen, Y., Zhang, Z., Tao, F., 2018. Improving regional winter wheat yield estimation through assimilation of phenology and leaf area index from remote sensing data. *Eur. J. Agron.* 101, 163–173. <https://doi.org/10.1016/j.eja.2018.09.006>.
- Cohen, J.L., Furtado, J.C., Barlow, M.A., Alexeev, V.A., Cherry, J.E., 2012. Arctic warming, increasing snow cover and widespread boreal winter cooling. *Environ. Res. Lett.* 7 (1), 14007. <https://doi.org/10.1088/1748-9326/7/1/014007>.
- De Bernardis, C., Vicente-Guijalba, F., Martinez-Marin, T., Lopez-Sanchez, J.M., 2016. Particle filter approach for real-time estimation of crop phenological states using time series of NDVI images. *Remote Sens.* <https://doi.org/10.3390/rs8070610>.
- Dengler, J., Birge, T., Bruun, H.H., Rašomaičius, V., Rūsiņa, S., Sickel, H., 2020. Grasslands of Northern Europe and the Baltic States. In: Goldstein, M.I., Dellasala, D. A. (Eds.), *Encyclopedia of the World's Biomes*. Elsevier, Oxford, pp. 689–702.
- Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.: Oceans* 99 (C5), 10143–10162. <https://doi.org/10.1029/94JC00572>.
- Evensen, G., 2004. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynam.* 54 (6), 539–560. <https://doi.org/10.1007/s10236-004-0099-2>.
- Graux, A., Bellocchi, G., Lardy, R., Soussana, J., 2013. Ensemble modelling of climate change risks and opportunities for managed grasslands in France. *Agr. For. Meteorol.* 170, 114–131. <https://doi.org/10.1016/j.agrformet.2012.06.010>.
- Guérif, M., Duke, C., 1998. Calibration of the SUCROS emergence and early growth module for sugar beet using optical remote sensing data assimilation. *Eur. J. Agron.* 9 (2), 127–136. [https://doi.org/10.1016/S1161-0301\(98\)00031-8](https://doi.org/10.1016/S1161-0301(98)00031-8).
- Hjelkrem, A.R., Höglind, M., van Oijen, M., Schellberg, J., Gaiser, T., Ewert, F., 2017. Sensitivity analysis and Bayesian calibration for testing robustness of the BASGRA model in different environments. *Ecol. Model.* 359, 80–91. <https://doi.org/10.1016/j.ecolmodel.2017.05.015>.
- Höglind, M., Bakken, A.K., Jørgensen, M., østrem, L., 2010. Tolerance to frost and ice encasement in cultivars of timothy and perennial ryegrass during winter. *Grass Forage Sci.* 65 (4), 431–445. <https://doi.org/10.1111/j.1365-2494.2010.00762.x>.
- Höglind, M., Jørgensen, M., østrem, L., 2006. Growth and development of frost tolerance in eight contrasting cultivars of timothy and perennial ryegrass during winter in Norway. *Proceedings of NJF Seminar 10–12*.
- Höglind, M., Thorsen, S.M., Semenov, M.A., 2013. Assessing uncertainties in impact of climate change on grass production in Northern Europe using ensembles of global climate models. *Agric. For. Meteorol.* 170, 103–113. <https://doi.org/10.1016/j.agrformet.2012.02.010>.
- Höglind, M., Van Oijen, M., Cameron, D., Persson, T., 2016. Process-based simulation of growth and overwintering of grassland using the BASGRA model. *Ecol. Model.* 335, 1–15. <https://doi.org/10.1016/j.ecolmodel.2016.04.024>.
- 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. Model.* 417, 108925 <https://doi.org/10.1016/j.ecolmodel.2019.108925>.
- Huang, J., et al., 2015. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agric. For. Meteorol.* 204, 106–121. <https://doi.org/10.1016/j.agrformet.2015.02.001>.
- Huang, J., et al., 2016. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to improve regional winter wheat yield estimation. *Agric. For. Meteorol.* 216, 188–202. <https://doi.org/10.1016/j.agrformet.2015.10.013>.
- Huang, X., et al., 2018. A dynamic agricultural prediction system for large-scale drought assessment on the Sunway TaihuLight supercomputer. *Comput. Electron. Agric.* 154, 400–410. <https://doi.org/10.1016/j.compag.2018.07.027>.
- Ines, A.V.M., Das, N.N., Hansen, J.W., Njoku, E.G., 2013. Assimilation of remotely sensed soil moisture and vegetation with a crop simulation model for maize yield prediction. *Remote Sens. Environ.* 138, 149–164. <https://doi.org/10.1016/j.rse.2013.07.018>.
- 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. <https://doi.org/10.1016/j.eja.2017.11.002>.
- Jing, Q., Bélanger, G., Baron, V., Bonesmo, H., Virkajarvi, P., Young, D., 2012. Regrowth simulation of the perennial grass timothy. *Ecol. Model.* 232, 64–77. <https://doi.org/10.1016/j.ecolmodel.2012.02.016>.
- Knyazikhin, Y., 1999. MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD 15) Algorithm Theoretical Basis Document. <http://eospso.gsfc.nasa.gov/atbd/modistables.html>.
- Liu, J.S., Chen, R., 1998. Sequential Monte Carlo methods for dynamic systems. *J. Am. Stat. Assoc.* 93 (443), 1032–1044. <https://doi.org/10.1080/01621459.1998.10473765>.
- Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* 11 (2), 431–441.
- Njoku, E.G., Jackson, T.J., Lakshmi, V., Chan, T.K., Nghiem, S.V., 2003. Soil moisture retrieval from AMSR-E. *IEEE T. Geosci. Remote* 41 (2), 215–229. <https://doi.org/10.1109/TGRS.2002.808243>.
- Østrem, L., Rapacz, M., Larsen, A., Dalmannsdottir, S., Jørgensen, M., 2015. Influences of growth cessation and photoacclimation on winter survival of non-native *Lolium–Festuca* grasses in high-latitude regions. *Environ. Exp. Bot.* 111, 21–31. <https://doi.org/10.1016/j.envexpbot.2014.10.008>.
- Parton, W.J., 1996. In: Powlson, D.S., Smith, P., Smith, J.U. (Eds.), *The CENTURY Model*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 283–291. [https://doi.org/10.1007/978-3-642-61094-3\\_23](https://doi.org/10.1007/978-3-642-61094-3_23).
- Patil, S.D., Stieglitz, M., 2015. Comparing spatial and temporal transferability of hydrological model parameters. *J. Hydrol.* 525, 409–417. <https://doi.org/10.1016/j.jhydrol.2015.04.003>.
- Rende, S., 2019. *What Can Be Learnt From the 2018 Drought and How to Adapt Swedish Agriculture to a Changing Climate?*
- Shah, S.H.H., Li, Y., Wang, J., Collins, A.L., 2020. Optimizing farmyard manure and cattle slurry applications for intensively managed grasslands based on UK-DNDC model simulations. *Sci. Total Environ.* 714, 136672 <https://doi.org/10.1016/j.scitotenv.2020.136672>.
- Suttie, J.M., Reynolds, S.G., Batello, C., 2005. *Grasslands of the World Food & Agriculture Org.*
- Tewes, A., Hoffmann, H., Krauss, G., Schäfer, F., Kerkhoff, C., Gaiser, T., 2020. New approaches for the assimilation of LAI measurements into a crop model ensemble to improve wheat biomass estimations. *Agronomy*. <https://doi.org/10.3390/agronomy10030446>.
- van der Linden, S., Woo, M., 2003. Transferability of hydrological model parameters between basins in data-sparse areas, subarctic Canada. *J. Hydrol.* 270 (3), 182–194. [https://doi.org/10.1016/S0022-1694\(02\)00295-0](https://doi.org/10.1016/S0022-1694(02)00295-0).

- Van Oijen, M., Höglind, M., 2016. Toward a Bayesian procedure for using process-based models in plant breeding, with application to ideotype design. *Euphytica* 207 (3), 627–643. <https://doi.org/10.1007/s10681-015-1562-5>.
- Varella, H., Guérif, M., Buis, S., 2010. Global sensitivity analysis measures the quality of parameter estimation: the case of soil parameters and a crop model. *Environ. Model. Softw.* 25 (3), 310–319. <https://doi.org/10.1016/j.envsoft.2009.09.012>.
- Virkajarvi, P., Hyrkas, M., Pakarinen, K., Rinne, M., 2012. Timotein ja ruokonadan erotsadonttuotoprosessissa. In: Hyrkas, M., Virkajarvi, P. (Eds.), Nurmen kasvu- ja kehitysprosessit. MIT Report 56 (In Finnish), pp. 22–46.
- Vuichard, N., et al., 2007. Estimating the greenhouse gas fluxes of European grasslands with a process-based model: 1. Model evaluation from in situ measurements. *Global Biogeochem. Cy.* 21 (1) <https://doi.org/10.1029/2005GB002611>.
- Wallach, D., 2011. Crop model calibration: a statistical perspective. *Agron. J.* 103 (4), 1144–1151. <https://doi.org/10.2134/agronj2010.0432>.
- Wang, J., Li, X., Lu, L., Fang, F., 2013. Estimating near future regional corn yields by integrating multi-source observations into a crop growth model. *Eur. J. Agron.* 49, 126–140. <https://doi.org/10.1016/j.eja.2013.03.005>.
- Wiréhn, L., 2018. Nordic agriculture under climate change: a systematic review of challenges, opportunities and adaptation strategies for crop production. *Land Use Policy* 77, 63–74. <https://doi.org/10.1016/j.landusepol.2018.04.059>.
- Woodward, S.J.R., Van Oijen, M., Griffiths, W.M., Beukes, P.C., Chapman, D.F., 2020. Identifying causes of low persistence of perennial ryegrass (*Lolium perenne*) dairy pasture using the Basic Grassland model (BASGRA). *Grass Forage Sci.* 75 (1), 45–63. <https://doi.org/10.1111/gfs.12464>.
- Yuan, H., Dai, Y., Xiao, Z., Ji, D., Shangguan, W., 2011. Reprocessing the MODIS Leaf Area Index products for land surface and climate modelling. *Remote Sens. Environ.* 115 (5), 1171–1187. <https://doi.org/10.1016/j.rse.2011.01.001>.
- Zhang, Z., Li, Z., Chen, Y., Zhang, L., Tao, F., 2020. Improving regional wheat yields estimations by multi-step-assimilating of a crop model with multi-source data. *Agric. For. Meteorol.* 290, 107993 <https://doi.org/10.1016/j.agrformet.2020.107993>.
- Zhao, Y., Chen, S., Shen, S., 2013. Assimilating remote sensing information with crop model using Ensemble Kalman Filter for improving LAI monitoring and yield estimation. *Ecol. Model.* 270, 30–42. <https://doi.org/10.1016/j.ecolmodel.2013.08.016>.