Enhancing field-level forecasting of crop growth status by incorporating the analytically estimated system uncertainties into a data assimilation procedure

Dong Wang

Affiliation not available

November 27, 2023

Dong Wang\textsuperscript{a,b,*}, Paul C. Struik\textsuperscript{a}, Lei Liang\textsuperscript{b} & Xinyou Yin\textsuperscript{a,*}

\textsuperscript{a} Centre for Crop Systems Analysis, Department of Plant Sciences, Wageningen University & Research, P.O. Box 430, 6700 AK Wageningen, The Netherlands

\textsuperscript{b} Shanghai Lankuaikei Technology Development Co. Ltd., No. 888 Huanhu West 2nd Road, Pudong New District, Shanghai, China

*Corresponding Authors: Dong Wang (dong.wang@wur.nl; dong.wang@lankuaikei.cn); Xinyou Yin (xinyou.yin@wur.nl)

Abstract

Accurate crop status forecasting benefits from assimilating remote sensing observations and crop model simulations. When conducting data assimilation (DA) using an Ensemble Kalman filter (EnKF), arbitrary inflation factors are normally adopted to account for unspecified uncertainties, thus avoiding filter divergence. Here, we developed a Bayesian methodology in which the uncertainties were systematically quantified by combining disparate methods in one framework. Its applicability and performance with crop model GECROS using the EnKF framework were tested against data collected from two years of field experiments, in which aboveground biomass ($W_{\text{above}}$), grain weight ($W_{\text{grain}}$), aboveground nitrogen (N) content ($N_{\text{above}}$), grain N content ($N_{\text{grain}}$) and leaf traits like leaf dry weight, leaf N content and leaf area index were measured for rice. Using only the observations from the first year, the uncertain parameters in GECROS were calibrated by a Markov Chain Monte Carlo approach, while the parameters in the assumed error model that describes the uncertainties of crop model simulations were estimated simultaneously. The calibrated model parameters performed well in the validation year, except for the simulated leaf traits (Normalized Root Mean Squared Error ($NRMSE$) $> 0.38$). Remotely sensed leaf traits predicted by a Gaussian Process Regression (GPR) model were more accurate ($NRMSE < 0.34$), with uncertainties of the remote sensing observations estimated from the GPR model itself. Assimilating simulated and predicted leaf traits with their estimated uncertainties into EnKF prevented filter divergence, and the forecast accuracy of crop model improved in the validation year. Compared with simulation without assimilating in-season remote sensing observations, the $NRMSE$ of updated whole-season $W_{\text{above}}$ and $N_{\text{above}}$ both decreased from 0.37 to 0.20; and those of updated $W_{\text{grain}}$ and $N_{\text{grain}}$ at harvest decreased from 0.40 and 0.28 to 0.22 and 0.19, respectively. The developed method contributes to systematic uncertainty analysis in DA and accurate forecasting of in-season and end-of-season crop carbon and N status for smart farming.

Keywords

Crop growth, crop model, uncertainty quantification, Markov Chain Monte Carlo approach, Gaussian Process
Regression, Ensemble Kalman Filter

1. Introduction

Accurate forecasting and mapping of spatiotemporal variabilities in aboveground biomass ($W_{\text{above}}$) and grain yield during the growing season are essential for informing farmers to perform field precision management under variable climatic conditions (Acevedo et al., 2020; Gao et al., 2017; Lobell and Azzari, 2017). Remote sensing science serves these purposes via connecting field measurements with sensor observations. For instance, crop yield forecasting based on regression models using field-measured yield and remote sensing features dates back to the 1970s (Idso et al., 1977). However, the poor exploration of model and data uncertainties has increasingly become a limitation for most remote sensing observations at high spatial resolution (Martínez-Ferrer et al., 2022). Although information like leaf area index (LAI, m$^2$ m$^{-2}$) and some weather variables can be incorporated into those regression methods to improve predictions (Johnson, 2014), the interactions within the continuum of atmosphere-crop-soil are widely overlooked. In this regard, it may be useful to explore dynamic crop models that have been developed since the 1960s to simulate crop growth and yield (e.g., de Wit (1965); de Wit and Penning de Vries (1985); Jones et al. (2003); Keating et al. (2003); Yin and van Laar (2005)) given that these models are based on in-depth understanding of crop physiological principles. However, as such crop models are generally developed and tested at the scale of a homogeneous plot, uncertainties are inevitable when applying them to heterogeneous farmers’ fields. Uncertainties are also caused by incomplete knowledge of physiological processes, parameter values, meteorological conditions, soil properties and management practices (Hansen and Jones, 2000). For predictions using either remote sensing or crop modelling, all these uncertainties propagate, leading to bias in simulated in-season crop growth and end-of-season crop yield.

Combined utilization of crop model simulations and remotely sensed observations is expected to produce a more accurate estimate than any of the two approaches on its own and attracts ever-increasing interests in smart farming (Houser et al., 2012; Jin et al., 2018). To this end, data assimilation (DA) methods have been developed (Jin et al., 2018). Monte Carlo-based Ensemble Kalman Filter (EnKF) (Evensen, 1994) is among the most popular methods for conducting DA (Carrassi et al., 2018), due to its simplicity, efficiency and adaptability to nonlinear and high-dimensional simulation models (Evensen, 2003; Kalnay et al., 2007). EnKF is an iterative procedure that keeps alternating between model forecasting and state updating. Each forecasting step produces an ensemble of different predictions that accounts for uncertainty about model inputs, parameter values and model structure. Each updating step uses observations, weighted by measurement uncertainty, to correct the ensemble forecast. Sampling error can be minimized by using a large ensemble size (Whitaker and Hamill, 2012). However, inappropriately estimated system errors may lead to filter divergence, in which subsequent ensemble forecasts drift further from the truth (Anderson and Anderson, 1999; Jazwinski, 1970) and distributions of forecasted states become too narrow. Forecasting uncertainty is thus underestimated relative to observational uncertainty, making the observations essentially irrelevant. To alleviate filter divergence, additive or multiplicative inflation factors are commonly used (Huang et al., 2019). For instance, while assimilating remotely sensed soil moisture and LAI within the crop model DSSAT, a variant of EnKF, Ensemble Square Root Filter, was applied, in which fixed small inflation factors (1.05 for soil moisture and 1.50 for LAI) were included to prevent filter divergence (Ines et al., 2013). Instead of using fixed values, Kivi et al. (2022) adaptively estimated dynamic inflation factors to assimilate in-situ observed daily soil moisture for updating soil water and nitrogen (N) dynamics in the crop model APSIM. However, as inflation factors are not physically constrained, their application to complicated dynamic models with many different outputs is not straightforward (Ying and Zhang, 2015). Quantifying parameterization errors in crop models and uncertainties of remotely sensed observations is indispensable when applying EnKF to achieve more accurate forecasts of crop growth status (Jin et al., 2018).

Parameter accuracy of a crop model significantly affects the performance of DA and yield forecasting (Kang and Özdoğan, 2019). To improve model parameter accuracy, various parameter inference methods have been developed and Bayesian approaches are becoming increasingly popular (e.g., Beven and Freer (2001); Vrugt et al. (2009b)). The interest in applying Bayesian approaches lies not only in inferring the most likely
parameter values, but also in estimating their underlying posterior probability distribution functions (pdf) and even in estimating model structural error (Huang et al., 2019). Markov Chain Monte Carlo (MCMC) methods are typically used in these Bayesian approaches to link crop model simulations with observations. Based on data probability quantified by a likelihood function and, commonly, the Metropolis-Hasting search strategy (Hastings, 1970; Metropolis et al., 1953), the prior probability distribution for the parameters of the crop model and residual error model is updated to a posterior distribution conditioned by the information in the data. Normally, residual errors are assumed independent and identically distributed (i.i.d.), following a normal distribution with zero mean and constant variance (Box and Tiao, 1973). However, in-field observations always have variable residuals throughout the growing season (Dumont et al., 2014). Thus, a likelihood function revised by observational variance was proposed for accounting for the heteroscedasticity in the crop model STICS (Dumont et al., 2014). A more generalized formal likelihood function based on a general error model was developed by Schoups and Vrugt (2010) for a hydrological model, which allows for the heteroscedasticity and non-Gaussian model residual errors. Their approach allows for diagnostic checking of residual error assumptions and does not require the i.i.d. assumption. As EnKF has been shown to be effective in cases with nonlinear or non-Gaussian errors (Han and Li, 2008; Katzfuss et al., 2016), investigating the applicability of integrating the calibrated uncertain parameters in a crop model with generalized error into the EnKF framework is in demand.

Errors in remote sensing data hamper the use of these data for predictions from nonparametric regression modelling, one of the most frequently used approaches to predict crop status from remote sensing data (Huang et al., 2019; Verrelst et al., 2019). Among the nonparametric models, the Gaussian Process Regression (GPR) model, developed within a Bayesian framework (Rasmussen and Williams, 2006), has been considered as a promising method, not only because of better prediction performance (Verrelst et al., 2012), but also because it quantifies predictive uncertainty (Berger et al., 2020a; Wang et al., 2019; Verrelst et al., 2019). Temporal and spatial transferability of GPR has been demonstrated by successfully transporting the GPR model to other images (Verrelst et al., 2013b). However, there is a need for comparison of DA from remote sensing (DA_{rs}) with DA from field measurements (DA_{fm}) (Huang et al., 2019). Due to destructive sampling, the sampling sites would not remain the same in ground observations but that inconsistency is normally neglected while conducting DA_{fm}. In contrast, in remote sensing predictions, although prediction errors always exist, temporal changes in crop growth can be predicted reasonably well. Thus, with those predominant merits of GPR, its performance upon further incorporating into the EnKF framework should be evaluated and compared with that of DA_{fm}.

Studies have been conducted to connect process-based simulations, field observations, and their uncertainties in order to have reliable forecasts using DA. In a hydrological modelling study, Vrugt et al. (2005) simultaneously estimated parameter uncertainties and structural errors as well as observational errors. In this approach, an inner EnKF loop for recursive state simulation and an outer global optimization loop for posterior estimation were incorporated in simultaneous parameter estimation and data assimilation. However, even though model predictive ability is supposed to be enhanced by improved parameterization, assessing model structural and input errors may be hindered when parameter values are not fixed (Schoups and Vrugt, 2010). In an observing system simulation experiment that assimilated LAI and soil moisture data into the crop model SWAP, Hu et al. (2017) found that simultaneously updating parameters tended to worsen the performance of grain yield prediction when the uncertain parameters that directly determine biomass and grain formation were incorporated. A method for systematically quantifying uncertainties in the crop model simulation and remotely sensed observations from a separate Bayesian process and applying them into an EnKF framework is strongly needed for better forecasting of crop growth status. Such a method can help for the careful approximation and application of uncertainties in other DA algorithms or frameworks (Huang et al., 2019), and can be a potential reference of the desired model-data fusion framework for better Earth system forecasting (Gettelman et al., 2022).

The objective of this study was to develop a Bayesian methodology that combines disparate quantitative methods into one framework, i.e. incorporating the systematically analyzed errors in crop model simulations and remote sensing observations into the data assimilation procedure of EnKF. We expect that this framework...

3
enhances the forecasting of the crop growth status. The methodology was validated in an actual case of a field experiment of rice. The crop model GECROS was selected for generating crop growth simulations, due to its generality and physiological robustness (Yin and Struik, 2017; Yin and van Laar, 2005) (see a brief description of the crop model GECROS in Supplement A). Our specific objectives were: 1) to calibrate and validate GECROS under field conditions in China given the heteroscedastic and non-Gaussian residual error assumption; 2) to evaluate the performance of the GPR model for remote sensing prediction and its uncertainty estimation; 3) to assess the applicability of estimated uncertainties of the crop model simulations and the remote sensing observations in EnKF.

2. Materials & Methods

The framework of the developed method and its summarized description are shown in Fig. 1.

2.1 Study site and field data

2.1.1 Experimental design

Field experiments for rice (Oryza sativa L., cv. Nanjing 46) were conducted at Chongming, Shanghai, China for two years (2019-2020). Basic soil information was collected before the experimentation (Table 1). The meteorological records of the closest weather station from our experimental site (ca. 5.7 km away) were downloaded from the website of the China meteorological data service centre (http://data.cma.cn). The weather variables include daily solar radiation, minimum and maximum temperatures, precipitation, vapour pressure, and wind speed. The mean air temperature and precipitation at the experimental site for the studied years are shown in Fig. S1, illustrating the typical weather conditions for rice cultivation in Yangtze River Delta.

Details about the field experiment were given previously (Wang et al., 2023). There were six treatments with varied N application amounts from 0 to 320 kg N ha⁻¹ per season in a complete randomized block design and three and four blocks in 2019 and 2020, respectively, in which each plot occupied an area of 180 m². Rice was sown with a row spacing of 20 cm and a within-row seed spacing of 2-3 cm on 14-Jun and 4-Jun in 2019 and 2020, respectively. The application time and amount of N were split according to the schedule in Table 2. Based on the local practice, sufficient phosphate (112.5 kg P₂O₅ ha⁻¹) and potash (112.5 kg K₂O ha⁻¹) fertilizers were applied per season to prevent phosphorus and potassium deficiencies. Other field managements were the same in each plot following local standard practices, including irrigation, and pest, weed and disease control.

2.1.2 Observations from field destructive sampling

Field destructive sampling was conducted in both years and its regimes in 2019 and 2020 are shown in Table 3. Seven types of observations at the canopy level were conducted: W_above, grain weight (W_grain), leaf weight (W_leaf), aboveground N content (N_above), grain N content (N_grain), leaf N content (N_leaf), and LAI. All observations in 2019 from different treatments were used for model calibration and those in 2020 were used for model validation.

Aboveground plants in an area of 1 and 2.25 m² in each plot were cut during each field measurement in 2019 and 2020, respectively. At harvest, two sampling sites were selected in each plot and their sampling areas were 2 and 4 m² in 2019 and 2020, respectively. The total fresh samples were bagged and weighed, while ca. 20% was dissected into component plant parts, including green leaves, yellow leaves, stems and leaf sheaths and panicle chaffs, and grains. The area of fresh green leaves and other green parts like stems was measured immediately using an LI-3100C Area Meter (Li-Cor, Lincoln, NE, USA). The dissected plant organs were weighed after oven-drying at 70 °C until constant weight and then ground to powder to assess N concentration using the Micro-Kjeldahl method. The N content in each plant organ was calculated as the product of their corresponding weight and N concentration. As all green parts are considered as functional leaves in GECROS (Yin and van Laar, 2005), the measured LAI, W_leaf and N_leaf were adjusted correspondingly following equations Eqns S1-S3 to make the crop model simulations and field measurements comparable.
2.2 Remote sensing observation and its uncertainty quantification

2.2.1 Collecting hyperspectral images

The field data collected in 2020 for validating GECROS were also utilized to validate the methodology proposed in this study. Thus, to obtain the remote sensing predictions, the hyperspectral imaging data were collected in 2020 at the same date before each destructive sampling (Table 3). Details about canopy hyperspectral reflectance measurements were previously given (Wang et al., 2023). In brief, a DJI M600 PRO hexacopter (DJI, Shenzhen, China), equipped with a Cubert S185 hyperspectral snapshot camera (Cubert GmbH, Ulm, Baden-Wurttemberg, Germany), was flown over the experimental field between 10 a.m. and 2 p.m. We captured 125 spectral bands in the range of 450-950 nm with a sampling interval of 4 nm. Compared with using reflectance only, the feature set combining reflectance, vegetation indices and texture information worked better when predicting leaf traits (Wang et al., 2023). Thus, the complete dataset used for remote sensing prediction in this study was consistent with the combined features from hyperspectral image and targeted three leaf traits, including \( W_{\text{leaf}}, N_{\text{leaf}} \) and LAI.

2.2.2 Gaussian Process Regression (GPR) model calibration and validation

GPR establishes the relationship between the input features \( x \in \mathbb{R}^B \) in the number of \( B \) and the output variables (leaf traits) \( y \in \mathbb{R}^\nu \) via the kernel function \( k \), which defines the relationship between the pair of data points. The output variable values \( Y \) and \( Y_s \) of all training \( x \) and testing \( x_s \) data points are considered to be from a joint multivariate normal distribution (Rasmussen and Williams, 2006):

\[
\begin{pmatrix}
Y \\
Y_s
\end{pmatrix} \sim N(0,
\begin{bmatrix}
k(x,x)+\sigma^2_d I & \text{amp}k(x,x_s) \\
\text{amp}k(x_s,x) & k(x_s,x_s)
\end{bmatrix}
\]

where \( k(x, x_s) \) denotes the matrix of the covariances evaluated at all pairs of training and testing data points; the same applies to other entries of \( \text{amp}k(x, x_s) \). The observed output variables are assumed with i.i.d. Gaussian noise \( (N(0, \sigma^2_{0*})) \) and \( I \) represents the identity matrix.

The posterior distribution of \( Y_s \) is estimated following (Rasmussen and Williams, 2006):

\[
Y_s | Y, x; x_s \sim N(y_s; \mu, y_s; \text{var})(2)
\]

where the predicted posterior mean \( y_{s; \mu} \) and variance \( y_s; \text{var} \) are calculated as \( k(x_s, x) [k(x,x) + \sigma^2_0 I]^{-1} Y \) and \( k(x_s, x_s) - k(x_s, x) [k(x,x) + \sigma^2_0 I]^{-1} k(x, x_s) \), respectively.

The commonly used anisotropic squared exponential kernel function is also adopted here (Verrelst et al., 2013a):

\[
k(x_i, x_j) = \nu \exp \left( -\sum_{b=1}^{B} \frac{(x_i^{(b)} - x_j^{(b)})^2}{2\sigma^2_b} \right) + \sigma^2_0 \delta_{ij}(3)
\]

where \( \nu \) is a scaling factor, \( \sigma_b \) is the length-scale per input feature \( b \), controlling the spread of the relations for each input feature, and \( \delta_{ij} \) is the Kronecker’s symbol. The hyperparameters, which are denoted as \( \theta_k = \{ \nu, \sigma_b, \sigma_0 \} \), are determined by maximizing the log likelihood in the training set (Rasmussen and Williams, 2006):

\[
\hat{\theta}_k = \arg \max \left( \frac{1}{2} \ln (2\pi) - \frac{1}{2} \ln \left| k(x,x) + \sigma^2_0 I \right| - \frac{1}{2} Y^T (k(x,x) + \sigma^2_0 I)^{-1} Y \right) (4)
\]

where \( n \) represents the size of training dataset.

For calibrating and validating the GPR model, the acquired complete dataset was split into a training (75%) and a testing (25%) dataset. To avoid local maxima, the values of the hyperparameters in the GPR model
were averaged from 100 iterations, and in each run, two-thirds of the training data were randomly selected from the whole training dataset (Verrelst et al., 2013a; Wang et al., 2019).

2.3 Crop model calibration and uncertainty estimation

2.3.1 Preset and uncertain parameters in GECROS

Before conducting the calibration procedure, key photosynthetic parameters in GECROS were preset with the values derived from our previous study (Wang et al., 2022). Meanwhile, phenological parameters, MTDV and MTDR, and parameters like HTMX and STEMNC were determined by the field measurements (Table 4).

Uncertain parameters in GECROS influencing crop carbon and N assimilation, organ formation, and soil water and N movement were set free and allowed to vary within the referenced ranges (Table 4). Moreover, there was a newly added parameter, rASSA, accounting for the uncertainty of estimated photosynthetic parameters. This was considered necessary because photosynthetic parameter values estimated in our previous study based on gas exchange and chlorophyll fluorescence measurements (Wang et al., 2022), had uncertainties. Firstly, the cultivars differed between our previous study and this study. Secondly, the data in our previous study was only limited to certain growing stages rather than in the whole growing season like this study. Thirdly, plant growth conditions differed, in greenhouse vs in field conditions, and photosynthetic parameters are known to acclimate to growth conditions (e.g., Cai et al. (2020); Yin et al. (2019)).

The remaining parameters in GECROS were set as default (Yin and van Laar, 2005).

2.3.2 The residual error assumption of the crop model

The relationship between the observations $Y$ and crop model simulations $\hat{Y}$ can be presented as,

$$Y_{t,i} = \hat{Y}_t(\theta) + \varepsilon_t + \sigma_{t,i}(5)$$

where $Y_{t,i}$ represents the $i$th replicated observation at the field destructive sampling date $t$, and $\hat{Y}_t(\theta)$ is the simulation given the crop model parameter set $\theta$ at $t$. $\varepsilon_t$ represents the residual error $\varepsilon$ of the simulation process at $t$ and $\sigma_{t,i}$ is the observation error for $Y_{t,i}$. As $\sigma_{t,i}$ cannot be separated, especially in circumstances where there are replications for observations (Kennedy and O’Hagan, 2001), for simplicity, the averages of field destructive sampling data $Y_t$ were applied instead and the estimation of $\sigma_{t,i}$ was not included. As local information, including weather, soil and field managements, has been recorded in great detail and treated as a priori knowledge, the errors caused by those forcing inputs are reduced (Hansen and Jones, 2000). Thus, the residual error $\varepsilon$ in this study represents mainly the structural error of the crop model. Accordingly, Eqn 5 is rewritten as,

$$Y_t = \hat{Y}_t(\theta) + \varepsilon_t(6)$$

To stabilize $\varepsilon_t$ and reduce heteroscedasticity, the Box-Cox transformation of the measurements and simulations are introduced at first (Box and Cox, 1964),

$$\tau(Y, \lambda) = \begin{cases} 
\frac{(Y^\lambda - 1)}{\lambda} & \text{if } \lambda \neq 0 \\
\ln(Y) & \text{if } \lambda = 0
\end{cases} (7)$$

where $\lambda$ is the transformation parameter and varies in the range from 0 to 1.

To further deal with the heteroscedasticity and nonnormality, residual errors are proposed as according to Schoups and Vrugt (2010):

$$\Phi_p \varepsilon_t = \eta_t a_t \sim \text{SEP}(0, 1, \xi, \beta) (8)$$

where $\Phi_p$ represents the $p$th order autoregressive model, $\eta_t$ and $a_t$ account for the heteroscedasticity and non-Gaussian residual error at $t$, respectively. The heteroscedasticity is assumed to increase linearly with $\mathbf{Y}$.
where $\sigma_1$ ranges from 0 to 1. $\alpha_i$ denotes a random error with zero mean and unit standard deviation (STD), described by a standardized skew exponential power (SEP) density with parameter $\xi$ and $\beta$ to account for the nonnormality. The skewness parameter $\xi$ ($\xi > 0$) affects the asymmetry of pdf and varies from 0.1 to 10, in which the density is symmetric when $\xi = 1$ and positively or negatively skewed when $\xi > 1$ or $\xi < 1$, respectively. The kurtosis parameter $\beta$ determines the peakedness of the pdf and varies between $-1$ and $+1$. In the case of a symmetric density, the pdf is a uniform distribution when $\beta = -1$, a Gaussian distribution when $\beta = 0$, and a Laplace or double-exponential distribution when $\beta = 1$. The marginal pdf for autoregressive models with Laplace innovations is typically quite complicated and the commonly used approximating approach is only valid for moderate to large sample sizes (Schoups and Vrugt, 2010), not suitable for the situations with limited field observations. Thus, in this study, the autoregressive models were excluded and $\beta$ was assumed to be zero. Consequently, the simplified pdf of the SEP(0, 1, $\xi$, 0) in Eqn 8 is expressed as:

$$p(a_i|\xi) = \frac{1}{\sqrt{2\pi}} \frac{2\sigma_1}{\xi + \xi^2} \exp \left\{ -\frac{1}{2} |a_{\xi,i}|^2 \right\} \tag{9}$$

where $a_{\xi,i} = \xi^{-\text{sign}(\mu_\xi + \sigma_\alpha)}(\mu_\xi + \sigma_\alpha \alpha_i)$, $\mu_\xi = M_1(\xi + 1)$ and $\sigma_\xi = \sqrt{(M_2 - M_1^2)(\xi^2 + \xi - 2) + 2M_1^2 - M_2}$. For the standardized exponential power pdf, $M_1$ and $M_2$ can be obtained as $\frac{2}{\sqrt{2\pi}}$ and 1, respectively (Fernández and Steel, 1998; Schoups and Vrugt, 2010).

The resulting expression for the log-likelihood function of Eqns 6-9 is:

$$\downarrow \left( \theta, \lambda, \sigma_1, \xi | Y \right) = -\frac{T}{2} \ln (2\pi) - \sum_{t=1}^{T} \ln \left( \frac{\tau(Y_t, \lambda) - \tau(\hat{Y}_t(\theta), \lambda)}{\sigma_1 Y_t(\theta)} \right) + T \ln \frac{2\sigma_1}{\xi + \xi^2} - \frac{1}{2} \sum_{t=1}^{T} |a_{\xi,t}|^2 \tag{10}$$

where $T$ is the times of field sampling across the whole crop growing season.

2.3.3 Methods of combining likelihood values

To calibrate crop model parameters with multivariate observations, the Bayes’ multiplication method was applied here (He et al., 2010). The combined log-likelihood $\downarrow_{\text{combined}}$ can be written as:

$$\downarrow_{\text{combined}}(\theta, \lambda, \sigma_1, \xi) = \sum_{m=1}^{M} \downarrow(\theta, \lambda_m, \sigma_{1m}, \xi_m | Y_m) \tag{11}$$

where $\lambda$, $\sigma_1$, and $\xi$ denote the parameter sets of the residual error model for the multivariate observations. $M$ represents the number of observation types. For the particular observation type $m$, $\lambda_m$, $\sigma_{1m}$, $\xi_m$ denote its residual error parameters, and $Y_m$ represents its averaged field destructive sampling data.

2.3.4 Calibration process and uncertainty estimation

After the specification of a prior parameter pdf, the uncertain parameters in crop model and residual error model can be calibrated by conditioning on the data through the combined log likelihood function of Eqn 11. In this study, the prior pdf of uncertain parameters was treated as a normal distribution. The prior means of crop model parameters were defined as midpoints of their identified range, while their prior STDs were derived from the fixed coefficient of variation (CV) at 0.25. Regarding residual error parameters for multivariate observations, the prior means of $\lambda$, $\sigma_1$ and $\xi$ were set as one, zero and one, respectively, and their prior STDs were 0.25. The DREAM-zs (DiffeRential Evolution Adaptive Metropolis) algorithm was adopted to generate a representative sample from the posterior distribution. DREAM-zs is designed to accelerate convergence for high-dimensional problems, by sampling from an archive of past parameter candidates. Furthermore, DREAM-zs increases the diversity of candidate points by generating jumps beyond parallel direction updates (Schoups and Vrugt, 2010). Although the original DREAM-zs does not require outlier detection, the step of detecting an outlier chain (Vrugt et al., 2009a) was included in this study, as outlier chains can significantly deteriorate the performance of the MCMC sampler, especially when there is more than one type of observation.

The number of chains was set as two times the number of calibrated parameters following the suggestion of Vrugt et al. (2009b) and in total, 100 000 evaluations were conducted in the process of MCMC sampling.
The outlier chain was detected until 60% evaluations were finished. The last 10% evaluations of each chain were compiled to calculate the mean and STD of the posterior pdf of each parameter.

As for the uncertainty estimation in the crop model, 2000 parameter sets were sampled from the posterior pdf to generate the corresponding simulations \( \hat{Y} \) and residual errors \( \varepsilon \). By calculating the 2.5% and 97.5% percentiles of simulations with or without \( \varepsilon \) for each type of observation, the uncertainties of the crop model caused by uncertain crop model parameters with or without model structural error were obtained. The estimation of \( \varepsilon \) involves generating independent samples from a SEP distribution. The followed sampling algorithm is described completely in Schoups and Vrugt (2010). With sampled residual error parameters, the corresponding residual errors \( \varepsilon \) were calculated using Eqn 8.

### 2.4 Data assimilation procedure

#### 2.4.1 Ensemble Kalman filter

In EnKF, the relation of the observation \( Y_t \) to the model simulated state \( \hat{Y}_t \) can be described as:

\[
Y_t = H \hat{Y}_t + \varepsilon_t \tag{12}
\]

where \( H \) is the measurement operator that maps the model state to the observation. The ensembles of observations and simulations of model state at \( t \) are stored in \( Y_t \) and \( \hat{Y}_t \), respectively. Both have a dimension of \( N_y \times N_{\text{ens}} \), in which \( N_y \) and \( N_{\text{ens}} \) represent the dimension of observed states and the ensemble size, respectively.

The observations \( Y_{t,j} \) in ensemble \( Y_t \) were drawn from a \( N_y \)-variate Gaussian distribution with mean equal to the observation, \( Y_f^t \), and covariance equal to \( R_t \)

\[
Y_{t,j} = Y_f^t + \sigma_y^t \tag{13}
\]

in which \( \sigma_y^t \sim N(0, R_t) \).

The model error covariance \( P_t^f \) of \( \hat{Y}_t \) is calculated, using:

\[
P_t^f = (N_{\text{ens}} - 1)^{-1} \sum_{j=1}^{N_{\text{ens}}} (\hat{Y}_{t,j} - Y_f^t)(\hat{Y}_{t,j} - Y_f^t)^T \tag{14}
\]

in which \( \hat{Y}_{t,j} \) and \( Y_f^t \) represent the single simulation trajectory and the mean of the ensemble \( \hat{Y}_t \), respectively and the superscript \( T \) denotes the transpose of the matrix. \( Y_f^t \) is calculated as \( \frac{\sum_{j=1}^{N_{\text{ens}}} \hat{Y}_{t,j}}{N_{\text{ens}}} \).

Under the linear assumptions, the updated analysis state \( Y_t^a \) and its error covariance \( P_t^a \) are calculated following:

\[
Y_t^a = \hat{Y}_t + K \left( Y_t - H \hat{Y}_t \right) \tag{15}
\]

\[
P_t^a = (I - KH) P_t^f \tag{16}
\]

in which \( I \) is the identity matrix, and \( K \) is the Kalman gain, defined as:

\[
K = P_t^f H^T \left( H P_t^f H^T + R_{t,\text{cov}} \right)^{-1} \tag{17}
\]

where \( P_t^f H^T = (N_{\text{ens}} - 1)^{-1} \sum_{j=1}^{N_{\text{ens}}} (\hat{Y}_{t,j} - Y_f^t)(HY_{t,j} - HY_f^t)^T \).

In summary, in the light of an ensemble of model trajectories, the EnKF approximates the probability density of the model states at each time step \( t \). The updated mean of this ensemble represents the “best” state estimate, whereas the spread of the updated ensemble members provides a measure of the output uncertainty (Evensen, 1994).

#### 2.4.2 Integration of the estimated uncertainties into Ensemble Kalman filter
In this study, the simulated states were analyzed and updated by the observations of $W$, $N_{\text{leaf}}$, $\xi_{\text{leaf}}$ and LAI, which were from the field measurements or from the remote sensing observations. Thus, $N_{y}$ was set at three. Meanwhile, $N_{\text{ens}}$ was set at 100 in this study, due to the reasonable trade-off between efficiency and pdf representation (de Wit and van Diepen, 2007). The residual error of the crop model simulations was generated following the abovementioned procedure (in Section 2.3.4). For the field observations, $Y_{y}$ was the measured data at each sampling date, and $R_{r}$ was derived from their replications, assuming that the CV for each single measurement was the same as the CV of the replications of each treatment at each sampling date. As for remotely sensed observations, both $Y_{y}$ and $R_{r}$ were predicted from the regressed GPR model. Moreover, the performances of crop model simulations without assimilating observations (open-loop) were also evaluated for comparison. To test the end-of-season forecast ability of the updated crop model, the assimilation of leaf traits ended at the grain-filling stage (20 days before maturity).

### 2.5 Evaluation metrics

With respect to the corresponding observations $y_{i}$, the performance of $\hat{y}_{i}$, simulations in model calibration and validation or updated simulations in DA, was evaluated by the coefficient of determination ($R^{2}$), the root mean square error ($RMSE$) and the normalized $RMSE (NRMSE)$:

$$R^{2} = 1 - \frac{\sum_{i=1}^{n} (y_{i} - \hat{y}_{i})^{2}}{\sum_{i=1}^{n} (y_{i} - \bar{y})^{2}}$$  \hspace{1cm} (18)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_{i} - \hat{y}_{i})^{2}}{(n-1)}}$$  \hspace{1cm} (19)

$$NRMSE = \frac{RMSE}{\bar{y}}$$  \hspace{1cm} (20)

where $n$ represents the number of evaluated data points, and $\bar{y}$ is the mean value of observations across the whole growing season.

Moreover, to assess the filter behavior of EnKF, we introduced $f_{rc}$, which represents the relative change of the states after and before updating compared with the measured $y_{t}$,

$$f_{rc} = \frac{Y_{t}^{n} - y_{t}}{Y_{t}^{n} - y_{t}}$$  \hspace{1cm} (21)

in which $Y_{t}^{n}$ is calculated as $\frac{(\sum_{i=1}^{N_{\text{ens}}} Y_{t}^{i})}{N_{\text{ens}}}$. The closer $f_{rc}$ is to one, the more likely does filter divergence occur.

### 3. Results

#### 3.1 Bayesian calibration of GECROS

The varying range of residual error parameter $\sigma_{1}$ affected the simulation performance in calibration using the data from the 2019 field experiment (Fig. 2). When the upper bound of $\sigma_{1}$ was set at 1.0, the simulated $W$, $N_{\text{leaf}}$, and LAI had the highest $NRMSE$. After decreasing the upper bound of $\sigma_{1}$, their $NRMSE$ tended to decrease, although the $NRMSE$ of simulated $N_{\text{grain}}$ tended to increase (Fig. 2a). Meanwhile, $R^{2}$ of simulations in calibration also varied with the upper bound of $\sigma_{1}$. Especially, the $R^{2}$ of simulations while setting the upper bound of $\sigma_{1}$ at 0.2 tended to be higher than those while setting upper bounds of $\sigma_{1}$ at other levels. Due to the overestimation (results not shown), even though simulated $N_{\text{grain}}$ achieved the lowest $NRMSE$ while setting the upper bound of $\sigma_{1}$ at 1.0, its $R^{2}$ values were rather low (Fig. 2b). The estimated means of posterior distribution of $\sigma_{1}$ for most of the observations were close to the fixed upper bound (Table 5, Fig. S2a-c). Although fixing the value of upper bound of $\sigma_{1}$ at 1.0 reduced this phenomenon, due to the unexpected residual error from the upper bound expansion of $\sigma_{1}$, the estimated $\sigma_{1}$ and $\xi$ did not converge well during the Bayesian calibration process (see Fig. S2 for the simulated $W_{\text{above}}$ as an example). Consequently, based on the sensitivity analysis of the upper bound of $\sigma_{1}$, the upper bound of $\sigma_{1}$ was set at 0.2 to conduct the following analysis.

While fixing the upper bound of $\sigma_{1}$ at 0.2, the simultaneously calibrated uncertain parameters in GECROS and residual error parameters are shown in Table 5. Compared with $W_{\text{above}}$ and $N_{\text{above}}$, the estimated
error parameters $\xi$ for $W_{\text{leaf}}, N_{\text{leaf}}, \text{LAI}, W_{\text{grain}}$ and $N_{\text{grain}}$ were farther away from one, indicating that the residual errors in those types of observations were supposed to be non-Gaussian (Table 5). Besides the residual error parameters, the uncertain parameters in GECROS were calibrated reasonably well (Table 5). Meanwhile, the uncertainty problem caused by the preset photosynthetic parameters was solved by introducing the parameter rASSA (Table 4), and the overestimation in the simulations was reduced accordingly (Fig. S3).

With the estimated posterior distributions of the uncertain parameters in the crop model GECROS (Fig. S4), the parameter uncertainty was analyzed for the field measurements in 2019, while the total uncertainty was calculated by further integrating the estimated residual error (Fig. 3). Taking no N input (0 kg N ha$^{-1}$) and locally common N input (240 kg N ha$^{-1}$) as examples, the simulations of different types of observations agreed well with the measured data and most of the points of the average measurements were within the range of simulated total uncertainty (Fig. 3).

### 3.2 Performance validating of the calibrated GECROS model

With the calibrated posterior distributions of uncertain parameters in GECROS, the performance of the simulations was validated against the data from the 2020 field experiment. The $\text{NRMSE}$ and $R^2$ are listed in Table 6.

Compared with that in the calibration, the performance of the crop model GECROS in the validation tended to become worse, especially for $W_{\text{leaf}}, N_{\text{leaf}}$ and LAI (Table 6). For simulated $W_{\text{grain}}$ and $N_{\text{grain}}$, however, the performance improved in the validation. This was probably caused by the extra field destructive sampling in 2020 at the stage of early grain filling (Fig. S5). The poor performance in simulating $W_{\text{leaf}}, N_{\text{leaf}}$ and LAI, probably stemmed from the underestimation when leaf growth approached the peak (Fig. S6).

### 3.3 Observations from remote sensing prediction

The leaf traits, like $W_{\text{leaf}}, N_{\text{leaf}}$ and LAI, were predicted using the GPR modelling of hyperspectral image data, as leaf traits are more likely observed and thus able to be robustly predicted from collected remote sensing features, compared with other observations such as $W_{\text{above}}, W_{\text{grain}}, N_{\text{above}}$ and $N_{\text{grain}}$. The performance of their site-specific predictions was evaluated in the training (Fig. 4a-c) and testing (Fig. 4d-f) datasets. We tried to reduce the overfitting by averaging the values of estimated hyperparameters from the repeated training subsets (Section 2.2.2). The predicted leaf traits in the training set performed better than those in the testing set. The $R^2$ values of the predicted leaf traits in the training set were higher than 0.95, while those in the testing set were lower than 0.88. The $\text{NRMSE}$ of predicted leaf traits in training set ranged from 0.110 to 0.172, while the $\text{NRMSE}$ in the testing test increased and varied from 0.175 to 0.336 (Fig. 4). Compared with $N_{\text{leaf}}$ (Fig. 4b, e), $W_{\text{leaf}}$ (Fig. 4a, d) and LAI (Fig. 4c, f) fitted with their measurements better in both the training set and the testing set.

Maps of the predicted leaf traits at the experimental site were generated, in which the temporal and spatial differences of crop growth at different growing stages were predicted (Fig. 5). In the treatments with low N input, the mean values of predicted leaf traits were low at the stem-elongating stage (Fig. 5a-c), and it later caused early senescence, which was reflected by the low values of the predicted leaf traits at the grain-filling stage (Fig. 5d-f). The mean and STD of the predicted leaf traits tended to be higher in experimental plots with higher N input at the stem-elongating stage (Fig. 5a-c). However, the predicted STD for leaf traits with lower predicted mean values at grain-filling stage tended to be even higher than those with higher predicted mean values at the stem-elongating stage (Fig. 5). The diverse uncertainties across the different growing seasons and the spatial variance between and within the experimental plots demonstrated the necessity of assimilating in situ observations into crop growth simulations for accurate forecasting of crop growth status.

### 3.4 Performance of assimilating observations into crop growth simulations

While assimilating observations into crop simulations, the filter divergence did not show up and the updated simulations agreed better with the measurements. In line with the illustrated trajectories of time series states from open-loop without assimilating observations and updated states by $\text{DA}_{\text{fm}}$ and $\text{DA}_{\text{rs}}$ (Fig. S7), the
differences in DAfm or DArs between updated states and measured states tended to diminish with progress of the growing season (Fig. S8). Moreover, the performance of updated states hardly changed with the varied ensemble size from 50 to 500 in EnKF (Fig. S9). The $f_{rc}$ for directly updated leaf traits by DAfm and DArs centered around 0.1 and 0.2 (Fig. 6a-b), respectively, indicating improved filter performance after assimilating observations and no filter divergence occurrence. Although the $f_{rc}$ for traits indirectly updated by DAfm and DArs centered around 0.8, there was no pattern that the $f_{rc}$ tended to be close to one with the progress of the growing season (Fig. 6c-d), which was another indication of the absence of filter divergence. Due to the method for simulating grain formation in the crop model GECROS, unlike with $W_{above}$ and $N_{above}$, $W_{grain}$ and $N_{grain}$ could not be updated immediately once the leaf traits were updated, but only could be updated gradually in the following growth days (Fig. S7). The $f_{rc}$ values of $W_{grain}$ and $N_{grain}$ were not included here.

The NRMSE of the updated states by DAfm and by DArs decreased, compared with those of open-loop (Fig. 7). Analyzed states like $W_{leaf}$, $N_{leaf}$, and LAI by DAfm and DArs were directly updated and unsurprisingly performed better than those by open-loop across the whole growing season, in which their NRMSE on average decreased from 0.468, 0.551 and 0.434 to 0.161, 0.228 and 0.136, respectively (Fig. 7a). More importantly, the NRMSE of the directly updated states, $W_{above}$ and $N_{above}$, by DAfm across the whole growing season decreased to 0.222 and 0.227, respectively, while those by DArs further decreased to 0.203 and 0.196 (Fig. 7b). Especially at the harvesting stage, those indirectly updated states of DAfm and DArs also agreed better with the field measurements than those of open-loop (Fig. 7c). Taking advantage of the in-situ prediction of crop growth status by the GPR modelling of remote sensing images, updated states of DArs tended to perform better than those of DAfm (Fig. 7b-c). Particularly, compared with arbitrarily assumed uncertainties of crop model simulations and remote sensing predictions, analyzed states of DArs based on the estimated uncertainties from the proposed Bayesian methodology showed better performance (Fig. 8a-b). Fixing the CVs of crop model simulations and remote sensing predictions at 0.1 and 0.01, respectively, resulted in a better performance than using other combinations of CVs. Thus, these CV values were selected for further evaluating the performance of varied inflation factor (Fig. 8), in which the inflation factor was expected to enlarge K in the DA process (Eqn 15) for better filter performance. The updated states of $N_{leaf}$ and $N_{above}$ benefited from the further incorporated inflation factor, on which ranged from 1.05 to 1.25. However, the NRMSE of updated states based on the assumed uncertainties combined with the inflation factor was in general higher than that based on the assumed uncertainties only, which was also higher than that based on the proposed Bayesian methodology (Fig. 8c).

3.5 Mapping carbon and nitrogen content of aboveground plant and grain at maturity

In view of the better performance of DArs, the maps of updated $W_{above}$, yield, $N_{above}$ and $N_{grain}$ at the experimental site at maturity were generated (Fig. 9) by assimilating the in-season crop growth simulation and leaf traits’ prediction (Fig. 5). Differences in the indirectly updated biomass and N content of aboveground plant and grain caused by the different N input were observed between the experimental plots. The variances of those updated simulations within each experimental plot were also shown in their corresponding maps (Fig. 9), in line with the existing variances in the maps of predicted leaf traits (Fig. 5).

4. Discussion

4.1 Impacts of residual error assumption for crop model calibration and data assimilation

The residual error assumption heavily affects the performance of the formal Bayesian approach (Beven et al., 2008). In our study, $\sigma_1$ denoted the heteroscedastic residual error, assuming that the residual error linearly increased with the crop model simulations, as the STDs in field measurements tended to vary with the averages (Fig. 3). Similarly, due to the increase of both the averages and the STDs of the $W_{above}$ throughout the seasons, the likelihood function was revised with the observational variances by Dumont et al. (2014) for the non-stationary residual errors. In line with their results, our study showed that $\sigma_1$ depicted the heteroscedasticity well for the crop model simulations (Fig. 3), although the calibration of the uncertain parameters tended to be sensitive to the range of $\sigma_1$ (Fig. 2). Instead of the heteroscedastic residual error,
Zhang et al. (2021a) used time-series variance across the growing season to investigate the heteroscedasticity of the uncertain parameters in the crop model AquaCrop, and showed that it significantly improved the effectiveness of the particle filter as well, when assimilating remotely sensed canopy cover into the crop model. In addition to the hypothesized heteroscedasticity, non-Gaussian errors were also introduced in our study and it was shown that the residual errors were more likely to be negatively skewed ($\xi < 1$), except for those of simulated $W_{\text{grain}}(\xi_{W_{\text{grain}}} = 2.86)$ and $N_{\text{grain}}(\xi_{N_{\text{grain}}} = 1.75)$ (Table 5).

Compared with the complex residual error assumptions, the i.i.d. Gaussian errors worked well in the MCMC approach for the parameter calibration of the crop model (Dumont et al., 2014; Kang and Özdoğan, 2019). However, while further assimilating crop model simulated and remotely sensed crop traits by EnKF, instead of the pre-assumed i.i.d. Gaussian errors, the inflation factor is introduced for accounting for model uncertainties (Kang and Özdoğan, 2019). Our results showed that the estimated uncertainties resulted in a better performance of the DA system, compared with the ones when the pre-assumed uncertainties was used alone or when it was combined with the inflation factor being introduced into EnKF (Fig. 8). Moreover, the inflation factor tends to not only be sensitive to the filter performance (Kang and Özdoğan, 2019; Whitaker and Hamill, 2012), but also differ between updated traits of crop and soil (Ines et al., 2013). With the rapid development of various satellites and unmanned aerial vehicles, a wide range of crop and soil information will be accessible to be incorporated into crop models (Jin et al., 2018). However, due to the uncertainties of crop and soil traits, rigorous determination of inflation factors in DA is difficult. Rather than the obscure adjustment of inflation factors, quantifying crop model uncertainties by the MCMC process together with the adapted residual error assumption in this study has great potential. With the Bayes’ multiplication method, we calibrated the crop model parameters well by exploiting multiple crop traits (Table 6, Fig. 3), in line with the results of He et al. (2010). The uncertainties of multiple crop traits in crop model simulation were estimated simultaneously (Table 5). The estimated crop model uncertainties worked well in EnKF while assimilating remotely sensed leaf traits into the crop model GECROS, while the $NRMSE$ of updated $W_{\text{grain}}$ and $N_{\text{grain}}$ increased slightly, compared with simulations using mean values of uncertain parameters (Table 6, Fig. 7). Recently, next to inflation factors and residual error models, the parameter and model structural errors of the simulation of rice phenology were determined by a multi-model ensemble method with assumed i.i.d. Gaussian errors, in which the different simulations generated from the different models were treated as the samples from the distribution of the unknown true model (Gao et al., 2021). Thus, more research is needed regarding the exploration of quantified crop model uncertainty and of their performance comparison in DA.

### 4.2 Uncertainty estimation of remote sensing predictions and its application in data assimilation

Like with the regression performance of GPR for leaf-level traits like specific leaf weight (Wang et al., 2019) and chlorophyll content (Verrelst et al., 2013a), our results showed high predictive performance for the traits $W_{\text{leaf}}, N_{\text{leaf}}$ and LAI, at canopy level, not only in the training dataset ($R^2 > 0.95$), but also in the testing dataset ($R^2 > 0.81$) (Fig. 4). Agreeing with the threshold proposed by Global Climate Observing System for the ecological application of remote sensing observations (GCOS, 2011), the relative uncertainties of predicted leaf traits in our results were below 20% (Fig. 5). The predicted low uncertainty was in line with other related GPR research (Verrelst et al., 2016; Wang et al., 2019). Similar to the results of Wang et al. (2019), relatively high uncertainties always came with high $N$ addition in the vegetative phase (Fig. 5a-c). We also found that the uncertainty tended to increase after entering the reproductive phase, especially for the treatments with low $N$ input, caused by early leaf senescence (Fig. 5d-e). Meanwhile, as GPR captured temporal and spatial variations in crop growth well, DA$_{a}$ performed better than DA$_{m}$ (Fig. 7), which agrees with the results of yield forecasting by assimilating LAI into the APSIM-Wheat model (Zhang et al., 2021b). The estimated uncertainty of remote sensing predictions likely affected inflation factor estimation as well, as the inflation factor is determined from the updated posteriors based on the observation errors (Whitaker and Hamill, 2012). The contribution of the GPR model should be further evaluated given that the uncertainty of remote sensing observations is commonly overlooked in DA (Huang et al., 2019) and arbitrarily set based on a general guess (e.g., Kang and Özdoğan (2019); Ines et al. (2013); Fig. 8b).
4.3 Performance of the analyzed system uncertainties in data assimilation

With the estimated uncertainties in crop model simulations and remote sensing observations, the inflation factor was no longer needed in this study. Filter divergence did not occur during the DA process with the analyzed system uncertainties (Fig. 6). 96% off \( r_c \) of leaf traits in DA\textsubscript{fm} and 76% in DA\textsubscript{sa} ranged from -0.9 to 0.9, indicating at least 10% uncertainty reduction of those leaf traits after conducting DA (Fig. 6a-b), and \( NRMSE \) of updated leaf traits on average decreased by 66% and 62% by DA\textsubscript{fm} and DA\textsubscript{sa}, respectively (Fig. 7a). This large improvement of directly updated states was in line with the results of Kivi et al. (2022) who estimated system uncertainties for EnKF while assimilating in situ observed daily soil moisture into the crop model APSIM. Even though the forecast accuracy of soil moisture was improved, the performance of updated LAI and yield tended to be worse than that of simulations without assimilating soil moisture (Kivi et al., 2022), which conflicts with previous studies (de Wit and van Diepen, 2007; Ines et al., 2013). As with the similar concerns of Schoups and Vrugt (2010) about the approach of simultaneous optimization and data assimilation proposed by Vrugt et al. (2005), crop growth simulation seems to be affected by measured soil moisture during joint estimation of system uncertainty, causing poorer performance in updated LAI and yield (Kivi et al., 2022). Without such tangled system uncertainties, our results showed that the performance of in-season updated and end-of-season forecasted crop carbon and N status improved (Fig. 7). However, there is still scope for further improvement. For instance, the distributions off \( r_c \) of updated aboveground status tended to be more diverged and closer to one (Fig. 6c-d), while those of updated leaf traits were more converged and closer to zero (Fig. 6a-b), implying a limitation in updating leaf traits for an accurate forecast of the \( W_{\text{above}} \) and \( N_{\text{above}} \). Meanwhile, the end-of-season forecast of \( W_{\text{above}} \) and \( N_{\text{above}} \) might also be impacted by their insufficiently updated in-season status (Fig. 7b-c, S8a-c).

4.4 Perspective of applying sophisticated crop models in data assimilation

Compared with applying simpler models like the crop model SAFY over large scales (Kang and Özdoğan, 2019), it is a challenge to use sophisticated crop models like GECROS for practical applications. For instance, GECROS is distinguished by a highly detailed photosynthetic process (Yin and Struik, 2017) and the related photosynthetic parameters were determined based on our previous study (Wang et al., 2022) and preset in this study, although the uncertainty still exists (Fig. S3). Also, the states updated in DA should be more carefully chosen, due to the complex feedback of the updated states in the sophisticated crop models that incorporate complex feedback mechanisms among biological processes. For instance, updating LAI in the crop model APSIM directly contributes to the improved performance of crop growth and yield formation (e.g., Zhang et al., 2022), due to its physiological mechanism that LAI influences the biomass accumulation by directly controlling the intercepted solar radiation in the canopy (Monsi and Saeki, 2005). However, as LAI is determined by both carbon and N status in the crop in the crop model GECROS (Yin et al., 2000), updating LAI alone hardly generated feedback for crop growth and yield forecast (results not shown) and thus updating states of \( W_{\text{leaf}} \) and \( N_{\text{leaf}} \) were incorporated together in this study as well (Fig. 7a). Similarly, Ines et al. (2013) indicated that updating states like \( W_{\text{leaf}} \) and specific leaf area, which feedback to LAI, might reduce the sensitivity of EnKF to model bias of the crop model DSSAT.

On the other hand, simple models have their own weaknesses. Due to the simplified physiological process, the key parameters in simple crop models like SAFY tend to be not only site-specific, but also year-specific (Claverie et al., 2012; Kang and Özdoğan, 2019). As the year-specific parameters should be calibrated with the actual in-season meteorological data, forecasting of crop growth and end-of-season yield is very uncertain (Kang and Özdoğan, 2019). In contrast to this, without any year-specific parameter calibration, GECROS performed reasonably well in the validation year (Table 6), especially for \( W_{\text{above}} \) (\( R^2 > 0.87, NRMSE = 0.28 \)). By only updating remotely sensed leaf traits, the performance of the simulated \( W_{\text{above}} \) improved and its \( NRMSE \) at maturity further decreased to 0.25 (Fig. 7b-c). Relying on the integrated simulation of physiological processes relating to N dynamics, \( N_{\text{above}} \) and \( N_{\text{grain}} \) were simulated by GECROS and further updated more accurately in EnKF than that of open-loop (Fig. 7b-c), which probably forms a better basis for predicting traits like grain quality (Ma et al., 2022).

4.5 Prospect of updated crop status forecast for field-level in-season management
Since most remote sensing algorithms have been adapted to large areas, making them applicable to the field-level is a priority now, especially in countries like China (Weiss et al., 2020) where the agricultural system is dominated by millions of smallholders and N is always overdosed (Cui et al., 2018). Spatial and temporal predictions of \(N_{\text{leaf}}\) (Fig. 5) are crucial for determining the management of fertilization timing to be performed at specific growth stages (Weiss et al., 2020). However, the use of remote sensing data solely is not sufficient to quantify top-dressing requirements (Weiss et al., 2020), due to the manifold interactions in the soil-crop-atmosphere continuum. Assimilating supplementary information from a crop model with remote sensing data has been identified as one of the most promising methods to make field management decisions (Jin et al., 2018; Weiss et al., 2020). For instance, Baret et al. (2007) demonstrated that an optimized in-season N application map can be generated by assimilating remotely sensed LAI and \(N_{\text{above}}\) into the crop model STICS, in which the N application rate in each map pixel (20×20 m\(^2\)) was optimized by maximizing the productivity, using the historical meteorological data between the time of decision and harvest as the unknown future weather conditions. With optimized real-time fertilizer management, severe environmental issues caused by overfertilization can be reduced, thereby enabling smart farming and sustainable agricultural production (Berger et al., 2020b).

Conclusion

In this study, we developed an approach that combines several disparate quantitative methods into one framework for explicitly quantifying the uncertainties of crop model simulations and remotely sensed observations, contributing to an accurate crop growth forecast while avoiding filter divergence in DA. Our results showed that calibration and uncertainty estimation of a complex crop model benefited from an MCMC approach using the adapted residual error model. The calibrated uncertain parameters in crop model performed reasonably well in validation. The GPR models for analyzing remote sensing data provided temporal and spatial predictions and corresponding uncertainties of leaf traits. Applying the quantified uncertainties into EnKF to update the leaf traits \(W_{\text{leaf}}\), \(N_{\text{leaf}}\) and LAI enabled the crop model simulation to agree better with the measurements, without filter divergence. More importantly, updated simulations of in-season and end-of-season \(W_{\text{above}}\), \(W_{\text{grain}}\), \(N_{\text{above}}\) and \(N_{\text{grain}}\) also performed better than those of simulations without assimilating observations. The developed method is geared toward cases where multiple crop traits are observed and in-situ crop and soil information becomes increasingly available with the rapid development of remote sensing technologies. Armed with the precise forecast of in-season crop carbon and N status, field management can be better optimized to support sustainable smart farming.

Acknowledgement

We greatly appreciate the expert input from Dr. Marcel van Oijen, whose critical comments, enlightening suggestions and detailed edits helped improve earlier drafts of this manuscript considerably.

References


Table 1. Basic information of initial soil conditions at the experimental site.

<table>
<thead>
<tr>
<th>Soil type</th>
<th>Clay percentage (%)</th>
<th>Total organic carbon (g C m⁻²)</th>
<th>Ammonium-N (g N m⁻²)</th>
<th>Nitrate-N (g N m⁻²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>loam</td>
<td>18.1</td>
<td>7039</td>
<td>0.19</td>
</tr>
<tr>
<td>2020</td>
<td>loam</td>
<td>21.0</td>
<td>6376</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Table 2. Split-applied nitrogen (N) fertilizer rates at different growth stages during the rice growing season in two experimental years.

<table>
<thead>
<tr>
<th>Year</th>
<th>Application stage</th>
<th>N rate (kg N ha⁻¹)</th>
<th>N rate (kg N ha⁻¹)</th>
<th>N rate (kg N ha⁻¹)</th>
<th>N rate (kg N ha⁻¹)</th>
<th>N rate (kg N ha⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>Beginning of tillering</td>
<td>0</td>
<td>40</td>
<td>80</td>
<td>160</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>Two weeks after tillering</td>
<td>0</td>
<td>16</td>
<td>32</td>
<td>64</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Panicle initiation</td>
<td>0</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Two weeks after flowering</td>
<td>0</td>
<td>10</td>
<td>20</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>2020</td>
<td>Seedling</td>
<td>0</td>
<td>12</td>
<td>24</td>
<td>48</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>Beginning of tillering</td>
<td>0</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Two weeks after tillering</td>
<td>0</td>
<td>8</td>
<td>16</td>
<td>32</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Panicle initiation</td>
<td>0</td>
<td>12</td>
<td>24</td>
<td>48</td>
<td>72</td>
</tr>
</tbody>
</table>

Table 3. Destructive sampling dates at their corresponding rice growth stages in the experimental years 2019 and 2020.

<table>
<thead>
<tr>
<th></th>
<th>Tillering</th>
<th>Stem-elongating</th>
<th>Booting</th>
<th>Flowering</th>
<th>Grain-filling</th>
<th>Grain-filling</th>
<th>Maturity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019</td>
<td>10-Jul</td>
<td>14-Aug</td>
<td>11-Sep</td>
<td>17-Sep</td>
<td>8-Oct</td>
<td>30-Oct</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. List of preset and uncertain parameters in the crop model GECROS with their definitions and references.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value or range</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preset SLNMIN</td>
<td>0.3</td>
<td>g N m$^{-2}$</td>
<td>Base leaf N for photosynthesis</td>
</tr>
<tr>
<td>MSR0</td>
<td>2.5</td>
<td>-</td>
<td>The ratio of mesophyll to stomatal resistance</td>
</tr>
<tr>
<td>DD0C3</td>
<td>0.38</td>
<td>-</td>
<td>The slope of the decrease of $C_v:C_a$ ratio with increasing VPD</td>
</tr>
<tr>
<td>XVN</td>
<td>210</td>
<td>µmol s$^{-1}$ g$^{-1}$</td>
<td>Linear slope of maximum Rubisco activity at 25°C versus leaf N content</td>
</tr>
<tr>
<td>XJN</td>
<td>255</td>
<td>µmol s$^{-1}$ g$^{-1}$</td>
<td>Linear slope of maximum PSII e$^-$ transport rate at 25 °C versus leaf N content</td>
</tr>
<tr>
<td>MTDV</td>
<td>80.86</td>
<td>days</td>
<td>Minimum thermal days for vegetative growth phase</td>
</tr>
<tr>
<td>MTDR</td>
<td>27.74</td>
<td>days</td>
<td>Minimum thermal days for reproductive growth phase</td>
</tr>
<tr>
<td>HTMX</td>
<td>0.8</td>
<td>m</td>
<td>Maximum plant height</td>
</tr>
<tr>
<td>STEMNC</td>
<td>0.005</td>
<td>g N g$^{-1}$</td>
<td>Minimum N concentration in structural stem material</td>
</tr>
<tr>
<td>Free NUPTX</td>
<td>0.095-0.230</td>
<td>g N m$^{-2}$ d$^{-1}$</td>
<td>Maximum daily crop nitrogen uptake</td>
</tr>
<tr>
<td>rASSA</td>
<td>0.5-1.0</td>
<td>-</td>
<td>The factor for carbon assimilates, to collectively correct for the uncertainty of photosynthetic parameters</td>
</tr>
<tr>
<td>LNCI</td>
<td>0.02-0.05</td>
<td>g N g$^{-1}$</td>
<td>Initial nitrogen concentration in living leaves</td>
</tr>
<tr>
<td>SLA0</td>
<td>0.0150-0.0166</td>
<td>m$^2$ leaf$^{-1}$</td>
<td>Specific leaf area constant</td>
</tr>
<tr>
<td>RDLV</td>
<td>0.01-0.10</td>
<td>-</td>
<td>Relative leaf death rate under conditions where seed fill is happened due to environmental stress</td>
</tr>
<tr>
<td>PMEH</td>
<td>0.6-0.9</td>
<td>-</td>
<td>Fraction of sigmoid curve inflexion in entire plant height growth period</td>
</tr>
<tr>
<td>CDMHT</td>
<td>400-750</td>
<td>g m$^{-2}$ m$^{-1}$</td>
<td>Stem dry weight per unit of plant height</td>
</tr>
<tr>
<td>RVNC</td>
<td>0.001-0.025</td>
<td>g N g$^{-1}$</td>
<td>Nitrogen concentration in the reserve part of stem</td>
</tr>
<tr>
<td>SDF</td>
<td>15-35</td>
<td>-</td>
<td>Factor for biomass determined total seed number</td>
</tr>
<tr>
<td>PNPRE</td>
<td>0.6-1.0</td>
<td>-</td>
<td>Proportion of seed nitrogen that comes from non-structural nitrogen</td>
</tr>
<tr>
<td>SEEDW</td>
<td>0.0243-0.0278</td>
<td>g seed$^{-1}$</td>
<td>Seed weight</td>
</tr>
<tr>
<td>SEEDNC</td>
<td>0.0058-0.0130</td>
<td>g N g$^{-1}$</td>
<td>Standard seed (storage organ) nitrogen concentration</td>
</tr>
<tr>
<td>WCMIN</td>
<td>0.05-0.15</td>
<td>m$^3$ m$^{-3}$</td>
<td>Minimum soil water content</td>
</tr>
<tr>
<td>WCFC</td>
<td>0.25-0.40</td>
<td>m$^3$ m$^{-3}$</td>
<td>Soil water content at holding capacity</td>
</tr>
<tr>
<td>WCMAX</td>
<td>0.40-0.55</td>
<td>m$^3$ m$^{-3}$</td>
<td>Soil water content at maximum holding capacity</td>
</tr>
<tr>
<td>BHC</td>
<td>2500-4000</td>
<td>g C m$^{-2}$</td>
<td>Initial value for microbial biomass and humified organic matter in the soil</td>
</tr>
<tr>
<td>HUMR</td>
<td>0.01-0.02</td>
<td>yr$^{-1}$</td>
<td>Decomposition rate constant for humified organic matter in the soil</td>
</tr>
<tr>
<td>BIOR</td>
<td>0.50-0.66</td>
<td>yr$^{-1}$</td>
<td>Decomposition rate constant for microbial biomass in the soil</td>
</tr>
</tbody>
</table>

Table 5. Estimated values of means and standard deviations (STDs) of posterior distributions of uncertain parameters in GECROS and parameters in error model while setting the upper bound of the heteroscedasticity parameter $\sigma_1$ at 0.2.

<table>
<thead>
<tr>
<th>Parameters in the GECROS model</th>
<th>Mean</th>
<th>STD</th>
<th>Parameters in the error model</th>
<th>Mean</th>
<th>STD</th>
</tr>
</thead>
<tbody>
<tr>
<td>NUPTX</td>
<td>0.206</td>
<td>0.018</td>
<td>$\sigma_{1,W_{above}}$</td>
<td>0.189</td>
<td>0.009</td>
</tr>
<tr>
<td>rASSA</td>
<td>0.705</td>
<td>0.024</td>
<td>$\xi_{W_{above}}$</td>
<td>0.976</td>
<td>0.171</td>
</tr>
<tr>
<td>LNCI</td>
<td>0.0203</td>
<td>0.0005</td>
<td>$\lambda_{W_{above}}$</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>SLA0</td>
<td>0.0164</td>
<td>0.0003</td>
<td>$\sigma_{1,W_{grain}}$</td>
<td>0.162</td>
<td>0.026</td>
</tr>
<tr>
<td>RDLV</td>
<td>0.017</td>
<td>0.013</td>
<td>$\xi_{W_{grain}}$</td>
<td>2.860</td>
<td>1.309</td>
</tr>
<tr>
<td>PMEH</td>
<td>0.873</td>
<td>0.028</td>
<td>$\lambda_{W_{grain}}$</td>
<td>0.999</td>
<td>0.001</td>
</tr>
<tr>
<td>CDMHT</td>
<td>627.3</td>
<td>89.0</td>
<td>$\sigma_{1,W_{leaf}}$</td>
<td>0.191</td>
<td>0.007</td>
</tr>
<tr>
<td>RVNC</td>
<td>0.008</td>
<td>0.006</td>
<td>$\xi_{W_{leaf}}$</td>
<td>0.427</td>
<td>0.180</td>
</tr>
</tbody>
</table>
Parameters in the GECROS model\textsuperscript{a} & Mean & STD & Parameters in the error model b & Mean & STD \\
\hline
SDF & 29.28 & 3.99 & $\lambda_{W_{\text{leaf}}}$ & 0.999 & 0.001 \\
PNPRE & 0.877 & 0.086 & $\sigma_1, N_{\text{above}}$ & 0.196 & 0.005 \\
SEEDW & 0.026 & 0.001 & $\xi_{N_{\text{above}}}$ & 0.952 & 0.122 \\
SEEDNC & 0.011 & 0.001 & $\lambda_{N_{\text{above}}}$ & 0.977 & 0.024 \\
WCMIN & 0.099 & 0.026 & $\sigma_1, N_{\text{grain}}$ & 0.149 & 0.034 \\
WCFC & 0.304 & 0.040 & $\xi_{N_{\text{grain}}}$ & 1.746 & 1.077 \\
WCMAX & 0.459 & 0.039 & $\lambda_{N_{\text{grain}}}$ & 0.718 & 0.229 \\
BHC & 3899 & 100 & $\sigma_1, N_{\text{leaf}}$ & 0.190 & 0.009 \\
HUMR & 0.018 & 0.001 & $\xi_{N_{\text{leaf}}}$ & 0.745 & 0.140 \\
BIOR & 0.629 & 0.026 & $\lambda_{N_{\text{leaf}}}$ & 0.944 & 0.063 \\
& & & $\sigma_1, \text{LAI}$ & 0.190 & 0.007 \\
& & & $\xi_{\text{LAI}}$ & 0.646 & 0.174 \\
& & & $\lambda_{\text{LAI}}$ & 0.943 & 0.069 \\
\hline
\textsuperscript{a} GECROS-model parameters and their units are defined in Table 4; \\
\textsuperscript{b} $\sigma_1$, $\xi$ and $\lambda$ denote the heteroscedasticity parameter, the skewness parameter and the Box-Cox transformation parameter in the error model, respectively (see the text). Definition of the traits used in the subscript of symbols: $W_{\text{above}}$, aboveground biomass; $W_{\text{grain}}$, grain weight; $W_{\text{leaf}}$, leaf weight; $N_{\text{above}}$, aboveground N content; $N_{\text{grain}}$, grain N content; $N_{\text{leaf}}$, leaf N content in the canopy; and LAI, leaf area index.

Table 6: \textit{NRMSE} and $R^2$ of simulations in the calibration (year 2019) and the validation (year 2020) using estimated mean values of uncertain parameters in the crop model GECROS\textsuperscript{a}.

| & $W_{\text{above}}$ & $W_{\text{grain}}$ & $W_{\text{leaf}}$ & $N_{\text{above}}$ & $N_{\text{grain}}$ & $N_{\text{leaf}}$ & LAI |
|---|---|---|---|---|---|---|---|
| Calibration (2019) $NRMSE$ & 0.162 & 0.171 & 0.332 & 0.280 & 0.266 & 0.431 & 0.362 |
| $R^2$ & 0.882 & 0.645 & 0.810 & 0.704 & 0.662 & 0.743 & 0.747 |
| Validation (2020) $NRMSE$ & 0.282 & 0.153 & 0.486 & 0.386 & 0.180 & 0.590 & 0.388 |
| $R^2$ & 0.873 & 0.893 & 0.515 & 0.732 & 0.849 & 0.725 & 0.579 |

\textsuperscript{a} To make the performance in the calibration and the validation comparable, the mean values of two years' field measurements for each type of observation were used here for calculating \textit{NRMSE}. $W_{\text{above}}$, aboveground biomass; $W_{\text{grain}}$, grain weight; $W_{\text{leaf}}$, leaf weight; $N_{\text{above}}$, aboveground N content; $N_{\text{grain}}$, grain N content; $N_{\text{leaf}}$, leaf N content in the canopy; and LAI, leaf area index.
Fig 1. Framework of the developed methodology in this study that applies systematically analyzed errors of observations and simulations in the data assimilation procedure to enhance crop status forecast. Field experiments in two successive years were conducted for the acquisition of the necessary dataset to validate this method. Daily weather data served as the forcing input of the crop model GECROS and all field observations in the first year were averaged for model calibration. Before conducting the calibration procedure, the parameters in GECROS were fixed or treated as uncertain and the error model that describes the uncertainty of crop model simulation was assumed. The uncertain parameters in GECROS and parameters in the error model were determined simultaneously by an efficient Markov Chain Monte Carlo approach (DREAMzs). To further improve the forecast ability of the crop model, the in-season observations of leaf traits in the second year were incorporated by the commonly used data assimilation procedure, Ensemble Kalman Filter (EnKF), which integrates the sequential observations into crop model simulations of crop growth processes. Two types of observations were collected. The first one was from field destructive measurements. The second type was from the remote sensing predictions, which were regressed from the machine learning method of Gaussian Process Regression (GPR). The uncertainties of field measurements were derived from the replications in the 2nd year’s field experiment, while that of remote sensing predictions were estimated from the GPR model itself. With the systematically analyzed uncertainties of crop model simulations and observations, in-season leaf traits at the canopy level, leaf weight ($W_{\text{leaf}}$), leaf nitrogen (N) content ($N_{\text{leaf}}$) and leaf area index (LAI) were updated directly by EnKF, and other crop carbon and N statuses, including aboveground biomass ($W_{\text{above}}$), grain weight ($W_{\text{grain}}$), aboveground N content ($N_{\text{above}}$) and grain N content ($N_{\text{grain}}$), were updated accordingly as well. The performance of the updated states in the second year was evaluated by their in-situ measurements, accordingly.

Fig. 2. Comparison of simulation performance in calibration, $NRMSE$ (a) and $R^2$ (b), in which the uncertain
parameters in GECROS were calibrated with different upper bounds of the heteroscedasticity parameter $\sigma_1$, $W_{\text{above}}$, aboveground biomass; $W_{\text{grain}}$, grain weight; $W_{\text{leaf}}$, leaf weight; $N_{\text{above}}$, aboveground N content; $N_{\text{grain}}$, grain N content; $N_{\text{leaf}}$, leaf N content in the canopy; and LAI, leaf area index.

Fig. 3. Estimation of parameter (dark gray region) and total (light gray region) uncertainty of GECROS with the nitrogen input of 0 kg ha$^{-1}$ (a-f) and 240 kg ha$^{-1}$ (g-l) in 2019. Red lines represent the simulations from estimated mean values of uncertain parameters. The solid circles with vertical bars denote the means and standard deviations of measurements. $W_{\text{above}}$, aboveground biomass; $W_{\text{grain}}$, grain weight; $W_{\text{leaf}}$, leaf weight; $N_{\text{above}}$, aboveground N content; $N_{\text{grain}}$, grain N content; and $N_{\text{leaf}}$, leaf N content in the canopy.
Fig. 4. Scatter plots of predicted leaf weight ($W_{\text{leaf}}$), leaf N content in the canopy ($N_{\text{leaf}}$) and leaf area index (LAI) from hyperspectral image data with respect to their measurements in the training (a-c) and testing (d-f) set, using the Gaussian Process Regression (GPR) algorithm. The vertical bars denote the predicted standard deviations from the GPR. The grey lines indicate a 1:1 relationship.

Fig. 5. Maps of the predicted leaf weight ($W_{\text{leaf}}$), leaf N content in the canopy ($N_{\text{leaf}}$) and leaf area index (LAI) at the stage of stem-elongating (15-Aug, 2020) (a-c) and grain-filling (10-Oct, 2020) (d-f) at the experimental site by regressed Gaussian Process Regression (GPR) models. The predicted value (mean) and its standard deviation (STD) were predicted from the GPR models and mapped in each panel. N0, N40, N80, N160, N240 and N320 denote varied nitrogen (N) application rates (see Table 2 for details).
Fig. 6. The distribution of $f_{rc}$ (see Eqn 21 for its definition) at each day conducting data assimilation (DA) by assimilating field measurements (DA$_{fm}$) (a, c) or remote sensing observations (DA$_{rs}$) (b, d) and its overall relative cumulative frequency. The $f_{rc}$ of directly updated states, leaf weight ($W_{leaf}$), leaf N content in the canopy ($N_{leaf}$) and leaf area index (LAI), were shown in Panel a and b, while that of indirectly updated states, aboveground biomass ($W_{above}$) and aboveground N accumulation ($N_{above}$), are shown in Panel c and d.

Fig. 7. Comparison of NRMSE of states simulated from the GECROS model without assimilating observations (open-loop) and of states analyzed from assimilating observations of field measurements (DA$_{fm}$) and of remote sensing predictions (DA$_{rs}$) in year 2020. The overall NRMSE in the whole growing season of the directly updated states, leaf weight ($W_{leaf}$), leaf N content in the canopy ($N_{leaf}$) and leaf area index (LAI), are shown in Panel a. The overall NRMSE in the whole growing season and NRMSE at the harvesting stage of the indirectly updated states, aboveground biomass ($W_{above}$), grain weight ($W_{grain}$), aboveground N
accumulation (\(N_{\text{above}}\)) and grain N content (\(N_{\text{grain}}\)), are shown in Panel b and c, respectively.

Fig. 8. Performance comparison of the data assimilation system with the assumed uncertainties alone, the assumed uncertainties combined with the introduced inflation factor in EnKF, and the quantified uncertainties. The \textit{NRMSE} of the updated states in the growing season was shown, based on different values of coefficient of variation (CV) of crop model simulations with the CV of remote sensing predictions fixed at 0.01 (a), different values of CV of remote sensing predictions with the CV of crop model simulations fixed at 0.1 (b), and different values of inflation factor in EnKF with the CVs of crop model simulations and remote sensing predictions fixed at 0.1 and 0.01, respectively (c). The inflation factor was fixed at 1.0 for panels (a) and (b), assuring the assumed uncertainties functioned alone. The red lines in each panel indicate the \textit{NRMSE} based on the estimated uncertainties of crop model simulations and remote sensing predictions from the proposed Bayesian methodology in this study. \(W_{\text{leaf}}\), leaf weight; \(N_{\text{leaf}}\), leaf N content in the canopy; LAI, leaf area index; \(W_{\text{above}}\), aboveground biomass; \(W_{\text{grain}}\), grain weight; \(N_{\text{above}}\), aboveground N content; and \(N_{\text{grain}}\), grain N content.
Fig. 9. Maps of the updated simulations of aboveground biomass ($W_{\text{above}}$) (a), yield (b), aboveground N content ($N_{\text{above}}$) (c) and grain N content ($N_{\text{grain}}$) (d) at maturity in year 2020 at the experimental site. Both mean and standard deviation (STD) of each state were derived from their updated ensemble values in EnKF. N0, N40, N80, N160, N240 and N320 denote varied nitrogen (N) application rates (see Table 2 for details).