European Journal of Agronomy 118 (2020) 126076

Contents lists available at ScienceDirect

European Journal of Agronomy

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

Analyzing uncertainty in critical nitrogen dilution curves

David Makowski[a,b,*], Ben Zhao[c], Syed Tahir Ata-Ul-Karim[d], Gilles Lemaire[e]

[a] University Paris-Saclay, INRAE, AgroParisTech, UMR 211, 78850, Thiverval-Grignon, France
[b] CIRED, 45bis Avenue de la Belle Gabrielle, 94130, Nogent-sur-Marne France
[c] Key Laboratory of Crop Water Use and Regulation, Ministry of Agriculture, Farmland Irrigation Research Institute, Chinese Academy of Agricultural Sciences, 380 Hongli Road, Xinxiang, Henan 453003, PR China
[d] Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, ChineseAcademy of Sciences, 71 East Beijing Road, Nanjing, Jiangsu 210008, PR China
[e] Honorary Director of Research INRAE (Retired), Lusignan, 86500, France

ARTICLE INFO

Keywords:
Bayesian statistics
Critical concentration
Critical N curve
Nitrogen
Probabilistic expert elicitation
Uncertainty

ABSTRACT

Nitrogen critical curves are frequently used to diagnose the N status of crops and grasslands. They play an important role in plant modelling and are frequently used in fertilizer management tools. During the last 20 years, a number of studies have been conducted for comparing critical curves obtained in various conditions (e.g., different cultivars) and understanding the origin of their differences. However, uncertainty in the determination of these curves is generally poorly analyzed in these studies, which increase the risk of false conclusions, in particular on the existence of differences between species, cultivars and cropping systems. Here, we present a Bayesian statistical model for estimating parameters of critical nitrogen dilution curve from experimental data. Contrary to standard methods commonly used for fitting critical nitrogen dilution curves, the proposed approach allows one to fit these curves in only one step, i.e., directly from the original biomass and nitrogen content measurements. Specifically, this method does not require the classification of nitrogen-limited data against non-nitrogen-limited data and does not necessitate the preliminary identification of critical nitrogen concentrations. Another advantage of the proposed method is that it can be easily implemented using freely-available software. We illustrate its practical interest using experimental data collected for winter wheat in France, and for maize and rice in China. We show that this method is useful for analyzing uncertainty in the fitted critical nitrogen curves and for comparing several curves obtained for different crop species and cultivars. The proposed method is based on the specification of prior probability distributions defining plausible ranges of values for the critical curve parameters, and we show here that it is preferable to use prior distributions that are not very informative if we want to limit their influence on the final result.

1. Introduction

Nitrogen (N) fertilization plays a major role in agricultural production but excess of N in agro-ecosystems has negative impacts on water pollution (Zhao et al., 2007), and increases risk of ammonia and N_2O emissions (Philibert et al., 2012; Ramanantenasoa et al., 2019). It is thus essential to precisely manage N fertilization and to develop operational tools helping farmers to determine optimal N fertilizer doses and times of application. A prerequisite for the development of such tools is to estimate crop nitrogen requirements as accurately as possible.

The concept of critical N concentration (N_C) is frequently used to diagnose the N status of crops (Lemaire et al., 2008). The value of N_C represents the minimum N concentration that is required for maximum biomass production. This concentration is usually computed as a function of biomass using a simple mathematical model, often called critical N curve. Although several variants of this model exist, the most common model is expressed as $N_C = A_1 W^{-A_2}$ where N_C is the critical N concentration for biomass W, and where A_1 and A_2 are two parameters that are estimated by fitting the model to a set of experimental data. Critical N curves relating N_C to W have become popular since the late 1990s, and they have been developed for a number of plant species, including winter wheat (Justes et al., 1994; Chen and Zhu, 2013), oilseed rape (Colnenne et al., 1998), maize (Plénet and Lemaire, 2000), ryegrass (Sandana et al., 2019), rice (Ata-Ul-Karim et al., 2017) among others.

It is essential to analyze uncertainty in fitted critical N curves in a rigorous manner. This is important for assessing risk of N deficit or N

* Corresponding author at: University Paris-Saclay, INRAE, AgroParisTech, UMR 211, 78850, Thiverval-Grignon, France.
E-mail address: david.makowski@inrae.fr (D. Makowski).

https://doi.org/10.1016/j.eja.2020.126076
Received 16 December 2019; Received in revised form 8 April 2020; Accepted 15 April 2020
Available online 18 May 2020
1161-0301/ © 2020 Elsevier B.V. All rights reserved.

excess and, also, for comparing critical curves obtained in different conditions (e.g., sites, years, different cultivars and/or crop managements) and understanding the origin of their difference. In some cases, two fitted curves are apparently different, but their difference does not always reflect a real difference due to the existence of a true effect of the factor studied, but may simply reflect errors in the estimated values of their parameters. In such cases, concluding that there is a real difference and a real effect of the factor under consideration will lead to a false conclusion. There is increasing concern that false findings may be very frequent in research (Ioannidis, 2005). To limit risk of false discoveries, a rigorous analysis of uncertainty is then crucial.

In practice, the parameters of critical N curves are estimated using a series of pairs of biomass and plant N concentration measurements obtained at different dates during the growing period for different N levels. Several methods have been proposed to estimate parameters of critical curves from this type of data (Greenwood et al. 1990; Justes et al., 1994; Chen and Zhu, 2013). All these methods require a classification of N-limited data vs. non-N-limited data at each date of measurement. The two groups of data (N-limited vs. non-N-limited) are then used to identify a so-called critical-N-concentration above which biomass is assumed to reach its maximum value. A critical N curve is finally fitted to the series of critical-N-concentration values obtained across the different dates of measurement. Although this approach has been successful in fitting critical N curves in a great diversity of contexts, it presents several limitations. This approach requires the definition of a classification rule to distinguish N-limited vs. non-N-limited data and of a method to define the N critical concentration from the two groups of data, and Chen and Zhu (2013) showed that the fitted critical N curve can be sensitive to these choices. Another issue is that the critical N curves are fitted to the selected critical N concentrations then considered as perfectly known, ignoring their own uncertainties. Once the critical nitrogen concentrations are estimated, the critical nitrogen curve is adjusted to their values without explicitly taking into account their uncertainties. A consequence is that the confidence intervals of the estimated parameters and of the fitted curves do not fully account for the uncertainties in the selected critical N concentrations.

Here, we present a Bayesian statistical model for estimating parameters of critical nitrogen dilution curve from experimental data. Contrary to standard methods commonly used for fitting critical curves, the proposed approach allows one to fit these curves in only one step, i.e., directly from the original biomass and nitrogen content measurements. Our approach is illustrated with experimental data collected for wheat, maize and rice. We show that the proposed method can be easily used to fit critical N curves, to analyze uncertainty, and to compare the parameter values estimated for different species and cultivars.

2. Materials and method

2.1. Data

2.1.1. Wheat

Data used in this study for wheat come from Justes et al. (1994). These data had been extracted from a large experimental network with different year-location combinations in France. In each experiment wheat crops were grown with at least four different levels of N fertilizer supply and several plant samples were regularly harvested all along the crop growth period until anthesis. Plant samples were analyzed to determine their aboveground biomass (t ha^{-1}) and plant shoot N concentration (%) at each sampling date. A series of pairs of N concentration and of biomass were thus obtained across the different N levels * dates combinations (Appendix A) and included in a single dataset for the statistical analysis. The dataset covers a large range of biomass values, from about 1 t ha^{-1} at the early stages of plant growth to about 12 t ha^{-1} at plant anthesis. Crop were managed with ample P and K supply, with irrigation when necessary for avoiding water stress, and with adequate plant disease control. The total number of pairs of biomass and N content is equal to 73 (Table 1).

Table 1
Number of dates of measurement (called 'dates' in the model), average number of N levels per date, and number of data (pairs of biomass and N content observations) for each dataset.

Crop/Cultivar	Number of dates of measurement	Average number of N levels per date	Number of data
Rice/Japonica	12	5	60
Rice/Indica	12	5	60
Maize/ZD958	10	4.5	45
Maize/DH605	10	4.5	45
Winter wheat	16	4.6	73

2.1.2. Maize

Four field experiments including each several N rates were conducted during the 2015 and 2016 growing seasons at Xinxiang (35.2 °N, 113.8 °E). The summer maize cultivars, N application rates, sowing and harvesting dates, as well as soil characteristics, are summarized in Appendix B. Soil samples were collected from 0 to 20 cm soil layer before sowing summer maize crops. The samples were air-dried, sieved, and then analyzed. All field experiments were arranged in a randomized complete block design with three replicates. The size of each plot was 60m^2 in all the experiments. N fertilizer was applied before sowing (50 %) and at the jointing stage (50 %). All plots received adequate quantities of triple super-phosphate and potassium-chloride before sowing. Summer maize was planted a density of 75,000 plants ha^{-1} with a row spacing of 60 cm. The total number of pairs of biomass (t ha^{-1}) and N content (%) is equal to 45 (Table 1).

2.1.3. Rice

Four multi-N rates (0–360 kg N ha^{-1}) field experiments using one Indica (Jingliangyou-534) and one Japonica rice (Jiahua-1) cultivar were conducted during 2017 and 2018 rice growing seasons in east China. Experiments were arranged with a randomized complete block design having three repeats. The size of every plot was 5 m × 5 m with the inter-row spacing of 30 cm. The planting density in all the experiments was approximately 22.2 × 10^4 plants ha^{-1}. Five N supply rates (0, 90, 180, 270, and 360 kg N ha^{-1} as urea) were applied. 40 % N was distributed before transplanting, 10 % at active tillering, 20 % at panicle initiation, and 30 % at booting. In all experiments, ample phosphate and potassium fertilizers were incorporated into the soil as monocalcium phosphate ($Ca(H_2PO_4)2$) and potassium chloride (KCl) before transplantation. Experiments were carried out according to local recommendations along with adequate plant pest and disease control measures to ensure optimal production. Plant samples were regularly harvested all along the crop growth period until heading (pre-anthesis growth period) for determination of plant dry mass and plant N concentration to provide a set of plant dry mass-%N data across the different N supply rates at each sampling stage. The total number of pairs of biomass (t ha^{-1}) and N content (%) is equal to 60 (Table 1).

2.2. Model

Our model is a Bayesian hierarchical model (Fig. 1). In this model, we consider that the response of biomass to nitrogen content follows a linear-plus-plateau function, as commonly considered in many studies (see for example Chen and Zhu, 2013 and Zhao et al., 2018). The variability of the parameters of the linear-plus-plateau function is described by probability distributions estimated from the whole set of available data using a Bayesian method. The parameters of the critical nitrogen dilution curve are then derived directly from the fitted probability distributions.

The first level of the model describes the biomass response to nitrogen content for a given date of measurement based on a linear-plus-

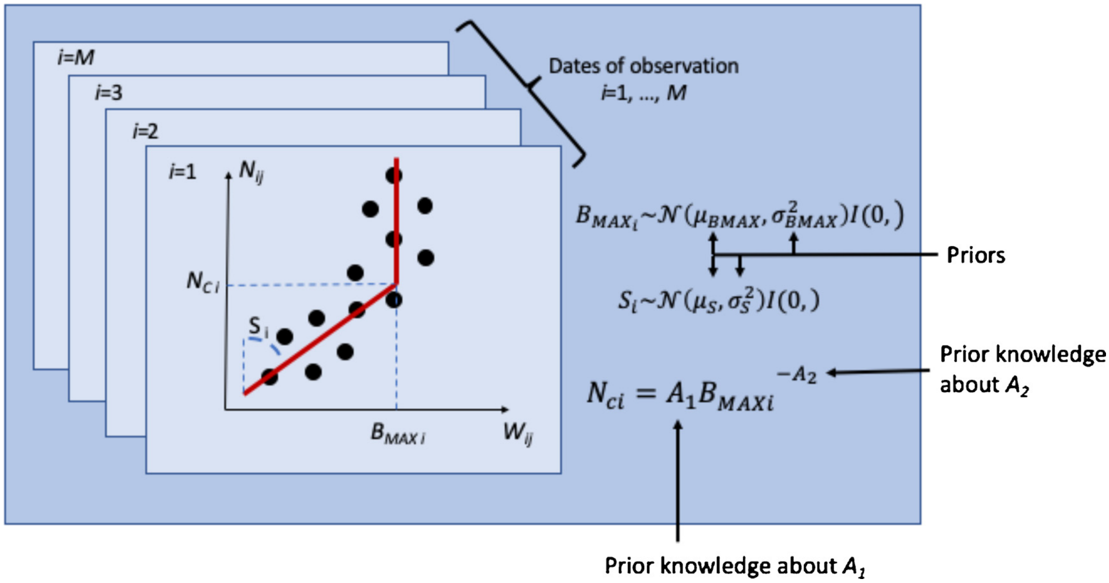


Fig. 1. Graphical description of the three-level hierarchical Bayesian model. Level 1 describes the relationship between biomass and nitrogen content observations for each date of measurement (light blue). Level 2 describes the critical nitrogen dilution curve and the variability of biomass and nitrogen content between measurement dates using several probability distributions (dark blue). Level 3 describes prior knowledge about parameter values (outside part of the graphic). The indices *i* and *j* correspond to the dates of observation and the supplied N fertilizer rates, respectively. See text for details. (For interpretation of the references to colour in the Figure, the reader is referred to the web version of this article).

plateau function. Each date of measurement corresponds to a specific crop growth stage in a given year at which biomass and N content are measured for different N fertilizer levels. The second level of the model describes the variability of the parameters of the linear-plus-plateau function across observation dates using probability distributions, and computes the critical nitrogen dilution curve. The third level describes prior knowledge about parameter values.

2.2.1. Level 1: Biomass response to nitrogen content

This part of the model describes the relationship between biomass (W_{ij}, t ha^{-1}) and nitrogen content (N_{ij}, %) measurements in the ith date of observation for the jth nitrogen dose. We assume that W_{ij} is distributed according to a Gaussian distribution whose mean value is specified through a linear-plus-plateau function of the nitrogen content as follows:

$$W_{ij} \sim \mathcal{N}(\mu_{ij}, \tau_b^2),\ \text{with}\ \mu_{ij} = B_{MAX_i} + S_i(N_{ij} - N_{Ci})\ \text{if}\ N_{ij} < N_{Ci}\ \text{and}\ \mu_{ij} = B_{MAX_i}\text{otherwise} \quad (1)$$

In Eq.(1), B_{MAX_i} is the maximum biomass in t ha^{-1} (non-N limited) in the ith date (i.e., B_{MAX_i} is equal to the mean value of W_{ij} when $N_{ij} \geq N_{Ci}$), S_i is the slope of the linear part of the function (i.e., the increase rate of biomass per unit increase of nitrogen content, t ha^{-1} % N^{-1}), N_{Ci} is the critical nitrogen content for the ith date (%), i.e., the value of nitrogen content (%) above which B_{MAX_i} is reached, and τ_b^2 is the residual variance. We also assume that the nitrogen content measurements (N_{ij}) is related to the critical nitrogen content (N_{Ci}) according to the distribution

$$N_{ij} \sim \mathcal{N}(N_{Ci}, \tau_n^2) \quad (2)$$

where τ_n^2 is a variance determining how much the observed nitrogen contents N_{ij} can vary around the critical nitrogen content N_{Ci} at the ith date. A high (low) value of τ_n^2 will reflect a strong (weak) variability of the measurements N_{ij} around the critical nitrogen content N_{Ci}. Part of this variability reflects the different levels of applied N considered in the field experiments and another part reflects measurement errors. With Eq. (2), the values of N_{Ci} are constrained to remain in the same order of magnitude as the measured nitrogen contents. This prevents the model from producing N_{Ci} values inconsistent with observed values. However, Eq. (2) might not offer enough flexibility as it assumes that the mean value of N_{ij} is equal to N_{Ci} for every date *i*. For this reason, we consider a second model based on a more flexible equation defined by:

$$N_{ij} \sim \mathcal{N}(N_{Ci} + \theta_i, \tau_n^2) \quad (3)$$

where θ_i is the deviation between the mean value of N_{ij} and the critical nitrogen content N_{Ci}. Eq.(3) does not assume that the mean value of N_{ij} is equal to N_{Ci} for every measurement date. Indeed, with Eq.(3), the mean value of N_{ij} is not strictly equal to N_{Ci} but to $N_{Ci} + \theta_i$, and each measurement date is characterized by a specific value of θ_i. In the remaining part of the text, models 1 and 2 refer to models based on Eqs. (2) and (3), respectively.

2.2.2. Level 2: Variability of biomass and nitrogen content between measurement dates

We assume that the critical nitrogen content N_{Ci} is related to the maximum biomass B_{MAXi} (i.e., to the mean value of W_{ij} when $N_{ij} = N_{Ci}$) by a power function defined as

$$N_{Ci} = A_1 B_{MAX_i}{}^{-A_2} \quad (4)$$

where A_1 and A_2 are two parameters. The values of B_{MAX_i} and of S_i are assumed to vary across dates according to two truncated Gaussian distributions defined by $B_{MAX_i} \sim \mathcal{N}(\mu_{BMAX}, \sigma^2_{BMAX}) I(0,)$ and $S_i \sim \mathcal{N}(\mu_S, \sigma_S^2) I(0,)$, where $I(a,b)$ is a truncation operator forcing values to fall within the range defined by *a* and *b*, $I(0,)$ thus indicating that B_{MAX_i} and S_i are forced to be positive. The use of a truncation is logical here as B_{MAX_i} represents a biomass (and is thus positive) and as the effect of nitrogen on biomass (measured by S_i) is expected to be positive when $N_{ij} < N_{Ci}$. The distributions of B_{MAX_i} and S_i determine the variability of the shape of the linear-plus-plateau function across measurement dates. We assume that the values of θ_i in Eq.(3) vary across measurement date as $\theta_i \sim \mathcal{N}(0, \tau_\theta^2)$. Here, a truncation is not required as there is no reason for θ_i to be strictly positive.

2.2.3. Level 3: prior

The model defined above included eight unknown quantities, namely A_1, A_2, μ_{BMAX}, σ^2_{BMAX}, μ_S, σ_S^2, τ_n^2, and τ_b^2. Prior knowledge on plausible values for these parameters are defined by specifying prior probability distributions. Two types of priors are used successively here, namely (i) weakly-informative priors and (ii) informative priors based on probabilistic expert elicitation. These two types of priors are further denoted to as prior 1 and prior 2, respectively.

The weakly-informative priors (priors 1) are designed to provide only little information about the plausible values of the eight unknown quantities while reducing the chance to get unrealistic values. These priors are defined by $\mu_{BMAX} \sim \mathcal{N}(6,10)$, $\mu_S \sim \mathcal{N}(0,10)$, $A_1 \sim Unif(2,6)$, $A_2 \sim Unif(0,0.5)$, $1/\sigma^2_{BMAX} \sim Gamma(0.001,0.001)$, $1/\sigma^2_B \sim Gamma(0.001,0.001)$, $1/\tau^2_b \sim Gamma(0.001,0.001)$, $1/\tau^2_n \sim Gamma(0.001,0.001)$, $1/\tau^2_\sigma \sim Gamma(0.001,0.001)$.

The informative priors (priors 2) are specified by expert elicitation. Probabilistic expert elicitation consists of extracting an expert's knowledge about the likely values of some unknown quantity of interest, and representing those beliefs with a probability distribution (Morris et al., 2014; Chen et al., 2019). Here, one expert with a thorough and international experience on critical N curves was elicited about the possible values of A_1, A_2, and μ_{BMAX}. The elicitation was conducted following the procedure described in details by Chen et al. (2019). As maize is a C4 crop, the expert chose to define two distributions for A_1, one for maize and one for the two C3 crop species considered (wheat and rice). For wheat and rice, the elicited prior distributions defined by the expert are $\mu_{BMAX} \sim Beta(2.31,\ 2.31,\ 1{,}15)$, $A_1 \sim \mathcal{N}(4.89,0.13)I(4{,}5.5)$, $A_2 \sim Beta(2.12{,}2.12{,}0.3{,}0.4)$. For maize, the prior for A_1 is $A_1 \sim Beta(2.03{,}1.5,\ 3{,}4)$, but the other priors are unchanged. Note that the last two parameters of the Beta distributions correspond to lower and upper bounds.

The two types of priors are shown in Fig. 2 for the parameters of A_1, A_2, μ_{BMAX}, and μ_S. Clearly, prior 2 covers narrower ranges of values than prior 1, in coherence with the fact that prior 2 is designed to be more informative than prior 1. The difference between the two types of prior is particularly strong for the two parameters of the critical N curve, i.e., A_1 and A_2 (Fig. 2AB). Noticeably, there is a marked difference between the priors of A_1 defined by the expert for maize vs. wheat + rice. The prior defined for maize covers lower values than the prior defined for wheat and rice (see the dotted vs. dashed lines in Fig. 2A).

2.3. Posterior distributions

The posterior distributions of the model parameters are estimated with a Markov chain Monte Carlo algorithm (MCMC) implemented with the R package rjags (Plummer, 2017) using both types of priors, successively. The R code used to fit models 1 and 2 is presented in Appendix D. With model 1 the convergence was achieved approximately after 10,000 iterations according to the Gelman-Rubin diagnosis. The first 10,000 iterations were discarded and the MCMC algorithm was run for 40,000 additional iterations which were then used to compute the median and 95 % credibility intervals for several quantities of interest, in particular A_1, A_2, μ_{BMAX}, μ_S, B_{MAXi} and S_i for all observation dates. The median values of B_{MAX_i} and S_i were used to fit a specific linear-plus-plateau function for each date separately. We also computed the median and 95 % credibility intervals of the critical nitrogen dilution curve $N_C = A_1 B_{MAX}{}^{-A_2}$. With model 2 (based on Eq.3), the convergence was achieved approximately after 10,000–50,000 iterations for the maize and wheat datasets. However, with model 2, we were not able to achieve convergence for rice, even with a large number of iterations, probably due to the fact that model 2 was overparametrized for rice. The results obtained with model 2 for maize and wheat were almost identical as those obtained with model 1 (see appendix C). For this reason, we only present the results obtained with model 1 in the next section. Nonetheless, values of A1 and A2 estimated with model 2 can be found in Appendix C.

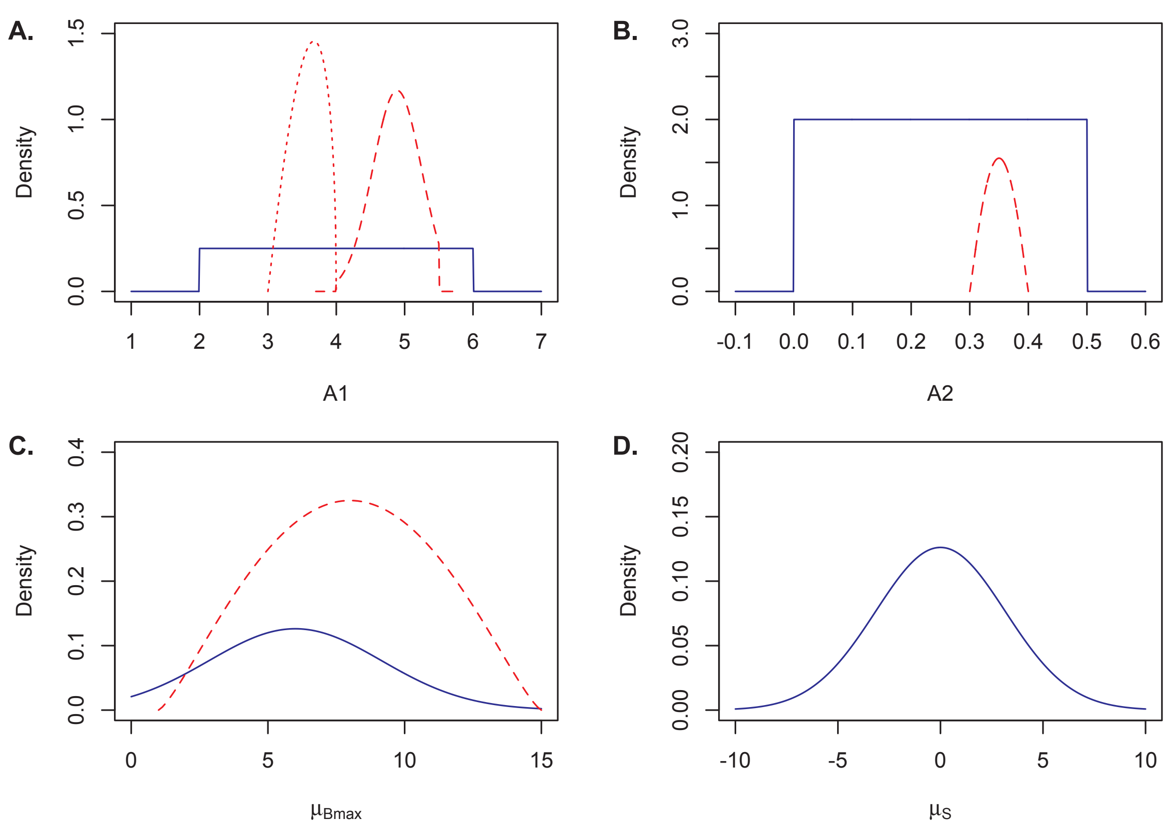


Fig. 2. Priors distributions defined for the parameters A_1 (A), A_2 (B), μ_{BMAX} (C), and μ_S (D). Weakly-informative priors are represented by the continuous blue lines. Informative priors defined by expert elicitation for A_1, A_2, and μ_{BMAX} are indicated in red dashed lines or in red dotted lines. For A_1, the informative prior is defined in dashed line for wheat and rice and in dotted line for maize. For A_2 and μ_{BMAX}, the informative priors are not differentiated among the three species by the expert and are all represented in dashed lines. (For interpretation of the references to colour in the Figure, the reader is referred to the web version of this article).

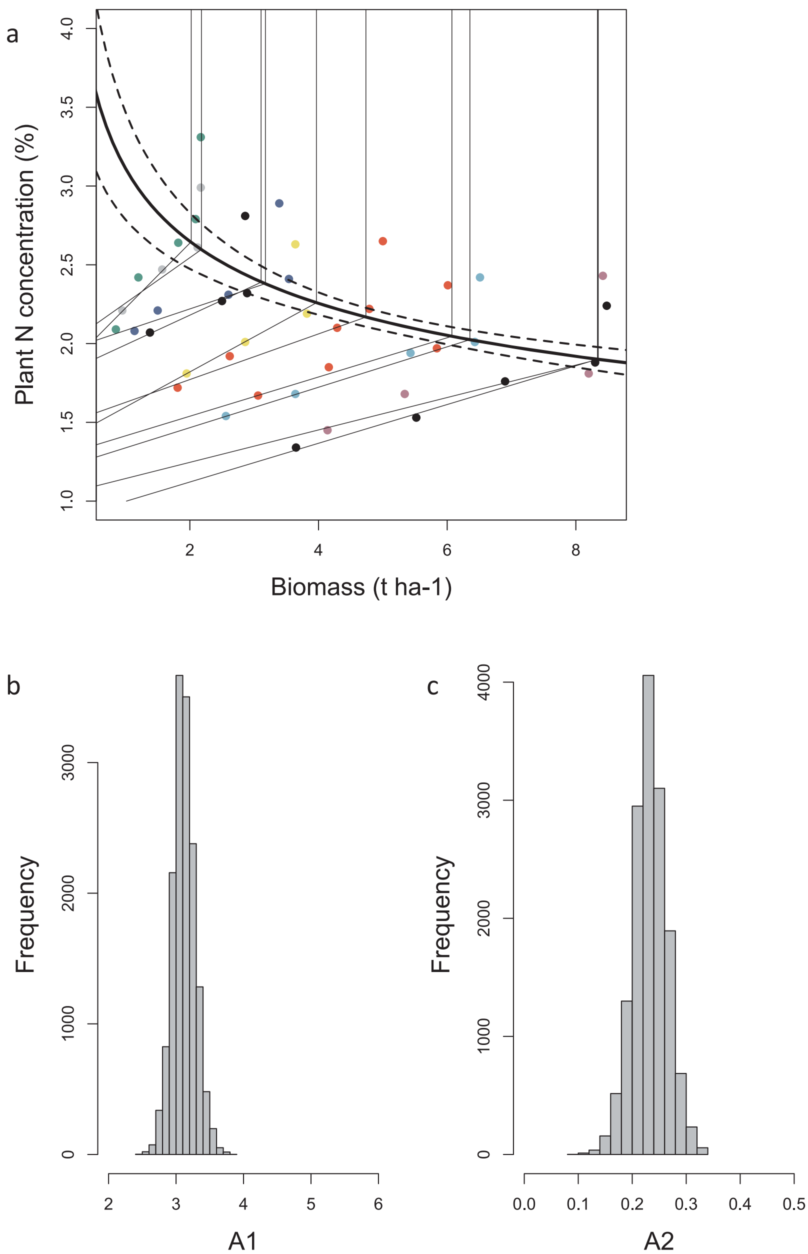


Fig. 3. Fitted critical N curve obtained for the maize cultivar DH605 and its 95 % credibility interval. In graphic 3a, continuous and dashed thick lines represent the posterior median and the 95 % credibility interval, respectively. The thin lines represent the linear-plus-plateau responses fitted for all combinations of measurement stage and year available in the dataset. Data collected for different stage*years are indicated by points of different colors. Posterior distributions of the parameters A1 and A2 (40,000 parameter values generated by MCMC) are presented in the histograms in graphics 3b and 3c, respectively. Results were obtained with prior 1. The ranges of the x-axis in b and c reflect the ranges of values covered by the prior distributions.

3. Results

3.1. Fitted critical N curves

The fitted critical N curve and its 95 % critical interval obtained for the maize cultivar DH605 with prior 1 are shown in Fig. 3a. The width of the critical interval describes the level of uncertainty in the fitted curve and directly reflects the distributions of the values of the parameters A1 and A2 (Fig. 3bc). Clearly, for the maize cultivar DH605, the level of uncertainty in critical N depends on the biomass value. The width of the critical interval is equal to about 0.5 % of plant N concentration when the biomass is lower than 2 t ha^{-1} and becomes lower than 0.25 % when the biomass is higher than 4 t ha^{-1}. The level of uncertainty is thus lower for high compared to low biomass values. This decreasing trend in the level of uncertainty is also observed for the second maize cultivar considered in this study (Appendix A1), but not for wheat or rice (Appendix A2-A4). For the latter two crop species, the uncertainty is small compared to maize, even for low biomass values, and its level does not show any substantial increasing or decreasing trend. The uncertainty is especially low for wheat for which the width of the 95 % credibility interval is close to 0.1 % over a wide range of biomass values. For rice, the level of uncertainty is intermediate between maize and wheat (Appendix A3-A4).

The whole set of fitted critical N curves obtained for the different crops and cultivars are compared in Fig. 4. With prior 1 (Fig. 4a), the critical N curve obtained for wheat is higher than those obtained for the other crop species when the biomass is lower than 4 t ha^{-1}, but it becomes lower than the critical N curve obtained for rice indica for higher values of biomass. The critical curve of indica is higher than the curve of japonica but the difference is very small for low biomass values (Fig. 4a). The curves obtained for the two maize cultivars are very similar and close to the curve of indica (Fig. 4a).

The curve obtained for wheat with prior 2 (Fig. 4b) is very similar to the curve obtained with prior 1. For wheat, the fitted critical N curve is thus relatively insensitive to the choice of prior. On the contrary, the curves obtained for rice indica and japonica are much higher with prior 2 compared to prior 1. With prior 2, the critical N curves of rice become very close to the critical N curve of wheat (Fig. 4b). For maize,

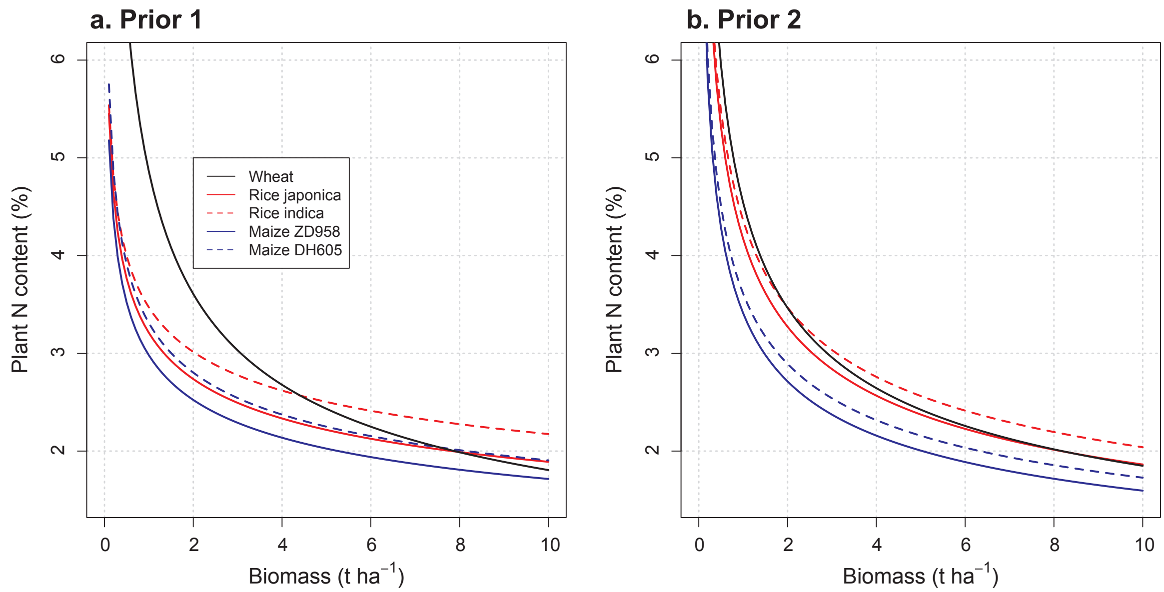


Fig. 4. Fitted critical N curve obtained for wheat, maize (two cultivars), and rice (two cultivars). Each curve corresponds to a posterior median obtained with prior 1 (a) or prior 2 (b).

compared to prior 1, the critical N curves obtained with prior 2 tend to be slightly higher and closer to the curve obtained for wheat.

Overall, the differences between the critical N curves obtained for the different species and cultivars are thus relatively small. These differences may simply reflect uncertainties in the values of the parameters A1 and A2 characterizing the critical N curves. In order to conclude, it is thus necessary to analyze the distribution of the parameter values. This is done in the next section.

3.2. Estimated parameter values

Fig. 5 shows the estimated parameter values obtained for the different crops and cultivars. The posterior medians correspond to point estimates that can be used to draw critical N curves, as shown in Figs. 3a and 4 ab. The 95 % credibility intervals presented in Fig. 5 cover 95 % of the values sampled in the posterior distributions using MCMC. Thus, the intervals shown in Fig. 5 for maize DH605 cover 95 % of the values of the histograms presented in Fig. 3bc. These intervals describe the levels of uncertainty in the values of the parameters A1 and A2.

With prior 1 (Fig. 5ac), the parameter values are significantly higher for wheat than for the other crop species. For wheat, the posterior median is equal to 4.86 (95 %CI = [4.61, 5.12]) for A1 and to 0.43 (95 %CI = [0.40, 0.46]) for A2. In comparison, for maize and rice, the

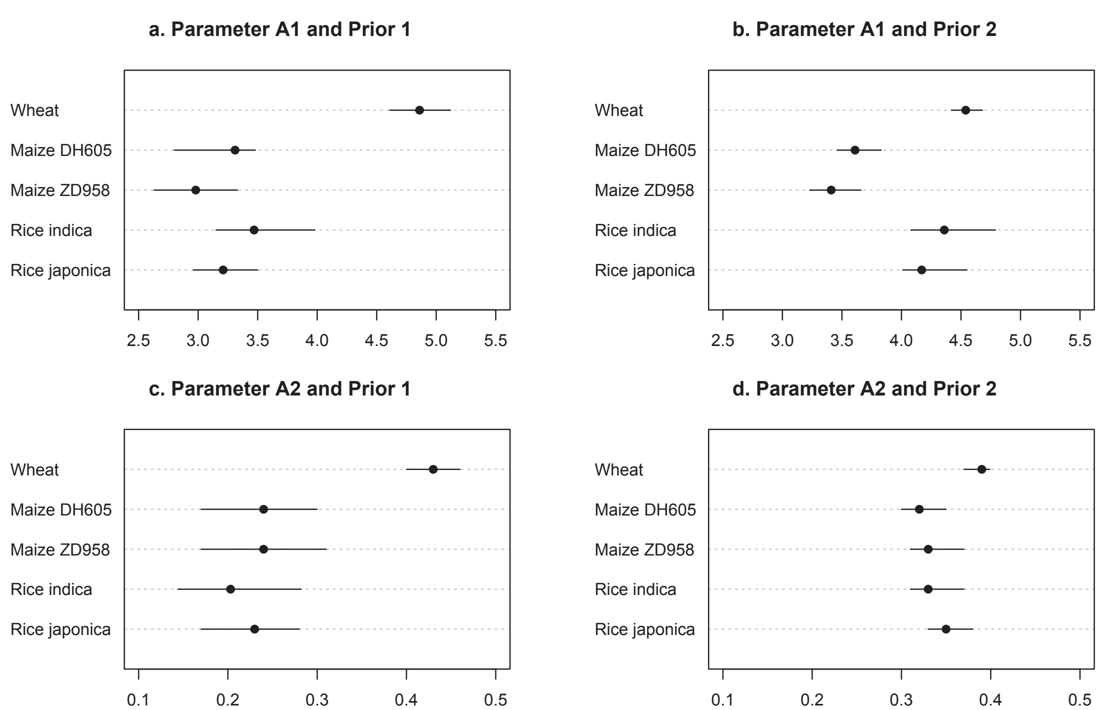


Fig. 5. Estimated values (posterior medians) and 95 % credibility intervals of parameters A1 (a, b, in %) and A2 (c, d) with priors 1 (a, c) and 2 (b, d), for wheat, maize (two cultivars), rice indica, and rice japonica.

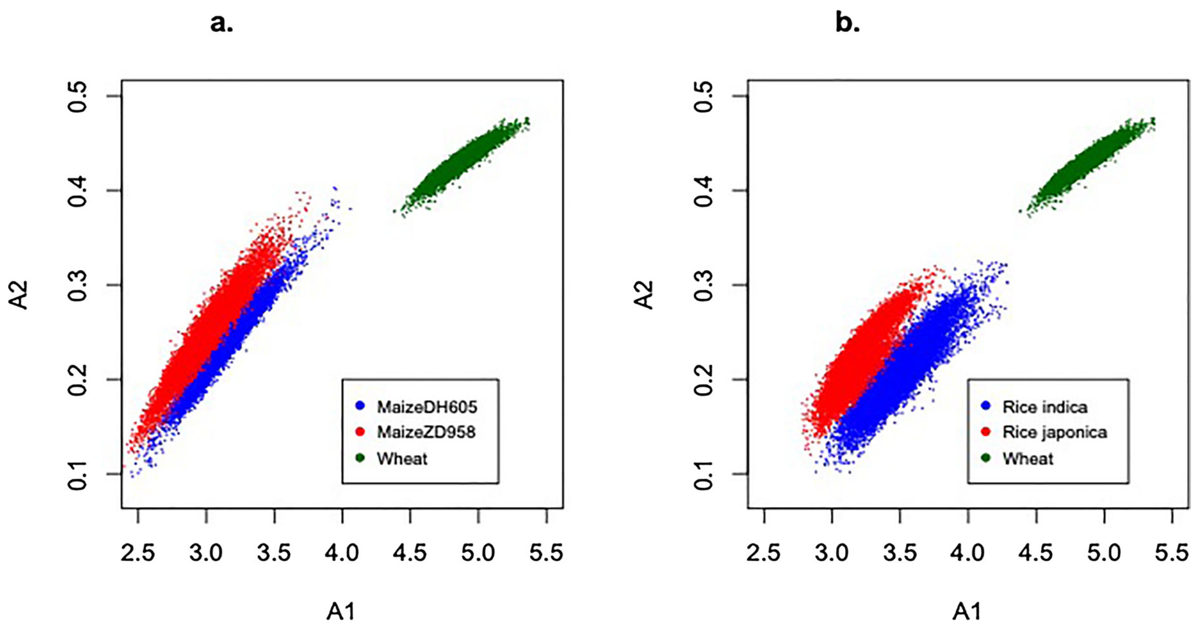


Fig. 6. Ensembles of parameter values drawn from the posterior distributions of A1 and A2 (obtained with model 1 and prior 1). Maize vs. Wheat (a) and Rice vs. Wheat (b).

posterior medians of A1 are always lower than 3.5 and those of A2 are always lower than 0.25. The uncertainty levels (reflected by the sizes of the credibility intervals) are substantially higher for maize and rice than for wheat. This is because the number of measurements used to fit the critical N curve is higher for wheat compared to maize and rice (see also Appendix A). Because of their large credibility intervals, the parameter values obtained for the two maize cultivars and for the two rice cultivars are not significantly different.

The whole ensembles of parameter values drawn by MCMC from the posterior distributions of A1 and A2 obtained with prior 1 are shown in Fig. 6. These values are those used to derived the posterior medians and credibility intervals presented in Figure 5ac. Parameter values presented in Fig. 6 show that A1 and A2 are positively correlated; high (low) values of A1 tend to be associated with high (low) values of A2. Fig. 6 also confirms that the parameter values obtained for the two cultivars of maize strongly overlap, that those obtained for the two rice cultivars partly overlap, and that the parameter values obtained for wheat are much higher than those obtained for maize and rice.

The use of prior 2 (informative priors) instead of prior 1 does not substantially impact the parameter values estimated for wheat (Fig. 5bd). For this crop, both priors lead to similar values for A1 and A2. On the contrary, for rice, the estimated parameter values are much higher with prior 2 than with prior 1. Consequently, the parameter values obtained for rice with prior 2 are not significantly different from those obtained for wheat anymore. This is because the informative priors (prior 2) defined for rice and wheat force the values of A1 to be higher than 4 and the values of A2 to be higher than 0.3 (Fig. 2ab). For wheat, prior 2 has no substantial impact because the posterior medians of A1 and A2 obtained with prior 1 were already higher than 4 and 0.3, respectively. For rice, the use of prior 2 has a strong impact on the estimated parameter values because the posterior medians of A1 and A2 obtained with prior 1 were much lower than 4 and 0.3.

For maize, the use of prior 2 has also a strong impact on the estimated values of A2 (Fig. 5d) but less on the values of A1. This is because the informative prior defined for A1 (prior 2) is different for maize than for rice and maize (Fig. 2a). For maize, the informative prior defined for A1 forces the values of this parameter to fall within the range 3–4, i.e., close to the posterior medians of A1 obtained with prior 1. The use of prior 2 instead of prior 1 has thus a limited impact on A1 for maize.

4. Discussion

The proposed Bayesian method has several advantages. First, it does not require any preliminary classification of N-limited data against non-N-limited data and does not necessitate the preliminary identification of critical N concentration values. This is an important advantage because it was shown that the critical N curve is sensitive to the method used for estimating critical N concentration values (i.e., type of statistical test and type I error level) and because there is no consensus on how these critical N concentration values should be estimated (Greenwood et al. 1990; Justes et al., 1994; Chen and Zhu, 2013). Our Bayesian approach can be implemented even with sparse data, i.e., when observations are available for a limited number of fertilizer doses only. This is made possible because the proposed model borrows strength from all dates of measurements and because parameters are estimated by combining data with prior information.

Second, the proposed method can be easily implemented with free software to fit critical nitrogen curves. The R code presented in appendix D can be easily run to estimate critical N dilution curves using a dataset including only three columns; a column with the biomass observations, a column with the associated nitrogen concentration observations, and a column with the indices identifying the different dates of the dataset. This R code produces chains of values for all parameters, including the two parameters defining the critical N curve. The generated chains of values can be easily summarized by standard quantities such as median, mean, and percentiles, and can be used to estimate the N critical curve and compute its credibility interval. In our applications, 10,000 iterations were sufficient to reach convergence in most cases and the computation time did not exceed 1 or 2 min using a standard commercial computer.

Finally, another advantage of the proposed approach is that it facilitates the analysis of the uncertainty of the fitted critical N curves. The proposed method is based on a Bayesian hierarchical model whose parameters are estimated taking into account the number of observation dates included in the data set, the number of data available per date and the variability of observations between and within dates. The estimation results are expressed by probability distributions from which the uncertainty of any quantity of interest can easily be analyzed. In particular, our approach allows us to calculate the credibility intervals of the N critical curves and their parameters. Our applications show that, because the dataset used for wheat has been constituted from a large number of experiments network across France, the uncertainty of critical N curve for wheat is relatively low. For maize and rice in China, where the numbers of available data are smaller, the widths of the credibility intervals are larger revealing a higher level of uncertainty.

From the probability distributions and credibility intervals computed by our method, it is possible to compare different species of crops, cultivars or cropping systems, taking into account uncertainties in parameter estimates. Since critical nitrogen curves are often included in mechanistic crop models, the probability distributions provided by our method could also be useful for performing uncertainty and sensitivity analyses with these models (Wallach et al., 2019). The importance of rigorous uncertainty analysis is illustrated by some of the results of our case studies where we found that parameter estimates for different cultivars of the same species could not be considered statistically significant when uncertainty is taken into account. Clearly, in our examples, the differences between the point estimates of the critical N curve parameters obtained for the different cultivars are small compared to the associated levels of uncertainty.

Like all Bayesian methods, our approach allows modellers to combine two sources of information to estimate the parameters. More specifically, it combines prior information based on expert knowledge and experimental data. Prior information is described using probability distributions that summarize the initial state of knowledge on parameter values before using the data. Here, we use two types of priors. The former are poorly informative and are designed to provide little information on plausible values of model parameters. They do not therefore strongly constrain the values of the parameters. The second priors are more informative and are specified by the probabilistic elicitation of an expert. This technique allows to represent the expert's knowledge on the value of a parameter through a probability distribution.

Probabilistic elicitation is relevant when you want to rely on both experimental data and expert knowledge for parameter estimation. The use of informative prior is useful when the number of observations available is low and insufficient to accurately estimate parameter values. However, this approach should be used with caution, as it can have a significant effect on parameter estimates, especially when the size of the data set is small. In our applications, the parameter estimates obtained for wheat are not substantially influenced by the choice of prior because the size of the dataset is relatively large in this case. For wheat, both priors lead to similar point estimates and credibility intervals. In contrast, for maize and rice, the sizes of the data sets are smaller and, in both cases, the values of the estimated parameters are sensitive to the choice of a prior; they take on larger values and their credibility intervals are narrower when calculations are made with informative a priori. When used, informative priors should therefore be defined by using qualified experts based on reliable information.

We believe that our approach opens new perspectives for the estimation of critical nitrogen dilution curves. Different variants of the model proposed here could be tested in the future in order to better take into account possible correlations between measurements, to better describe the a priori information available on the values of the parameters, or to handle larger networks of experiments.

CreditAuthorsStatement

DM: Study design, statistical analysis, writing of the paper.
GL: Study design, review of the paper.
BZ: Data collection, review of the paper.
STA: Data collection, review of the paper.

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.

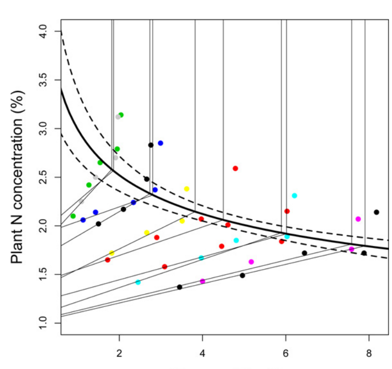
Acknowledgments

This work is part of the project Licite (Institut Carnot Plant2Pro) and of the project CLAND (16-CONV-0003).

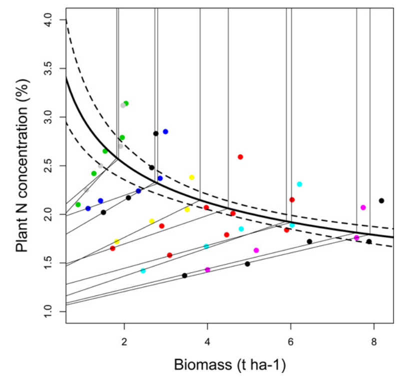
Appendix A. Individual fitted critical N curves

In this appendix, we present the posterior median, the 95 % credibility intervals and the experimental data obtained for maize ZD958 m (see main text for the other maize cultivar), rice (Japonica and Indica) and wheat.

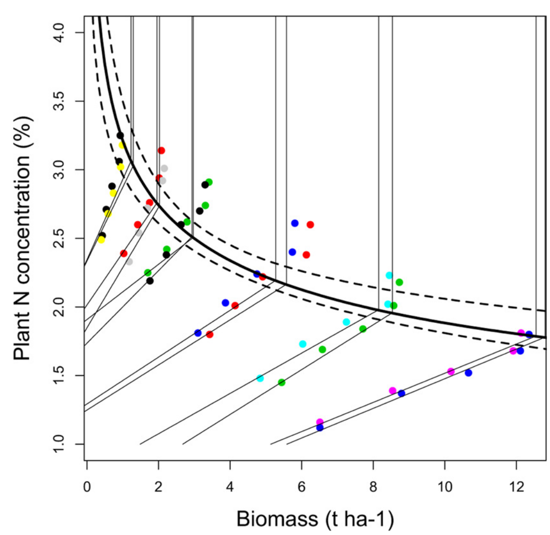
A1. Maize ZD958m



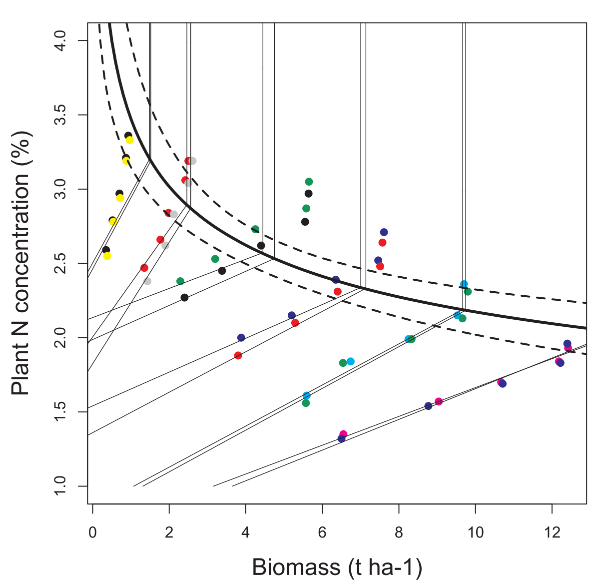
A2. Wheat



A3. Rice Japonica



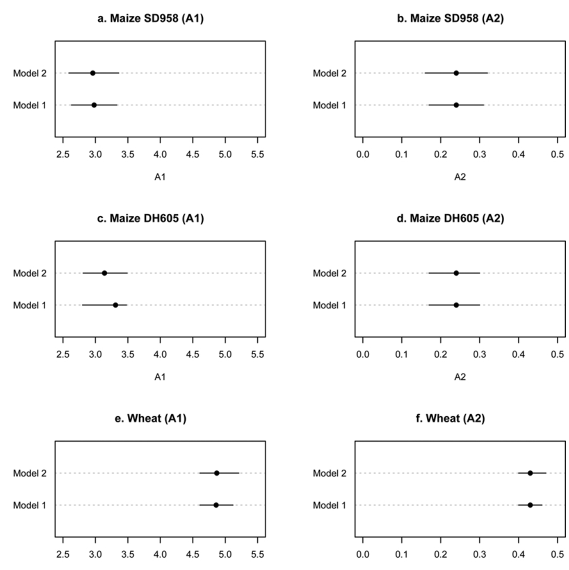
A4. Rice Indica



Appendix B. Characteristics of the maize experiments

Experiment No.	Sowing/Harvesting date	Soil characteristics	Cultivar	N(kg N ha – 1)	Sampling stage
Experiment 1	8-Jun	Type: light loam soil	Zhengdan958	0 (N0)	Elongation stage
(2015 Xinxiang)	25-Sep	Organic matter: 12.26 g kg – 1	(ZD958)	75 (N1)	Bell stage
		Total N: 0.74 g kg – 1		150 (N2)	Tasseling stage
		Olsen-P: 35.67 mg kg – 1		225 (N3)	Anthesis stage
		NH4oAc-K+: 84 mg kg – 1		300 (N4)	Silking stage
Experiment 2	8-Jun	Type: sandy light loam soil	Denghai605	0 (N0)	Elongation stage
(2015 Xinxiang)	25-Sep	Organic matter: 10.43 g kg – 1	(DH605)	75 (N1)	Bell stage
		Total N: 0.61 g kg – 1		150 (N2)	Tasseling stage
		Olsen-P: 33.94 mg kg – 1		225 (N3)	Anthesis stage
		NH4oAc-K+: 76 mg kg – 1		300 (N4)	Silking stage
Experiment 3	6-Jun	Type: light loam soil	Zhengdan958	0 (N0)	Elongation stage
(2016 Xinxiang)	22-Sep	Organic matter: 14.2 g kg – 1	(ZD958)	90 (N1)	Bell stage
		Total N: 0.83 g kg – 1		180 (N2)	Tasseling stage
		Olsen-P: 44 mg kg – 1		270 (N3)	Anthesis stage
		NH4oAc-K+: 90 mg kg – 1			Silking stage
Experiment 4	6-Jun	Type: light loam soil	Denghai605	0 (N0)	Elongation stage
(2016 Xinxiang)	22-Sep	Organic matter: 9.5 g kg – 1	(DH605)	90 (N1)	Bell stage
		Total N: 0.57 g kg – 1		180 (N2)	Tasseling stage
		Olsen-P: 23.51 mg kg – 1		270 (N3)	Anthesis stage
		NH4oAc-K+: 58.45 mg kg – 1			Silking stage

Appendix C. Estimated values of A1 and A2 for maize and wheat obtained with model 1 (based on Eq.(2)) and model 2 (based on Eq.(3))



Appendix D. R Code

Model 1

Model 1

```
###Data

# Q=total number of biomass observations

# K=number of dates

# W=column of data including biomass observations

# N=column of data including observations of nitrogen concentrations

# Date=column with the indices identifying the different dates of the dataset


###Model parameters

# Nc=Critical nitrogen concentration

# Bmax=maximum biomass value in a specific date

# S=slope of the linear-plus-plateau function

# W=biomass increase per unit of nitrogen concentration

# A1 and A2 = parameters of the critical N curve

# tau_b and tau_n = 1/residual variances for biomass and nitrogen content observations

# Mu_Bmax,Prec_Bmax = parameters defining the between-date variability of Bmax

# Mu_S,Prec_S = parameters defining the between-date variability of S


Q<-length(Date)
K<-length(unique(Date))

modelstring= "
model {

 for (i in 1:Q)
 {
        W[i]~dnorm(mu[i], tau_b)
```

Model 2

```
        N[i]~dnorm(Nc[Date[i]], tau_n)
        mu[i]<-min(Bmax[Date[i]], Bmax[Date[i]]+S[Date[i]]*(N[i]-Nc[Date[i]]))
 }

 for (j in 1:K)
 {
        Nc[j]=A1*Bmax[j]^(-A2)
        Bmax[j]~dnorm(Mu_Bmax,Prec_Bmax)T(0,)
        S[j]~dnorm(Mu_S,Prec_S)T(0,)
                }

                #Weakly informative
                Mu_Bmax~dnorm(6,0.1)
                Mu_S~dnorm(0,0.1)
                A1~dunif(2,6)
                A2~dunif(0,0.5)

                #Informative prior C3
                #A1~dnorm(4.89,7.72)T(4,5.5)
                #ZA2~dbeta(2.12,2.12)
                #A2=(0.4-0.3)*ZA2+0.3
                #ZMu_Bmax~dbeta(2.31,2.31)
                #Mu_Bmax=(15-1)*ZMu_Bmax+1
                #Mu_S~dnorm(0,0.1)

                #Informative prior C4
                #ZA1~dbeta(2.03,1.5)
                #A1=(4-3)*ZA1+3
                #ZA2~dbeta(2.12,2.12)
                #A2=(0.4-0.3)*ZA2+0.3
                #ZMu_Bmax~dbeta(2.31,2.31)
                #Mu_Bmax=(15-1)*ZMu_Bmax+1
                #Mu_S~dnorm(0,0.1)

                Prec_Bmax~dgamma(0.001,0.001)
                Prec_S~dgamma(0.001,0.001)
                tau_b~dgamma(0.001,0.001)
                tau_n~dgamma(0.001,0.001)

}
"
writeLines(modelstring, con="model.txt")

model<-jags.model('model.txt', data=list('W'=W, 'N'=N, 'Date'=Date, 'Q'=Q,'K'=K),
 n.chains=3, n.adapt=10000)
```

Model 2

```
modelstring= "
model {

 for (i in 1:Q)
 {
        W[i]~dnorm(mu[i], tau_b)
        N[i]~dnorm(Nc[Date[i]]+Theta[Date[i]], tau_n)
        mu[i]<-min(Bmax[Date[i]], Bmax[Date[i]]+S[Date[i]]*(N[i]-Nc[Date[i]]))
 }

 for (j in 1:K)
 {
        Nc[j]=A1*Bmax[j]^(-A2)
        Bmax[j]~dnorm(Mu_Bmax,Prec_Bmax)T(0,)
        S[j]~dnorm(Mu_S,Prec_S)T(0,)
        Theta[j]~dnorm(0,tau_t)
                }

                #Weakly informative
                Mu_Bmax~dnorm(6,0.1)
                Mu_S~dnorm(0,0.1)
                A1~dunif(2,6)
                A2~dunif(0,0.5)

                #Informative prior C3
                #A1~dnorm(4.89,7.72)T(4,5.5)
                #ZA2~dbeta(2.12,2.12)
                #A2=(0.4-0.3)*ZA2+0.3
                #ZMu_Bmax~dbeta(2.31,2.31)
                #Mu_Bmax=(15-1)*ZMu_Bmax+1
                #Mu_S~dnorm(0,0.1)
```

References

Ata-Ul-Karim, S.T., Zhu, Yan, Liu, Xiaojun, Cao, Qiang, Tian, Yongchao, Cao, Weixing, 2017. Comparison of different critical nitrogen dilution curves for nitrogen diagnosis in rice. Sci. Rep. 7, 42679. https://doi.org/10.1038/srep42679.

Chen, M., Brun, F., Raynal, M., Debord, C., Makowski, D., 2019. Use of probabilistic expert elicitation for assessing risk of appearance of grape downy mildew. Crop. Prot. 126, 104926.

Chen, P., Zhu, Y., 2013. A new method for winter wheat critical nitrogen curve determination. Agron. J. 105, 1839–1846.

Colnenne, C., Meynard, J.-M., Reau, R., 1998. Determination of a critical nitrogen dilution curve for winter oilseed rape. Ann. Bot. 81, 311–317. https://doi.org/10.1006/anbo.1997.0557.

Ioannidis, J.P., 2005. Why most published research findings are false. PLoS Med. 2, e124.

Justes, E., Mary, B., Machet, J.M., 1994. Determination of a critical nitrogen dilution curve for winter wheat crops. Ann. Bot. 74 (397–407), 1. https://doi.org/10.1006/anbo.1994.1133.

Lemaire, G., Jeuffroy, M.-H., Gastal, F., 2008. Diagnosis tool for plant and crop N status in vegetative stage: theory and practices for crop N management. Eur. J. Agron. 28, 614–624. https://doi.org/10.1016/j.eja.2008.01.005.

Morris, D.E., Oakley, J.E., Crowe, J.A., 2014. A web-based tool for eliciting probability distributions from experts. Environ. Model. Softw. 52, 1–4. https://doi.org/10.1016/j.envsoft.2013.10.010.

Philibert, A., Loyce, C., Makowski, D., 2012. Quantifying uncertainties in N2O emission due to N fertilizer application in cultivated areas. PLoS One 7 (11), e50950. https://doi.org/10.1371/journal.pone.0050950.

Plénet, D., Lemaire, G., 2000. Relationships between dynamics of nitrogen uptake and dry matter uptake in maize crops: determination of critical N concentration. Plant Soil 216, 65–82. https://doi.org/10.1023/A:1004783431055.

Plummer, M., 2017. JAGS User Manual.

Ramanantenasoa, M.M.J., Génermont, S., Gilliot, J.-M., Makowski, D., 2019. Meta-modeling methods for estimating ammonia volatilization from nitrogen fertilizer and manure applications. J. Environ. Manage. 236, 195–205. https://doi.org/10.1016/j.jenvman.2019.01.066.

Sandana, P., Lobos, I.A., Pavez, P.B., Moscoso, C.J., 2019. Validation of a critical nitrogen dilution curve for hybrid ryegrass. Grass Forage Sci. 2019, 1–11.

Wallach, D., Makowski, D., Jones, J.W., Brun, F., 2019. Working With Dynamic Crop Models, third edition. Elsevier Academic Press.

Zhao, B.Z., Zhang, J.B., Flury, M., Zhu, A.N., Jiang, Q.A., Bi, J.W., 2007. Groundwater contamination with NO3–N in a wheat–corn cropping system in the North China Plain. Pedosphere 17, 721–731. https://doi.org/10.1016/S1002-0160(07)60087-3.

Zhao, B., Ata-Ul-Karim, Syed Tahir, Duan, Aiwang, Liu, Zhandong, Wang, Xiaolong, Xiao, Junfu, Liu, Zugui, Qin, Anzhen, Ning, Dongfeng, Zhang, Weiqiang, Lian, Yanhao, 2018. Determination of critical nitrogen concentration and dilution curve based on leaf area index for summer maize. Field Crop Research 228, 195–203.