

Analyzing uncertainty in critical nitrogen dilution curves

David Makowski, Ben Zhao, Syed Tahir Ata-Ul-Karim, Gilles Lemaire

▶ To cite this version:

David Makowski, Ben Zhao, Syed Tahir Ata-Ul-Karim, Gilles Lemaire. Analyzing uncertainty in critical nitrogen dilution curves. European Journal of Agronomy, 2020, 118, pp.126076. 10.1016/j.eja.2020.126076 . hal-03117535

HAL Id: hal-03117535 https://hal.science/hal-03117535v1

Submitted on 14 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Analyzing uncertainty in critical nitrogen dilution curves

2	David Mal	kowski ^{1,2,*} , Ben Zhao ³ , Syed Tahir Ata-Ul-Karim ⁴ , Gilles Lemaire ⁵
3	1.	University Paris-Saclay, INRAE, AgroParisTech, UMR 211 78850 Thiverval-Grignon France
4	2.	CIRED, 45bis Avenue de la Belle Gabrielle, 94130 Nogent-sur-Marne France
5	3.	Key Laboratory of Crop Water Use and Regulation, Ministry of Agriculture, Farmland Irrigation
6		Research Institute, Chinese Academy of Agricultural Sciences, 380 Hongli road, Xinxiang, Henan
7		453003, PR China
8	4.	Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese
9		Academy of Sciences, 71 East Beijing Road, Nanjing, Jiangsu 210008, P. R. China
10	5.	Honorary Director of Research INRAE (retired), Lusignan, 86500, France.
11	* Correspo	onding author
12	david.mak	kowski@inrae.fr
13		
14	Highligh	nts
15	• We	e present a Bayesian method for estimating parameters of critical nitrogen dilution
16	cu	rve
17	• Ito	does not require the classification of nitrogen-limited against non-nitrogen-limited
18	da	ta
19	• Ito	can be easily implemented using freely-available software
20	• Th	is method is useful for analyzing uncertainty in the fitted critical nitrogen curves
21	• It o	can take prior knowledge about parameter values into account.
22		

23 Abstract

Nitrogen critical curves are frequently used to diagnose the N status of crops and grasslands. 24 25 They play an important role in plant modelling and are frequently used in fertilizer 26 management tools. During the last 20 years, a number of studies have been conducted for 27 comparing critical curves obtained in different conditions (e.g., different cultivars) and 28 understanding the origin of their difference. However, uncertainty in the determination of 29 coefficient of these curves is generally poorly analyzed in these studies, which increase the 30 risk of false conclusions, in particular on the existence of differences between species, 31 cultivars and cropping systems. Here, we present a Bayesian statistical model for estimating 32 parameters of critical nitrogen dilution curve from experimental data. Contrary to standard 33 methods commonly used for fitting critical nitrogen dilution curves, the proposed approach 34 allows one to fit these curves in only one step, i.e., directly from the original biomass and 35 nitrogen content measurements. Specifically, this method does not require the classification 36 of nitrogen-limited data against non-nitrogen-limited data and does not necessitate the 37 preliminary identification of critical nitrogen concentrations. Another advantage of the proposed method is that it can be easily implemented using freely-available software. We 38 illustrate its practical interest using experimental data collected for winter wheat in France, 39 40 and for maize and rice in China. We show that this method is useful for analyzing uncertainty 41 in the fitted critical nitrogen curves and for comparing several curves obtained for different 42 crop species and cultivars. The proposed method is based on the specification of prior 43 probability distributions defining plausible ranges of values for the critical curve parameters, 44 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. 45

46 Keywords: Bayesian statistics, critical concentration, critical N curve, nitrogen, probabilistic
47 expert elicitation, uncertainty.

48

49 1. Introduction

Nitrogen (N) fertilization plays a major role in agricultural production but excess of N in agroecosystems has negative impacts on water pollution (Zhao et al. 2007), and increases risk of ammonia and N₂O 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.

57 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 58 59 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. 60 Although several variants of this model exist, the most common model is expressed as $N_c =$ 61 $A_1W^{-A_2}$ where N_c is the critical N concentration for biomass W, and where A_1 and A_2 are 62 two parameters that are estimated by fitting the model to a set of experimental data. Critical 63 64 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 65 66 and Zhu, 2013), oilseed rape (Colnenne et al., 1998), maize (Plénet and Lemaire, 2000), 67 ryegrass (Sandana et al. 2019), rice (Ata-Ul-Karim et al. 2017) among others.

68 It is essential to analyze uncertainty in fitted critical N curves in a rigorous manner. This is 69 important for assessing risk of N deficit or N excess and, also, for comparing critical curves 70 obtained in different conditions (e.g., sites, years, different cultivars and/or crop 71 managements) and understanding the origin of their difference. In some cases, two fitted 72 curves obtained in two different conditions are apparently different, but their difference 73 does not reflect a real difference due to the existence of a true effect of the factor studied, 74 but simply reflects errors in the estimated values of their parameters. In such cases, 75 concluding that there is a real difference and a real effect of the factor under consideration 76 will lead to a false conclusion. There is increasing concern that false findings may be very 77 frequent in research (Ioannidis, 2005). To limit risk of false discoveries, a rigorous analysis of 78 uncertainty is then crucial.

79 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 80 81 growing period for different N levels. Several methods have been proposed to estimate 82 parameters of critical curves from this type of data (Greenwood et al. 1990; Justes et al., 83 1994; Chen and Zhu, 2013). All these methods require a classification of N-limited data vs. 84 non-N-limited data at each date of measurement. The two groups of data (N-limited vs. non-85 N-limited) are then used to identify a so-called critical-N-concentration above which biomass 86 is assumed to reach its maximum value. A critical N curve is finally fitted to the series of 87 critical-N-concentration values obtained across the different dates of measurement. 88 Although this approach has been successful in fitting critical N curves in a great diversity of 89 contexts, it presents several limitations. This approach requires the definition of one 90 classification rule to distinguish N-limited vs. non-N-limited data and of a method to define 91 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.

107 2. Materials and Method

108 2.1. Data

109 2.1.1. Wheat

110 Data used in this study for wheat are those used by Justes et al. (1994). These data had been 111 extracted from a large experimental network with different year-location combinations in 112 France. In each experiment wheat crops were grown with at least four different levels of N 113 fertilizer supply and several plant samples were regularly harvested all along the crop 114 growth period until anthesis. Plant samples were analyzed to determine their aboveground 115 biomass (t ha⁻¹) and plant shoot N concentration (%) at each sampling date. A series of pairs 116 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 117 dataset covers a large range of biomass values, from about 1 t ha⁻¹ at the early stages of 118 119 plant growth to about 12 t ha⁻¹ at plant anthesis. Crop were managed with ample P and K 120 supply, with irrigation when necessary for avoiding water stress, and with adequate plant 121 disease control. The total number of pairs of biomass and N content is equal to 73 (Table 1).

122 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 used to measure.
All field experiments were arranged in a randomized complete block design with three
replicates. The size of each plot was 60m² 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).

134 2.1.3. Rice

135 Four multi-N rates (0 to 360 kg N ha⁻¹) field experiments using one Indica (Jingliangyou-534) 136 and one Japonica rice (Jiahua-1) cultivar were conducted during 2017 and 2018 rice growing 137 seasons in east China. Experiments were arranged with a randomized complete block design 138 having three repeats. The size of every plot was 5 m × 5 m with the inter-row spacing of 30 139 cm. The planting density in all the experiments was approximately 22.2×10⁴ 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 140 141 before transplanting, 10% at active tillering, 20% at panicle initiation, and 30% at booting. In 142 all experiments, ample phosphate and potassium fertilizers were incorporated into the soil as 143 monocalcium phosphate (Ca(H₂PO₄)2) and potassium chloride (KCl) before transplantation. 144 Experiments were carried out according to local recommendations along with adequate plant 145 pest and disease control measures to ensure optimal production. Plant samples were regularly 146 harvested all along the crop growth period until heading (pre-anthesis growth period) for 147 determination of plant dry mass and plant N concentration to provide a set of plant dry mass-148 %N data across the different N supply rates at each sampling stage. The total number of pairs 149 of biomass (t ha⁻¹) and N content (%) is equal to 60 (Table 1).

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

Crop/Cultivar	Number of dates	Average number of N	Number of data
	of measurement	levels per date	
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

154

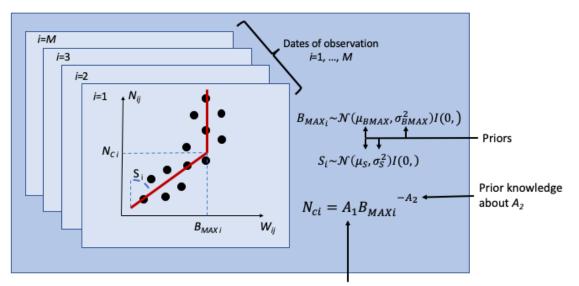
165

155 2.2. Model

156 Our model is a Bayesian hierarchical model including three levels (Figure 1). The basic 157 principle of this model is to consider that the response of biomass to nitrogen content 158 follows a linear-plus-plateau function, as commonly considered in many studies (see for 159 example Chen and Zhu, 2013 and Zhao et al. 2018). The variability of the parameters of the 160 linear-plus-plateau function are described by probability distributions estimated from the 161 whole set of available data using a Bayesian method. The parameters of the critical nitrogen 162 dilution curve are then derived directly from the fitted probability distributions. 163 The first level of the model describes the biomass response to nitrogen content for a given 164 date of measurement based on a linear-plus-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. The critical nitrogen dilution curve is computed from these
 parameters. The third level describes prior knowledge about parameter values.
- 170
- 171 Figure 1. Graphical description of the three-level hierarchical Bayesian model. Level 1 describes the
- 172 relationship between biomass and nitrogen content observations for each date of measurement (light blue).
- 173 Level 2 describes the critical nitrogen dilution curve and the variability of biomass and nitrogen content
- 174 between measurement dates using several probability distributions (dark blue). Level 3 describes prior
- 175 knowledge about parameter values (outside part of the graphic). The indices *i* and *j* correspond to the dates of
- 176 observation and the supplied N fertilizer rates, respectively. See text for details.



Prior knowledge about A1

177

178 Level 1: Biomass response to nitrogen content

- 179 This part of the model describes the relationship between biomass (W_{ij} , t ha⁻¹) and nitrogen
- 180 content (N_{ij} , %) measurements in the ith date of observation for the jth nitrogen dose. We

181 assume that W_{ij} is distributed according to a Gaussian distribution whose mean value is 182 specified through a linear-plus-plateau function of the nitrogen content as follows:

183

184
$$W_{ij} \sim \mathcal{N}(\mu_{ij}, \tau_b^2)$$
, with $\mu_{ij} = B_{MAX_i} + S_i(N_{ij} - N_{Ci})$ if $N_{ij} < N_{Ci}$

185 and
$$\mu_{ij} = B_{MAX_i}$$
 otherwise (1)

186

187 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 188 equal to the mean value of W_{ij} when $N_{ij} \ge N_{Ci}$), S_i is the slope of the linear part of the 189 function (i.e., the increase rate of biomass per unit increase of nitrogen content, t ha⁻¹ %N⁻¹), 190 N_{Ci} is the critical nitrogen content for the ith date (%), i.e., the value of nitrogen content (%) 191 above which B_{MAX_i} is reached, and τ_b^2 is the residual variance. We also assume that the 192 nitrogen content measurements (N_{ij}) is related to the critical nitrogen content (N_{Ci}) 193 according to the distribution

194

195
$$N_{ij} \sim \mathcal{N}(N_{Ci}, \tau_n^2),$$
 (2)

196

197 where τ_n^2 is a variance determining how much the observed nitrogen contents N_{ij} can vary 198 around the critical nitrogen content N_{Ci} in the ith date. In Eq.(2), nitrogen content 199 measurements are assumed to vary around the critical nitrogen with a variance equal to τ_n^2 . 200 A high (low) value τ_n^2 will reflect a strong (weak) variability of the measurements N_{ij} around 201 the critical nitrogen content N_{Ci} . Part of this variability reflects the different levels of applied 202 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 more flexible equation defined by:

$$208 \qquad N_{ij} \sim \mathcal{N}(N_{Ci} + \theta_i, \tau_n^2), \tag{3}$$

209 where θ_i is the deviation between the mean value of N_{ij} and the critical nitrogen content 210 N_{Ci} . Eq.(3) does not assume that the mean value of N_{ij} is equal to N_{Ci} for every 211 measurement date. Indeed, with Eq.(3), the mean value of N_{ij} is not strictly equal to N_{Ci} but 212 to $N_{Ci} + \theta_i$, and each measurement date is characterized by a specific value of θ_i . In the 213 remaining part of the text, models 1 and 2 refer to models based on Eqs.(2) and (3),

214 respectively.

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

216 We assume that the critical nitrogen content N_{Ci} is related to the maximum biomass B_{MAXi}

217 (i.e., to the mean value of W_{ij} when $N_{ij} = N_{Ci}$) by a power function defined as

218
$$N_{Ci} = A_1 B_{MAX_i}^{-A_2}$$
 (4)

219 where A_1 and A_2 are two parameters. The values of B_{MAX_i} and of S_i are assumed to vary

220 across dates according to two truncated Gaussian distributions defined by

221
$$B_{MAX_i} \sim \mathcal{N}(\mu_{BMAX}, \sigma_{BMAX}^2) I(0,)$$
 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*, *l*(0,) thus indicating that
- 223 B_{MAX_i} and S_i are forced to be positive. The use of a truncation is logical here as B_{MAX_i}
- 224 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.

230 Level 3: Prior

The model defined above included eight unknown quantities, namely A_1 , A_2 , μ_{BMAX} , σ_{BMAX}^2 ,

232 $\mu_s, \sigma_s^2, \tau_n^2$, and τ_b^2 . Prior knowledge on plausible values for these parameters are defined by

233 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,

236 respectively.

237 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 $\mu_{BMAX} \sim \mathcal{N}(6,10)$, $\mu_S \sim \mathcal{N}(0,10)$,

240 $A_1 \sim Unif(2,6), A_2 \sim Unif(0,0.5), 1/\sigma_{BMAX}^2 \sim Gamma(0.001,0.001), 1/$

241 $\sigma_B^2 \sim Gamma(0.001, 0.001), 1/\tau_b^2 \sim Gamma(0.001, 0.001), 1/\tau_n^2 \sim Gamma(0.001, 0.001),$

242 $1/\tau_{\theta}^2 \sim Gamma(0.001, 0.001).$

243 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

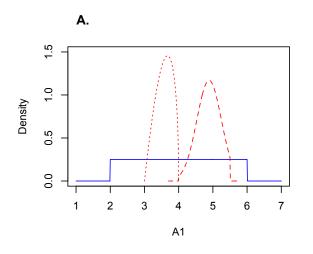
245 unknown quantity of interest, and representing those beliefs with a probability distribution

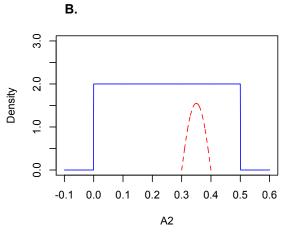
246 (Morris et al., 2014; Chen et al., 2019). Here, one expert with a thorough and international

experience on critical N curves was elicited about the possible values of A_1 , A_2 , and μ_{BMAX} .

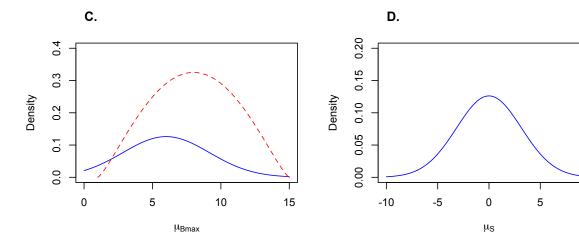
248	The elicitation was conducted following the procedure described in details by Chen et al.
249	(2019). As maize is a C4 crop, the expert chose to define two distributions for A_1 , one for
250	maize and one for the two C3 crop species considered (wheat and rice). For wheat and rice,
251	the elicited prior distributions defined by the expert are $\mu_{BMAX} \sim Beta(2.31, 2.31, 1, 15)$,
252	$A_1 \sim \mathcal{N}(4.89, 0.13)I(4, 5.5), A_2 \sim Beta(2.12, 2.12, 0.3, 0.4)$. For maize, the prior for A_1 is
253	$A_1 \sim Beta(2.03, 1.5, 3, 4)$, but the other priors are unchanged. Note that the last two
254	parameters of the Beta distributions correspond to lower and upper bounds.
255	The two types of priors are shown in Figure 2 for the parameters of A_1, A_2, μ_{BMAX} , and μ_S .
256	Clearly, prior 2 covers narrower ranges of values than prior 1, in coherence with the fact that
257	prior 2 is designed to be more informative than prior 1. The difference between the two
258	types of prior is particularly strong for the two parameters of the critical N curve, i.e., A_1 and
259	A_2 (Figure 2AB). Noticeably, there is a marked difference between the priors of A_1 defined
260	by the expert for maize vs. wheat + rice. The prior defined for maize covers lower values
261	than the prior defined for wheat and rice (see the dotted vs. dashed lines in Figure 2A).

Figure 2. Priors distributions defined for the parameters A_1 (A), A_2 (B), μ_{BMAX} (C), and μ_S (D). Weaklyinformative 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.





10





271

272 2.3. Posterior distributions

273 The posterior distributions of the model parameters are estimated with a Markov chain 274 Monte Carlo algorithm (MCMC) implemented with the R package rjags (Plummer, 2017) 275 using both types of priors, successively. The R code used to fit models 1 and 2 is presented in 276 Appendix D. With model 1 the convergence was achieved approximately after 10,000 277 iterations according to the Gelman-Rubin diagnosis. The first 10,000 iterations were 278 discarded and the MCMC algorithm was run for 40,000 additional iterations which were 279 then used to compute the median and 95% credibility intervals for several quantities of 280 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 281 separately. We also computed the median and 95% credibility intervals of the critical 282 nitrogen dilution curve $N_C = A_1 B_{MAX}^{-A_2}$. With model 2 (based on Eq.3), the convergence 283 284 was achieved approximately after 10,000 to 50,000 iterations for the maize and wheat 285 datasets. However, with model 2, we were able to achieve convergence for rice, even with a 286 large number of iterations, probably due to the fact that model 2 was overparametrized for 287 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 288 289 obtained with model 1 in the next section. Nonetheless, values of A1 and A2 estimated with 290 model 2 can be found in Appendix C.

3. Results

293 3.1. Fitted critical N curves

294 The fitted critical N curve and its 95% critical interval obtained for the maize cultivar DH605 295 with prior 1 are shown in Figure 3a. The width of the critical interval describes the level of 296 uncertainty in the fitted curve and directly reflects the distributions of the values of the 297 parameters A1 and A2 (Figure 3bc). Clearly, for the maize cultivar DH605, the level of 298 uncertainty in critical N depends on the biomass value. The width of the critical interval is 299 equal to about 0.5% of plant N concentration when the biomass is lower than 2 t ha⁻¹ and 300 becomes lower than 0.25% when the biomass is higher than 4 t ha⁻¹. The level of uncertainty 301 is thus lower for high compared to low biomass values. This decreasing trend in the level of 302 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 303 304 uncertainty is small compared to maize, even for low biomass values, and its level does not 305 show any substantial increasing or decreasing trend. The uncertainty is especially low for 306 wheat for which the width of the 95% credibility interval is close to 0.1% over a wide range 307 of biomass values. For rice, the level of uncertainty is intermediate between maize and 308 wheat (Appendix A3-A4).

309

Figure 3. Fitted critical N curve obtained for the maize cultivar DH605 and its 95% credibility interval. In graphic

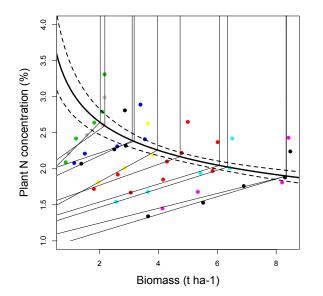
312 3a, continuous and dashed thick lines represent the posterior median and the 95% credibility interval,

313 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
- 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
- 318 distributions.

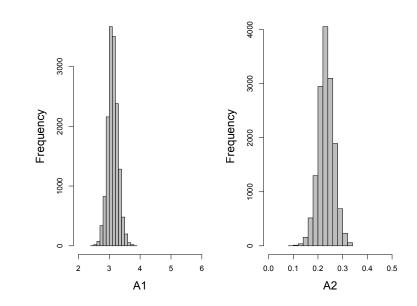
319 a.











c.

322

The whole set of fitted critical N curves obtained for the different crops and cultivars are compared in Figure 4. With prior 1 (Figure 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 (Figure 4A). The curves obtained for the two maize cultivars are very similar and close to the curve of indica (Figure 4A).

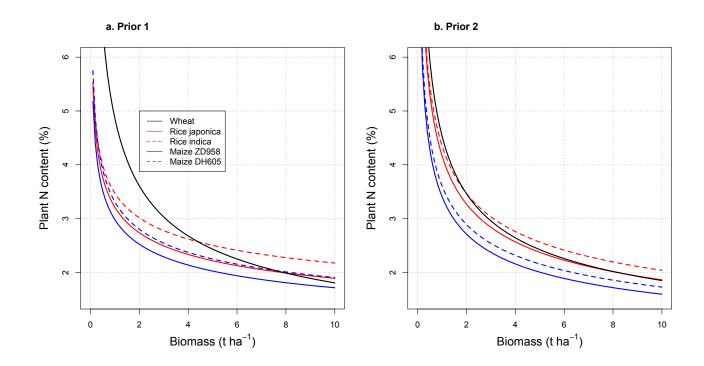
The curve obtained for wheat with prior 2 (Figure 4B) is very similar to the curve obtained

332 with prior 1. For wheat, the fitted critical N curve is thus insensitive to the choice of prior. On 333 the contrary, the curves obtained for rice indica and japonica are much higher with prior 2 334 compared to prior 1. With prior 2, the critical N curves of rice become very close to the 335 critical N curve of wheat (Figure 4B). For maize, compared to prior 1, the critical N curves 336 obtained with prior 2 tend to be slightly higher and closer to the curve obtained for wheat. 337 Overall, the differences between the critical N curves obtained for the different species and 338 cultivars are thus relatively small. These differences may simply reflect uncertainties in the 339 values of the parameters A1 and A2 characterizing the critical N curves. In order to conclude, 340 it is thus necessary to analyze the parameter values. This is done in the next section.

341

331

- 343 Figure 4. Fitted critical N curve obtained for wheat, maize (two cultivars), and rice (two cultivars). Each curve
- 344 corresponds to a posterior median obtained with prior 1 (A) or prior 2 (B).





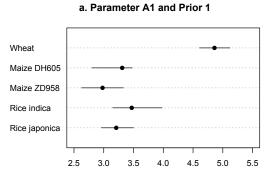
346

347 3.2. Estimated parameter values

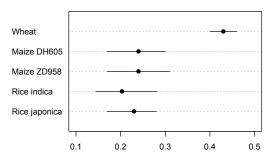
Figure 5 shows the estimated parameter values obtained for the different crops and
cultivars. The posterior medians correspond to point estimates that can be used to draw
critical N curves, as shown in figures 3a and 4ab. The 95% credibility intervals presented in
Figure 5 cover 95% of the values sampled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

355

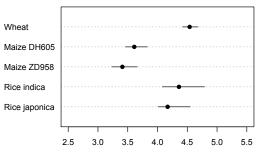
Figure 5. Estimated values (posterior medians) and 95% credibility intervals of parameters A1 (a, b, in %) and
A2 (c, d) with priors 1 (a, c) and 2 (b, d), for wheat, maize (two cultivars), rice indica, and rice japonica.

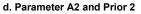


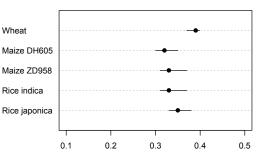




b. Parameter A1 and Prior 2







358

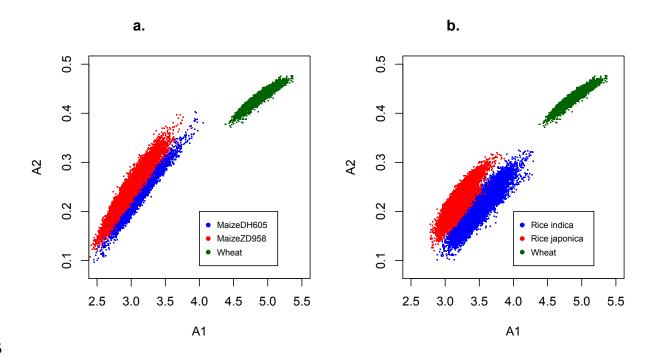
361 With prior 1 (Figure 5ac), the parameter values are significantly higher for wheat than for 362 the other crop species. For wheat, the posterior median is equal to 4.86 (95%CI=[4.61, 5.12]) 363 for A1 and to 0.43 (95%CI=[0.40, 0.46]) for A2. In comparison, for maize and rice, the 364 posterior medians of A1 are always lower than 3.5 and those of A2 are always lower than 365 0.25. The uncertainty levels (reflected by the sizes of the credibility intervals) are 366 substantially higher for maize and rice than for wheat. This is because the number of 367 measurements used to fit the critical N curve is higher for wheat compared to maize and rice 368 (see also Appendix A). Because of their large credibility intervals, the parameter values 369 obtained for the two maize cultivars and for the two rice cultivars are not significantly 370 different. 371 The whole ensembles of parameter values drawn by MCMC from the posterior distributions 372 of A1 and A2 obtained with prior 1 are shown in Figure 6. These values are those used to 373 derived the posterior medians and credibility intervals presented in Figure 5ac. Parameter 374 values presented in Figure 6 show that A1 and A2 are positively correlated; high (low) values 375 of A1 tend to be associated with high (low) values of A2. Figure 6 also confirms that the 376 parameter values obtained for the two cultivars of maize strongly overlap, that those 377 obtained for the two rice cultivars partly overlap, and that the parameter values obtained 378 for wheat are much higher than those obtained for maize and rice.

379

380

381

Figure 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).



385

386

387 The use of prior 2 (informative priors) instead of prior 1 does not substantially impact the 388 parameter values estimated for wheat (Figure 5bd). For this crop, both priors lead to similar 389 values for A1 and A2. On the contrary, for rice, the estimated parameter values are much 390 higher with prior 2 than with prior 1. Consequently, the parameter values obtained for rice 391 with prior 2 are not significantly different from those obtained for wheat anymore. This is 392 because the informative priors (prior 2) defined for rice and wheat force the values of A1 to 393 be higher than 4 and the values of A2 to be higher than 0.3 (Figure 2ab). For wheat, prior 2 394 has no substantial impact because the posterior medians of A1 and A2 obtained with prior 1 395 were already higher than 4 and 0.3, respectively. For rice, the use of prior 2 has a strong 396 impact on the estimated parameter values because the posterior medians of A1 and A2 397 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 (Figure 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 (Figure 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.

404

405 4. Discussion

406 The proposed Bayesian method has several advantages. First, it does not require any 407 preliminary classification of N-limited data against non-N-limited data and does not 408 necessitate the preliminary identification of critical N concentration values. This is an 409 important advantage because it was shown that the critical N curve is sensitive to the 410 method used for estimating critical N concentration values (i.e., type of statistical test and 411 type I error level) and because there is no consensus on how these critical N concentration 412 values should be estimated (Greenwood et al. 1990; Justes et al., 1994; Chen and Zhu, 2013). 413 Our Bayesian approach can be implemented even with sparse data, i.e., when observations 414 are available for a limited number of fertilizer doses only. This is made possible because the 415 proposed model borrows strength from all dates of measurements and because parameters 416 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 minutes using a
standard commercial computer.

428 Finally, another advantage of the proposed approach is that it facilitates the analysis of the 429 uncertainty of the fitted critical N curves. The proposed method is based on a Bayesian 430 hierarchical model whose parameters are estimated taking into account the number of 431 observation dates included in the data set, the number of data available per date and the 432 variability of observations between and within dates. The estimation results are expressed 433 by probability distributions from which the uncertainty of any quantity of interest can easily 434 be analyzed. In particular, our approach allows us to calculate the credibility intervals of the 435 N critical curves and their parameters. Our applications show that, because the dataset used 436 for wheat has been constituted from a large number of experiments network across France, 437 the uncertainty of critical N curve for wheat is relatively low. For maize and rice in China, 438 where the numbers of available data are smaller, the widths of the credibility intervals are 439 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.

451 Like all Bayesian methods, our approach allows modellers to combine two sources of 452 information to estimate the parameters. More specifically, it combines prior information 453 based on expert knowledge and experimental data. Prior information is described using 454 probability distributions that summarize the initial state of knowledge on parameter values 455 before using the data. Here, we use two types of priors. The former are poorly informative 456 and are designed to provide little information on plausible values of model parameters. They 457 do not therefore strongly limit the values of the parameters. The second priors are more 458 informative and are specified by the probabilistic elicitation of an expert. This technique 459 allows to represent the experts' knowledge on the value of a parameter through a 460 probability distribution.

Probabilistic elicitation is useful 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

467 the choice of prior because the size of the dataset is relatively large in this case. For wheat, 468 both priors lead to similar point estimates and credibility intervals. In contrast, for maize and 469 rice, the sizes of the data sets are smaller and, in both cases, the values of the estimated 470 parameters are sensitive to the choice of a prior; they take on larger values and their 471 credibility intervals are narrower when calculations are made with informative a priori. 472 When used, informative priors should therefore be defined by using qualified experts based 473 on relevant information. 474 We believe that our approach opens new perspectives for the estimation of critical nitrogen

476 in order to better take into account possible correlations between measurements, to better477 describe the a priori information available on the values of the parameters, or to handle

dilution curves. Different variants of the model proposed here could be tested in the future

478 larger networks of experiments.

479

475

480 Acknowledgments

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

483

484 References

485 Ata-Ul-Karim S.T., Yan Zhu, Xiaojun Liu, Qiang Cao, Yongchao Tian, Weixing Cao., 2017.

486 Comparison of different critical nitrogen dilution curves for nitrogen diagnosis in rice.

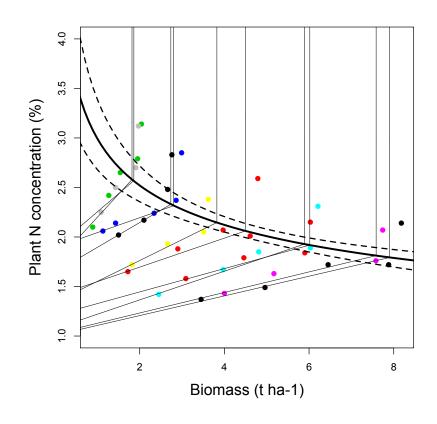
487 Scientific reports 7, 42679. DOI: 10.1038/srep42679.

- 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 protection 126,
 104926.
- 491 Chen, P., Zhu., Y. 2013. A new method for winter wheat critical nitrogen curve
- 492 determination. Agron. J. 105, 1839–1846
- 493 Colnenne, C., Meynard J-M., Reau, R. 1998. Determination of a critical nitrogen dilution
- 494 curve for winter oilseed rape. Ann. Bot. 81, 311–317. doi:10.1006/anbo.1997.0557
- 495 Ioannidis, J.P., 2005 Why most published research findings are false. PLoS Med 2, e124.
- 496 Justes, E., Mary, B., Machet, J.M. 1994. Determination of a critical nitrogen dilution curve for
- 497 winter wheat crops. Ann. Bot. 74, 397–407 1. doi:10.1006/anbo.1994.1133
- 498 Lemaire, G., Jeuffroy, M-H., Gastal, F. 2008. Diagnosis tool for plant and crop N status in
- 499 vegetative stage: Theory and practices for crop N management. Eur. J. Agron. 28, 614–624.
- 500 doi:10.1016/j.eja.2008.01.005
- 501 Morris, D.E., Oakley, J.E., Crowe, J.A., 2014. A web-based tool for eliciting probability
- 502 distributions from experts. Environ. Model. Softw 52, 1–4.
- 503 https://doi.org/10.1016/j.envsoft.2013.10.010.
- 504 Philibert, A., Loyce, C., Makowski, D., 2012. Quantifying uncertainties in N2O emission due to
- 505 N fertilizer application in cultivated areas. Plos One 7(11), e50950.
- 506 doi:10.1371/journal.pone.0050950
- 507 Plénet, D., Lemaire, G. 2000. Relationships between dynamics of nitrogen uptake and dry
- 508 matter uptake in maize crops: Determination of critical N concentration. Plant Soil 216, 65–
- 509 82. doi:10.1023/A:1004783431055

- 510 Plummer, M. 2017. JAGS user manual. <u>http://mcmc-jags.sourceforge.net</u>. <u>https://cran.r-</u>
- 511 project.org/web/packages/rjags/index.html
- 512 Ramanantenasoa, M.M.J., Génermont, S., Gilliot, J-M., Makowski, D. 2019. Meta-modeling
- 513 methods for estimating ammonia volatilization from nitrogen fertilizer and manure
- applications. Journal of Environmental Management 236, 195-205. DOI:
- 515 10.1016/j.jenvman.2019.01.066
- 516 Sandana, P., Lobos, I.A., Pavez, P.B., Moscoso, C.J. 2019. Validation of a critical nitrogen
- 517 dilution curve for hybrid ryegrass. Grass Forage Sci. 2019, 1–11.
- 518 Wallach, D., Makowski, D., Jones, J.W., Brun, F. 2019 Working with dynamic crop models.
- 519 Third edition. Elsevier Academic Press.
- 520 Zhao B., Syed Tahir Ata-Ul-Karim, Aiwang Duan, Zhandong Liu, Xiaolong Wang, Junfu Xiao,
- 521 Zugui Liu, Anzhen Qin, Dongfeng Ning, Weiqiang Zhang, Yanhao Lian. 2018. Determination of
- 522 critical nitrogen concentration and dilution curve based on leaf area index for summer
- 523 maize. Field Crop Research 228, 195-203.
- 524 Zhao, B.Z., J.B. Zhang, M. Flury, A.N. Zhu, Q.A. Jiang, Bi, J.W. 2007. Groundwater
- 525 contamination with NO3–N in a wheat–corn cropping system in the North China Plain.
- 526 Pedosphere 17, 721–731. doi:10.1016/S1002-0160(07)60087-3

527 Appendix A: Individual fitted critical N curves

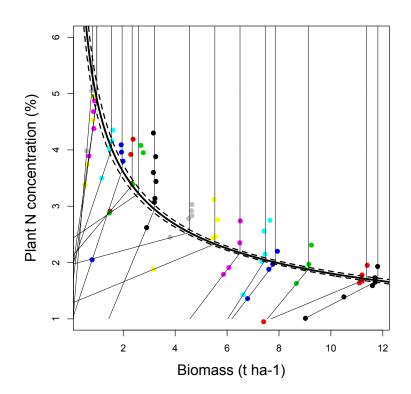
- 528 In this appendix, we present the posterior median, the 95% credibility intervals and the
- 529 experimental data obtained for maize ZD958m (see main text for the other maize cultivar),
- 530 rice (Japonica and Indica) and wheat.
- 531
- 532 A1. Maize ZD958m



- 533
- 534
- 535

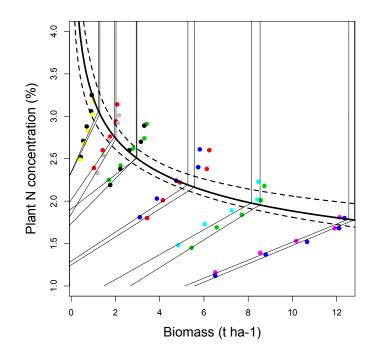
536

539 A2. Wheat



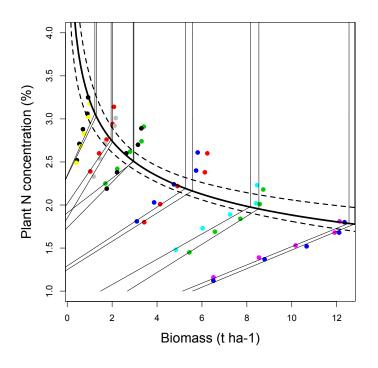








544 A4. Rice Indica

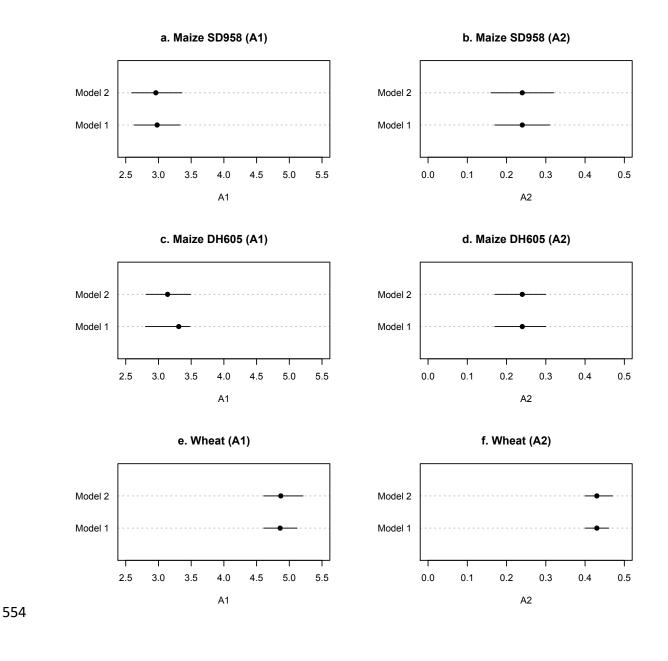




548 Appendix B: Characteristics of the maize experiments

Experiment No.	Sowing/Harvesting		Cultivar	N(kg N ha ⁻¹)	Sampling stage
Experiment No.	date	Soil characteristics	Cultivar		
Experiment 1	8-Jun	Type: light loam soil	Zhengdan958	3 0 (NO)	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
		$NH_4 oAc-K^+$: 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	3 0 (NO)	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
		NH₄0Ac-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)).



557 Appendix D: R Code

- 558 Model 1
- 559 ###Data
- 560 # Q=total number of biomass observations
- 561 # K=number of dates
- 562 # W=column of data including biomass observations
- 563 # N=column of data including observations of nitrogen concentrations
- 564 # Date=column with the indices identifying the different dates of the dataset
- 565
- 566 ###Model parameters
- 567 # Nc=Critical nitrogen concentration
- 568 # Bmax=maximum biomass value in a specific date
- 569 # S=slope of the linear-plus-plateau function
- 570 # W=biomass increase per unit of nitrogen concentration
- 571 # A1 and A2 = parameters of the critical N curve
- 572 # tau_b and tau_n = 1/residual variances for biomass and nitrogen content observations
- 573 # Mu_Bmax, Prec_Bmax = parameters defining the between-date variability of Bmax
- 574 # Mu_S,Prec_S = parameters defining the between-date variability of S
- 575
- Q<-length(Date) 576 577 K<-length(unique(Date)) 578 579 modelstring=" model { 580 581 582 for (i in 1:Q) 583 { W[i]~dnorm(mu[i], tau_b) 584

585	N[i]~dnorm(Nc[Date[i]], tau_n)
586	mu[i]<-min(Bmax[Date[i]], Bmax[Date[i]]+S[Date[i]]*(N[i]-Nc[Date[i]]))
587	}
588	
589	for (j in 1:K)
590	
591	Nc[j]=A1*Bmax[j]^(-A2)
592	Bmax[j]~dnorm(Mu_Bmax,Prec_Bmax)T(0,)
593	S[j]~dnorm(Mu_S,Prec_S)T(0,)
594	}
595	
596	#Weakly informative
597	Mu Bmax~dnorm(6,0.1)
598	— • • • •
	Mu_S~dnorm(0,0.1)
599	A1~dunif(2,6)
600	A2~dunif(0,0.5)
601	
602	#Informative prior C3
603	#A1~dnorm(4.89,7.72)T(4,5.5)
604	#ZA2~dbeta(2.12,2.12)
605	#A2=(0.4-0.3)*ZA2+0.3
606	#ZMu Bmax~dbeta(2.31,2.31)
	— • • • • • • • • • • • • • • • • • • •
607	#Mu_Bmax=(15-1)*ZMu_Bmax+1
608	#Mu_S~dnorm(0,0.1)
609	
610	#Informative prior C4
611	#ZA1~dbeta(2.03,1.5)
612	#A1=(4-3)*ZA1+3
613	#ZA2~dbeta(2.12,2.12)
614	#A2=(0.4-0.3)*ZA2+0.3
615	#ZMu Bmax~dbeta(2.31,2.31)
616	#Mu_Bmax=(15-1)*ZMu_Bmax+1
617	#Mu_S~dnorm(0,0.1)
618	
619	Prec_Bmax~dgamma(0.001,0.001)
620	Prec_S~dgamma(0.001,0.001)
621	tau b~dgamma(0.001,0.001)
622	tau_n~dgamma(0.001,0.001)
623	
624	}
625	J II
	An et the estimated at the second state of the second state of the
626	writeLines(modelstring, con="model.txt")
627	
628	model<-jags.model('model.txt', data=list('W'=W, 'N'=N, 'Date'=Date, 'Q'=Q,'K'=K),
629	n.chains=3, n.adapt=10000)
630	
631	

632	Model 2
633	modelstring= "
634	model {
635	
636	for (i in 1:Q)
637	{
638	W[i]~dnorm(mu[i], tau_b)
639	N[i]~dnorm(Nc[Date[i]]+Theta[Date[i]], tau_n)
640	mu[i]<-min(Bmax[Date[i]], Bmax[Date[i]]+S[Date[i]]*(N[i]-Nc[Date[i]]))
641	}
642	for (i in 1.14)
643 644	for (j in 1:K) {
645	۱ Nc[j]=A1*Bmax[j]^(-A2)
646	Bmax[j]~dnorm(Mu_Bmax,Prec_Bmax)T(0,)
647	S[j]~dnorm(Mu S,Prec S)T(0,)
648	Theta[j]~dnorm(0,tau_t)
649	}
650	
651	#Weakly informative
652	Mu_Bmax~dnorm(6,0.1)
653	Mu_S~dnorm(0,0.1)
654	A1~dunif(2,6)
655	A2~dunif(0,0.5)
656	
657	#Informative prior C3
658 650	#A1~dnorm(4.89,7.72)T(4,5.5)
659 660	#ZA2~dbeta(2.12,2.12) #A2=(0.4-0.3)*ZA2+0.3
661	#ZMu Bmax~dbeta(2.31,2.31)
662	#Mu Bmax=(15-1)*ZMu Bmax+1
663	#Mu_S~dnorm(0,0.1)
664	
665	#Informative prior C4
666	#ZA1~dbeta(2.03,1.5)
667	#A1=(4-3)*ZA1+3
668	#ZA2~dbeta(2.12,2.12)
669	#A2=(0.4-0.3)*ZA2+0.3
670	#ZMu_Bmax~dbeta(2.31,2.31)
671	#Mu_Bmax=(15-1)*ZMu_Bmax+1
672	#Mu_S~dnorm(0,0.1)
673 674	$\operatorname{Proc} \operatorname{Proc}^{\circ} \operatorname{Proc}^{\circ} \operatorname{dramma}(0.001, 0.001)$
674 675	Prec_Bmax~dgamma(0.001,0.001) Prec_S~dgamma(0.001,0.001)
676	tau b~dgamma(0.001,0.001)
677	tau n~dgamma(0.001,0.001)
	(

678	tau_t~dgamma(0.001,0.001)
679	}
680	II Contraction of the second se
681	writeLines(modelstring, con="model.txt")
682	model<-jags.model('model.txt', data=list('W'=W, 'N'=N, 'Date'=Date, 'Q'=Q,'K'=K),
683	n.chains=3, n.adapt=10000)
684	
685	