Rethinking crop nutrition diagnosis models: methods, inference and practical applications in crop production and breeding

by

Josefina Lacasa

B.S., Universidad de Buenos Aires, 2020

M.S., Kansas State University, 2023

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Abstract

Fertilizer management is one of the most important aspects of agronomic management, affecting the system's sustainability, and relies heavily on the statistical models for nutritional diagnoses. While under-fertilizing may penalize crop yields, excessive fertilization is linked to negative environmental externalities. It is thus desirable that the fertilizer rates approximately match the crop nutrient requirements. Because crop nutrient requirements are usually associated to crop growth, it is convenient to monitor the nutritional status to determine the amount of nutrient required according to crop growth. With the increase in computational power of mainstream computers, many popular crop nutrition models are undergoing changes that leverage such advances in technology. This dissertation is organized in six chapters (Chapter 1, Introduction, and Chapter 6, Final remarks) that revise, rethink, and expand some of the most popular models applied to crop nutrition management. Chapters 2-3 are oriented to the statistical inference, and portray methods developments that may help improve inference from crop nutrition models. Chapter 2 portrays some advantages enabled by modern statistical computing tools by comparing a standard statistical framework introduced in the 1990s, versus a modern statistical framework introduced in 2020. Chapter 3 follows up on the findings of Chapter 2 and elaborates the model and establishes prospects for statistical modeling of crop nutrition models with current statistical tools. Chapters 4-5 are oriented to the practical application of crop nutrition models and the integration of modern measuring hardware, and portray methods for applications in phenotyping and plant breeding settings. Chapter 4 compares different metrics for quantifying crop nutritional status for breeding applications in wheat (*Triticum aestivum* L.). Chapter 5 identifies avenues for research for further developing methods and measuring devices for crop nutritional status phenotyping. While most of the crop nutrition problems presented in this dissertation consider nitrogen management, these finding are relevant for other nutrients as well.

Keywords: Nitrogen; Fertilization; Fertilizer use efficiency; Statistical methods; Bayesian modeling; Phenotyping.

Rethinking crop nutrition diagnosis models: methods, inference and practical applications in crop production and breeding

by

Josefina Lacasa

B.S., Universidad de Buenos Aires, 2020

M.S., Kansas State University, 2023

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2023

Approved by:

Major Professor
Ignacio Ciampitti

Abstract

Fertilizer management is one of the most important aspects of agronomic management, affecting the system's sustainability, and relies heavily on the statistical models for nutritional diagnoses. While under-fertilizing may penalize crop yields, excessive fertilization is linked to negative environmental externalities. It is thus desirable that the fertilizer rates approximately match the crop nutrient requirements. Because crop nutrient requirements are usually associated to crop growth, it is convenient to monitor the nutritional status to determine the amount of nutrient required according to crop growth. With the increase in computational power of mainstream computers, many popular crop nutrition models are undergoing changes that leverage such advances in technology. This dissertation is organized in six chapters (Chapter 1, Introduction, and Chapter 6, Final remarks) that revise, rethink, and expand some of the most popular models applied to crop nutrition management. Chapters 2-3 are oriented to the statistical inference, and portray methods developments that may help improve inference from crop nutrition models. Chapter 2 portrays some advantages enabled by modern statistical computing tools by comparing a standard statistical framework introduced in the 1990s, versus a modern statistical framework introduced in 2020. Chapter 3 follows up on the findings of Chapter 2 and elaborates the model and establishes prospects for statistical modeling of crop nutrition models with current statistical tools. Chapters 4-5 are oriented to the practical application of crop nutrition models and the integration of modern measuring hardware, and portray methods for applications in phenotyping and plant breeding settings. Chapter 4 compares different metrics for quantifying crop nutritional status for breeding applications in wheat (*Triticum aestivum* L.). Chapter 5 identifies avenues for research for further developing methods and measuring devices for crop nutritional status phenotyping. While most of the crop nutrition problems presented in this dissertation consider nitrogen management, these finding are relevant for other nutrients as well.

Keywords: Nitrogen; Fertilization; Fertilizer use efficiency; Statistical methods; Bayesian modeling; Phenotyping.

Table of Contents

List of Figures

List of Tables

Acknowledgments

The completion of this dissertation marks the end of a long journey that would had been impossible without the help and support of some critical people I am truly grateful to.

I am immensely grateful to Ignacio Ciampitti for his trust and mentorship, for granting me the greatest intellectual freedom, and for always pushing my boundaries yet in a safe space for making -and learning to solve- mistakes.

I also want to acknowledge my committee members for their patient dedication in the formats of advise, feedback and generous time: Trevor Hefley, David Makowski, and Vara Prasad.

Thanks to my first student mentor, Lu, for her vision and for being a true inspiration in how to elevate others.

My office mate Fran came in for this last year-and-a-half with excellent discussions about crop physiology, statistics, life, and of course with Caro's budincitos. Thanks to Caro for her constant support that often came in the shape of delightful budincitos.

My lab mates have taught me a lot and also, definitely made this journey enjoyable: thank you to KSU Crops of 2019, 2021, and Ciampitti Lab of 2022, 2023.

A long list of people stood by my side (physically or in spirit) through ups and downs, encouraging me to keep going. This would have been really hard without their support. Thanks to my family and my friends from school, Club Ciudad, FAUBA, Manhattan, and life.

Finally, I want to pause and thank especially Juan Carlos, Karin, Victoria, Juan Martin, Mariana, and Juan Francisco for inspiring me to deliver my best at all times with their unconditional support, from the very start to the very end.

Dedication

To Karin and Juan Carlos, my first agronomy teachers.

Chapter 1

Introduction

Modern agricultural farming systems face major challenges to achieve food security while minimizing environmental externalities. In particular, the accessibility to nutrients like nitrogen (N) and phosphorous fertilizers has improved the yield of staple crops (Stewart and Roberts 2012), but excessive fertilizer applications increase the risk of triggering undesirable ecosystem disruptions (Davidson et al. 2012). For example, N fertilization of cereal crops like maize (*Zea mays* L.) and wheat (*Triticum aestivum* L.) is considered key for crop success but can often be a source of air and water pollution, potentially leading to eutrophication, contamination of groundwater and of the atmosphere (Eickhout et al. 2006). Adequate N fertilizer management practices are thus crucial for maintaining sustainable gains in food security over time.

The consensus for improving sustainability, from a nutrient management standpoint, is to reduce external N inputs as much as possible to reduce the environmental footprint (Stevens 2019). The ultimate goal is to maintain food security by achieving maximum productivity (e.g., grain yield per hectare) with similar quality (e.g., grain protein content) with the minimum amount of N inputs added to the system. Several efforts during the past decades have proposed strategies to increase productivity with the minimum amount of N inputs. Some of the most notable strategies proposed thus far target the crop agronomic management (e.g., 4Rs, right place, timing, source, and rate for fertilization) or the crop genetic improvement (e.g., selection of genotypes with a more efficient conversion). Consequently, large investments have been destined globally to developing technology for nutritional diagnoses and fertilizer recommendations, as well as sensors and equipment for determining the crop nutritional status (Kanter et al. 2020).

In order to recommend preferred agricultural practices, scientists create conceptual frameworks (i.e., models) of agricultural systems for predicting the behaviour of a system and thus establishing and communicating said 'best practices'. Models describing nutrient dynamics and relationships are usually synthesized in mathematical equations, where most unobserved variables (e.g., describing the relationship between yield and fertilizer rate) are estimated with statistical methods. It is thus critical to design sound mathematical models and to use statistical methods that are appropriate for the problem in hand. Both the mathematical model and the statistical methods play a central role in estimating which are the best practices in agriculture and quantifying crop nutrient requirements. For example, the determination of the best agronomic practices, as well as the selection of the best genotypes for crop improvement relies heavily on the metrics and the methods considered to estimate the ranking of the best practices or genotypes.

The technological development that took place during the last decades has expanded the range of possible tools for aiding crop nutrition management for both the data processing and the measurement tasks. For example, current equipment has made high-throughput phenotyping accessible, allowing breeders to explore more traits related to plant N use. Novel tools and techniques for high-throughput phenotyping are under constant development (Pinto et al. 2023), eventually leading towards being able to select crops that make use available N more efficiently. In addition to new measuring equipment, the processing power of modern computers has enabled intensive statistical analyses that would not have been possible with mechanical calculators (Efron and Hastie 2016). The joint development of high-performing computers with high-throughput phenotyping platforms may create a synergistic relationship to make the data interpretable and useful. However, data processing is still considered a bottleneck in plant breeding programs (Araus and Cairns 2014), demanding updates in the current statistical methods applied to agriculture.

This dissertation thus reflects on the technological transition that occurred over the last decades and provides a revisionist analysis regarding current conceptual frameworks for quantifying nutrient use in crop nutrition models. More specifically, the dissertation focuses on the statistical methods, metrics, equipment, and conceptual (physiological) framework of current popular crop nutrition models. The global objective of this dissertation is to revise standard methods, from a perspective that considers recent advances in science and technology, for obtaining sound inference for providing more trustworthy recommendations to the agricultural stakeholders. The specific objectives are to: (i) compare statistical frameworks for estimating a specific crop nutrition model, the nitrogen dilution curve, especially focusing on the standard framework versus a novel framework that was enabled by modern statistical computing tools (Chapter 2), (ii) characterize the current status of statistical framework in applied statistics in agriculture, and establish prospects for developing statistical methods, with a focus on crop nutrition models (Chapter 3), (iii) compare different metrics for quantifying crop nutritional status for breeding applications in wheat (Chapter 4), and (iv) revisit the current status of the methods for phenotyping crop nutritional status and provide future prospective (Chapter 5). Last, it is worthwhile mentioning that most of the crop nutrition problems presented in this dissertation consider only N nutrient, but these finding are relevant and easily transferable to other macronutrients.

Chapter 2

Comparison of statistical methods to fit critical nitrogen dilution curves

Published in European Journal of Agronomy

Lacasa, J., Makowski, D., Hefley, T., Fernandez, J., van Versendaal, E., Lemaire, G., & Ciampitti, I. (2023). Comparison of statistical methods to fit critical nitrogen dilution curves. European Journal of Agronomy, 145, 126770.

/https://doi.org/10.1016/j.eja.2023.126770

Abstract

The critical nitrogen dilution curve (CNDC) is one of the most important models underpinning diagnostic tools of crop nitrogen (N) status. The CNDC relates plant N content ($\%N$) to plant biomass under optimal nutritional conditions, using a simple nonlinear function with two parameters. More precisely, the CNDC estimates the critical %N ($\%N_C$) – the amount of $\%N$ needed to achieve maximum biomass at a certain moment. Thus, the comparison of a measured $\%N$ to $\%N_C$ diagnoses if a plant is N deficient and if application of N fertilizer is necessary. Because of the practical importance of the CNDC, its parameters should be estimated as accurately as possible from field data. Two contrasting frameworks have been proposed in the literature for parameter estimation of the CNDC. The most popular approach is a sequential framework that consists of fitting two statistical models sequentially. Recently, an alternative framework has been introduced to estimate the parameters in a single step, using a hierarchical Bayesian model (hierarchical framework). In this study, we compare both methods and evaluate their performances for a large range of experimental designs. We consider datasets with different numbers of observation times (4, 8, 16), N fertilizer rates (3, 4, 5), and levels of measurement accuracy (low, medium, high). Our results show that the hierarchical framework outperforms the sequential framework under most scenarios. The sequential framework often resulted in poor statistical properties of the parameter estimates (e.g., increased bias and poorly calibrated confidence intervals) and to overestimated $\%N_C$ values. The estimates for the amount of N needed to sustain N sufficiency levels were between 30 and 100 kg N ha^{-1} greater under the sequential framework, compared to the hierarchical framework. Thus, adopting the hierarchical framework could contribute to decreasing the risk of overfertilization and increasing the sustainability of agricultural systems.

Keywords: Hierarchical models; Bayesian modeling; N fertilization; fertilizer use efficiency.

2.1 Introduction

Managing crop nitrogen (N) fertilization is critical to improve on-farm efficiency while reducing environmental externalities (Cassman et al. 2002). Many N fertilization management tools were developed based on the minimum N concentration ($\%N$) needed to maximize crop growth, called the critical N concentration ($\%N_C$, Justes et al. 1994). This concept became popular among agronomists because the difference between the actual $\%N$ and the value of $\%N_C$ can be used to determine whether a crop is N deficient. This approach is useful to optimize N application timing and rates, and reduce environmental impact while improving fertilizer use efficiency (Lemaire and Ciampitti 2020). The critical N concentration at a given observation time i, $\%N_{Ci}$, is related to the maximum crop biomass that can be achieved at the same time (W_{maxi}, in Mg ha^{-1}) by a nonlinear function with two parameters A_1 and A_2 defined by Salette and Lemaire (1981):

$$\%N_{Ci} = A_1 \cdot W_{maxi}^{-A_2}. \tag{2.1}$$

The nonlinear function (2.1) is named critical N dilution curve (CNDC) because it describes the $\%N_C$ dilution that takes place through crop development. Since values of $\%N_C$ cannot be observed directly, they are estimated from biomass and $\%N$ measurements collected at several dates during crop development (further indexed by the subscript i). Normally, experiments aiming to estimate CNDC have many experimental plots with different levels of N fertilization (further indexed by the subscript j). To this end, the relationship between W and $\%N$ is commonly modeled with a linear-plus-plateau function (Salette and Lemaire 1981):

$$W_{ij} = \begin{cases} W_{maxi} + S_i \cdot (\%N_{ij} - \%N_{Ci}) & \text{if } \%N_{ij} < \%N_{Ci} \\ W_{maxi} & \text{if otherwise} \end{cases}, \tag{2.2}$$

where W_{ij} is the biomass of the jth level of N fertilization at the ith observation time, S_i is the increase in biomass per increase of $\%N$ (in percentage unit) at the ith observation time (i.e., the response slope of W versus $\%N$), $\%N_{Ci}$ is the $\%N_C$ at the ith observation time, and W_{maxi} is the maximum achievable biomass at this time. By fitting (2.2) to data collected at different times, it is possible to estimate $\%N_{Ci}$ for each observation time, and then to estimate A_1 and A_2 by using (2.1) and the pairs of $\%N_{Ci}$ and W_{maxi}.

In a systematic review, Fernandez et al. (2022) identified two contrasting frameworks that are frequently used to estimate A_1 and A_2 from field data. The main difference between both frameworks is how they connect Equations 2.1 and 2.2. The oldest and most commonly used approach (Justes et al. 1994), herein termed "sequential framework", consists in fitting two different statistical models sequentially. First, (2.2) is fitted to data observed at different observation times covering different crop growth stages and biomass levels, obtaining point estimates of $\%N_C$ and W_{max} for each observation time. Then, (2.1) is fitted to the point estimates of $\%N_C$ and Wmax to obtain estimates for A1 and A2. Recently, a new framework, herein termed "hierarchical framework", was introduced to fit Eqs. (1-2) simultaneously using a hierarchical Bayesian model (Makowski et al. 2020). The authors pointed out that the greatest advantages of the model are linked to (i) accounting for estimation of the uncertainty of all unknown quantities (e.g., $\%N_C$ and W_{max}) and (ii) easy implementation to fit the CNDC to field data (Ciampitti et al. 2021; Jégo et al. 2022; Fernandez et al. 2021). The hierarchical framework is consistently gaining users because of these reasons (Fernandez et al. 2022).

It is still unclear which statistical approach is more relevant. Characteristics like unbiasedness and variance of the estimators (i.e., the statistical properties) have never been formally tested and compared between two CDNC fitting methods. In the statistical sciences, statistical properties must be evaluated using simulation studies, as proposed by Fernandez et al. (2022) and Makowski et al. (2002) in the field of agronomy. The objectives of this study are to (i) compare the results of both frameworks using observed field data,

(ii) assess the ability of both frameworks in estimating A1, A2 and $\%N_C$, for a large range of experimental designs using simulated data, and (iii) assess the consequences in terms of N fertilizer recommendations of choosing one framework over the other.

2.2 Methods

2.2.1 Analysis of maize and rice data

We used two published datasets for maize (Plénet and Lemaire 2000; Zhao et al. 2018) and one for rice (Yao et al. 2021) accessible through a large review presenting an open dataset (Ciampitti et al. 2022). These crops have contrasting photosynthesis metabolisms (maize as a C4 crop and rice as a C3 crop) and are among the most important crops cultivated worldwide.

Before analyzing the data, we performed the standard data filtering procedure. Briefly, this means we discarded the data collected at observation times for which the maximum biomass value Wmax was not reached. Indeed, in such situations, it was not possible to obtain reliable estimates of W_{max} and $\%N_C$, and thus impossible to implement the sequential framework. To determine if the data reached the biomass plateau W_{max}, we fitted two models and determined which one described the data better. First, we fitted a simple linear model (without plateau) and a linear-plus-plateau model (Eq. 2) to the biomass and $\%N$ data, collected at each observation time. We fitted the models using ordinary least squares with the 'lm' and 'nls' functions in R (R Core Team 2019), and tested both models against each other at each observation time using AIC. We removed the data collected at a given observation time when the algorithm used to fit the linear-plateau did not converge for that time, or when the simple linear model had the lowest AIC score (indicating a better fit). This approach for data filtering corresponds to the current standard practice for data curation to fit CNDC. Testing whether this technique is appropriate is beyond the scope of this paper. After filtering, the three final maize and rice datasets included 14, 8, and

6 observation times respectively, for which reliable Wmax and $\%N_C$ estimates could be obtained.

In order to compare the performance of the two estimation frameworks with datasets including a smaller number of observation times, we defined smaller datasets including half the amount of observation times randomly chosen from the original dataset. We thus obtained 6 datasets; 3 large datasets including 14, 8 and 6 observation times for maize and rice, and 3 smaller datasets including half the number of observation times only (i.e., 7, 4 and 3).

We used these datasets to estimate the parameters A_1 and A_2 of the maize and rice CNDC applying the sequential and hierarchical frameworks, successively. For the sequential framework, (12.1) was fitted to the values of $\%N_C$ and Wmax derived from the linear-plus-plateau model (2.2) at each observation time using the 'nls' R function. Equation (2.1) was fitted to the $\%N_C$ and Wmax values obtained for each of the six datasets, successively. This allowed us to estimate A_1 and A_2 with the large and small datasets. The 95% confidence intervals (CI) for the parameters A_1 and A_2 were obtained using the 'confint' function in R, that assumes a normal distribution of the data. This two-step procedure (step 1: estimation of $\%N_C$ and W_{max}, step 2: estimation of A_1 and A_2) corresponds to the most commonly used method to fit CNDC curves (Fernandez et al. 2022). Under the hierarchical framework, we fitted the Bayesian model defined by Makowski et al. (2020) to the datasets using the least informative priors proposed in this study (see Makowski et al. 2020). Briefly, the Bayesian model is fitted to $\%N$ and W field data, producing estimated values for $\%N_C$, W_{max}, A_1, and A_2 directly, in one step. In this way, the uncertainty in the values of $\%N_C$ and W_{max} is taken into account when estimating A_1 and A_2. The Bayesian model is fitted using the R package 'rjags' (Plummer et al. 2016). We produced 3 chains of parameter values in order to check the convergence of the Markov chain Monte Carlo (MCMC) fitting algorithm implemented by rjags, as commonly done with this type of algorithm. We assumed chain convergence when Rhat values were between 1 and 1.02 [see Gelman and Rubin

(1992) for more details]. The posterior means and 95% credible intervals of A_1 and A_2 were obtained from the posterior samples obtained from the posterior samples obtained from the MCMC algorithm for each model and data combination (i.e., rice and maize). In addition, we obtained posterior distributions of $\%N_C$ for 10 levels of W between 2 and 10 Mg ha^{-1}. For additional information on the code, please consult the GitHub repository "github.com/jlacasa/CNDC_methods_comparison".

Finally, we calculated the difference in estimated $\%N_C$ according to both frameworks. Each fitted curve was used to estimate $\%N_C$ for a range of W from 2 to 10 Mg ha^{-1}. For a given biomass value W, the difference between the crop critical N uptake under both frameworks was estimated in

General description

We performed a simulation study to compare the performances of both frameworks to estimate the parameters A_1 and A_2, for a large range of experimental designs. Data simulation is a standard approach to compare statistical methods and has been used in various contexts in agriculture (Makowski et al. 2007, 2002; Makowski and Wallach 2001; Wallach 2011; Spilke et al. 2005; Tanaka 2021) and several other fields (Boulesteix et al. 2020; Hefley et al. 2017, 2016, 2014, 2013; Morris et al. 2019). In this approach, data are simulated with previously defined true parameter values and thus, it is possible to check if an estimation method is able to obtain parameter estimates that are close to the true values, both in terms of bias and variance.

Data simulation

Our simulation study was performed based on the CNDC obtained for maize from observed field data, where $A_1 = 3.8$, and $A_2 = 0.44$. We assumed that this CNDC represents the true CNDC and we used it to generate 1,000 simulated datasets of different sample sizes. We defined three amounts of observation times (4, 8, or 16), three levels of N rates (3, 4, or 5, equivalent to the number of $\%N$ observation per observation time), and three types of distributions of $\%N$ data (very low, low, and medium proportion of $\%N$ data greater than $\%N_C$). The distribution of $\%N$ is an important factor because it defines the ability to estimate Wmax correctly. If the data include enough $\%N$ values below and above $\%N_C$, there is a high chance that the plateau will be easily identified. In this case, $\%N$ and W_{max} could be accurately estimated. To explore contrasting situations, the proportions of $\%N$ greater than $\%N_C$ were tuned in order to be able to obtain reliable individual linear-plus-plateau fits in approximately one-third (for very low proportions of $\%N$ ¿ $\%N_C$), two-thirds (for low proportions of $\%N$ ¿ $\%N_C$) and all (for medium proportions of $\%N$ ¿ $\%N_C$) of the observation times. The combination of these design characteristics (i.e., varying number of observation times, N rates and $\%N$ data distribution) led to 333=27 different scenarios. The following simulation procedure was then implemented 1,000 times for each design type:

1. Generate one pair of Wmax and S for each observation time (by sampling randomly from the values produced by the individual linear-plus-plateau fits obtained using a maize dataset),

2. Calculate $\%N_C$ for each observation time from Wmax and the true values of A1 and A2 using Eq.(1),

3. Generate 3, 4, or 5 (depending on the design) values of $\%N$ from a range of values specified in order to generate a very low, low, or medium proportion of values greater than $\%N_C$,

4. Generate 3, 4, or 5 virtual W data by plugging the values of $\%N$ generated at the previous step in Eq.(2) and adding a random Gaussian noise with variance equal to 0.08 (value consistent with variation observed in field data),

5. Fit the CNDC, estimate A1 and A2 using each one of the two frameworks successively (as explained in 2.1), and compute the 95% confidence (standard framework) and credible (hierarchical framework) intervals of the estimates of A1 and A2.

Why do we need a simulation study in agronomy?

First, let us clarify the concept “simulation study”. In agriculture, this term is traditionally employed when referring to the use of crop growth models and for developing different scenarios of interactions of field practices under different environments. In this work, we relate this term to its usage in the statistical sciences. A simulation study is traditionally implemented to compare statistical methods in terms of their inference properties (e.g., bias and variance). They are considered a standard practice in statistics because they are a practical way to characterize the estimator properties. Usually, they are the only existing technique to compare different statistical methods in these terms. In this approach, data are simulated with previously defined ‘true’ parameter values and data generating process and thus, it is possible to check if an estimation method is able to obtain parameter estimates that are close to the true values. Using field data, one cannot assess the quality of an estimator due to the before mentioned ‘truth’ is unknown. In agronomy, a simulation study can be used to determine the bias and mean square error of a specific type of the estimator of a key parameter of practical interest, considering specific field experiment designs. Therefore, the utilization of a simulation study is critical to compare statistical methods and to implement more reliable data assessment to more effectively evaluate the impacts of experimental designs, treatment effect, and for model testing of the data.

2.2.2 Framework evaluation

At the end of this procedure, we got 1,000 simulated datasets for each scenario. We then estimated the parameters A_1 and A_2 of the CNDC from these datasets using the sequential and hierarchical frameworks successively. Thus, we had 1000 estimates of A_1 and A_2 for each of the 27 designs and each framework. Last, we compared the estimates $\hat{A}_{\mathrm{X}}$ (X = 1, 2) to the true parameters A_{X} by computing the percentage bias error ($PBE = 100 \cdot \frac{\frac{1}{n}\Sigma_{i=1}^{n}\widehat{A_{X}}_{i} - A_X}{A_X}$) and the relative root mean squared error ($RRMSE = 100 \cdot \frac{\sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(\widehat{A_{X}}_{i} - A_X)^2}}{A_X}$), the lower, the better. Moreover, we evaluated the confidence and credible intervals for both frameworks by computing their coverage probabilities. A coverage probability indicates the proportion of the 1,000 intervals that include the true parameter values. Thus, a value close to 95% indicates properly calibrated intervals. In addition, we estimated $\%N_C$ with W=10, and computed the point estimate bias and root mean squared error (RMSE). Last, we estimated the difference in estimated $\%N_C$ as defined in section (2.1.), in order to identify possible systematic differences between frameworks that may result in changes in fertilizer recommendations. All the details of the simulation procedure are provided in github.com/jlacasa/CNDC_methods_comparison.

2.3 Results

2.3.1 Critical nitrogen dilution curves estimated from the maize and rice datasets

Overall, the size of the dataset reduced the magnitude of the point estimates and their associated uncertainty. Maize and rice showed a similar pattern in terms of the response to increasing number of observation times (Figure 2.2D).

With the large and medium-sized maize datasets (Plénet and Lemaire 2000; Zhao et al. 2018), the estimated values of A1 and A2 were slightly lower with the hierarchical framework (Figure 2.2D) than with the sequential framework.

Point estimates (95% confidence intervals) obtained for A_1 with the sequential framework were around 3.6 (3.06, 4.18) and 3.4 (3.07, 3.73), versus 3.32 (3.09, 5.82) and 3.2 (2.85, 3.56) for the hierarchical framework, for the large and medium-sized maize datasets, respectively. For A_2, the estimated values obtained with the sequential framework were 0.39 (0.29, 0.49)

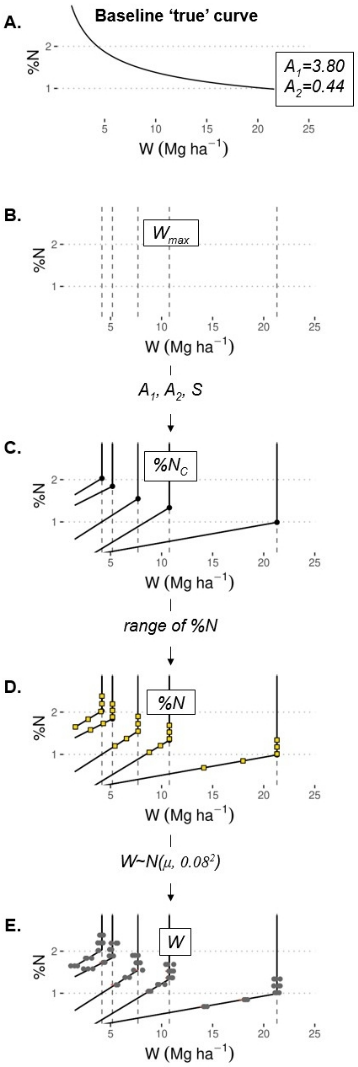

Figure 2.1: *Conceptual framework of the data simulation process for nitrogen (N) concentration, %N, and aboveground biomass (W). An example to generate a dataset with five observation times and five N rates, where maximum biomass, W_{max}, is achieved at all times. A. True reference curve with parameters A_1 = 3.80 and A_2 = 0.44. B. Values for W_{max}. Each vertical dashed line indicates the W_{max} value for an observation time and will be used to simulate the data. C. Determination of $\%N_C$. Points show the values of $\%N_C$ corresponding to each observation time, obtained using (2.1) and parameters A_1, A_2 (A), and W_{max} (B). Lines indicate the linear-plus-plateau functions for each observation time, obtained with each set of Wmax, $\%N_C$, and S. D. %N data. For a given range of %N, the simulated values for %N data are obtained at equal intervals minimum and maximum %N. E. Simulated biomass observations. W values are obtained using (2.2) and parameters W_{max} (B), S (B), $\%N_C$ (C) and %N data (D), with a Gaussian noise with σ=0.08.*

and 0.29 (0.22, 0.36), versus 0.36 (0.34, 0.63) and 0.26 (0.20, 0.32) for the hierarchical framework, for the large and medium-sized datasets, respectively. Although relatively small, the differences between the two frameworks were great enough to have an impact on the fitted CNDC (Figure 2.2). The estimated values of $\%N_C$ were lower with the hierarchical framework compared to the sequential framework. For the medium maize dataset, biomass lower than 7.3 Mg ha^{-1} corresponded to $\%N_{Csequential}$ ¿ $\%N_{Chierarchical}$, and slightly opposite for W ¿ 7.3 Mg ha^{-1} (see curve intersection in Figure 2.3). The curve intersection was at 7.4 Mg ha^{-1} for the large dataset. We found that the estimated N uptake requirements for W between 2 and 10 Mg ha^{-1} (computed as the area between both curves in Figure 2.3), were 9.6 kg ha^{-1} greater for the sequential framework with the medium maize dataset, and were 5.9 kg ha^{-1} with the large dataset.

For the complete rice dataset, the estimated values of A_1 and A_2 were greater with the hierarchical than with the sequential framework (Figure 2.2D). Point estimates (95% CIs) of A1 obtained with the two frameworks were similar [$\hat{A}_1^{RICESEQUENTIAL}$ = 3.69 (3.38, 4.01) and Â1RICE HIERARCHICAL = 3.72 (3.27, 4.23)], but the relative difference between the estimates of A2 provided by the two frameworks was larger [Â2RICE SEQUENTIAL = 0.37 (0.32, 0.41) and Â2RICE HIERARCHICAL = 0.40 (0.34, 0.44)]. Because of the difference in the estimated values of A2, the CNDC obtained with the hierarchical model was lower for a large range of biomass values, resulting in lower estimated N requirements. Differences in critical N (calculated as sequential – hierarchical, expressed as kg N ha^{-1}) ranged from 4.73 kg N ha^{-1} (for W = 5 Mg ha^{-1}) to 18.7 kg N ha^{-1} (for W = 15 Mg ha^{-1}) (Figure 2.4). Total N uptake requirements, computed as the area between both curves (Figure 2.3), were 29.6 kg N ha^{-1} greater for the sequential framework, for W between 2 and 10 Mg ha^{-1}. For rice, the credible intervals (corresponding to the hierarchical framework) were wider than the confidence intervals (corresponding to the sequential framework), while an opposite result was obtained for maize.

Results obtained with the subsets of the maize and rice data (including half their ob-

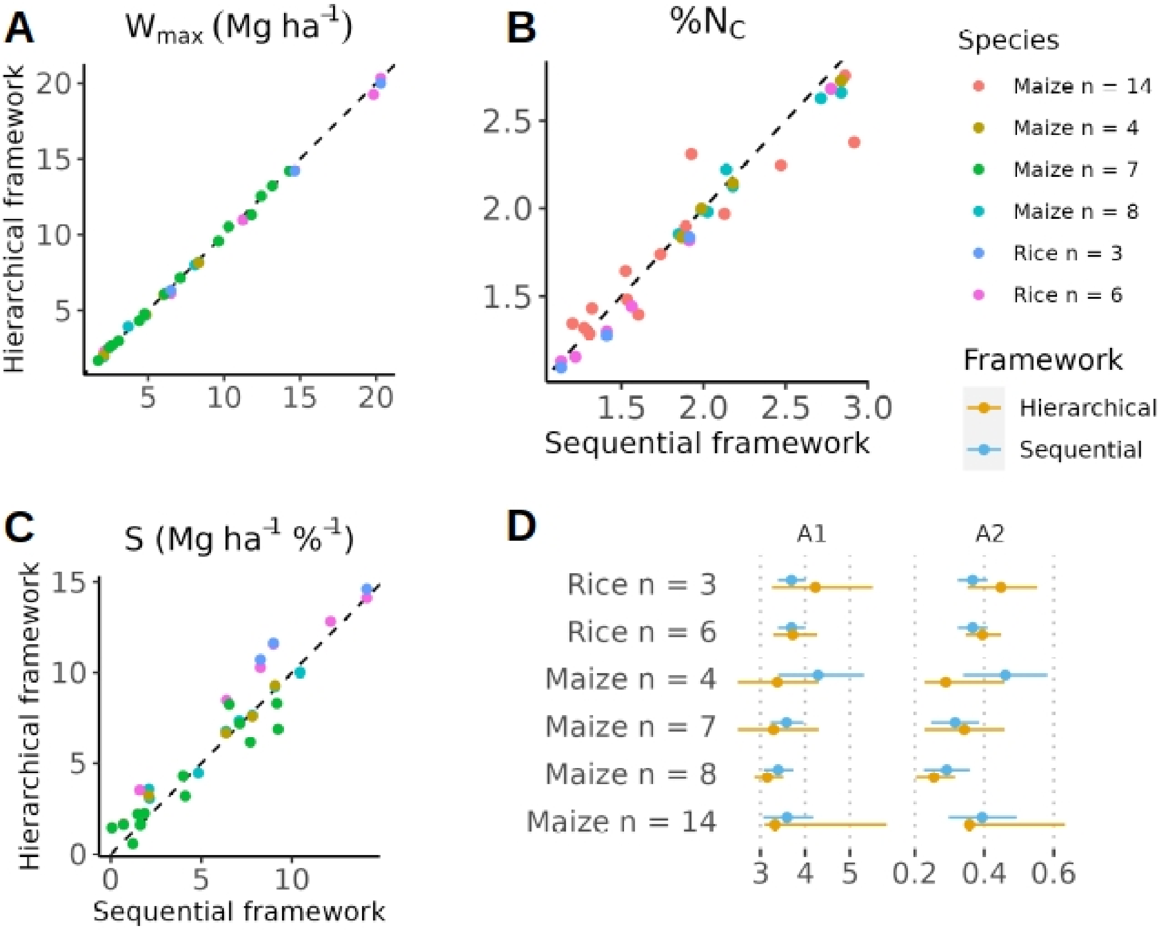

Figure 2.2: *Comparison of the values of maximum biomass, Wmax, (Mg ha^{-1}) (A), critical nitrogen concentration, $\%N_C$ (B), the slope of biomass versus $\%N$, S (Mg ha$^{-1}\%^{-1}$) (C), and parameters A_1 and A_2 (D) estimated with the sequential and hierarchical frameworks, for both all small-sized and complete maize and rice datasets. The diagonal dashed line in A, B and C indicates the 1:1 relationship.*

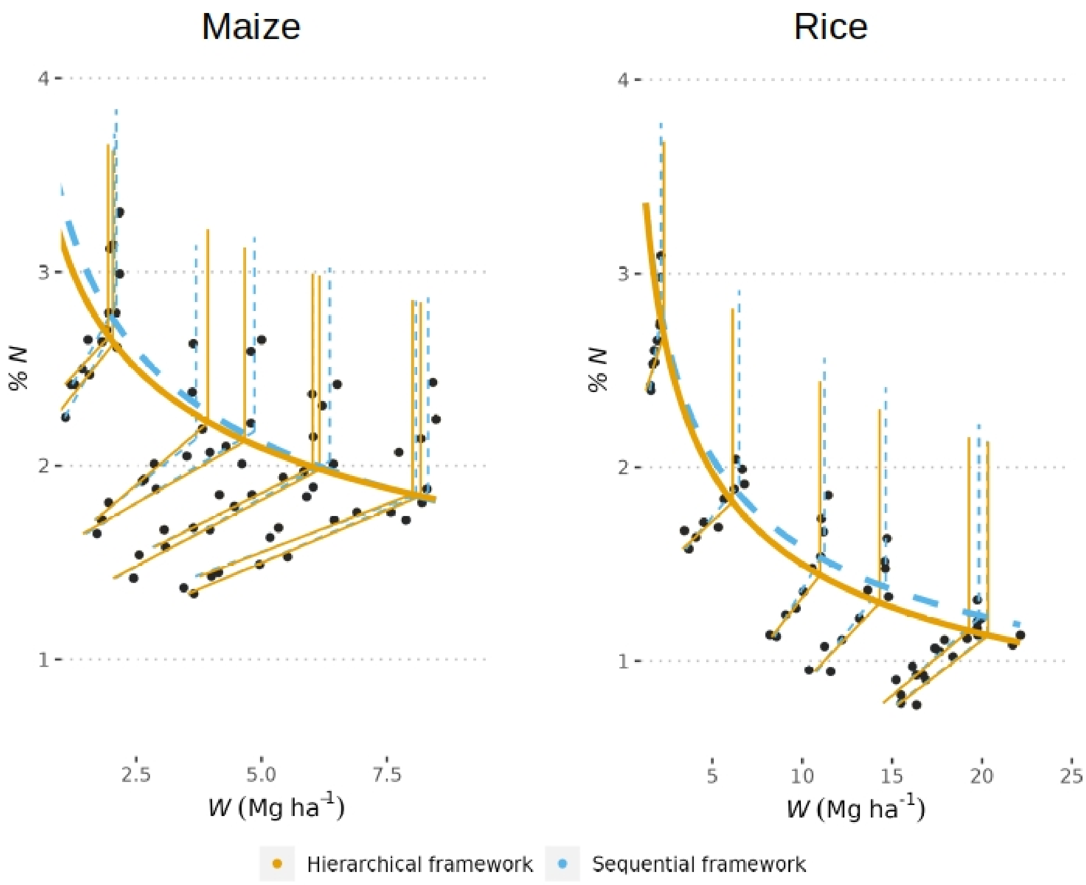

Figure 2.3: *Critical nitrogen (N) dilution curves portraying changes in N concentration (%N) at varying biomass (W) levels, Mg ha*$^{-1}$*, for both maize (Zhao et al. 2018) and rice (Yao et al. 2021) under the sequential (dashed blue lines) and hierarchical (full orange lines) frameworks (full datasets).*

servation times only) revealed that the number of observation times has an impact on the point estimates and on the level of uncertainty. With only 3, 4, or 7 observation times, the credibility and confidence intervals were wider compared to those obtained with the large datasets (Figure 2.2). Point estimates were also different when obtained with smaller datasets, especially for the rice due to a greater variation in biomass.

A more detailed analysis of the other parameter estimates (Wmax, S and $\%N_C$) revealed that stronger differences were found between the two methods for $\%N_C$ and S than for W_{max} (Figure 2.2). For both crops, estimated values of $\%N_C$ were higher with the

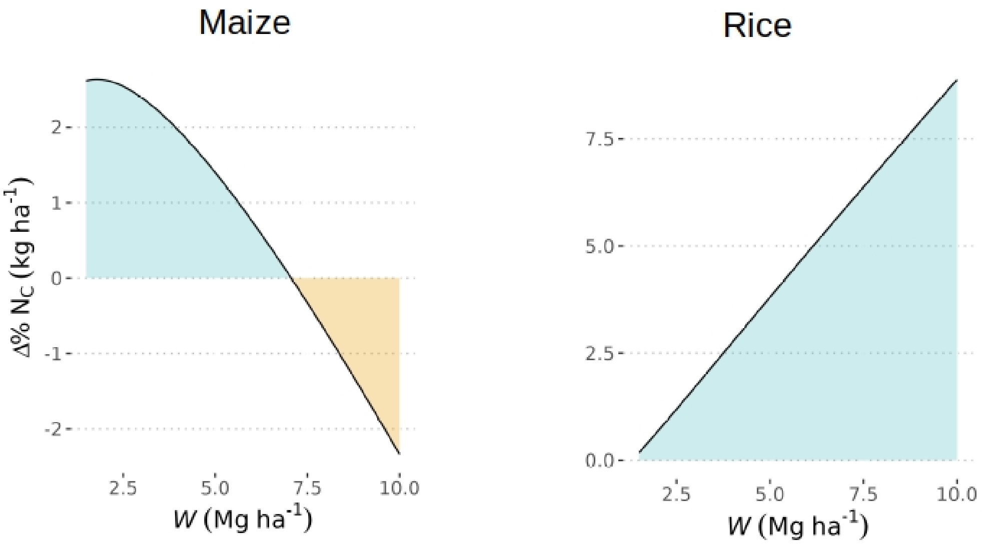

Figure 2.4: *Differences in estimated critical nitrogen (N) concentration,* $\%N_C$ *(calculated as* $\%N_{sequential} - \%N_{Chierarchical}$*, expressed in kg N* ha^{-1}*) as a function of crop biomass (W, Mg* ha^{-1}*), for both maize (medium dataset) and rice crops. Blue areas indicate positive differences (i.e., sequential* $\%N_C$ *estimates greater than hierarchical ones), and orange indicate negative differences.*

sequential framework, while the estimated slopes S were higher with the hierarchical framework, especially for rice. These results are consistent with the negative correlation existing between $\%N_C$ and S.

2.3.2 Simulation study

The analysis of maize and rice observed field data indicated that $\%N_C$ tended to be lower for a large range of biomass values when estimated with the hierarchical model, resulting in lower estimated N requirements. However, in order to know if these differences are systematic, and can be generalized to other situations, it is necessary to analyze the results obtained from the simulated data presented below. Another advantage of using simulated data is that they allow us to evaluate the accuracy of the estimated values of A_1 and A_2 and the reliability of their confidence intervals, which is not possible with observed field data because the true values of the parameters are unknown when observed data are used.

The simulation study showed that the hierarchical framework often outperformed the sequential framework with datasets of low quality, especially when the number of observation times is low or when the proportion of $\%N$ data exceeding $\%N_C$ is low or very low. With high quality datasets, both frameworks led to more similar performances, and the sequential framework could even sometimes lead to slightly less biased estimates than the Bayesian framework. Considering A_1 and A_2 together and all scenarios, the bias was lower in most (18/27) cases with the hierarchical framework, while the RRMSE was lower in all (27/27) cases with the hierarchical framework (Table 2.1). Moreover, while the hierarchical framework never failed to converge, the sequential framework failed in roughly 20% of the simulated datasets, which is an important practical limitation. Lack of convergence of the sequential framework was more frequent when the size of the dataset was small.

The hierarchical framework clearly outperformed the sequential framework for datasets including few (4) observation times or a low or very low proportion of $\%N$ exceeding $\%N_C$ ($\%N$¿$\%N_C$). For example, with datasets with only 4 observation times and the high proportion of $\%N$¿$\%N_C$, the RRMSE of A1 was reduced from 38.32% (sequential framework) to 10.14% (hierarchical framework) and the bias from 13.49% to -0.06%. This improvement was even greater when the number of data with $\%N$¿$\%N_C$ was very low; in this case, RRMSE of A1 was reduced from 92.30 (sequential framework) to 15.24% (hierarchical framework), and the bias from -84.57% to -1.2% (Table 2.1). The differences between frameworks were smaller with large datasets. In fact, the bias was slightly lower with the sequential framework for large datasets and medium proportion of $\%N$¿$\%N_C$, but the RRMSE was always lower with the Bayesian framework, in all scenarios (Table 2.1).

With the sequential framework, roughly 2% to 70% of the observation dates (depending on the design) were removed on average (Table 2.1), either because the optimization algorithm did not converge trying to fit the linear-plus-plateau (Equation 2.1), or AIC indicated that the simple linear model was better. Under the same range of $\%N$ data, the simulations with three N fertilizer rates had the greatest proportion of data removal, i.e., always equal or greater than 52.7%. Thus, simulations with three N rates had the greatest reductions in sample size to estimate A1 and A2. The number of observation times had a great impact on inference properties, because of its effect on sample size. Bias was thus reduced tenfold by only doubling the number of observation times. Although less influential, higher number of N rates led to lower RRMSE and bias of A1, A2 and $\%N_C$. Increasing from three to four fertilizer rates led to a stronger reduction of RRMSE and bias than increasing from four to five N rates (Table 2.1).

			PBE				RRMSE			
(l)4-11 Proportion of %N>%Nmax	% of data removed in the filtering process	Number of observation times	A1		A2		A1		A2	
			Hierarchical framework	Sequential framework	Hierarchical framework	Sequential framework	Hierarchical framework	Sequential framework	Hierarchical framework	Sequential framework
(r)1-3 Medium	0.17	4	-0.06	13.49	1.07	14.2	10.14	38.32	10.14	38.32
	0.16	8	2.58	0.61	2.13	0.88	6.15	7.15	6.15	7.15
Medium	0.16	16	3.23	1.04	2.6	1.17	3.69	4.95	3.69	4.95
	0.49	4	-1.92	65.82	2.15	66.18	13.91	81.63	13.91	81.63
Low	0.44	8	-1.16	5.69	0.45	6.36	12.56	27.4	12.56	27.4
	0.44	16	0.81	0.51	0.76	0.58	4.55	6.08	4.55	6.08
	0.61	4	-1.2	84.57	4.76	84.7	15.24	92.03	15.24	92.03
Very low	0.61	8	0.75	32.95	1.53	33.32	12.99	58.36	12.99	58.36
	0.62	16	2.24	3.56	1.48	3.58	8.19	19.06	8.19	19.06
(l)2-11										

Table 2.1: *Relative root mean squared error (RRMSE) and percentage bias error (PBE, %) for A_1 and A_2 with the sequential and hierarchical frameworks, for different types of datasets with five N rates.*

Proportion of %N>%NC	Number of observation times	Number of N fertilizer rates	NS-H (kg ha-1)	%NC bias (%) Sequential framework	 Hierarchical framework	%NC RRMSE (%) Sequential framework	 Hierarchical framework
Very low	N = 4	3	1.70 (-8.10, 10.02)	1.29	-0.15	1.3	0.21
		5	0.88 (-6.68, 9.69)	1.17	-0.09	1.27	0.2
	N = 16	3	1.79 (-3.40, 8.74)	0.23	0.02	0.02	0.03
		5	0.79 (-3.29, 7.18)	0.05	0.01	0.24	0.03
Medium	N = 4	3	2.37 (-9.94, 11.01)	1.23	-0.08	1.33	0.28
		5	-0.57 (-6.95, 6.57)	0.18	-0.01	0.51	0.04
	N = 16	3	2.85 (-0.41, 6.48)	-0.01	0.02	0.56	0.06
		5	0.95 (-1.39, 3.78)	0	0.01	0.01	0.01

Table 2.2: *Difference between required kg ha1 N to attain* $\%N_C$ $(\Delta NS - H)$, *estimated as* $\frac{\%N_C^{SEQUENTIAL} - \%N_C^{HIERARCHICAL}}{100} \cdot W$, *and bias and root mean squared error (RMSE) in* $\%N_C$ *estimation (%). The biomass* W *was set equal to 10 Mg* ha^1. *Results are presented for contrasted characteristics of simulated datasets.*

The sequential framework estimated higher $\%N_C$ and N requirement values in the simulation study for 22 out of the 27 designs (Figure 2.5, Table 2.2), thus confirming the results obtained with the real maize and rice datasets. The difference ranged from -0.57 kg plant N ha-1 up to 2.85 kg plant N ha-1 at a given moment, depending on the dataset characteristics. At the same time, the estimated values of $\%N_C$ had larger bias and root mean squared error (RMSE; Table 2.2). Overall, our simulations showed that the sequential framework led to higher and less reliable N requirement estimates in most cases.

Compared to the sequential framework, the hierarchical framework had better coverage probability (i.e., closer to 0.95) with low-quality datasets (Table 2.3). The number of observation times was always positively related to coverage probability, and designs including higher proportions of data with $\%N > \%N_C$ showed better coverage probabilities for both frameworks. The sequential framework was slightly favored over the hierarchical framework with datasets including the medium proportion of $\%N¿\%N_C$ and at least eight observation dates. However, the sequential framework showed too low coverage probabilities in all other the scenarios, especially when the number of observation times was equal to four. In this case, the coverage probability of the sequential framework was ranging from 0.08 to 0.58, and was thus much lower than the target value 0.95. Under these scenarios, the hierarchical framework was able to achieve better coverage probabilities, although sometimes several points lower than the nominal value (95%), thus showing that coverage was not always optimal with this framework.

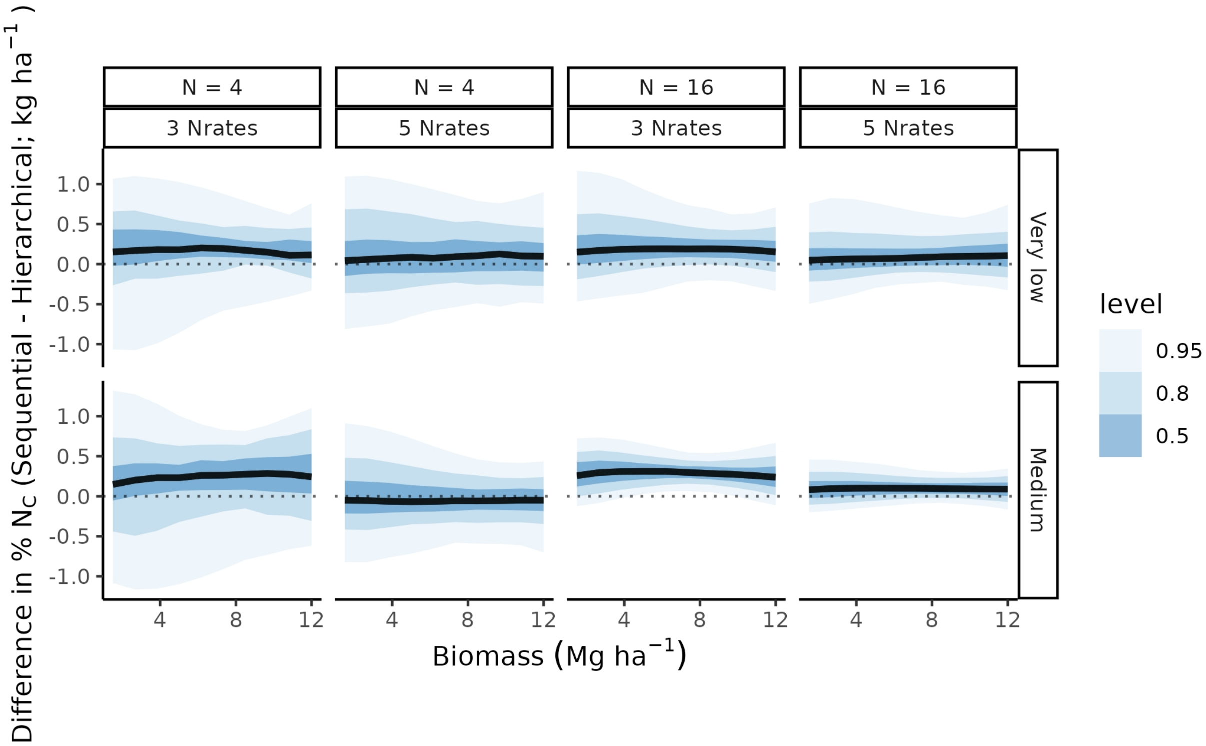

Figure 2.5: *Difference between the two frameworks in the estimated N requirement to attain $\%N_C$ between the two frameworks, computed as sequential – hierarchical (kg plant N ha-1), for a selection of simulated dataset characteristics: very low and medium proportions of $\%N > \%N_C$, 4 or 16 observation times ($N = 4$, $N = 16$), and three or five N fertilizer rates (3 Nrates, 5 Nrates). The colored bands indicate the quantiles 0.50, 0.80, and 0.95 of the simulations. The solid line indicates the median.*

Proportion of %N>%NC	% removed data	Number of observation times	*A1* Sequential framework	Hierarchical framework	*A2* Sequential framework	Hierarchical framework
Medium	0.17	4	0.58	0.91	0.58	0.92
	0.16	8	0.92	0.86	0.95	0.88
	0.16	16	0.9	0.77	0.94	0.83
Low	0.49	4	0.19	0.91	0.19	0.92
	0.44	8	0.74	0.81	0.77	0.84
	0.44	16	0.85	0.9	0.91	0.9
Very low	0.61	4	0.08	0.92	0.08	0.92
	0.61	8	0.48	0.85	0.5	0.85
	0.62	16	0.82	0.81	0.86	0.84

Table 2.3: *Coverage probability of the 95% confidence/credible intervals for A_1 and A_2 parameters, for the sequential and hierarchical frameworks fitted to simulated datasets with different experimental designs based on five N fertilization rates.*

2.4 Discussion

The statistical framework selected for CNDC estimation has a strong impact on the estimated critical N dilution curves and on the associated estimated crop $\%N_C$ and N requirements. The implementation of the sequential framework could be a potential source of inaccurate N diagnosis, economic losses, and negative environmental impacts because it tends to produce unreliable CNDC parameter estimates and $\%N_C$ estimates, sometimes resulting in greater estimations of crop N requirements. These results are supported by our analyses based on both real field data and simulated data. Thus, we recommend using the hierarchical framework based on the Bayesian model of (Makowski et al. 2020), or one of its variants. This framework provides more robust results, especially with small datasets and/or when the proportion of $\%N$ observations exceeding $\%N_C$ is low. In concordance with previous studies (Fernandez et al. 2021; Jégo et al. 2022), the implementation of this approach will tend to estimate lower N requirements. The consequences for fertilizer rate recommendations are difficult to quantify because they depend on the method of fertilizer rate calculation and decision support tools used by farmers and their advisors. However, our results show that the $\%N_C$ values calculated with the Bayesian framework tend to be lower in most of the cases considered, which may lead to lower fertilizer rates regardless of the decision support tool implemented. In fact, the area between both curves, as reported in the results, can be interpreted as a proxy for the difference in N fertilizer recommendations. Our estimates for the required amount of N needed for the crop to sustain N sufficiency levels would increase to about a difference of 30-100 kg N fertilizer ha^{-1} more under the sequential framework, considering a fertilizer N use efficiency of 30% in cereals (Raun and Johnson 1999). The hierarchical framework could thus potentially contribute to increasing sustainability in agricultural systems.

We found that the sequential framework provides unreliable estimates under low-quality data scenarios, mainly because of small sample size bias of maximum likelihood estimators and as a consequence of ignoring uncertainty for $\%N_C$ and W_{max} estimation. The worst scenario for the sequential framework corresponds to datasets with small sample size (e.g., three N fertilizer treatments), or with a narrow range of $\%N$ that does not allow exploring high levels of biomass and obtaining reliable estimates for $\%N_C$ and W_{max}. In this scenario, the sequential framework provides highly biased estimations and very poor coverage probability (i.e., there is a low probability a 95% confidence interval covers the true value). Perhaps the major factor causing these issues is the practice of taking point estimates (in this case, $\%N_C$ and W_{max}) obtained from one statistical model and "plugging" these estimates into another statistical model, without properly accounting for uncertainty. Even though $\%N_C$ and W_{max} have an associated estimation uncertainty, they are routinely used as if they were collected as data in (2.1) (i.e., treated as a known measurable quantity). Consequently, estimation uncertainty is neglected and this tends to yield ill results (Anupreet and Raftery 2022). The hierarchical framework does not fit the linear-plus-plateau model for each observation time independently, but finds the best combination of parameter values to fit to the data as a whole and considers the uncertainty associated to all parameters. Moreover, the hierarchical model has the advantage of using the Bayesian theorem and can thus integrate previous experts' knowledge, summarized in the prior distributions that allow to fit the model even when the data are few. For a long time, Bayesian methods have had the reputation of being more difficult to apply than classical statistical methods. However, with the increase in computing power and the development of new R packages, Bayesian methods can now be easily coded and run in a few seconds with standard computers. Moreover, the Bayesian framework tested here did not encounter any convergence problem, even with small datasets, whereas the sequential framework often encountered convergence issues and therefore cannot be used in all circumstances. The issues of the sequential framework outlined here are expected in any other setting that fits two statistical models sequentially, which is a common, but perhaps poor, practice in agronomy (Pagano and Maddonni 2007; Vega et al. 2001). In a nutshell, the most important factors that can explain the low performance of the sequential framework are small sample size bias of maximum likelihood estimators, and disregarding the estimation uncertainty.

An additional benefit of the hierarchical framework for the case of the CNDC is that it does not require removing the data collected at the observation times where Wmax is not reached. Although the two frameworks were compared here using exactly the same number of observation times after data filtering, the hierarchical framework could be applied to unfiltered datasets, given that it does not require to remove the data from those observation times where Wmax is not reached (Fernandez et al. 2021; Jégo et al. 2022; Makowski et al. 2020). Thus, the adoption of the hierarchical framework may use much larger datasets than the sequential framework, and consequently reduce estimation uncertainty that could lead to potentially more accurate results. In fact, in our simulations, a relatively low level of observation errors was considered to simulate data. Real-world experiments may face conditions of high variance of observation errors and that can be even more detrimental for the sequential framework. Thus, our results may be slightly favoring the sequential method. We expect that the performances reported here for the Bayesian framework are actually underestimated; the bias and RRMSE and probability coverage could probably be further improved by applying the hierarchical framework to the full datasets without filtering. This idea will be explored in future works, as the sole objective of this paper is to revise the practice of fitting statistical models sequentially in depth.

Further investigations should address questions related to experimental design, data filtering methods, and further expansions of the model that may help improve inference possibilities. First, the optimization of the experimental design to increase accuracy in CNDC is still an open question. Fernandez et al. (2022) point at the best design characteristics featured in the literature, but their optimality still needs to be validated. It is still unknown how much could be gained using optimal design compared to standard designs. This is an open question. In addition, optimal designs are likely to be different with a Bayesian framework than with a sequential framework because the Bayesian approach is more flexible and can be applied to diverse datasets. Second, as mentioned before, the more efficient use of the data with a Bayesian hierarchical model should improve the inference. Last, the Bayesian model described in Makowski et al. (2020) could be expanded by including additional levels that would allow making inference for groups of cultivars, species, or locations with unbalanced data. This may contribute improving CNDC estimation with small datasets, especially for rare cultivars or understudied species.

2.5 Conclusions

Although it is very commonly used by agronomists, the sequential framework performs worse than the more recent hierarchical framework in most cases. The latter has generally lower root mean squared error and bias and better coverage probability, and should thus be preferred for estimating critical N dilution curves, especially with small datasets. The sequential framework may sustain appropriate inference properties for large datasets but, with small datasets, its use could lead to inaccurate diagnosis of plant N status and overestimation of crop N requirements. The hierarchical framework provides more reliable critical N dilution curve, better estimated values of plant N requirements, and allows for a thorough analysis of uncertainty.

Chapter 3

Rethinking statistical practices fitting nitrogen dilution curves as a case study for crop nutrition models

Abstract

Crop nutritional diagnostic models are critical for fertilizer management in sustainable agriculture. Usually, nutritional diagnostics rely heavily on the estimates of the statistical models, which highlights the importance of the statistical methods and the quantification of uncertainty. This important role of the statistical methods has motivated the recent revision of widely implemented crop nutrition models, and many users now consider Bayesian statistics an accessible platform for propagating uncertainty. We present the critical nitrogen (N) dilution curve (CNDC) as a case study of a nutrition model that underwent significant changes regarding the application of statistical methods. The quantitative methods for estimating CNDC have shifted to a Bayesian hierarchical modeling framework, allowing us to re-think the practices that made sense under the traditional framework but may not anymore. Thus, the objectives of this study were (i) to rethink data pre-processing techniques, and (ii) to demonstrate the utility of propagating and quantifying estimation uncertainty in model fitting and interpretation tasks. We show that the standard method for pre-processing (i.e., *cleaning*) the data should be avoided because it may lead to risks of biased estimates and produces wider credible intervals for the parameter estimates. We conclude saying that we need to build statistical workflows that propagate, quantify, and communicate uncertainty in statistical models to obtain the best inference possible.

Keywords: NNI; nitrogen nutrition index; Bayesian models.

3.1 Introduction

Crop nutritional diagnostic models are critical for fertilizer management in sustainable agriculture. Nutritional diagnostics allow providing enough, yet not excessive, nutrients by allowing to adjust the rate and timing of fertilizer applications according to the crop's current requirements. Crop nutrient requirements accompany crop growth (Lemaire et al. 2008) and thus, current guidelines recommend periodic diagnostics of crop nutritional status (Ravier et al. 2021). Said recommendation guidelines rely heavily on estimates of statistical models that provide reference values, expected under conditions of nutritional sufficiency. Any condition below the expected value for nitrogen (N) sufficient crops is usually considered deficient and would require agronomic intervention (i.e., fertilizer application). For example, some reference models represent the allometric relationships between W accumulation and nutrient uptake and are usually used for determining the crop nutritional status (Plénet and Lemaire 2000; Justes et al. 1994). Thus, crop nutritional status is usually described by current biomass (W) and tissue nutrient content ($\%N$). If the $W - \%N$ combination does not correspond to a N sufficient scenario, that crop is considered deficient.

In this article we focus on one particular reference model used to assist N management: the critical nitrogen dilution curve (CNDC). The CNDC is used to determine the nitrogen nutrition index (NNI), one of the most important metrics to describe the crop N status. Since the first introduction of the CNDC, much of the methods to estimate the model parameters have changed. Originally, Justes et al. (1994) introduced a methodology that required fitting two statistical models sequentially, using the estimated parameters of the first one, as data in the subsequent model. Following, Makowski et al. (2020) introduced a Bayesian hierarchical model that estimates all parameters in a single step. Specific details about the CNDC statistical model are explained more in detail below. At this point, it is worth mentioning that the number of users of the hierarchical model (i.e., Makowski et al. 2020) is increasing steadily, and is replacing the other method (Fernandez et al. 2022).

In this study, we consider the recent paradigm shift is an opportunity to rethink old practices for data management and analysis. Our specific objectives include rethinking the standard workflow to fit CNDC. More specifically, we aim (i) to rethink data pre-processing techniques, and (ii) to demonstrate the utility of propagating and quantifying estimation uncertainty in model fitting and interpretation tasks.

3.2 The statistical model behind the nitrogen nutrition index

As mentioned above, the NNI is derived from the CNDC, a model that describes the allometric relationship between W and the N concentration ($\%N$) needed to maximize biomass (the critical $\%N$, $\%N_C$). More specifically, the CNDC describes the $\%N_C$ of the ith observation moment as a function of maximum W at the same observation moment ($W_{max\ i}$). In this way, the CNDC establishes the reference values for a crop that has sufficient N to sustain maximum growth:

$$\%N_{C\ i} = A_1 \cdot W_{max\ i}^{-A_2}, \tag{3.1}$$

where A_1 and A_2 are two parameters that must be estimated with data of biomass and nitrogen content. The CNDC thus allows finding the $\%N_C$ for any level of W (i.e., at any point during vegetative crop growth) by computing $\%N_{C\ i} = A_1 \cdot W_i^{-A_2}$. The NNI is the ratio between the crop actual $\%N$ ($\%N_{actual\ ij}$) and $\%N_{C\ i}$, and is used to indicate when a crop is N deficient (i.e., if $N_{actual\ ij} < \%N_{C\ i}$, then NNI<1 and the crop is considered N deficient).

The hierarchical model in Makowski et al. (2020) can be conceptualized as the connection of three simple models (or levels): the data model, the process model and the parameter model. The data model describes how the data arise in the field, following a linear-plateau relationship between W and $\%N_C$:

$$W_{ij} \sim N(\mu_{ij}, \sigma^2) \tag{3.2}$$

$$\mu_{ij} = \begin{cases} W_{max\ i} & \text{if} \%N_{ij} \geq \%N_{C\ i} \\ W_{max\ i} - s_i \cdot (\%N_{ij} - \%N_{C\ i}) & \text{if} \%N_{ij} < \%N_{C\ i} \end{cases}, \tag{3.3}$$

where σ^2 is the variance and μ_{ij} ($i = 1, ..., J$ and $j = 1, ..., n$, where J is the total number of observation moments and n is the total number of observations) is the expected W of the ith observation at the jth time. Then, the expected value μ_{ij} is described with a linear-plateau model with maximum biomass W_{maxi}, slope s_i and critical $\%N$ value $\%N_{Ci}$. The process model describes the non-linear relationship between $\%N_{C\ i}$ and $W_{max\ i}$ (i.e., the CNDC in (3.1)). Last, the parameter model describes the prior knowledge regarding parameters A_1 and A_2 in (3.1), $W_{max\ i}$ and s_i in (3.3), etc. Note the practical importance of correctly estimating parameters A_1 and A_2: they will determine N status and thus define the need for N fertilizer.

The major contribution of Makowski et al. (2020) (i.e., the main difference between the standard framework and the modern framework) is the correct propagation of uncertainty in the statistical model. Previously, the standard method in Justes et al. (1994) required fitting two models sequentially, using the estimates of one model (i.e., $\%N_{C\ i}$ and $W_{max\ i}$ in (3.3)) as data in another model (i.e., (3.1)). This sequential approach to fitting the CNDC required filtering out all i observations where $\%N_{C\ i}$ and $W_{max\ i}$ were not clearly detectable. In contrast, Makowski et al. (2020) connected all the parameters in a single hierarchical model. It is beyond the scope of this paper to provide an in-depth explanation of the standard framework and its differences to the modern framework (see Fernandez et al. 2021, and Lacasa et al. 2023). The hierarchical structure of the model thus provides a framework to avoid wasting efforts in unnecessary data removal, and increasing the dimensionality of the statistical model. In what follows, we rethink the standard workflow for fitting CNDC and providing nutritional diagnostics.

3.3 Field Data

Some of the following sections include data examples, that are publicly available and contain information on experiments studying critical nitrogen dilution curves for several crops (Ciampitti et al. 2022). We selected all data on maize (*Zea mays* L.), and selected only observations with $W > 1$, according to Plénet and Lemaire (2000). The data include observed pairs of W_{ij} and $\%N_{ij}$ in (3.1) and (3.2), for each treatment and observation moment.

We divided the maize data into fitting data and testing data. One experiment where most observation moments clearly reached a W_{max} plateau (Zhao et al. 2018) was selected for testing (Figure 3.1b). That specific experiment was selected because, since the W_{max} plateau for most observation moments was estimable, it was appropriate to classify treatments as 'N deficient' or 'N sufficient' for subsequent evaluations. To obtain observations classified as deficient or non-deficient to use as reference for testing, we fitted a linear-plateau model (outlined in (3.3)) to each observation moment using the package 'nls' in R (R Core Team 2019). Then, we obtained estimates for $W_{max,j}$, s_j and $\%N_C, j$ and classified each observation as N sufficient when $W_{ij} \geq W_{max,j}$, or N deficient otherwise. These classifications were used as the 'golden standards' for evaluating the model fitting techniques.

We examine the influence of standard 'data pre-processing' procedures on statistical inference. We fitted the model described above to the fitting data set twice: first using traditional data pre-processing methods and then without said step, as proposed above.

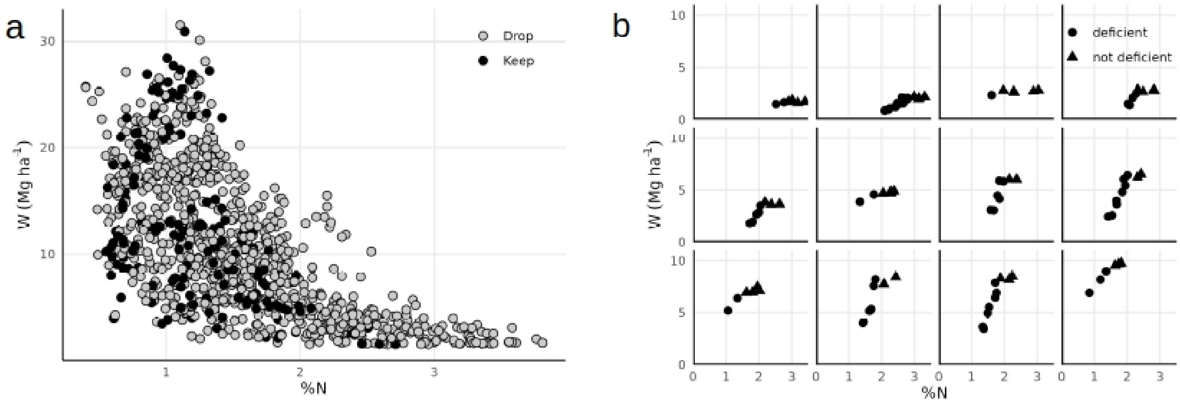

Figure 3.1: *Visual description of training data (a) and testing data (b). In the left panel, each point represents an observation of N concentration (%N) and biomass (W). Point color indicates wheter the observation will be dropped in a data pre-processing step. In the right panel, each subplot represents an observation moment of biomass (W) and N concentration (%N) of the testing data. Point shapes indicate whether the observation has reached maximum W (not deficient) or not (deficient).*

3.4 Current methods for pre-processing the data before model fitting

We review the standard practice of pre-processing the data to demonstrate how Bayesian hierarchical models are more suitable for applications with low-informative data (i.e., some observation moments do not have much information about W_{max} in (3.2). Recall that the standard approach for pre-processing the data for CNDC estimation involves removing all observation times where $W_{max,j}$ is not detectable (Fernandez et al. 2022; Fernández et al. 2021). It is worth noting that the previous standard method required such a filtering step because it needed W_{max} and $\%N_C$ for each observation time to be estimable, since they were used as the independent variable (i.e., as data) in a different model fitting (3.1). However, a hierarchical framework does not require the pre-processing data step because information is shared among different levels of the model.

We demonstrate the difference between two approaches for handling the field data: the standard criteria pre-processing the data, versus using the complete data. Following the standard workflow, we fit two models to the data from each observation moment, and compared them using the Akaike Information Criterion (AIC). We fit a simple intercept and slope model, and a linear-plateau model. We removed all data where the linear-plateau model did not converge or was not preferred according to AIC. Thus, we obtained two datasets: one filtered and one complete. We estimated A_1 and A_2 for each data set by fitting the model described by Makowski et al. (2020), using uninformative priors.

We found that the pre-processing data has a strong influence on both the expected value and variance of the estimates. The estimates of A_1 and A_2 were distributed along greater values for the model fitted using the complete dataset (Figure 3.2). The width of the credible interval for both parameters was also larger. For A_1, we found a mean (95% credible interval) of 3.70 (3.49, 3.92) for the model fitted to the complete dataset, versus 2.88 (2.56, 3.23) for the filtered dataset. Likewise, for A_2, we found 0.43 (0.40, 0.45) for the model fitted to the complete dataset, versus 0.36 (0.32, 0.40) for the filtered dataset.

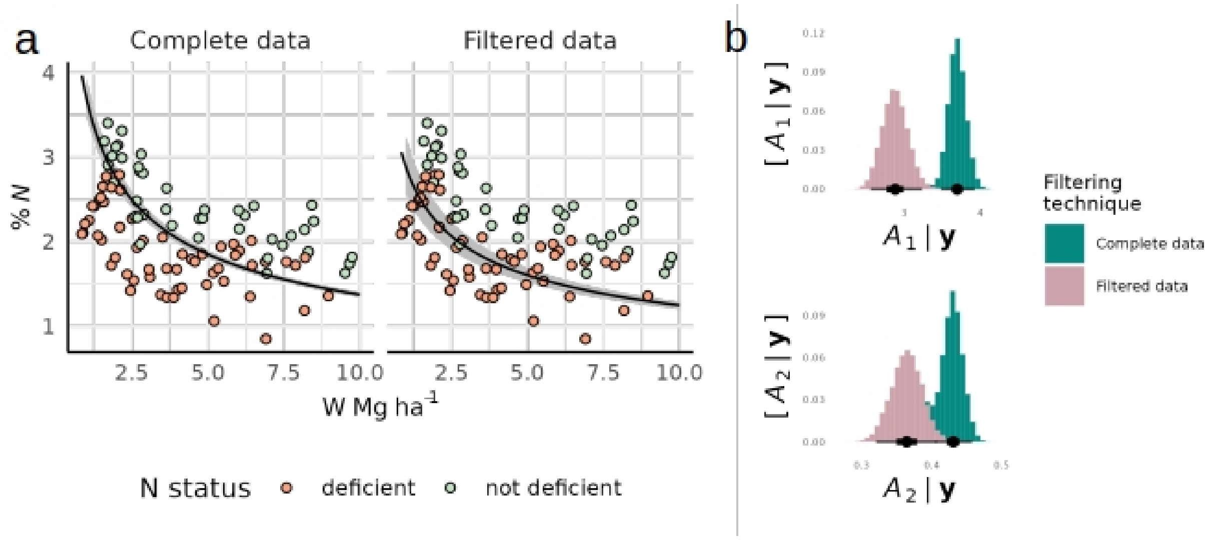

Figure 3.2: *Effect of data pre-processing on results. (a) Fitted curves using the complete dataset (left) versus the filtered dataset (right). (c) Posterior distributions of parameters A_1 and A_2 of the Critical Nitrogen Dilution Curve for models fitted to the complete dataset (green) versus the filtered dataset (pink)*

3.5 On the propagation of uncertainty for nutritional diagnoses and yield predictions

We illustrate the utility of hierarchical models for propagating and quantifying uncertainty easily. First, we focus on quantifying the uncertainty of crop nutritional status. We then elaborate the hierarchical model adding an additional grain yield level, to show diagnostics of having a nutritional deficit with impacts on grain yield.

First, we compare computing a point estimate of NNI, versus the posterior distribution of NNI. The output of a Bayesian model consists of samples from the marginal posterior distribution of A_1 and A_2, and thus a set of probable values of $\%N_C$ can be easily derived. An essential attribute of the Bayesian model is the propagation and quantification of uncertainty. In this case, that means NNI with uncertainty incorporates the estimation uncertainty for providing nutritional diagnostics. We compute the point estimate of NNI and the posterior distribution of NNI:

- **Point estimate of NNI**: This is the Bayesian equivalent to the NNI proposed in Justes et al. (1994). NNI_{ij} was calculated as $\frac{\%N_{actual,ij}}{\widehat{\%N_{C\ i}}}$. The crop was declared N deficient if NNI < 1. Note that the denominator, $\widehat{\%N_{C\ i}}$, is the posterior mean of $\%N_{C\ i}|\mathbf{y}$, the samples from the posterior distribution.

- **Posterior distribution of NNI**: Samples from the NNI posterior probability distribution were calculated as $\frac{\%N_{actual,ij}}{\%N_{C\ i}|\mathbf{y}}$. The crop was declared N deficient if $\mathrm{NNI}_{\mathrm{p0.975}} < 1$. Note that $\%N_{actual,ij}$ is a single scalar, while $\%N_{C\ i}|\mathbf{y}$ are samples from the posterior distribution.

Our results show that the quantification of uncertainty to estimate NNI provides fairer estimates to describe N status. For example, about 24% observations that would be classified as ‘N deficient’ under a deterministic framework fell within the 95% CI that indicates sufficiency (Figure 3.2).

In our second example of uncertainty propagation, we include a yield level. Grain yield is usually the most important outcome for stakeholders, yet it is not present in the CNDC (only biomass, which is a proxy for yield). To establish a connection between the CNDC and yield, many studies fit models that use NNI at anthesis for making yield forecasting and fertilizer prescriptions. The NNI around flowering is usually positively related to yield until reaching a plateau at NNI values of 0.9 (Ravier et al. 2017). For simplicity, the relationship between NNI and yield is modeled with a linear-plateau equation (Ziadi et al. 2008). Note that other NNI thresholds or models may be considered when targeting a combination of grain yield and quality (Ravier et al. 2017). However, the issue with the NNI-yield models is that the do not propagate uncertainty and may limit the inference. Thus, we incorporate yield into the model to demonstrate how uncertainty is automatically propagated by simply adding a level in a hierarchical model.

Integrating the yield formation components is very relevant for agronomic management because it enables early assessments of N limitations for yield production, with a realistic quantification of uncertainty, alowing risk assessment. Early vegetative assessments may be associated to greater uncertainty than assessments close to flowering (Yao et al. 2023). Moreover, maintaining the estimations under the same hierarchical model allow the propagation of the uncertainty for a more realistic assessment on the confidence of observing yield penalties.

The addition of the yield level to the hierarchical model results in a more throughout propagation of uncertainty. It is no surprise that the uncertainty estimates were relatively large. Note that the threshold for almost surely no yield loss is close to a NNI of 0.9 (Fig. 3.3).

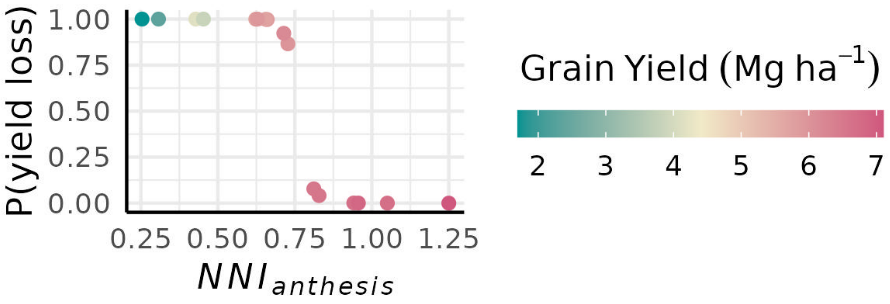

Figure 3.3: *An assessment of NNI impacts on yield. Point color indicates observed grain yield (Mg ha^{-1}). This plot represents an example for the probability of yield loss versus NNI at anthesis for wheat crops.*

3.6 Discussion

In our context of a shifting paradigm in statistical modeling, we propose rethinking common methods in crop science with the final objective of improving the quality of our inference. We present the case of the CNDC as an example of a model that underwent several methodological changes in a narrow time frame, challenging crop scientists to adjust their quantitative skills. The CNDC is a particularly good example because it consists of several components connected in a rather complex way (e.g., the connection of (3.2) and (3.3)), tempting users to take methodological shortcuts (e.g., fitting two models in sequence instead of a hierarchical model), yet at the cost of reducing the quality of the statistical inference (Lacasa et al. 2023). In this article we demonstrate that an honest quantification and reporting of uncertainty estimates is crucial in crop nutrition models. This means, not only reporting NNI uncertainty estimates, but also propagating errors properly.

The CNDC is a good example portraying scientific progress as a sequence of reviewing, improving, and testing a given topic (e.g., methods in crop nutrition). Considering the time frame since the first development of methods for estimating the CNDC in 1994 (Justes et al. 1994), the increase in computational power of mainstream commercial computers has enabled the development of computationally-intensive analyses. Thus, Makowski et al. (2020) introduced a Bayesian hierarchical model to estimate CNDC, as a technique to propagate uncertainty across the estimands. Two years later, Fernandez et al. (2022) reported its vast application and success. For this reason, Lacasa et al. (2023) compared it to the older method and showed that the Bayesian model is usually better calibrated, more efficient, and more robust to noisy data. Later on, Bohman et al. (2023) elaborated the model under a popular R package, incorporating also further levels. Likewise, in this article, we illustrate other advantages of estimating CNDC under a hierarchical Bayesian framework.

We demonstrated that the hierarchical model does not require an initial data filtering step and allows both a more efficient use of field data, and an adequate and straight-forward framework for propagating uncertainty. In the CNDC, there are several 'latent variables' that are not observed, like W_{max} and N_C. Some data contain more information for estimating W_{max} and N_C than other. For example, experiments with more N rates and covering both areas below and above N_{Cj} in (3.2), contain more information for estimating the parameters in (3.2). Before the hierarchical model in Makowski et al. (2020), these 'latent variables' were estimated and then used as data fitting a model with (3.3). That case depended to have enough information in each observation moment for estimating W_{max} and N_C, or else all data from that observation moment had to be discarded. Under a hierarchical framework, observation moments that are rich in information regarding W_{max} and N_C (i.e., those that clearly reach the plateau) improve inference for observation moments lacking information to detect W_{max} and N_C (i.e., those that do not reach the plateau), by sharing this information among levels of the model. The problem in hand can be defined also as a missing data problem, where the latent variables W_{max} and N_C in (3.2) are always missing (at random). Then, the standard method of discarding the data is one way of handling missing data. This method, however, may waste an important amount of field effort and may introduce bias in the estimation of $A-1$ and A_2 in (4.1). The newer method including a Gibbs sampler thus represents a more efficient way of performing multiple imputation (Little and Rubin 2019).

The new Bayesian hierarchical framework has provided researchers with a framework for a straightforward propagation and quantification of uncertainty. For example, we have elaborated the Bayesian model with a further level (i.e., yield). As yield response to N is highly influenced by seasonal weather conditions (Correndo et al. 2021), a Bayesian model provides a robust framework for quantifying and reporting risk to be considered for management decisions. With the amount of claims for assessing risk in agricultural production, a more honest quantification of uncertainty allows a more honest risk assessment for stakeholders.

The joint development of a Bayesian framework with an adequate propagation of uncertainty has paved the way for embracing uncertainty in crop nutrition models. In our experience, it is uncommon to consider uncertainty for providing nutritional diagnostics. For example, NNI estimates are usually compared to single points (e.g., a threshold of 0.9, Ravier et al. 2017). Some experts have suggested the notion of uncertainty in NNI estimation (e.g., Lemaire et al. (2008) stating that values 'close to 1' indicate N sufficiency), yet criteria for evaluating nutritional status does not contemplate this point yet. However, uncertainty is a core variable in decision theory, especially in the context of agriculture and the interplay between both the farmer's background knowledge (Huffman 2020) and the environmental conditions. Furthermore, recent studies have shown the importance of considering the probability of failure as the risk in technology adoption (Chavas and Nauges 2020) and how this portrays the need to go beyond mean–variance analysis and to examine the effects of new technologies on the lower tail of the payoff distribution. This information is currently lacking, and Bayesian assessments may be powerful tools to help filling this gap. With the shift of the statistical framework we may be moving towards a workflow that accounts for this point. Bayesian models provide an accessible framework to quantify uncertainty even in complex processes like the CNDC. We envision methodological developments taking place in the following years, where not only uncertainty quantification but a correct propagation of uncertainty will be more relevant.

3.7 Conclusions

The recent shift of paradigm in statistical methods in crop nutrition has enabled changes that may help embracing uncertainty in crop nutrition models. We elaborate the main benefits of a recently introduced model to fit CNDC, that arise from its hierarchical and Bayesian features. There may exist paths for gaining efficiency in the use of field data. We emphasize the importance of quantifying and considering uncertainty for statistical inference, but also showing how propagating uncertainty is fundamental for correct statistical inference.

Chapter 4

Revisionist analysis of relevant metrics for nitrogen phenotyping in wheat crop

Abstract

Modern breeding programs aim to increase the effective use of nitrogen (N) in cereal grain crops such as wheat (*Triticum aestivum* L.) for reducing the environmental footprint of agriculture. Metrics for assessing the efficiency in crop N use include N utilization efficiency (NUtE, grain yield to N uptake ratio) and N nutrition index (NNI, actual to critical N concentration ratio). Our main objective was comparing NUtE and NNI (determined at anthesis as a predictive method) to evaluate their performance as phenotyping metrics for effective use of N in wheat. A dataset from 11 field experiments testing wheat cultivars in contrasting N environments was explored in this study, including a revisionist analysis with historical cultivars. Briefly, NUtE has nearly no relationship with yield and is negatively related to grain protein concentration. Contrastingly, NNI is positively related to yield and grain number, and is slightly associated to grain protein concentration. The NUtE increased but NNI at anthesis remained stable over time. Thus, NUtE provides overoptimistic results on phenotyping and should be replaced by allometric-based metrics like NNI.

Keywords: breeding-driven changes; N anthesis status; N nutrition index; N use efficiency.

4.1 Introduction

The efficient use of nitrogen (N) fertilizer plays a critical role in achieving food security while minimizing the environmental footprint of agriculture (Eickhout et al. 2006). Industrial N fertilizers have historically enhanced global food security by supplying a major yield limiting factor (Spiertz 2009). However, inefficient N fertilization increases the risk of detrimental effects to the environment such as groundwater pollution and greenhouse gas emissions (Cameron et al. 2013). Thus, reducing N waste has become a major goal for agricultural stakeholders and researchers are incentivized to study strategies for reducing N inputs without penalizing crop yields, i.e., increasing the effective use of nutrients (e.g. Sutton et al. 2021).

Recent studies envision genetic improvement of commercial cultivars as a cost-effective strategy for achieving sustainable agricultural systems (e.g., Ciampitti et al. 2022; Cormier et al. 2016; Foulkes et al. 2009). Plant breeders have since started considering evaluating the genotypes regarding their efficiency in using nutrients like N for grain yield production. Given that breeders do not receive clear incentives yet, the use of metrics quantifying said 'nutrient use efficiency' is still incipient in breeding programs and most are borrowed from production agriculture. Fertilizer management guidelines in production agriculture have historically evaluated the efficiency of agricultural systems with a set of metrics that measure the amount of grain yield obtained per unit of N (Dobermann 2007). For example, one of the first metrics used to quantify the sustainability of an agricultural system was the N use efficiency (NUE, crop yield per unit of nitrogen available Moll et al. 1982). The NUE can be further analyzed by dissecting it into the nitrogen uptake efficiency (NupE; ratio of N taken up by the crop on N soil plus the fertilizer supply) and the N utilization efficiency component (NUtE, yield to N uptake ratio), as well as an apparent N recovery efficiency (AFR or NRE, N uptake to N fertilizer ratio, Foulkes et al. 1998). The NUtE is many times considered an appropriate metric for genotypic evaluations targeting gains in internal physiological processes affecting photosynthesis or partitioning (Cormier et al. 2016; Moll et al. 1982; Lammerts van Bueren and Struik 2017). Thus, plant breeders have borrowed the NUtE concept for N phenotyping and screening purposes.

Interestingly, studies in production agriculture have proposed new metrics that are more robust than NUE-derived metrics (e.g., NUtE) for comparing contrasting scenarios (e.g. NupE in Cohan et al. 2019). A very relevant model is the critical N dilution curve (Justes et al. 1994; Lemaire et al. 1998), which disrupted the scientific paradigm for diagnosing wheat N nutritional status with applications for evaluating the effective use of N. The most relevant characteristic of the N dilution curve is that it considers the allometry between N uptake and biomass accumulation (W). This allometry indicates that plant N uptake is co-regulated both by soil mineral N and by plant growth rate capacity (Devienne-Barret et al. 2000). It is possible to estimate the critical plant N concentration ($\%N_C$) as the minimum plant $\%N$ needed to obtain maximum crop W: $\%N_C = A_1 \cdot W_{max}^{-A_2}$ (Lemaire et al. 2008). The allometric coefficient $A_2 > 0$ thus indicates a decline in plant N concentration ($\%N$) as crop W increases, leading then to a N dilution curve. These critical N dilution curves can be used for evaluating crop N status under field conditions as the actual %N relative to the $\%N_C$ value corresponding to the actual crop mass: the N nutrition index (NNI, actual to critical $\%N$ ratio) indicates whether the crop is N deficient (NNI<1), N sufficient (NNI=1), or under luxury consumption of N (NNI>1). The NNI represents a viable alternative to NUtE for evaluating efficient use of N in agricultural systems (Lemaire and Ciampitti 2020), but until this point, no study has compared these two N metrics.

Unlike production agriculture, most evaluation criteria in breeding programs are still based on NUE-derived metrics (e.g., Martre et al. 2023), and thus face the risk of confounding effects of N dilution product of crop growth with true N use efficiency (Lemaire and Ciampitti 2020; Li et al. 2022). Recent studies flag NUE and its derived components as possibly misleading due to the lack of consideration of the allometric relationship W-N. In fact, the crop N dilution process leads automatically to an increase in NUtE as lower plant %N corresponds to a higher crop W per unit of N uptake. Therefore, the NUtE of different genotypes must be compared under similar crop W (Lemaire and Ciampitti 2020). If the comparison is done under different crop W, any apparent difference in NUE must be analyzed carefully since changes in this N metric can simply be explained by changes in crop growth capacity (Lemaire and Ciampitti 2020). In consequence, any progress in NUE driven by increasing crop growth capacity is obtained with lower plant %N that leads to the well-known trade-off between yield and %N (Lemaire et al. 2021) but also for grain crops like wheat (*Triticum aestivum* L.). In fact, historical studies report an apparent improvement in NUE accompanying historical yield gains (Cormier et al. 2013; Ortiz-MonasterioR et al. 1997; Peng et al. 2022; Sylvester-Bradley and Kindred 2009), mainly reflecting changes in crop W but without direct impacts on N metabolism. Previous research has pointed out that breeding programs may be missing insights from analyses based on the N-W allometry (Ciampitti et al. 2022).

Following this rationale, the aim of this study was to provide a quantitative comparison for NUtE versus NNI as metrics to evaluate the efficient use of N for grain yield/quality in breeding programs, with potential for developing new N phenotyping metrics. The specific objectives were to (i) establish functional relationships between grain yield, grain protein concentration all associated with NNI or with NUtE, (ii) analyze the use of NNI for crop phenotyping, and (iii) revise historical changes in NUtE versus NNI to provide insights on the direct impact of N changes on wheat yield increases over time.

4.2 Materials and Methods

4.2.1 Field experiments

A total of 11 wheat experiments were performed during the seasons 1999/2000, 2000/2001, 2003/2004, 2004/2005, 2006/2007, 2007/2008, and 2008/2009 in six locations in France. These experiments correspond to previously published studies, namely Cormier et al. (2013), Laperche et al. (2006), Bogard et al. (2010), and Gaju et al. (2011) and provide large enough genetic variability to explore contrasting relationships between N status, yield, and grain protein concentration (GPC). Cormier et al. (2013) used a set of 228 cultivars released between 1969 and 2010 to study changes in NUE that may explain yield gains and changes in grain protein. Laperche et al. (2006) screened a mapping population (n = 235) of French germplasm for their NUE characteristics. Bogard et al. (2010) studied a set of 30 French and British modern wheat cultivars, while Gaju et al. (2011) studied 16 modern French cultivars. All experiments presented at least two N rates (i.e., low and high N), and measured both W and %N at anthesis and harvest time for wheat crop. Thus, the pairs of actual W and actual $\%N_C$ at anthesis allowed the estimation of NNI at this crop growth stage. Further measurements were grain yield, GPC at harvest (%), grain number per m^2 (GN, grains m^{-2}), and thousand grain weight (TGW, g). Protein Yield (PY, Mg ha^{-1}) was calculated as grain yield by GPC. A detailed description and further information about the experiments can be found in Cormier et al. (2013), Laperche et al. (2006), Bogard et al. (2010), and Gaju et al. (2011) and Table 4.1.

4.2.2 Computations and Statistical Analyses

For the first objective, we investigated the relationships between yield, grain protein, NNI and NUE via clustering and principal component analyses (PCA). We used the data in Bogard et al. (2010) and Gaju et al. (2011) corresponding to modern cultivars. First, we obtained the means of each genotype in each environment. We clustered the genotype-environment specific data according to the scaled values for grain yield and grain protein using a k-means algorithm and a set of 23 indices to obtain the optimum number of clusters (Charrad et al. 2014). Then, we performed three PCAs: (1) using data from all N environments in Bogard et al. (2010) and Gaju et al. (2011) (2) using only high N supply data, (3) using only low N supply data. In all cases, we considered the scaled values of the variables grain yield (GY), grain protein concentration (GPC), protein yield (PY), N utilization efficiency (NUtE), and NNI. We computed NUtE = W/Nupt. In this case,

Article	**Number of site-years**	**Genotypic characteristics**	**N treatments**	**Measurements**
Laperche et al. (2006)	5 (4 in 99/00, and 3 in 00/01)	Mapping population (n = 235)	High N, Low N	N uptake, Biomass, Grain Protein Content, Grain Yield, Grain Number
Bogard et al. (2010)	4 (2 in 03/04, and 2 in 04/05)	Modern wheat cultivars (n = 30)	High N, Low N	
Gaju et al. (2011)	8 (4 in 06/07, and 4 in 07/08)	Modern wheat cultivars (n = 16)	High N, Low N	
Cormier et al. (2013)	4 (2 in 07/08, 2 in 08/09)	Wheat cultivars (1969–2010) (n = 228)	High N, Low N	

Table 4.1: *Summary of experiments (article, number of site-years, genotypic characteristics, and fertilizer nitrogen levels, treatments) and measurements. (n = ...) indicates the total number of cultivars tested in the experiment. Units: N uptake (Mg ha^{-1}), Biomass (Mg ha^{-1}), Grain Protein Concentration (%), Grain Yield (Mg ha^{-1}), Grain number (grains m^{-2}).*

NUtE represents the internal efficiency of producing W per unit of absorbed N (Lemaire and Ciampitti 2020). For the second objective, we analyzed the relationship between NNI, the yield components (i.e., grain number and thousand grain weight), and GPC, with an emphasis on the implications for phenotyping applications. We used the data in Laperche et al. (2006) to fit a mixed-effects linear model describing yield as a function of grain number and thousand grain weight:

$$y_{ijl} \sim N(\mu_{ijk}, \sigma^2), \tag{4.1}$$

$$\mu_{ijk} = \beta_0 + x_{1,ijk}\beta_1 + x_{2,ijk}\beta_2 + u_{j|k} + v_k, \tag{4.2}$$

where y_{ijk} is yield (in Mg ha^{-1}) for the ith repetition (i = 1, 2, 3) of the jth cultivar (j = 1, 2, . . . , 235) under the kth N environment (k = 1, 2), with expected value μ_{ijk} and variance

σ^2. The expected value is described as the sum of a global intercept β_0, grain number per m^2 of the ith repetition, jth cultivar and kth N environment ($x_{1,ijk}$) times its effect on yield β_1, thousand grain weight of the ith repetition, jth cultivar and kth N environment ($x_{2,ijk}$, in g) times its effect on yield β_2, the random effect of the N environment v_k and the random effect of cultivar nested within the N environment $u_{j|k}$. Moreover, we analyzed the relationship between relative yield and relative GN in low N and high N environments. For that, we computed the relative yield of the jth cultivar as $ry_{ijk} = \frac{y_{ijk}}{y_{kmax}}$, where ry_{ijk} is the expected relative yield of the ith observation of the jth cultivar in environment k, y_{ijk} is the observed yield of the ith observation of the jth cultivar in the kth environment, and y_{kmax} is the expected maximum yield of a cultivar in the kth environment.

For the third objective, rethinking historical changes in NUE versus NNI, we performed three regression analyses of NUtE, NNI efficiency (NNIe, yield to NNI ratio), NNI, and the ratio GN to NNI (i.e., GN $\cdot$ NNI^{-1}), as a function of the year of release. We fitted the data in Cormier et al. (2013) to mixed-effects linear models that describe yield as a function of the year of release (fixed), variety (random), and N environment (random). The models can be expressed as:

$$y_{ijk} \sim N(\mu_{ijk}, \sigma^2), \tag{4.3}$$

$$\mu_{ijk} = \beta_0 + x_{1,ijk}\beta_1 + u_{j|k} + v_k, \tag{4.4}$$

where y_{ijk} is NUtE, NNIe, NNI, or GN$\cdot NNI^{-1}$ for the ith repetition (i = 1, 2, 3) of the jth cultivar (j = 1, 2, …, 235) under the kth N environment (k = 1, 2), with expected value μ_{ijk} and variance σ^2. The expected value is described as the sum of a global intercept β_0, year of release of the ith repetition, jth cultivar and kth N environment ($x_{1,ijk}$) times its effect on yield β_1, the random effect of N environment v_k and the random effect of cultivar nested within N environment $u_{j|k}$. All analyses were performed using R software (R Core Team 2019), 'tidyverse' (Wickham et al. 2019) for data wrangling and package 'lme4' (Bates et al. 2014) for fitting mixed-effects models.

Cluster	Grain yield (Mg ha^{-1}	Protein concentration (%)
High Yield	8.34 (4.77, 11.61)	9.52 (5.80, 13.83)
Medium Yield	8.52 (6.37, 10.52)	10.05 (6.88, 12.84)
Low Yield, High Protein	6.23 (4.36, 9.00)	11.93 (7.52, 16.21)

Table 4.2: *Description of 45 wheat cultivars clustered according to their grain yield and grain protein concentration. The data represent cultivars from Bogard et al. (2010) and Gaju et al. (2011). Values represent means, brackets include 95% high density intervals.*

4.3 Results

4.3.1 NNI is better related to grain yield than NUtE

Three clusters were obtained by dividing the 45 cultivars into the groups 'high yield', 'medium yield' (both with a wide range of protein levels), and 'low yield, high protein' (Table 4.2). The PCA revealed that, in all cases, grain yield and NUtE had near to no relationship (Figure 4.1). Oppositely to NUtE, NNI at anthesis was always positively related to grain yield. Moreover, there was a positive relationship between NNI and GPC, although not as strong as the one observed for NNI and yield (Figure 4.1).

For N environment-specific analyses (Figures A.1 and A.2), the relationships between the variables were similar: the null relationship between NUtE-yield was preserved, as well as the positive relationship between NNI and yield. However, NNI was not related to GPC for the environment-specific analyses, suggesting large-scale effects of N status on GPC that are no longer identified at higher resolution.

4.3.2 Breeding for NNI is breeding for grain number

The NNI at anthesis was positively related to grain number m-2, grain weight, and GPC (Figure 4.2). Yields were positively associated with grain number [slope = 0.373 (0.369, 0.376) Mg $\mathrm{ha}^{-1}\mathrm{g}1000^{-1}$] and grain weight [slope = 0.147 (0.144, 0.149) Mg $\mathrm{ha}^{-1}\mathrm{g}1000\mathrm{grains}^{-1}$], although the relationship was stronger for grain number (Figure 4.2a).

The relative performance of the genotypes was similar for both low N and high N envi-

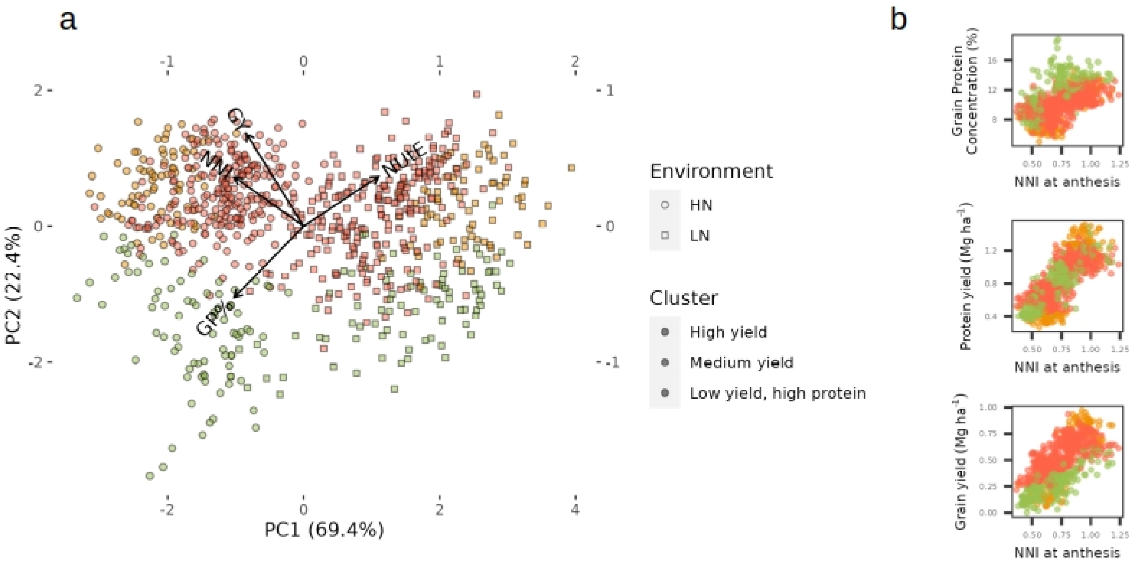

Figure 4.1: *a. Relationship between grain yield (GY), grain protein (GPC%), Nitrogen Utilization Efficiency (NUtE) and Nitrogen Nutrition Index at anthesis (NNI) on the first two axes of a principal component (PC) analysis for 45 wheat varieties. b. Relationship between NNI at anthesis and grain protein content(GPC, in %), grain yield (GY, in Mg ha^{-1}), and Protein Yield (PY, in Mg ha^{-1}). Point shapes indicate N environment: high N (HN), or low N (LN). Point colors indicate cluster: 'high yield' (154 observations), 'medium yield' (474 observations), or 'low yield, high protein' (170 observations).*

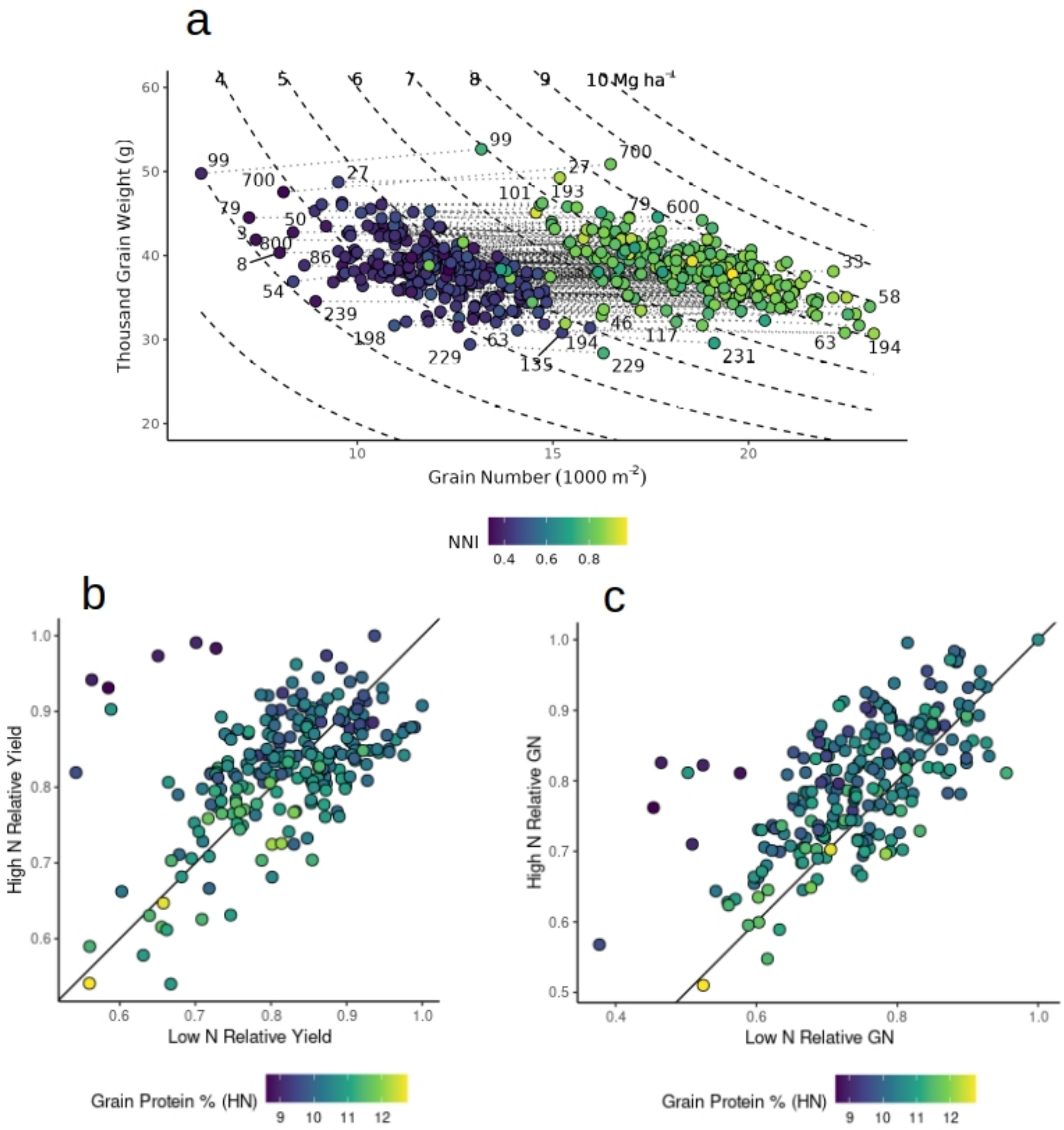

Figure 4.2: *a. Grain yield components: thousand grain weight (g) and grain number (1000 m-2). Dashed lines indicate isoyield lines (yield indicated in Mg ha*$^{-1}$*). Point color indicates NNI at anthesis. b. Relative yield under high N environments versus Relative yield under low N environments. Colors indicate grain protein concentration (%). c. Relative grain number under high N environments versus Relative grain number under low N environments. Colors indicate grain protein concentration (%). The solid line represents the 1:1 line.*

ronments. While the major factor driving N status (i.e., NNI) was the N environment, the GW-GN relationship remained similar for the same genotype, across different N environments (Figure 4.2 b and c). A few genotypes were considered “less desirable” given their sensitivity to N stress (top left corner in Figure 4.2 b). Otherwise, the regions along the 1:1 line contain similar genotypes, suggesting their similar performance under high N and low N environments (Figure 4.2 b and c).

4.3.3 Breeding-driven changes in NUtE are not equivalent to changes in NNI at anthesis

Historical gains in NUtE and NNIe (yield to NNI ratio) over time (Figure 4.3). Overall, the increase in NUtE was 0.2 g yield g Nupt year^{-1} ($p<0.01$), and the increase in NNIe was 0.07 Mg yield NNI^{-1}year^{-1} ($p<0.01$), while the absolute value of NNI changed only about -0.001 year^{-1} ($p=0.22$). In relative terms, the historical gain estimates were greater for NUtE than for NNIe. Interestingly, we found that the historical changes followed the same trend under both low- and high- N environments ($p=0.4$ for NUtE; $p=0.7$ for NNI). Moreover, we found that new varieties were able to produce more grains under a same N status (i.e., NNI) at anthesis ($p<0.01$).

We examined historical trends in other variables that have functional relationships with yield, to better understand changes in NUtE. We found that the increases in NUtE were accompanied by increases in grain yield (0.26 Mg ha^{-1}year^{-1}), harvest index (0.00096 year^{-1}), and dilutions of grain protein concentration (-0.032 %year^{-1}).

4.4 Discussion

This study portraits the benefits of using NNI for quantifying the role of efficient crop N utilization, especially to avoid misleading conclusions confounding effects between improvements in N use and N dilution with direct impacts on yield gain. Our results show that

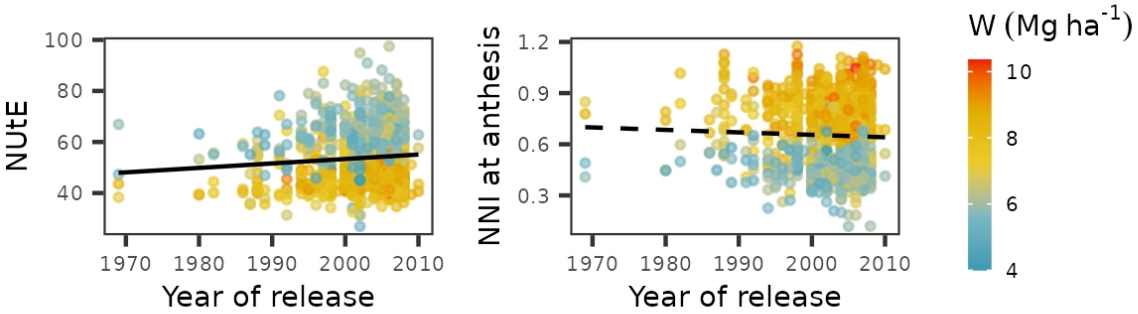

Figure 4.3: *Historical changes in NUtE (a) and NNI (b) for 228 wheat varieties registered between 1969 and 2010. Point colors indicate grain yield (Mg ha^{-1}), and the regression lines indicate historical trends, significant at $p<0.05$ (solid line) and non-significant (dashed line).*

using allometric-based criteria avoids confounding true improvements in crop N use for N dilution driven by the W-N allometry (i.e., greater W, greater N dilution, more apparent efficiency; Lemaire and Ciampitti 2020). In our revisionist analysis, we found that traditional NUtE-based analyses overestimate the influence of the efficient use of N associated to yield gains. In some cases, yield gains may be driven by increased overall biomass or harvest index (Austin 1999; Hay 1995; Brancourt-Hulmel et al. 2003). In such cases, crops are expected to maintain their N-W allometry, leading to apparent gains in NUE (Lemaire and Ciampitti 2020). However, it is possible to expect yield gains at similar N status that may be related, for example, to the vertical distribution of foliar N (Drouet and Bonhomme 2004) or increases in early vegetative N uptake and late N remobilization (Fernandez et al. 2022) or other processes that increase growth or grain number per unit N (Reynolds et al. 2000). Usually, breeding progress is a product of both driving forces, but detecting real changes in N use requires directed analyses considering the N-W allometry.

While modern breeding programs are built on the progress obtained by decades of (mostly) yield-driven advances, there may be some underexploited traits remaining (Brancourt-Hulmel et al. 2003; Long et al. 2006). Even when the yield-driven approach to genotypic selection has affected canopy architecture and harvest index (Trethowan et al. 2007), other processes for yield formation may remain underexploited due to the difficulty to modify them without them being intentionally targeted for modifications. For example, retrospective analyses show that the vegetative:reproductive growing cycle partitioning, and biomass partitioning, among other traits, have experienced improvements over time (Lopes et al. 2012; Reynolds et al. 2012). Consequently, the additional benefit from improving the same traits is expected to be relatively low in comparison to other yet 'underexploited' traits (Fischer and Edmeades 2010). Some of these 'underexploited' traits mentioned in the literature include photosynthetic metabolic pathways (Austin 1999; Long et al. 2006; Fischer and Edmeades 2010), fruiting efficiency (Pretini et al. 2021), N storage capacity for increasing N remobilization during grain filling (Foulkes et al. 2009; Gaju et al. 2014), and root architecture for N capture (Lynch 2019). However, complexity and trade-off might be evident between modifying root N capture and improving crop W (and yield), as highlighted for maize crop the relationship between root types and W are not proportional and allometric (York and Lynch 2015). Therefore, the pathway for increasing nutrient capturing is complex due to regulation of genes on root architecture and signals controlling expression of root types for effective foraging (Jia and von Wirén 2020; Krouk et al. 2010), but recognized as a key component for boosting crop productivity (Reynolds et al. 2021). More recently, inclusion of biological nitrification inhibition (BNI) to wheat cultivars is a new potential pathway to reduce N fertilizer use over time for major crop cereals (Subbarao et al. 2021; Subbarao and Searchinger 2021). To measure the efficacy of modifying these traits, current leading studies are using NUE-based metrics for evaluating agricultural systems (Martre et al. 2023) possibly leading, as we mentioned, to misleading conclusions due to unaccounted W-%N allometries. This study clearly demonstrated that an allometric-based metric like NNI may be a robust metric for genotypic evaluations to better select for traits that can more effectively increase nutrient capture for yield formation. We have demonstrated this convenience using NNI as a measure related with N uptake, and GN NNI^{-1} as related to fruiting efficiency and general productivity under the same N status.

Future breeding strategies may benefit from allometric-based phenotyping metrics like NNI also as a strategy to alleviate screening labor. Phenotyping corresponds to one of the greatest bottlenecks in breeding programs (Araus and Cairns 2014). In this context, the adoption of allometric-based phenotyping may relieve labor and monetary resources by providing an early screening of genotypes for concurrently targeting improvements in both yield and GPC. We argue that the resources saved may be reallocated towards increasing the range (i.e., the difference in available N between the extreme-low and extreme-high N treatments) and resolution (i.e., the number of N treatments) of N scenarios. Allocating more efforts into exploring contrasting N conditions (instead of the regular two) for an increased range and resolution may help disentangle other crop processes because the economy of N in the canopy is a complex process. These contrasting N conditions applied to different genotypes have already shown that the ability to tolerate N deficiency depends on the time when N deficiency occurs (Ravier et al. 2017). The concept of NNI path along the growing cycle and tolerable deficiency could be another trait to select more tolerant and efficient genotypes and to propose combinations of genotypes with different N deficiency strategies according to the risk of environmental N supply (soil N mineralization reduction, N stress induced by water stress, etc.). Such analyses may bring insights about mechanics in grain protein formation under different anthesis and grain filling N conditions. Rapid phenotyping NNI at anthesis may thus help avoiding indirectly selecting cultivars with the typical yield-driven protein dilution (Calderini et al. 1995).

The adoption of NNI metrics in breeding programs may be accompanied with remote sensing techniques borrowed from production agriculture for additional savings in time and labor devoted to crop phenotyping tasks. Recent efforts have shown practical applications of remote sensing techniques for both breeding programs (Yang et al. 2020) and for phenotyping crop N nutritional status Berger et al. (2020). The compatibility of the NNI concept with remote sensing applications (Lemaire et al. 2008; Zhao et al. 2021; Li et al. 2022) supports the notion that NNI may represent a very relevant metric deserving further developments for breeding applications. Such efforts may be carried out almost exclusively in trans-disciplinary teams that combine domains including crop physiology, remote sensing, and sensor engineering. We envision that the products produced by trans-disciplinary teams for phenotyping N status with remote sensing techniques and allometric-based metrics may save time, labor, and help target under-exploited traits in plant breeding programs. All things considered, however, the success of such phenotyping platforms and even the success of breeding programs considering N-related traits depends on the exploitation of genetic variability related to N uptake and metabolism. While genetic variability is considered one of the major bottlenecks in plant breeding (Lopes et al. 2015), other new opportunities arise to explore new traits more directly linked to N (Wu et al. 2019; Subbarao et al. 2021; Fernandez et al. 2022) that could provide an alternative for increasing yields via reduction on the dependency of external N inputs (Ciampitti et al. 2022). Future efforts should be focused on developing trans-disciplinary plant breeding programs to accelerate high-throughput phenotyping techniques to sustain future yield gains for reducing dependency to inputs under varying climate environments.

4.5 Conclusions

Allometric-based (i.e., NNI-based) approach for phenotyping holds great potential to aid breeding programs. We found that NUtE, one of the most popular metrics used in plant breeding programs, has little to no relationship with yield and grain protein concentration. Thus, selection via NNI criteria may help producing higher-yielding cultivars by indirect selection of greater grain number and protein concentration. Methods borrowed from remote sensing applications in production agriculture may be beneficial to accelerate high-throughput phenotyping techniques that include sound metrics that quantify the efficiency in the use of N. Future research will benefit from integrating this N metric into breeding programs via more directly targeting soil-plant N processes.

Chapter 5

Challenges in phenotyping crop nutritional status – a bibliometric perspective

5.1 Introduction

Reducing the dependency of synthetic fertilizers applied to modern agricultural systems is one of the major objectives for increasing sustainability. Nitrogen (N) fertilizer application, especially when in excess to crop N demand, is one of the largest sources of air and water pollution. To reduce the environmental impact, external N inputs can be reduced via two major strategies: enhancing agronomic management (e.g., right rate, timing, source, and placement of external nutrient inputs) and crop genetic selection to produce similar yield levels under reduced N inputs (Lemaire and Ciampitti 2020). Thus, evaluating the crop N status is important in both crop production settings (herein referred to as production agriculture) and pre-commercial breeding programs. Previous studies have reported that systems including periodic assessments of crop N status are able to adjust their yield expectations and thus, the amount of fertilizer adequate for the cropping system. However, phenotyping the crop N status is still a challenging task, especially at early stages of crop improvement when hundreds of materials are tested under field conditions, as there is a trade off between accuracy and labor intensity. This perspective article intends to provide a summary of the current and prospective techniques available for quantifying the crop N status. Finally, we aim to convey a strong message regarding possible opportunities for filling the knowledge gap between production agriculture and plant breeding in terms of the advances in phenotyping crop nutritional status, and propose future research directions and opportunities.

5.2 Technology transfer from production agriculture and pre-commercial breeding stages to early stages of the plant breeding pipeline

Phenotyping crop traits is one of the most important bottlenecks in plant breeding programs (Furbank and Tester 2011; Araus and Cairns 2014). Phenotyping is crucial for understanding plant mechanisms to help breeders improve specific targets for crop traits, like those related to N metabolism (e.g., N concentration, Feng et al. 2016). However, while some crop traits may facilitate selection, their measurements can be labor intensive, expensive, and time consuming (Ghanem et al. 2015). Thus, there is a growing demand for developing novel low-cost phenotyping methods (Nguyen and Kant 2018), like proximal and remote sensing techniques that alleviates field labor (Tao et al. 2022).

Studies in production agriculture have demonstrated that the best metrics for describing crop N status are those considering the allometric relationship between crop growth and N uptake (Lemaire and Ciampitti 2020). In this context, the nitrogen nutrition index (NNI) may be considered the golden standard for determining the crop N status. Recent efforts have pursued the estimation of NNI using proximal and remote sensors (Lemaire et al. 2008).

The current-state-of-the-art techniques for proximal and remote sensing for determining crop N status in major field crops can be studied by exploring advantages and limitations for breeding, including both equipment and metrics (Figure 5.1). In terms of the equipment, the first efforts for proximal sensing crop N status were hand-held sensors, like the broadly implemented Soil-Plant Analysis Development (SPAD) chlorophyll meters (e.g., Ma et al. 1996). The SPAD meter was first implemented for measuring chlorophyll content in production, and then adopted by breeders to better understand plant N distribution (e.g., Babar et al. 2006). Even though SPAD reading need a full-N reference (plots with high N fertilization), it also provides a flexible tool to take measurements with very high spatial resolution (e.g., characterize the strata of the plant canopy; Silva-Perez et al. 2018) and highly relevant for genotypic characterization in breeding programs. Currently, proximal sensing applications are expanding with the use of ground-based robotic applications (e.g., rovers; Atefi et al. 2021). For remote sensing, applications for crop N status phenotyping started with the newest generation of satellites with greater spatial resolution (i.e., 0.5 to 5 m) and hold great expectations for future applications given the low amount of field efforts required (Pinto et al. 2023). In addition, remote sensing applications expanded with the introduction of commercial unmanned aerial vehicles (UAV), yielding data with higher spatial resolution than current commercially available satellite imagery. More recently, some airborne (i.e., airplanes) hyperspectral imagery is being used for remote sensing of crop N status (Wang et al. 2021). Most remote sensing applications aim to reduce field labor but face their challenge in the data processing step (Araus and Cairns 2014).

Broadly, there are two general approaches to quantifying crop N status with proximal or remote sensing: (i) via determination of spectral bands and calculation of a vegetation index (Prey et al. 2020), or (ii) via measurement of crop physiological parameters (e.g., nitrogen nutrition index; Mistele and Schmidhalter 2008). For example, sensors placed in both satellite and UAV obtain spectral bands and permit calculation of vegetation indices such as the normalized difference vegetation index (NDVI) and the normalized difference vegetation

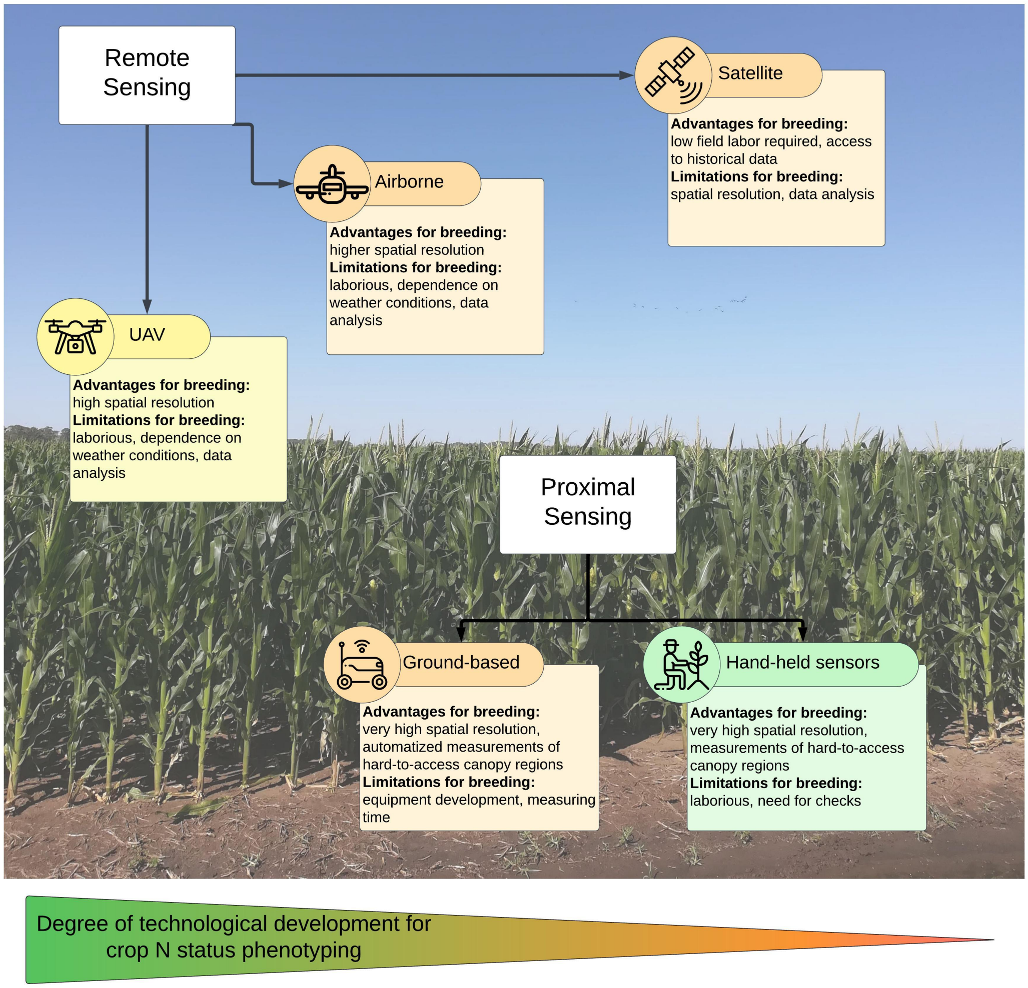

Figure 5.1: *Overview of current state-of-the-art techniques for proximal and remote sensing crop N status. Box colors indicate the degree of technological development for crop N status phenotyping, i.e., more developed (green) to recently introduced (red).*

index on greenness (GNDVI). These indices may be used as a proxy for crop biomass (Xu et al. 2022), or to estimate chlorophyll content or nitrogen nutrition index (NNI; Chen 2015). The NNI is thus far the only index that describes the N nutritional status considering the allometric relationship between N uptake and crop growth (Lemaire and Ciampitti 2020). On the other hand, proxies for chlorophyll content may be implemented when targeting a more homogeneous vertical distribution of foliar N along the canopy in crops like corn (Reyes Ponce 2016).

5.3 Bibliometric characterization of the scientific production and opportunities

We performed a bibliometric analysis to describe the scientific production of studies including remote sensing applications for crop N status field determinations and shed some light on the current knowledge gap between production agriculture and plant breeding. We obtained 1042 peer-reviewed publications from a search of the Web of Science database, using the terms nitrogen (or N), phenotyping, and crop. The data were analyzed using the R package 'bibliometrix'. Recent advances in remote sensing applications for N status phenotyping in both production agriculture and breeding have been summarized in review articles (estimation of NNI in Li et al. 2022; N status sensing in Berger et al. 2020; applications to phenomics in Tao et al. 2022; satellite imagery applications in plant breeding in Pinto et al. 2023). Note that the reviews on proximal and remote sensing applications to breeding programs only vaguely mention N status phenotyping. Perhaps because N use is not fully considered in breeding programs yet, the link between these disciplines is still incipient (Figure 5.2). Most of the proximal and remote sensing techniques in plant breeding seem to be somewhat less developed (Figure 5.2a; Chawade et al. 2019), and standing at least a decade behind to those from production agriculture (Figure 5.2b). In other words, the technology transfer seems to take about a decade and preserves most original methods.

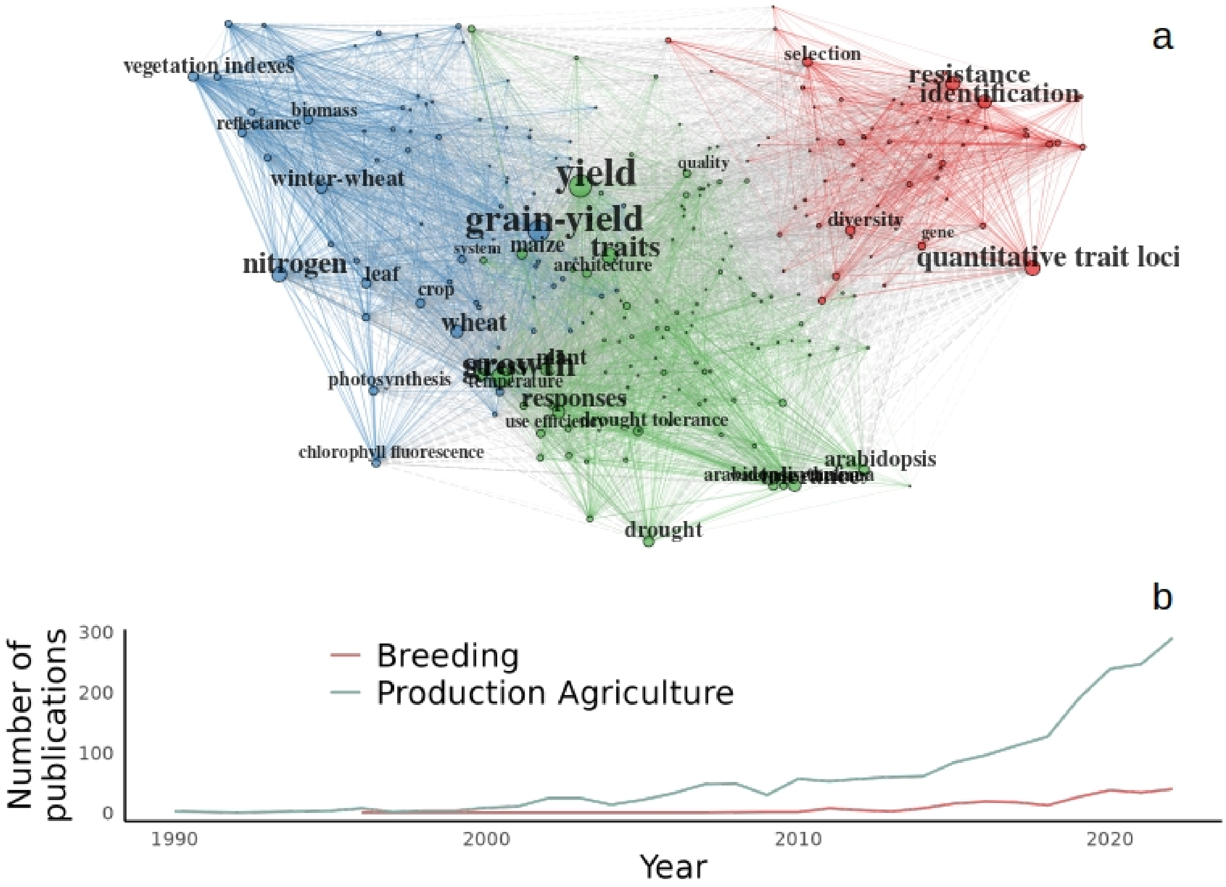

Figure 5.2: *Scientific production on crop N status phenotyping. Top panel (a) portrays a network for the keywords from a search of the Web of Science database, using the terms nitrogen (or N), phenotyping, and crop. Colors indicate the network community (i.e., a group of highly connected keywords) and is divided into concepts related to genetics (red), proximal and remote sensing terms(blue), and crop physiology terms (green). Bottom panel (b) shows the number of publications investigating crop N status phenotyping for applications in production agriculture (turquoise) or for application in breeding programs (red).*

For example, recent efforts for quantifying crop N status (Berger et al. 2020) are considered promising in breeding scenarios, but there is no clear evidence that they have been applied yet. Likewise, crop remote sensing using satellite imagery has been around in production agriculture, but is a novelty for breeding programs, perhaps due to contrasting requirements in terms of spatial resolution.

5.4 Prospective views and conclusions

We envision the development of new equipment, traits, and criteria, developed exclusively for enhancing information for plant breeders that do not necessarily stem for production agriculture. First, metrics and equipment for quantifying crop N status that are borrowed from production agriculture should be analyzed carefully for breeding applications (Chawade et al. 2019). Many of the developments of new low-cost phenotyping methods for production agriculture need adjustments to be transferred to breeding programs, due to different spatio-temporal focuses of the research problems. Production agriculture focuses mostly on the upper leaf strata for N status characterization (Berger et al. 2020), but breeders require characterizing additional sections of the vertical canopy profile that may be targeted by crop breeding. For example, corn (*Zea mays* L.) yield gains are associated to improved N status of the lower leaf strata (Reyes Ponce 2016; Fernandez et al. 2022) and, in wheat (*Triticum aestivum* L.), basal N content is also associated to higher responsiveness to N fertilization (Nguyen et al. 2016). It is thus worth considering the lower leaf strata for canopy characterization. New crop N status phenotyping techniques should be developed, in addition to mainstream technology borrowed from production agriculture. Some examples for inspiration are present, for example, for phenotyping soluble carbohydrates in the stem with a spectroradiometer mounted on a 4-wheel drive motorbike (Dreccer et al. 2014).

We have provided a portrait of the current state-of-the-art technologies, available and missing, for rapid phenotyping crop N status in breeding programs. Most breeding programs rely heavily on borrowing proximal and remote sensing phenotyping techniques and metrics from production agriculture. Future research should focus on exploring new sensor techniques, metrics, timing for determination, aim for phenotyping novel crop traits. Expanding the number of variables for crop phenotyping may lead to a better understanding of the potential impact for directly targeting yield gains derived from impact on plant N economy.

Chapter 6

Final Remarks

This dissertation comprises a set of studies aiming to maximize the utility provided by crop nutrition models such as the critical nitrogen (N) dilution curve, via rethinking and revising methods in the workflow associated with said models. The general consensus for managing crop nutrition is summarized in conceptual frameworks that are expressed as mathematical models. These models synthesize the current beliefs regarding the relationship between nutrient availability, crop nutrient uptake, and crop growth. Frequently, the coefficients quantifying the strength of the relationships between variables (e.g., grain yield and N concentration) are estimated from data using statistical models. Thus, the selection of the components considered in the general conceptual framework (i.e., the model) matters. That includes the selection of metrics considered, the equipment to measure those metrics, and the statistical methods used for obtaining inference. All the factors listed above affect results, inference, conclusions and ultimately define the course of resource allocation. This dissertation proposes rethinking some common practices and metrics, considering both statistical and crop physiology aspects, and indicates future prospects for developing sound models for determining crop N status for production and breeding applications.

The research presented in this dissertation should convey the notion that the conceptual frameworks for quantifying N use in agricultural systems relies heavily on the methods and metrics implemented. We first showed that the selection of the statistical framework for analyzing those data is crucial. Chapter 2 shows that the statistical framework can greatly affect the results and thus ultimately may have an impact on the sustainability of an agricultural system. Chapter 3 reflects on the selection of certain methods, and proposes rethinking some common practices in crop nutrition models. The path forward should include models that are as realistic as possible, with data acquired as reliable as possible, that adequately account for (i.e., propagate and quantify) the associated uncertainty.

Moreover, the preferred frameworks for measuring nutrient use in agricultural systems are based on sound crop physiology concepts. Chapter 4 showcases the importance that the metrics for measuring nutrient use consider the allometric relationship between N uptake and crop growth. Quite surprisingly, there is quite some room for improvement on the breeding side in terms of technology adoption and special measuring devices. Chapter 5 reports current advancements, knowledge gaps, and suggests possible ways to move crop N status phenotyping forward.

All things considered, future research should aim for (i) developing statistical models with realistic assumptions that can propagate and quantify uncertainty, and (ii) developing new measurement equipment that explicitly correspond to the target plant traits. Stakeholders working on both points should hold a tight communication, as the effective impact on real world problems will depend on a synergistic development of the two aspects of crop phenotyping. This calls again for the development of trans-disciplinary organizations with clear goals: the goal in this case is improving sustainability in agriculture while preserving (or improving) current food security conditions, and calls for teams of physicists, engineers, crop physiologists, agronomists, statisticians and applied mathematicians.

Bibliography

Anupreet, P. and Raftery, A. E. (2022). Comparing methods for statistical inference with model uncertainty. *Proceedings of the National Academy of Sciences*, 119(16):e2120737119.

Araus, J. L. and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in Plant Science*, 19(1):52–61.

Atefi, A., Ge, Y., Pitla, S., and Schnable, J. (2021). Robotic Technologies for High-Throughput Plant Phenotyping: Contemporary Reviews and Future Perspectives. *Frontiers in Plant Science*, 12.

Austin, R. B. (1999). Yield of Wheat in the United Kingdom: Recent Advances and Prospects. *Crop Science*, 39(6):1604–1610.

Babar, M. A., Reynolds, M. P., van Ginkel, M., Klatt, A. R., Raun, W. R., and Stone, M. L. (2006). Spectral Reflectance to Estimate Genetic Variation for In-Season Biomass, Leaf Chlorophyll, and Canopy Temperature in Wheat. *Crop Science*, 46(3):1046–1057.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4.

Berger, K., Verrelst, J., Féret, J.-B., Wang, Z., Wocher, M., Strathmann, M., Danner, M., Mauser, W., and Hank, T. (2020). Crop nitrogen monitoring: Recent progress and principal developments in the context of imaging spectroscopy missions. *Remote Sensing of Environment*, 242:111758.

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J.-M., Jeuffroy, M.-H., Gate, P., Martre, P., and Le Gouis, J. (2010). Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. *Journal of Experimental Botany*, 61(15):4303–4312.

Bohman, B. J., Culshaw-Maurer, M. J., Ben Abdallah, F., Giletto, C., Bélanger, G., Fernández, F. G., Miao, Y., Mulla, D. J., and Rosen, C. J. (2023). Quantifying critical N dilution curves across G × E × M effects for potato using a partially-pooled Bayesian hierarchical method. *European Journal of Agronomy*, 144:126744.

Boulesteix, A.-L., Groenwold, R. H. H., Abrahamowicz, M., Binder, H., Briel, M., Hornung, R., Morris, T. P., Rahnenführer, J., and Sauerbrei, W. (2020). Introduction to statistical simulations in health research. *BMJ Open*, 10(12):e039921.

Brancourt-Hulmel, M., Doussinault, G., Lecomte, C., Bérard, P., Le Buanec, B., and Trottet, M. (2003). Genetic Improvement of Agronomic Traits of Winter Wheat Cultivars Released in France from 1946 to 1992. *Crop Science*, 43(1):37–45.

Calderini, D. F., Torres-León, S., and Slafer, G. A. (1995). Consequences of Wheat Breeding on Nitrogen and Phosphorus Yield, Grain Nitrogen and Phosphorus Concentration and Associated Traits. *Annals of Botany*, 76(3):315–322.

Cameron, K. C., Di, H. J., and Moir, J. L. (2013). Nitrogen losses from the soil/plant system: a review. *Annals of Applied Biology*, 162(2):145–173.

Cassman, K. G., Dobermann, A., and Walters, D. T. (2002). Agroecosystems, Nitrogen-use Efficiency, and Nitrogen Management. *AMBIO: A Journal of the Human Environment*, 31(2):132–140.

Charrad, M., Ghazzali, N., Boiteau, V., and Niknafs, A. (2014). NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, 61(6 SE - Articles):1–36.

Chavas, J.-P. and Nauges, C. (2020). Uncertainty, Learning, and Technology Adoption in Agriculture. *Applied Economic Perspectives and Policy*, 42(1):42–53.

Chawade, A., van Ham, J., Blomquist, H., Bagge, O., Alexandersson, E., and Ortiz, R. (2019). High-Throughput Field-Phenotyping Tools for Plant Breeding and Precision Agriculture.

Chen, P. (2015). A Comparison of Two Approaches for Estimating the Wheat Nitrogen Nutrition Index Using Remote Sensing.

Ciampitti, I., van Versendaal, E., Rybecky, J. F., Lacasa, J., Fernandez, J., Makowski, D., and Lemaire, G. (2022). A global dataset to parametrize critical nitrogen dilution curves for major crop species. *Scientific Data*, 9(1):277.

Ciampitti, I. A., Fernandez, J., Tamagno, S., Zhao, B., Lemaire, G., and Makowski, D. (2021). Does the critical N dilution curve for maize crop vary across genotype x environment x management scenarios? - a Bayesian analysis. *European Journal of Agronomy*, 123:126202.

Cohan, J.-P., Le Souder, C., Guicherd, C., Lorgeou, J., Du Cheyron, P., Bonnefoy, M., Decarrier, A., Piraux, F., and Laurent, F. (2019). Combining breeding traits and agronomic indicators to characterize the impact of cultivar on the nitrogen use efficiency of bread wheat. *Field Crops Research*, 242:107588.

Cormier, F., Faure, S., Dubreuil, P., Heumez, E., Beauchêne, K., Lafarge, S., Praud, S., and Le Gouis, J. (2013). A multi-environmental study of recent breeding progress on nitrogen use efficiency in wheat (Triticum aestivum L.). *Theoretical and Applied Genetics*, 126(12):3035–3048.

Cormier, F., Foulkes, J., Hirel, B., Gouache, D., Moënne-Loccoz, Y., and Le Gouis, J. (2016). Breeding for increased nitrogen-use efficiency: a review for wheat (T. aestivum L.). *Plant Breeding*, 135(3):255–278.

Correndo, A. A., Tremblay, N., Coulter, J. A., Ruiz-Diaz, D., Franzen, D., Nafziger, E., Prasad, V., Rosso, L. H. M., Steinke, K., Du, J., Messina, C. D., and Ciampitti, I. A. (2021). Unraveling uncertainty drivers of the maize yield response to nitrogen: A Bayesian and machine learning approach. *Agricultural and Forest Meteorology*, 311:108668.

Davidson, E. A., David, M. B., Galloway, J. N., Goodale, C. L., Haeuber, R., Harrison, J. A., Howarth, R. W., Jaynes, D. B., Lowrance, R. R., Nolan, B. T., Peel, J. L., Pinder, R. W., Porter, E., and Snyder, C. S. (2012). Excess Nitrogen in the U.S. Environment: Trends, Risks, and Solutions. *Issues in ecology*, (15).

Devienne-Barret, F., Justes, E., Machet, J. M., and Mary, B. (2000). Integrated Control of Nitrate Uptake by Crop Growth Rate and Soil Nitrate Availability under Field Conditions. *Annals of Botany*, 86(5):995–1005.

Dobermann, A. (2007). Nutrient use efficiency–measurement and management.

Dreccer, M. F., Barnes, L. R., and Meder, R. (2014). Quantitative dynamics of stem water soluble carbohydrates in wheat can be monitored in the field using hyperspectral reflectance. *Field Crops Research*, 159:70–80.

Drouet, J.-L. and Bonhomme, R. (2004). Effect of 3D nitrogen, dry mass per area and local irradiance on canopy photosynthesis within leaves of contrasted heterogeneous maize crops. *Annals of botany*, 93(6):699–710.

Efron, B. and Hastie, T. (2016). *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge University Press, Cambridge.

Eickhout, B., Bouwman, A. F., and van Zeijts, H. (2006). The role of nitrogen in world food production and environmental sustainability. *Agriculture, Ecosystems & Environment*, 116(1):4–14.

Feng, W., Zhang, H.-Y., Zhang, Y.-S., Qi, S.-L., Heng, Y.-R., Guo, B.-B., Ma, D.-Y., and Guo, T.-C. (2016). Remote detection of canopy leaf nitrogen concentration in winter wheat by using water resistance vegetation indices from in-situ hyperspectral data. *Field Crops Research*, 198:238–246.

Fernández, J. A., Lemaire, G., Bélanger, G., Gastal, F., Makowski, D., and Ciampitti, I. A. (2021). Revisiting the critical nitrogen dilution curve for tall fescue: A quantitative synthesis. *European Journal of Agronomy*, 131:126380.

Fernandez, J. A., Messina, C. D., Rotundo, J. L., and Ciampitti, I. A. (2021). Integrating nitrogen and water-soluble carbohydrates dynamics in maize: A comparison of hybrids from different decades. *CROP SCIENCE*, 61(2):1360–1373.

Fernandez, J. A., van Versendaal, E., Lacasa, J., Makowski, D., Lemaire, G., and Ciampitti, I. A. (2022). Dataset characteristics for the determination of critical nitrogen dilution curves: From past to new guidelines. *European Journal of Agronomy*, 139:126568.

Fischer, R. A. and Edmeades, G. O. (2010). Breeding and cereal yield progress. *Crop Science*, 50(April):S–85–S–98.

Foulkes, M. J., Hawkesford, M. J., Barraclough, P. B., Holdsworth, M. J., Kerr, S., Kightley, S., and Shewry, P. R. (2009). Identifying traits to improve the nitrogen economy of wheat: Recent advances and future prospects. *Field Crops Research*, 114(3):329–342.

Foulkes, M. J., Sylvester-Bradley, R., and Scott, R. K. (1998). Evidence for differences between winter wheat cultivars in acquisition of soil mineral nitrogen and uptake and utilization of applied fertilizer nitrogen. *The Journal of Agricultural Science*, 130(1):29–44.

Furbank, R. T. and Tester, M. (2011). Phenomics – technologies to relieve the phenotyping bottleneck. *Trends in Plant Science*, 16(12):635–644.

Gaju, O., Allard, V., Martre, P., Le Gouis, J., Moreau, D., Bogard, M., Hubbart, S., and Foulkes, M. J. (2014). Nitrogen partitioning and remobilization in relation to leaf senescence, grain yield and grain nitrogen concentration in wheat cultivars. *Field Crops Research*, 155:213–223.

Gaju, O., Allard, V., Martre, P., Snape, J. W., Heumez, E., LeGouis, J., Moreau, D., Bogard, M., Griffiths, S., Orford, S., Hubbart, S., and Foulkes, M. J. (2011). Identification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crops Research*, 123(2):139–152.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457–472.

Ghanem, M. E., Marrou, H., and Sinclair, T. R. (2015). Physiological phenotyping of plants for crop improvement. *Trends in Plant Science*, 20(3):139–144.

Hay, R. K. M. (1995). Harvest index: a review of its use in plant breeding and crop physiology. *Annals of Applied Biology*, 126(1):197–216.

Hefley, T. J., Baasch, D. M., Tyre, A. J., and Blankenship, E. E. (2014). Correction of location errors for presence-only species distribution models. *Methods in Ecology and Evolution*, 5(3):207–214.

Hefley, T. J., Hooten, M. B., Drake, J. M., Russell, R. E., and Walsh, D. P. (2016). When can the cause of a population decline be determined? *Ecology Letters*, 19(11):1353–1362.

Hefley, T. J., Hooten, M. B., Hanks, E. M., Russell, R. E., and Walsh, D. P. (2017). The Bayesian Group Lasso for Confounded Spatial Data. *Journal of Agricultural, Biological and Environmental Statistics*, 22(1):42–59.

Hefley, T. J., Tyre, A. J., and Blankenship, E. E. (2013). Fitting population growth models in the presence of measurement and detection error. *Ecological Modelling*, 263:244–250.

Huffman, W. E. (2020). Human Capital and Adoption of Innovations: Policy Implications. *Applied Economic Perspectives and Policy*, 42(1):92–99.

Jégo, G., Sansoulet, J., Pattey, E., Beaudoin, N., Bélanger, G., Ziadi, N., Tremblay, N., Grant, C., Tremblay, G., O'Donovan, J., Harker, K. N., Blackshaw, R., Johnson, E., and Justes, E. (2022). Determination of nitrogen dilution curves of corn, canola, and spring wheat in Canada using classical and Bayesian approaches. *European Journal of Agronomy*, 135:126481.

Jia, Z. and von Wirén, N. (2020). Signaling pathways underlying nitrogen-dependent changes in root system architecture: from model to crop species. *Journal of Experimental Botany*, 71(15):4393–4404.

Justes, E., Mary, B., Meynard, J.-M., Machet, J.-M., and Thelier-Huche, L. (1994). Determination of a Critical Nitrogen Dilution Curve for Winter Wheat Crops. *Annals of Botany*, 74(4):397–407.

Kanter, D. R., Chodos, O., Nordland, O., Rutigliano, M., and Winiwarter, W. (2020). Gaps and opportunities in nitrogen pollution policies around the world. *Nature Sustainability*, 3(11):956–963.

Krouk, G., Lacombe, B., Bielach, A., Perrine-Walker, F., Malinska, K., Mounier, E., Hoyerova, K., Tillard, P., Leon, S., Ljung, K., Zazimalova, E., Benkova, E., Nacry, P., and Gojon, A. (2010). Nitrate-Regulated Auxin Transport by NRT1.1 Defines a Mechanism for Nutrient Sensing in Plants. *Developmental Cell*, 18(6):927–937.

Lacasa, J., Makowski, D., Hefley, T., Fernandez, J., van Versendaal, E., Lemaire, G., and Ciampitti, I. (2023). Comparison of statistical methods to fit critical nitrogen dilution curves. *European Journal of Agronomy*, 145:126770.

Lammerts van Bueren, E. T. and Struik, P. C. (2017). Diverse concepts of breeding for nitrogen use efficiency. A review. *Agronomy for Sustainable Development*, 37(5):50.

Laperche, A., Brancourt-Hulmel, M., Heumez, E., Gardet, O., and Gouis, J. L. (2006). Estimation of genetic parameters of a DH wheat population grown at different N stress levels characterized by probe genotypes. *Theoretical and Applied Genetics*, 112(5):797–807.

Lemaire, G. and Ciampitti, I. (2020). Crop Mass and N Status as Prerequisite Covariables for Unraveling Nitrogen Use Efficiency across Genotype-by-Environment-by-Management Scenarios: A Review. *Plants*, 9(10):1309.

Lemaire, G., Gastal, F., and Salette, J. (1998). Analysis of the effect of N nutrition on dry matter yield of a sward by reference to potential yield and optimum N content. In *XVI International Grassland Congress*, pages 179–180, Nice, FR.

Lemaire, G., Jeuffroy, M.-H., and Gastal, F. (2008). Diagnosis tool for plant and crop N status in vegetative stage: Theory and practices for crop N management. *European Journal of Agronomy*, 28(4):614–624.

Lemaire, G., Tang, L., Bélanger, G., Zhu, Y., and Jeuffroy, M.-H. (2021). Forward new paradigms for crop mineral nutrition and fertilization towards sustainable agriculture. *European Journal of Agronomy*, 125:126248.

Li, X., Ata-UI-Karim, S. T., Li, Y., Yuan, F., Miao, Y., Yoichiro, K., Cheng, T., Tang, L., Tian, X., Liu, X., Tian, Y., Zhu, Y., Cao, W., and Cao, Q. (2022). Advances in the estimations and applications of critical nitrogen dilution curve and nitrogen nutrition index of major cereal crops. A review. *Computers and Electronics in Agriculture*, 197:106998.

Little, R. J. A. and Rubin, D. B. (2019). *Statistical analysis with missing data*, volume 793. John Wiley \& Sons.

Long, S. P., Zhu, X. G., Naidu, S. L., and Ort, D. R. (2006). Can improvement in photosynthesis increase crop yields?

Lopes, M. S., El-Basyoni, I., Baenziger, P. S., Singh, S., Royo, C., Ozbek, K., Aktas, H., Ozer, E., Ozdemir, F., Manickavelu, A., Ban, T., and Vikram, P. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *Journal of Experimental Botany*, 66(12):3477–3486.

Lopes, M. S., Reynolds, M. P., Manes, Y., Singh, R. P., Crossa, J., and Braun, H. J. (2012). Genetic Yield Gains and Changes in Associated Traits of CIMMYT Spring Bread Wheat in a "Historic" Set Representing 30 Years of Breeding. *Crop Science*, 52(3):1123–1131.

Lynch, J. P. (2019). Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. *New Phytologist*, 223(2):548–564.

Ma, B. L., Morrison, M. J., and Dwyer, L. M. (1996). Canopy Light Reflectance and Field Greenness to Assess Nitrogen Fertilization and Yield of Maize. *Agronomy Journal*, 88(6):915–920.

Makowski, D., Doré, T., and Monod, H. (2007). A new method to analyse relationships between yield components with boundary lines. *Agronomy for Sustainable Development*, 27(2):119–128.

Makowski, D. and Wallach, D. (2001). How to improve model-based decision rules for nitrogen fertilization. *European Journal of Agronomy*, 15(3):197–208.

Makowski, D., Wallach, D., and Tremblay, M. (2002). Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods. *Agronomie*, 22(2):191–203.

Makowski, D., Zhao, B., Ata-Ul-Karim, S. T., and Lemaire, G. (2020). Analyzing uncertainty in critical nitrogen dilution curves. *European Journal of Agronomy*, 118:126076.

Martre, P., Dueri, S., Guarin, J., Ewert, F., Webber, H., Calderini, D., Molero, G., Reynolds, M., Miralles, D., Garcia, G., and Others (2023). The nitrogen price of improved wheat yield under climate change.

Mistele, B. and Schmidhalter, U. (2008). Estimating the nitrogen nutrition index using spectral canopy reflectance measurements. *European Journal of Agronomy*, 29(4):184–190.

Moll, R. H., Kamprath, E. J., and Jackson, W. A. (1982). Analysis and Interpretation of Factors Which Contribute to Efficiency of Nitrogen Utilization1. *Agronomy Journal*, 74(3):562–564.

Morris, T. P., White, I. R., and Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11):2074–2102.

Nguyen, G. N. and Kant, S. (2018). Improving nitrogen use efficiency in plants: effective phenotyping in conjunction with agronomic and genetic approaches. *Functional Plant Biology*, 45(6):606–619.

Nguyen, G. N., Panozzo, J., Spangenberg, G., and Kant, S. (2016). Phenotyping approaches to evaluate nitrogen-use efficiency related traits of diverse wheat varieties under field conditions[*]. *Crop and Pasture Science*, 67(11):1139–1148.

Ortiz-MonasterioR, J. I., Sayre, K. D., Rajaram, S., and McMahon, M. (1997). Genetic Progress in Wheat Yield and Nitrogen Use Efficiency under Four Nitrogen Rates. *Crop Science*, 37(3):cropsci1997.0011183X003700030033x.

Pagano, E. and Maddonni, G. A. (2007). Intra-specific competition in maize: Early established hierarchies differ in plant growth and biomass partitioning to the ear around silking. *Field Crops Research*, 101(3):306–320.

Peng, C., Zhang, Z., Li, Y., Zhang, Y., Dong, H., Fang, Y., Han, L., Xu, W., and Hu, L. (2022). Genetic improvement analysis of nitrogen uptake, utilization, translocation, and distribution in Chinese wheat in Henan Province. *Field Crops Research*, 277:108406.

Pinto, F., Zaman-Allah, M., Reynolds, M., and Schulthess, U. (2023). Satellite imagery for high-throughput phenotyping in breeding plots. *Frontiers in Plant Science*, 14.

Plénet, D. and Lemaire, G. (2000). Relationships between dynamics of nitrogen uptake and dry matter accumulation in maize crops. Determination of critical N concentration. *Plant and Soil*, 216(1):65–82.

Plummer, M., Stukalov, A., and Denwood, M. (2016). rjags: Bayesian graphical models using MCMC. *R package version*, 4(6).

Pretini, N., Alonso, M. P., Vanzetti, L. S., Pontaroli, A. C., and González, F. G. (2021). The physiology and genetics behind fruiting efficiency: a promising spike trait to improve wheat yield potential. *Journal of Experimental Botany*, 72(11):3987–4004.

Prey, L., Hu, Y., and Schmidhalter, U. (2020). High-Throughput Field Phenotyping Traits of Grain Yield Formation and Nitrogen Use Efficiency: Optimizing the Selection of Vegetation Indices and Growth Stages. *Frontiers in Plant Science*, 10.

R Core Team (2019). R: A Language and Environment for Statistical Computing.

Raun, W. R. and Johnson, G. V. (1999). Improving Nitrogen Use Efficiency for Cereal Production. *Agronomy Journal*, 91(3):357–363.

Ravier, C., Meynard, J.-M., Cohan, J.-P., Gate, P., and Jeuffroy, M.-H. (2017). Early nitrogen deficiencies favor high yield, grain protein content and N use efficiency in wheat. *European Journal of Agronomy*, 89:16–24.

Ravier, C., Sabatier, R., Beillouin, D., Meynard, J.-M., Trépos, R., and Jeuffroy, M.-H. (2021). Decision rules for managing N fertilization based on model simulations and viability assessment. *European Journal of Agronomy*, 125:126247.

Reyes Ponce, A. (2016). The physiological basis of genetic improvement in maize (Zea mays L.) yield in the US Corn Belt.

Reynolds, M., Atkin, O. K., Bennett, M., Cooper, M., Dodd, I. C., Foulkes, M. J., Frohberg, C., Hammer, G., Henderson, I. R., Huang, B., Korzun, V., McCouch, S. R., Messina, C. D., Pogson, B. J., Slafer, G. A., Taylor, N. L., and Wittich, P. E. (2021). Addressing Research Bottlenecks to Crop Productivity. *Trends in Plant Science*, 26(6):607–630.

Reynolds, M., Foulkes, J., Fyrbank, R., Griffiths, S., King, J., Murchie, E., Parry, M., and Slafer, G. (2012). Achieving yield gains in wheat. *Plant, Cell & Environment*, 35(10):1799–1823.

Reynolds, M. P., van Ginkel, M., and Ribaut, J. (2000). Avenues for genetic modification of radiation use efficiency in wheat. *Journal of Experimental Botany*, 51(suppl_1):459–473.

Salette, J. and Lemaire, G. (1981). The variation of nitrogen content during the growth of grasses: a dilution equation [tall fescue, cocksfoot]. *Comptes Rendus Hebdomadaires des Seances de l'Academie des Sciences. Serie 3. Sciences de la Vie (France).*

Silva-Perez, V., Molero, G., Serbin, S. P., Condon, A. G., Reynolds, M. P., Furbank, R. T., and Evans, J. R. (2018). Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. *Journal of Experimental Botany*, 69(3):483–496.

Spiertz, J. H. J. (2009). Nitrogen, Sustainable Agriculture and Food Security: A Review BT - Sustainable Agriculture. pages 635–651. Springer Netherlands, Dordrecht.

Spilke, J., Piepho, H.-P., and Hu, X. (2005). A simulation study on tests of hypotheses and confidence intervals for fixed effects in mixed models for blocked experiments with missing data. *Journal of Agricultural, Biological, and Environmental Statistics*, 10(3):374–389.

Stevens, C. J. (2019). Nitrogen in the environment. *Science*, 363(6427):578–580.

Stewart, W. M. and Roberts, T. L. (2012). Food Security and the Role of Fertilizer in Supporting it. *Procedia Engineering*, 46:76–82.

Subbarao, G. V., Kishii, M., Bozal-Leorri, A., Ortiz-Monasterio, I., Gao, X., Ibba, M. I., Karwat, H., Gonzalez-Moro, M. B., Gonzalez-Murua, C., Yoshihashi, T., Tobita, S., Kommerell, V., Braun, H.-J., and Iwanaga, M. (2021). Enlisting wild grass genes to combat nitrification in wheat farming: A nature-based solution. *Proceedings of the National Academy of Sciences*, 118(35):e2106595118.

Subbarao, G. V. and Searchinger, T. D. (2021). A "more ammonium solution" to mitigate nitrogen pollution and boost crop yields. *Proceedings of the National Academy of Sciences*, 118(22):e2107576118.

Sutton, M. A., Howard, C. M., Kanter, D. R., Lassaletta, L., Móring, A., Raghuram, N., and Read, N. (2021). The nitrogen decade: mobilizing global action on nitrogen to 2030 and beyond. *One Earth*, 4(1):10–14.

Sylvester-Bradley, R. and Kindred, D. R. (2009). Analysing nitrogen responses of cereals to prioritize routes to the improvement of nitrogen use efficiency. *Journal of Experimental Botany*, 60(7):1939–1951.

Tanaka, T. S. T. (2021). Assessment of design and analysis frameworks for on-farm experimentation through a simulation study of wheat yield in Japan. *Precision Agriculture*, 22(5):1601–1616.

Tao, H., Xu, S., Tian, Y., Li, Z., Ge, Y., Zhang, J., Wang, Y., Zhou, G., Deng, X., Zhang, Z., Ding, Y., Jiang, D., Guo, Q., and Jin, S. (2022). Proximal and remote sensing in plant phenomics: 20 years of progress, challenges, and perspectives. *Plant Communications*, 3(6):100344.

Trethowan, R. M., Reynolds, M. P., Ortiz-Monasterio, J. I., and Ortiz, R. (2007). The Genetic Basis of the Green Revolution in Wheat Production. In *Plant Breeding Reviews*, pages 39–58.

Vega, C. R., Andrade, F. H., Sadras, V. O., Uhart, S. A., and Valentinuz, O. R. (2001). Seed number as a function of growth. A comparative study in soybean, sunflower, and maize. *Crop Science*, 41(3):748–754.

Wallach, D. (2011). Crop Model Calibration: A Statistical Perspective. *Agronomy Journal*, 103(4):1144–1151.

Wang, R., Rejesus, R. M., Tack, J. B., and Aglasan, S. (2021). Do higher temperatures influence how yields respond to increasing planting density? *Agricultural and Resource Economics Review*, 50(2):273–295.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43):1686.

Wu, L., Zhu, X., Lawes, R., Dunkerley, D., and Zhang, H. (2019). Comparison of machine learning algorithms for classification of LiDAR points for characterization of canola canopy structure. *International Journal of Remote Sensing*, 40(15):5973–5991.

Xu, Y., Yang, Y., Chen, X., and Liu, Y. (2022). Bibliometric Analysis of Global NDVI Research Trends from 1985 to 2021.

Yang, W., Feng, H., Zhang, X., Zhang, J., Doonan, J. H., Batchelor, W. D., Xiong, L., and Yan, J. (2020). Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Molecular Plant*, 13(2):187–214.

Yao, B., Ata-Ul-Karim, S. T., Li, Y., Ye, T., Zhu, Y., Cao, W., Cao, Q., and Tang, L. (2023). Plant nitrogen status at phenological stages can well estimate wheat yield and its components. *Field Crops Research*, 297:108950.

Yao, B., He, H.-b., Xu, H.-c., Zhu, T.-z., Liu, T., Ke, J., You, C.-c., Zhu, D.-q., and Wu, L.-q. (2021). Determining nitrogen status and quantifying nitrogen fertilizer requirement using a critical nitrogen dilution curve for hybrid indica rice under mechanical pot-seedling transplanting pattern. *Journal of Integrative Agriculture*, 20(6):1474–1486.

York, L. M. and Lynch, J. P. (2015). Intensive field phenotyping of maize (Zea mays L.) root crowns identifies phenes and phene integration associated with plant growth and nitrogen acquisition. *Journal of Experimental Botany*, 66(18):5493–5505.

Zhao, B., Ata-Ul-Karim, S. T., Duan, A., Liu, Z., Wang, X., Xiao, J., Liu, Z., Qin, A., Ning, D., Zhang, W., and Lian, Y. (2018). Determination of critical nitrogen concentration and dilution curve based on leaf area index for summer maize. *Field Crops Research*, 228:195–203.

Zhao, Y., Wang, J.-w., Chen, L.-p., Fu, Y.-y., Zhu, H.-c., Feng, H.-k., Xu, X.-g., and Li, Z.-h. (2021). An entirely new approach based on remote sensing data to calculate the nitrogen nutrition index of winter wheat. *Journal of Integrative Agriculture*, 20(9):2535–2551.

Ziadi, N., Brassard, M., Bélanger, G., Claessens, A., Tremblay, N., Cambouris, A. N., Nolin, M. C., and Parent, L.-É. (2008). Chlorophyll Measurements and Nitrogen Nutrition Index for the Evaluation of Corn Nitrogen Status. *Agronomy Journal*, 100(5):1264–1273.

Appendix A

Supplementary Figures

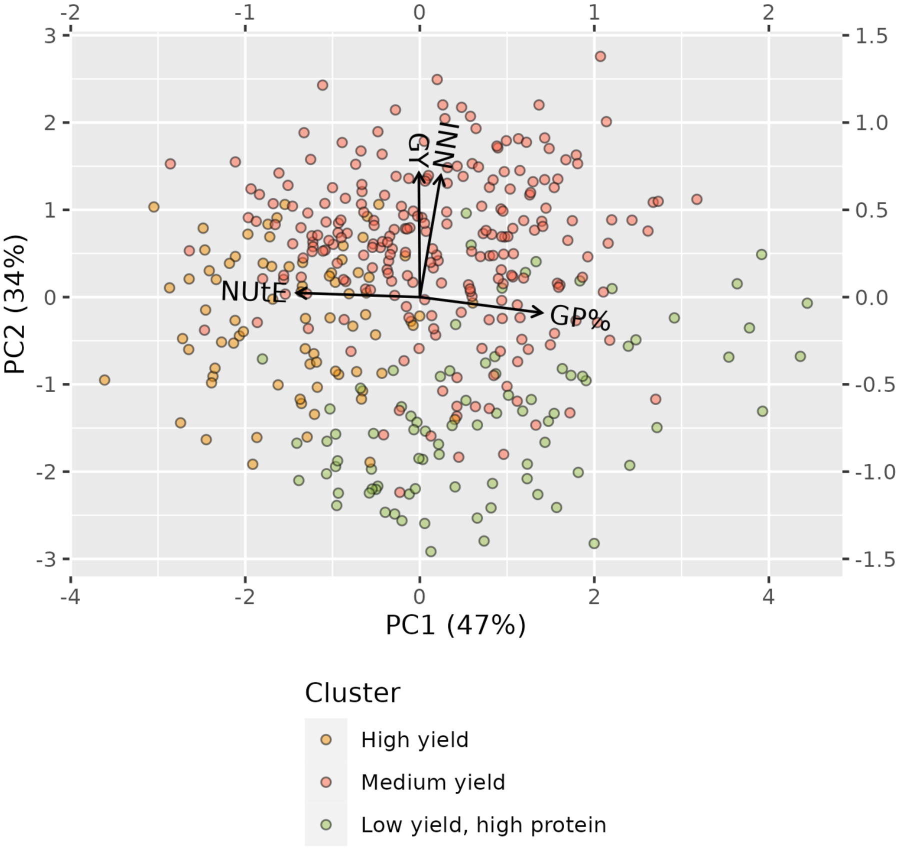

Figure A.1: *Relationships between grain yield (GY),grain protein (GP%), Nitrogen Utilization Efficiency (NUtE) and Nitrogen Nutrition Index at anthesis (NNI) on the first two axes of a principal component (PC) analysis for data of wheat cultivars grown under low-N environments.*

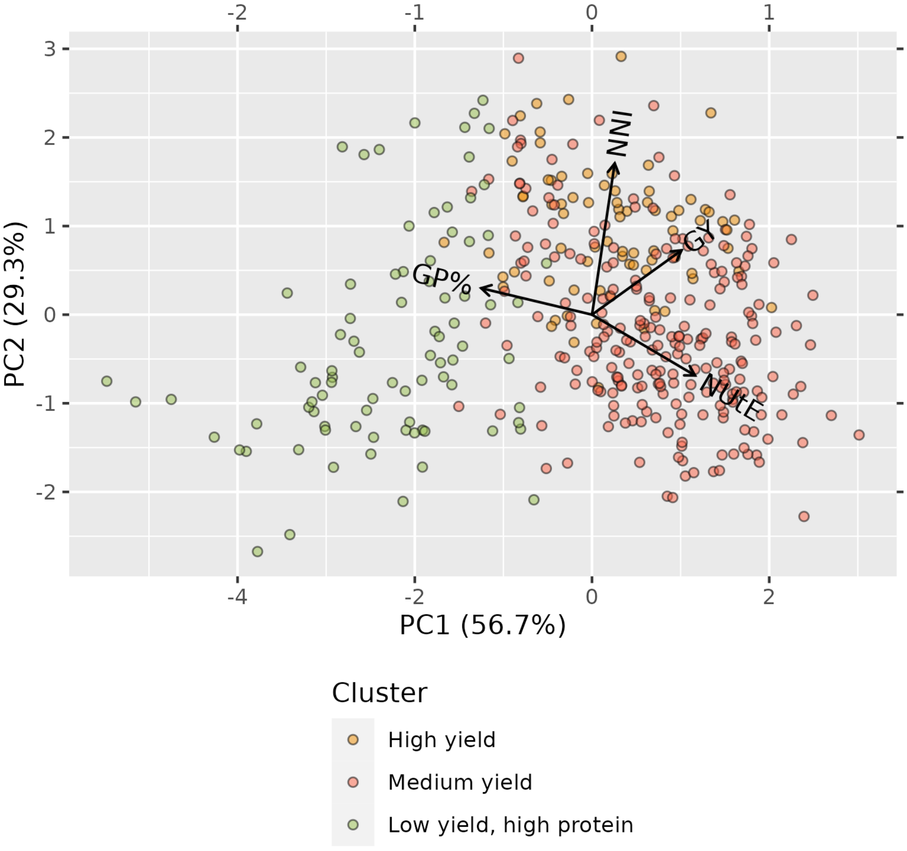

Figure A.2: *Relationships between grain yield (GY),grain protein (GP%), Nitrogen Utilization Efficiency (NUtE) and Nitrogen Nutrition Index at anthesis (NNI) on the first two axes of a principal component (PC) analysis for data of wheat cultivars grown under high-N environments.*