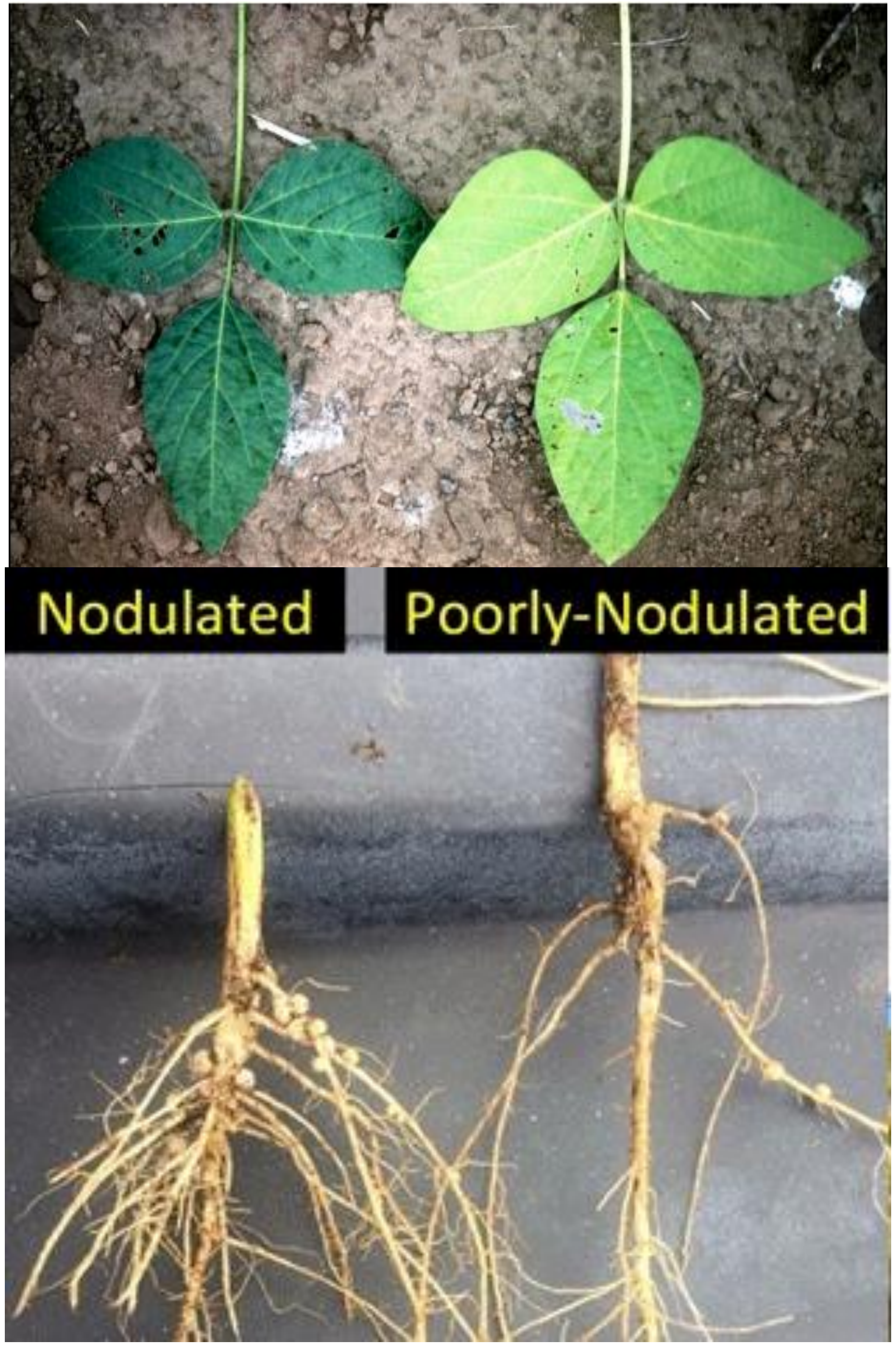
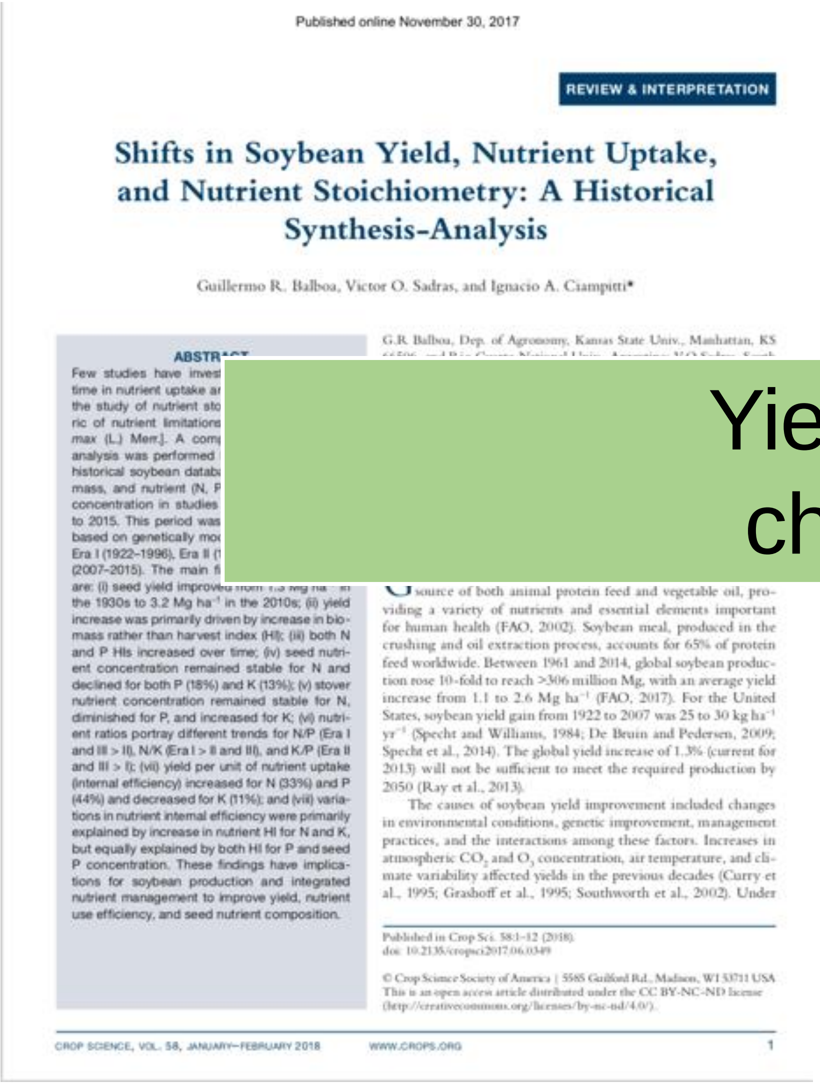


Unlocking the linkage between breeding and production research for N fixation, protein, and yields

Ignacio Ciampitti
Professor, Digital Ag Farming Systems
Department of Agronomy

Ciampitti Lab

Historical characterization



REVIEW & INTERPRETATION

Shifts in Soybean Yield, Nutrient Uptake, and Nutrient Stoichiometry: A Historical Synthesis-Analysis

Guillermo R. Balboa, Victor O. Sadras, and Ignacio A. Ciampitti*

ABSTR

Few studies have invest
time in nutrient uptake ar
the study of nutrient sto
ric of nutrient limitations
max (L.) Merr.]. A comp
analysis was performed
historical soybean databa
mass, and nutrient (N, P
concentration in studies
to 2015. This period was
based on genetically mod
Era I (1922–1996), Era II (1
(2007–2015). The main fi
are: (i) seed yield improved
the 1930s to 3.2 Mg ha^{-1} in the 2010s; (ii) yield increase was primarily driven by increase in biomass rather than harvest index (HI); (iii) both N and P HIs increased over time; (iv) seed nutrient concentration remained stable for N and declined for both P (18%) and K (13%); (v) stover nutrient concentration remained stable for N, diminished for P, and increased for K; (vi) nutrient ratios portray different trends for N/P (Era I and III > II), N/K (Era I > II and III), and K/P (Era II and III > I); (vii) yield per unit of nutrient uptake (internal efficiency) increased for N (33%) and P (44%) and decreased for K (11%); and (viii) variations in nutrient internal efficiency were primarily explained by increase in nutrient HI for N and K, but equally explained by both HI for P and seed P concentration. These findings have implications for soybean production and integrated nutrient management to improve yield, nutrient use efficiency, and seed nutrient composition.

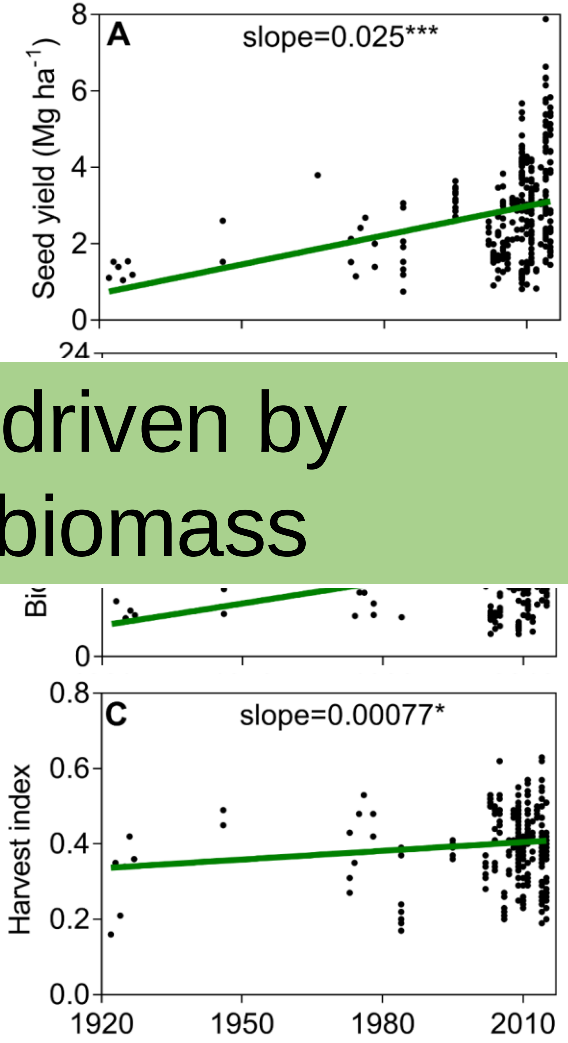
G.R. Balboa, Dep. of Agronomy, Kansas State Univ., Manhattan, KS

source of both animal protein feed and vegetable oil, providing a variety of nutrients and essential elements important for human health (FAO, 2002). Soybean meal, produced in the crushing and oil extraction process, accounts for 65% of protein feed worldwide. Between 1961 and 2014, global soybean production rose 10-fold to reach >306 million Mg, with an average yield increase from 1.1 to 2.6 Mg ha^{-1} (FAO, 2017). For the United States, soybean yield gain from 1922 to 2007 was 25 to 30 kg ha^{-1} yr^{-1} (Specht and Williams, 1984; De Bruin and Pedersen, 2009; Specht et al., 2014). The global yield increase of 1.3% (current for 2013) will not be sufficient to meet the required production by 2050 (Ray et al., 2013).

The causes of soybean yield improvement included changes in environmental conditions, genetic improvement, management practices, and the interactions among these factors. Increases in atmospheric CO_2 and O_3 concentration, air temperature, and climate variability affected yields in the previous decades (Curry et al., 1995; Grashoff et al., 1995; Southworth et al., 2002). Under

Published in Crop Sci. 58:1–12 (2018).
doi: 10.2135/cropsci2017.06.0349

+360%



Yield

Yield mostly driven by
changes in biomass

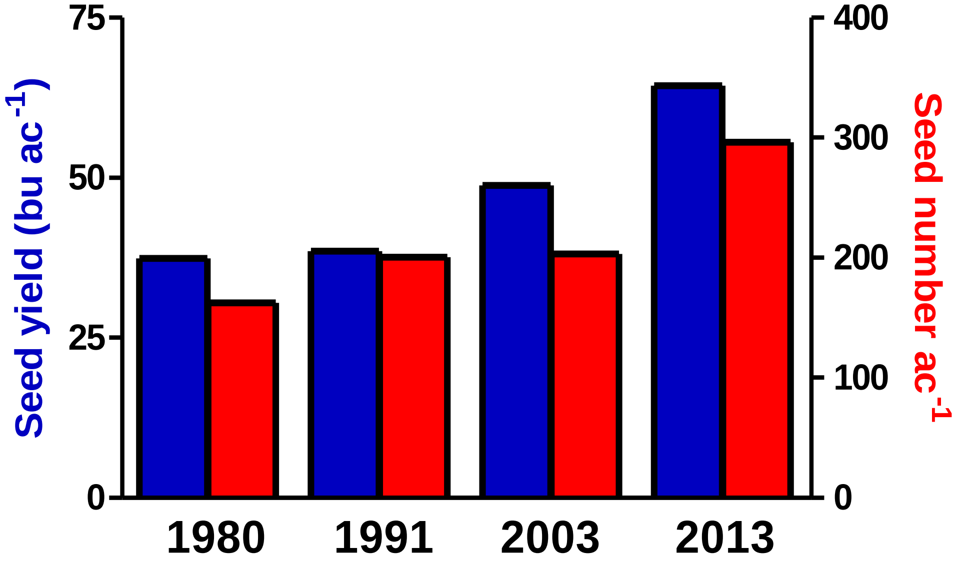
+23%

Harvest Index

(Balboa, Sadras, and Ciampitti, 2018)

Historical characterization

Ciampitti Lab



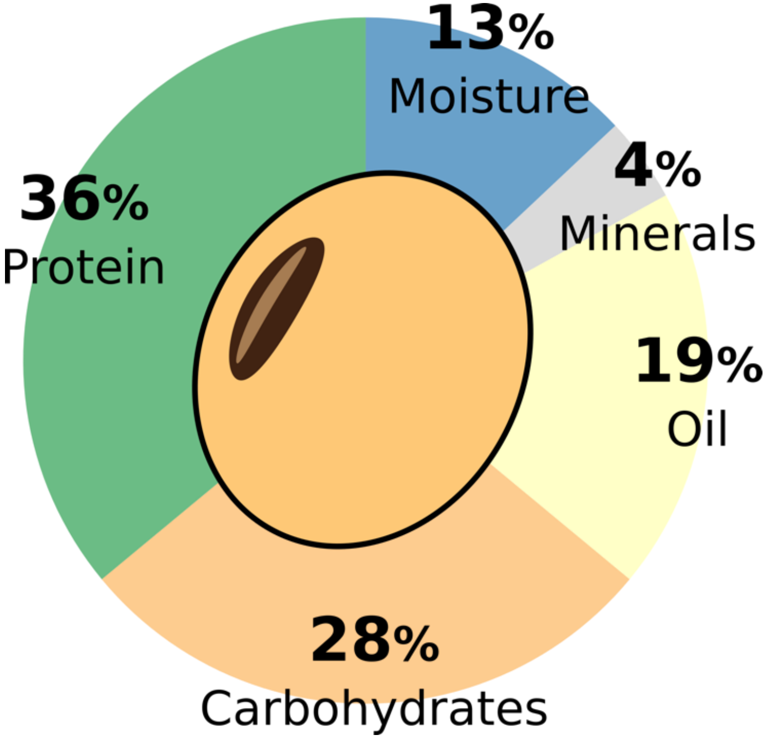
Seed number followed the same trend that seed yield. Improving with genetic gain.

Seed weight did not follow a similar trend being more variable through the years.

(Tamagno, Ciampitti, et al. 2018)

Linkage between breeding, production for N fixation, protein, and yields

History: Shift in soybean seed composition

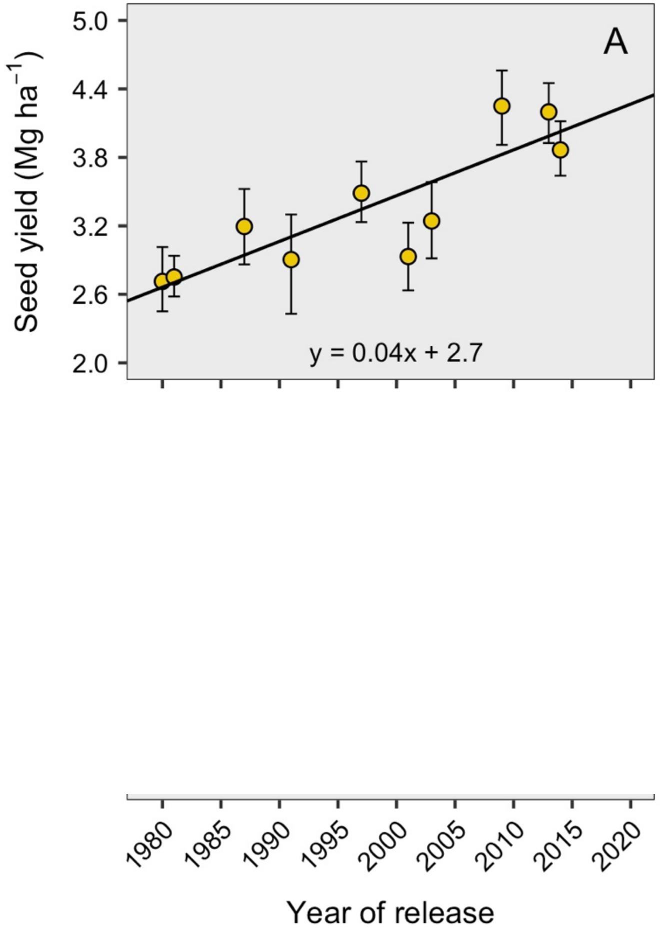


Yield increased **by**

0.6 bu/a/yr

Protein decreased **by**

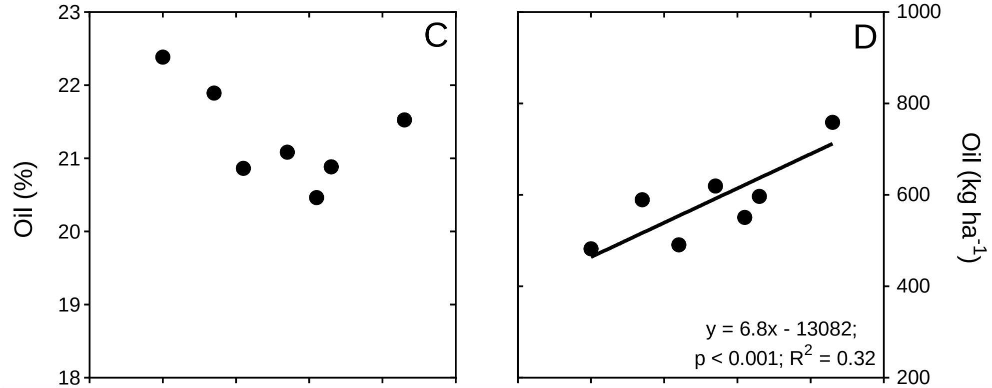
-0.12 g/kg/yr

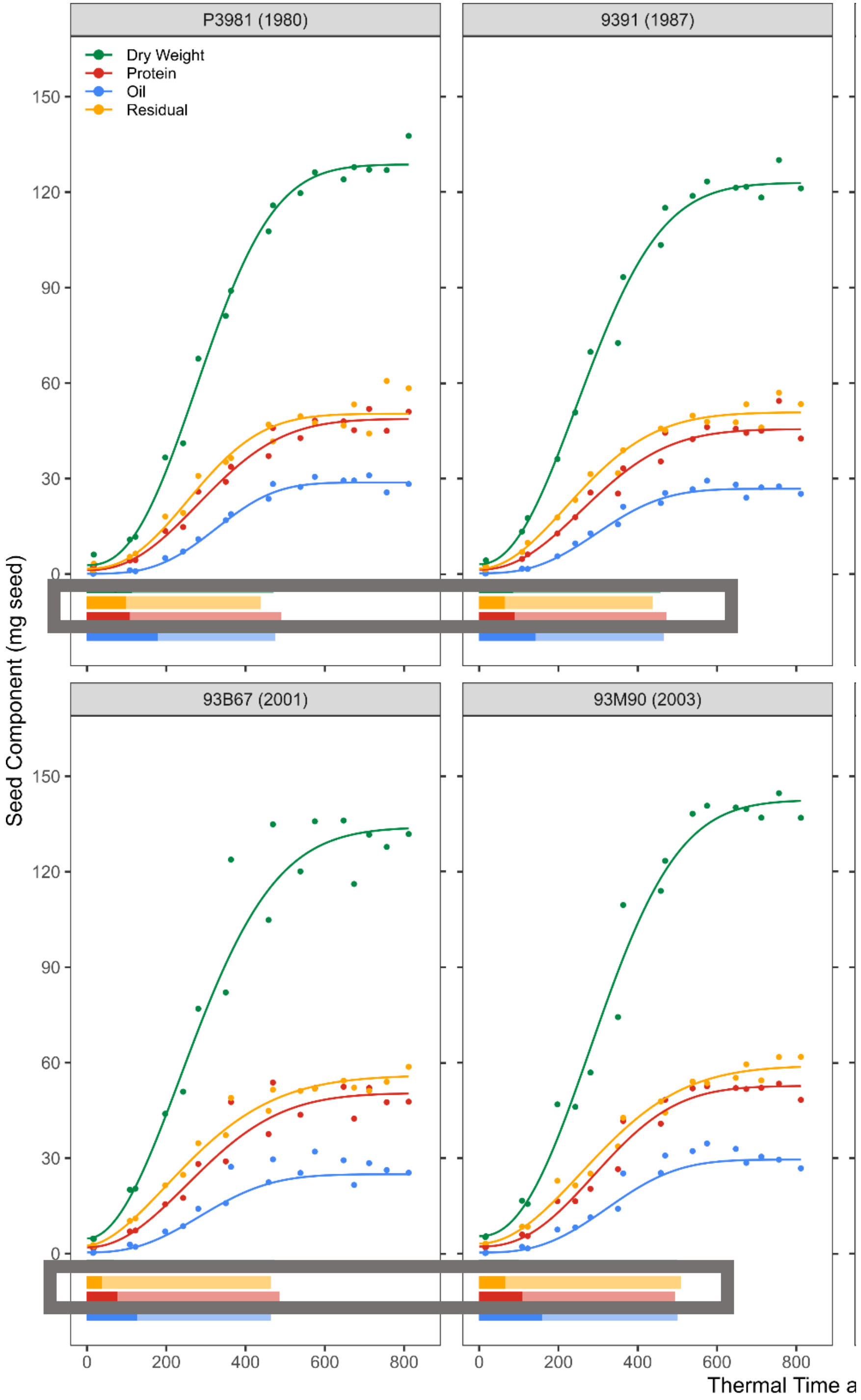


de Borja Reis, Moro Rosso, Tamagno and Ciampitti, 2020 (Scientific Reports)

Shift on seed composition

Changes in seed composition during seed filling period for "older" and "modern" genotypes

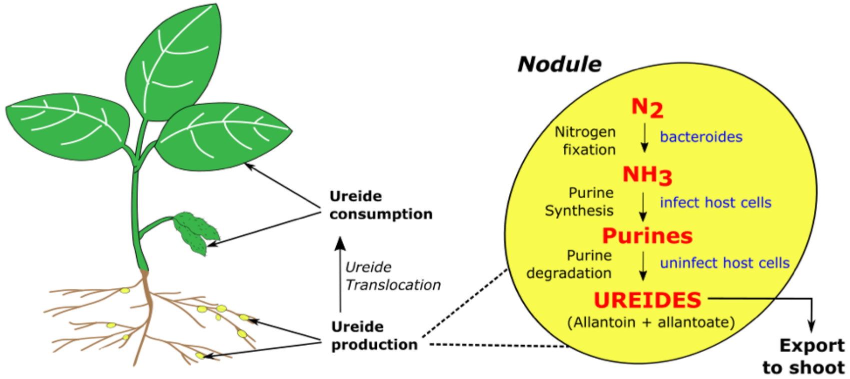




Ciampitti Lab

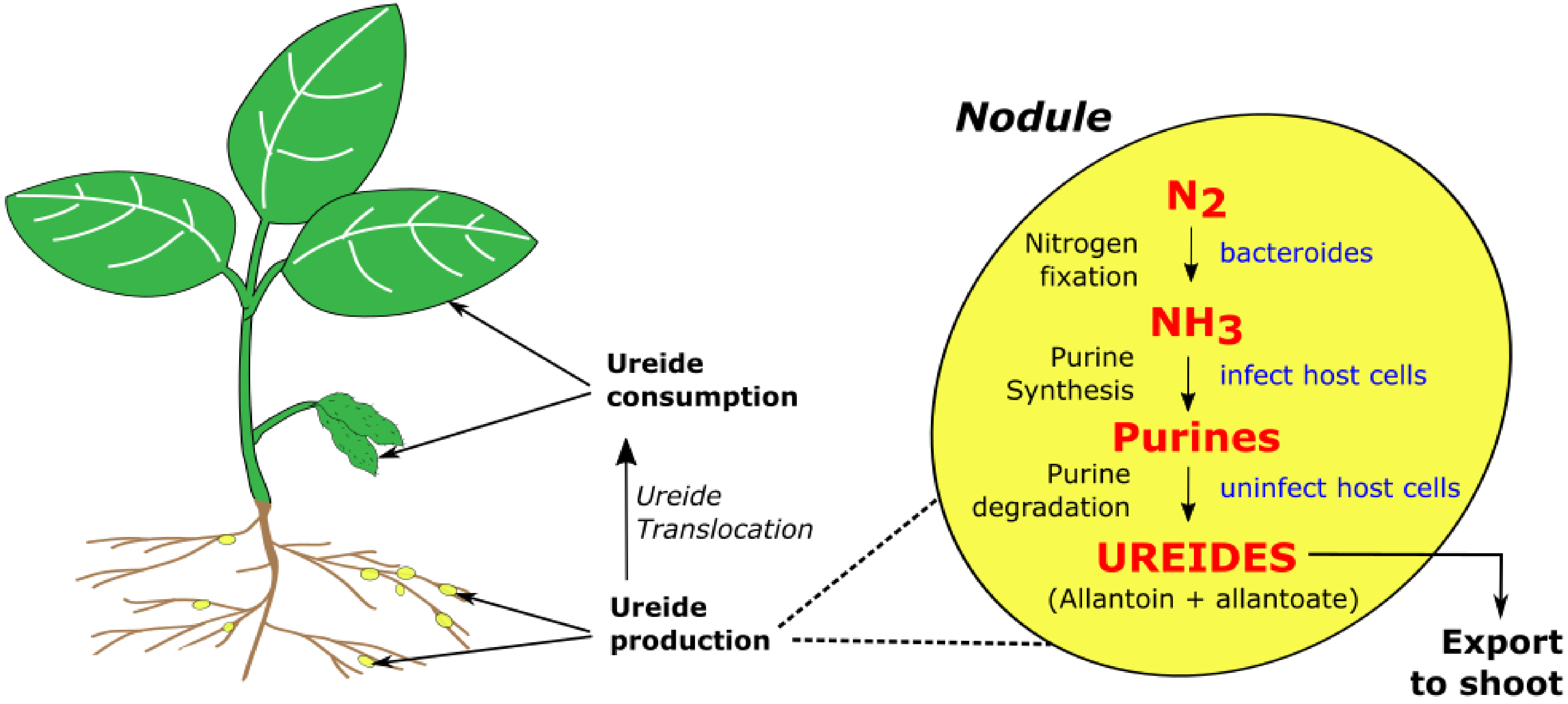
Biological N Fixation and Plant N Demand

Nitrogen deficiency, soybeans can present deficiency to N in situations in which the nodulation was not well established, and N fixation is not functional.



Extension Publication on Soybean N fixation:
https://bookstore.ksre.ksu.edu/pubs/MF3462.pdf

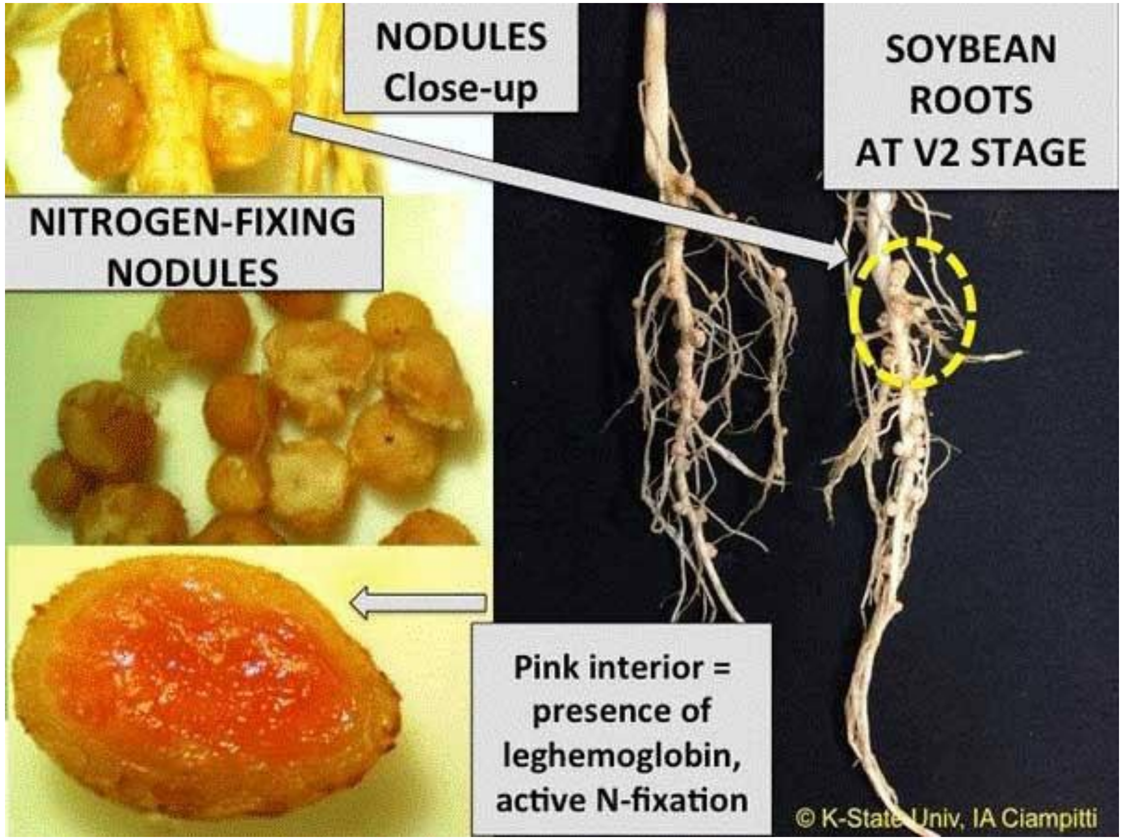
Biological N fixation

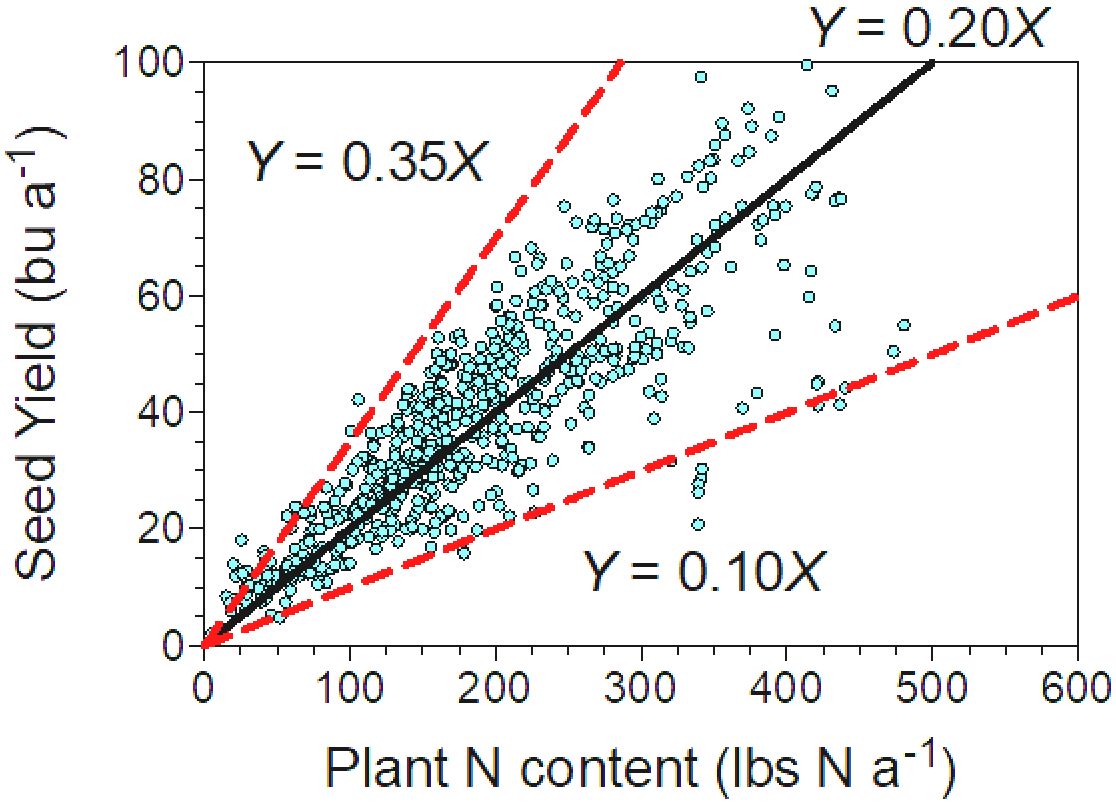


Soybean plant N demand is primarily satisfied by biological N fixation (BNF). Recent studies analyzing heavy-fertilized soybean crops suggest the presence of an "**N-gap**" (difference between plant N demand and N provided by BNF + soil).

Ciampitti and Salvagiotti, 2018 (Agron. J.)

Biological N Fixation and Plant N Demand





Plant N demand increases with yield, **100 lbs N/acre per 20 bu/acre**

Ciampitti and Salvagiotti, 2018 (Agron. J.)

Nitrogen and impact on yields

Published August 9, 2018

SOIL FERTILITY AND CROP NUTRITION

Exploring Nitrogen Limitation for Historical and Modern Soybean Genotypes

O. A. Ortez, F. Salvagiotti, J. M. Enrico, P. V. V. Prasad, P. Armstrong, and I. A. Ciampitti*

ABSTRACT

The United States (USA) and Argentina (ARG) account for over 50% of the global soybean [*Glycine max* (L.) Merr.] production. Soybean N demand is partially met (50–60%) by the biological nitrogen fixation (BNF) process; however, an unanswered scientific knowledge gap exists on the ability of the BNF process to fulfill soybean N demand at varying yield levels. The overall objective of this study is to explore the potential N limitation using different N strategies for historical and modern soybean genotypes. Four field experiments were conducted during 2016 and 2017 growing seasons in Kansas (USA) and Santa Fe (ARG). Twenty-one historical and modern soybean genotypes released from the 1980s to 2010s were tested under three N treatments: (i) control, without N application (Zero-N); (ii) 56 kg N ha^{-1} applied at R3-R4 growth stages (Late-N); and (iii) 670 kg ha^{-1} equally split at planting, R1, and R3–R4 growth stages (Full-N). Historical soybean yield gains, from the 1980s to 2010s, were 29% in the USA and 21% in ARG. Following the yield trend, seed N content increased for modern genotypes in parallel to the reduction on seed protein concentration. Regarding N treatments, Full-N produced 12% yield increase in the USA and 4% in ARG. Yield improvement was mainly related to increases in aboveground biomass, seed number (genotype effect), and to a lesser extent, to seed weight (N effect). This study suggests a potential N limitation for soybean, although there are still questions about the way in which N must be provided to the plant.

Core Ideas

- Yields (seed number) increased over time with modern soybean genotypes.
- Seed protein concentration decreased over time.
- Nitrogen fertilization impacted yield via changes on the seed weight.
- Nitrogen limited yields for high-yielding modern soybean genotypes.

Soybean [*Glycine max* (L.) Merr.] is considered as the main source for vegetable oil and animal protein feed in the world (FAO, 2002). The United States (USA) and Argentina (ARG) account for more than 50% of the global soybean production (USDA-NASS, 2017). In the USA, more than 85% of the soybean area is in the Corn Belt region, where it is mainly planted in rotation with corn (*Zea mays* L.) (>60%). In ARG, soybean is primarily planted in the Rolling Pampas and Chaco regions, mainly after wheat (*Triticum aestivum* L.), and after corn to a lesser extent.

Soybean yield potential (Yp) is genetically determined and attained under ideal conditions (genotype × environment × management practices, G × E × M), assuming no limitations in resources (e.g., water and nutrient supply) and in absence of any biotic (e.g., insects, diseases) and abiotic (e.g., temperature, drought, salinity) yield-limiting factors (Evans, 1993). Yield gap between Yp and actual farmer yield (YA) is primarily defined by the interacting effect between genotypes (material selection), the environment (soil + weather), and management practices (e.g., planting date, nutrient and pest management).

A historical yield analysis for soybean showed that seed yield improved by 246% (1300 vs. 3200 kg ha^{-1}) from the 1930s to 2010s (Balboa et al., 2018). Annual seed yield increases of 31 kg ha^{-1} in the USA (Specht et al., 1999) and 28 kg ha^{-1} globally (Wilcox, 2004) were reported from the 1970s to 2000s. As yield increased, a negative effect on seed protein was recorded by Rowntree et al. (2013), with a 0.19 g kg^{-1} yr^{-1} decrease in seed protein for maturity group (MG) II and 0.24 g kg^{-1} yr^{-1} decrease for MG III, from the 1920s and 2000s. Changes in seed yield and seed protein concentration were a consequence of both genetic (Boerma, 1979; Specht and Williams, 1984; Voldeng et al., 1997; Wilson et al., 2014; de Felipe et al., 2016) and management practices (Frederick et al., 1991; Heatherly and Elmore, 2004; Bastidas et al., 2008; Bradley and Sweets, 2008).

Published in Agron. J. 110:2080–2090 (2018)
doi:10.2134/agronj2018.04.0271
Available freely online through the author-supported open access option

O.A. Ortez, P.V.V. Prasad, and I.A. Ciampitti, Dep. of Agronomy, Kansas State Univ., Manhattan, Kansas 66506; F. Salvagiotti, J.M. Enrico, Crops, Soils, and Water Management Group, EEA INTA Oliveros, Route 11 km 353 (C 2206), Santa Fe Province, Argentina; P.V.V. Prasad, Sustainable Intensification Innovation Lab, Kansas State Univ., Manhattan, KS 66506; P. Armstrong, Center for Grain and Animal Health Research, USDA-ARS, Manhattan, KS 66502. Received 20 Apr. 2018. Accepted 11 June 2018. *Corresponding author (ciampitti@ksu.edu).

Abbreviations: ADM, aboveground dry biomass; ANOVA, analysis of variance; ARG, Argentina; BNF, biological nitrogen fixation; Full-N, 670 kg ha^{-1} equally split at planting, R1, and R3–R4 growth stages; HI, harvest index; Late-N, 56 kg N ha^{-1} applied at R3-R4 growth stages; MG, maturity group; UAN, urea ammonium nitrate; USA, United States; Zero-N, control, without N application.

Trials	Genotypes	Released Year
KS, US (13)	P3981, Williams82, 9391, 9392, P93B82, 93B67, 93M90, 93Y92, 94Y23, P35T58R, P39T67R, P31T11R, and P34T43R2	1980, 1981, 1987, 1991, 1997, 2001, 2003, 2009, 2013, 2013, 2013, 2014, and 2014
ARG (8)	Williams, A4422, DM49, A3910, DM4800, DM3700, NS4955, and SRM3988	1984, 1988, 1990, 1994, 2000, 2003, 2014, and 2015

N Fertilizer Rates

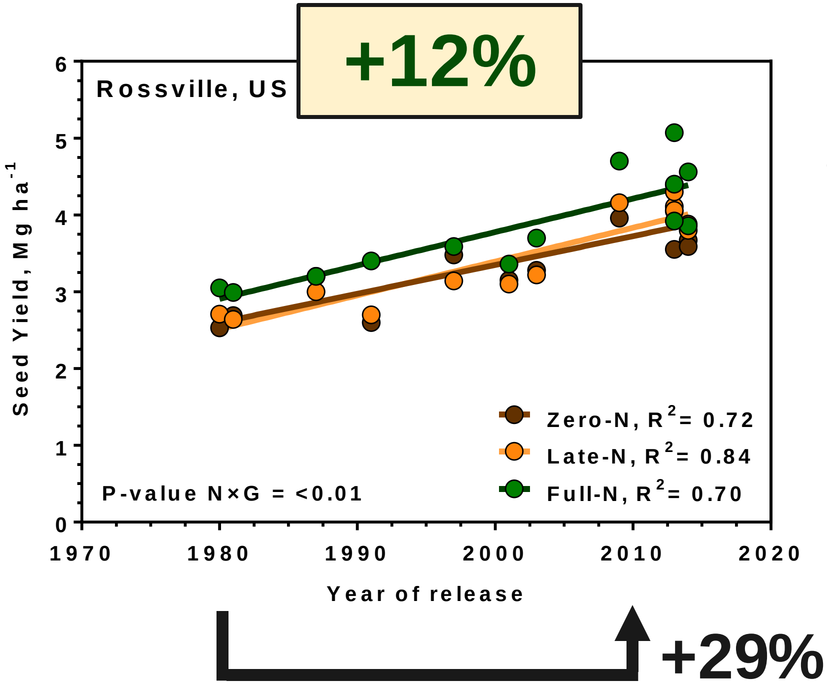
Zero-N: control

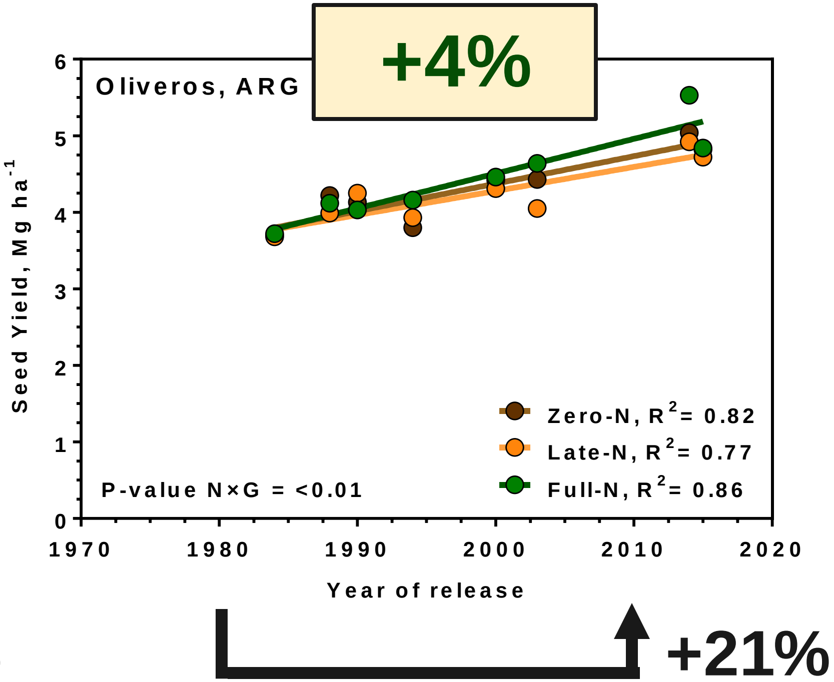
Late-N: 56 kg ha^{-1}; applied at the R3 stage

Full-N: 670 kg ha^{-1}, non-limiting; equal split at planting, R1, and R3

Ortez, Ciampitti et al., 2018 (Agron. J.)

Historical Yield Gains: Yield x nitrogen





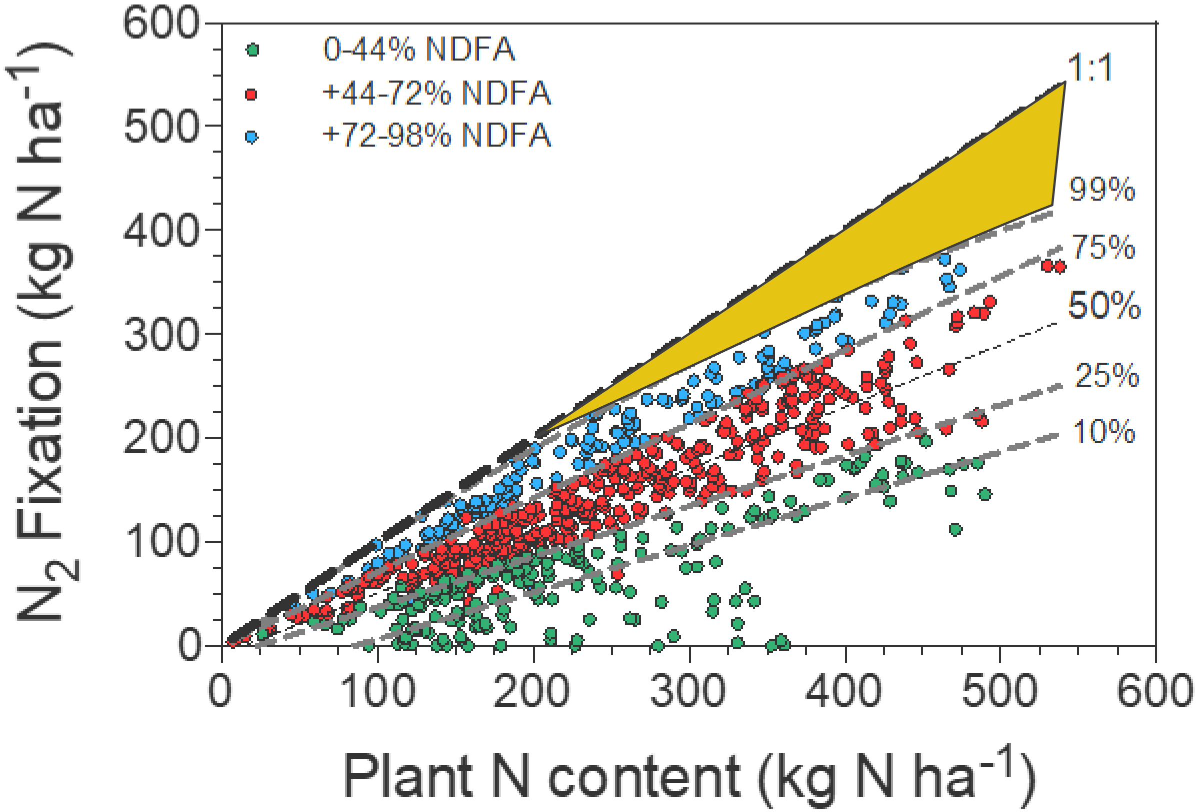
Greater yields:

Full-N, +12% in US and +4% in ARG

Modern genotypes, +29% in US and +21% in ARG

Ortez, Ciampitti et al., 2018 (Agron. J.)

Seasonal changes in N fixation

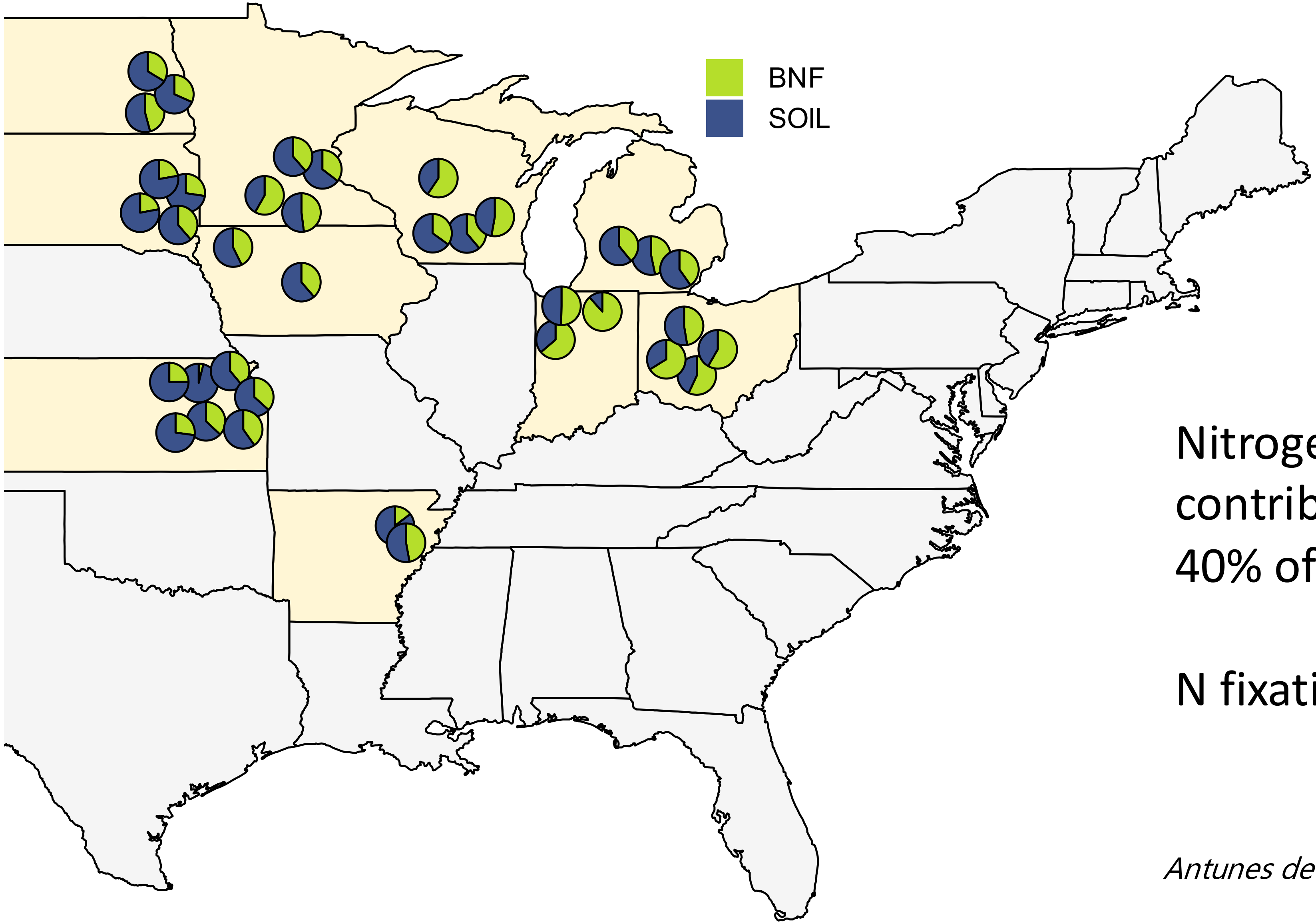


N-gap, plant N demand not satisfied by N fixation

N fixation satisfies N demand until **225 kg N/ha or 200 lbs/acre**, increasing the N-gap as the plant N demand increases

Ciampitti and Salvagiotti, 2018 (Agron. J.)

US Soybean N Fixation map



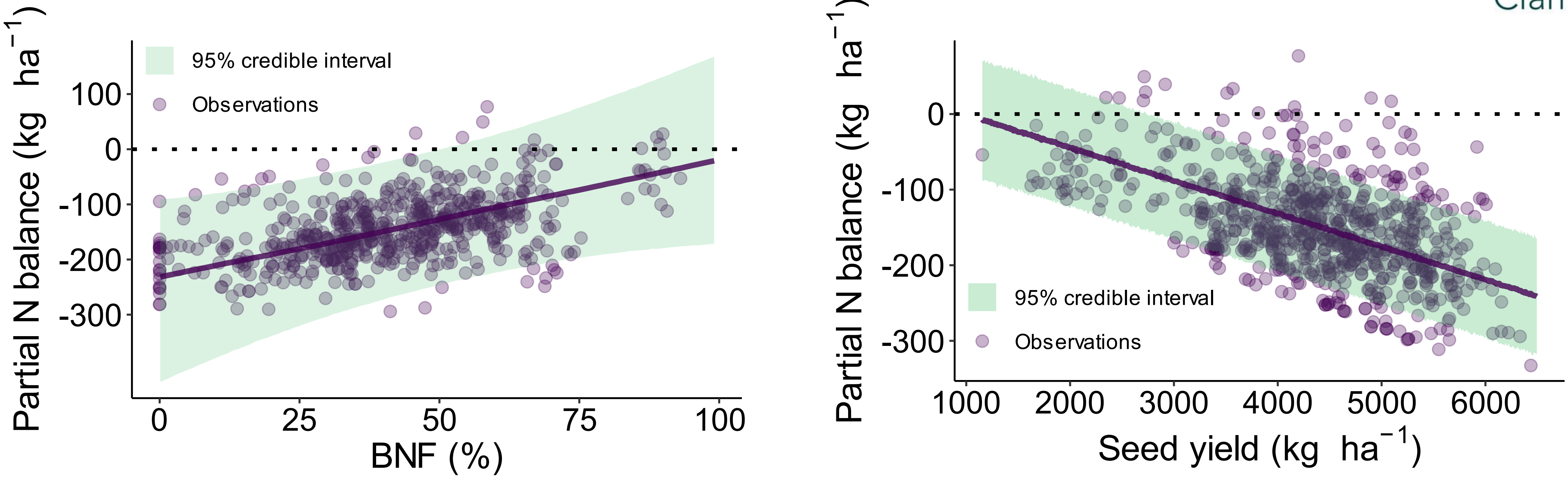
Nitrogen fixation contributed to an average of 40% of the total N demand

N fixation: from 5 % to 90%

Antunes de Almeida, Ciampitti et al. (2024)

Partial N balance in the US

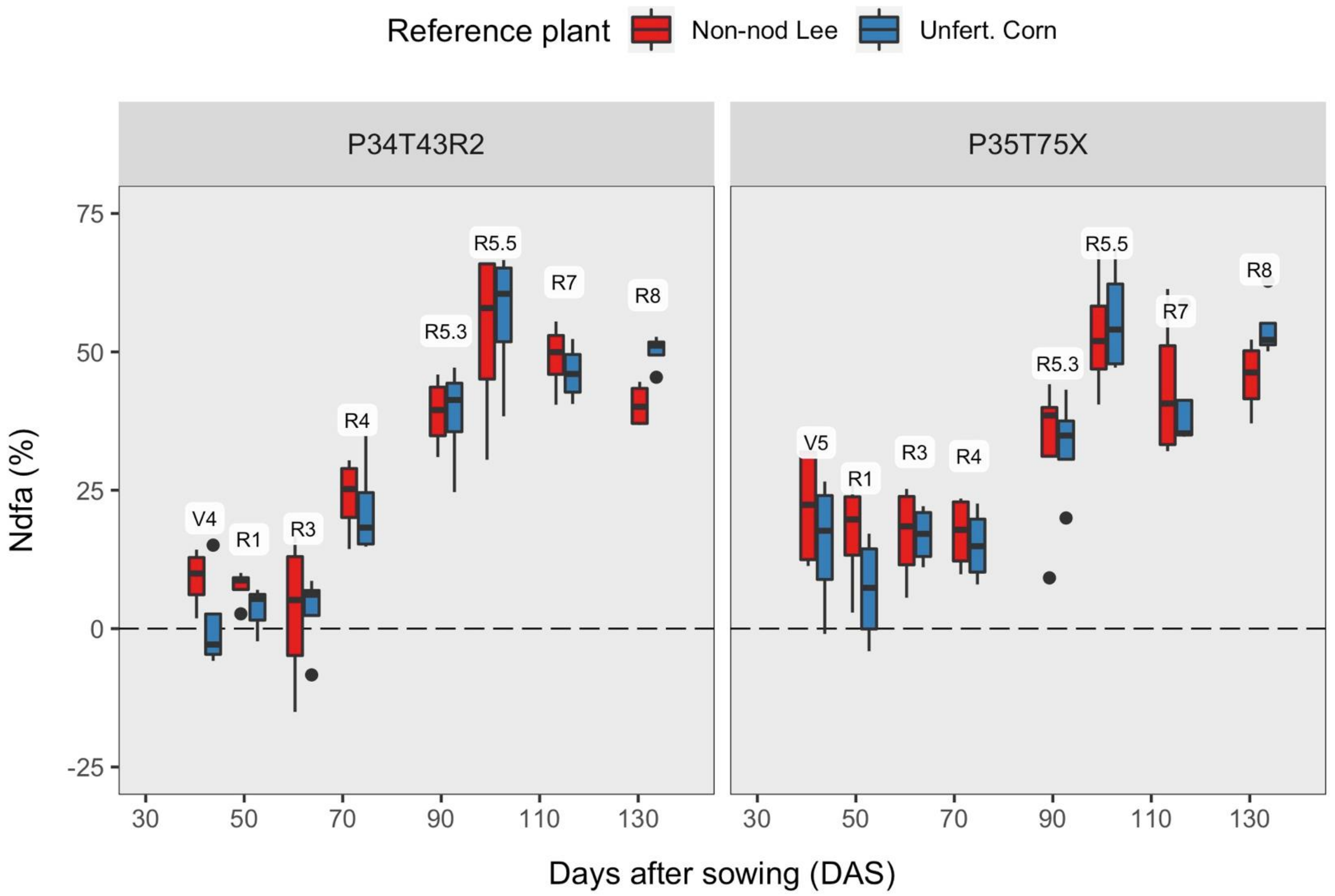
Ciampitti Lab



The assessment of N fixation is crucial for a sustainable soybean-based farming system

Antunes de Almeida, Ciampitti et al. (2025)

Seasonal changes in N fixation

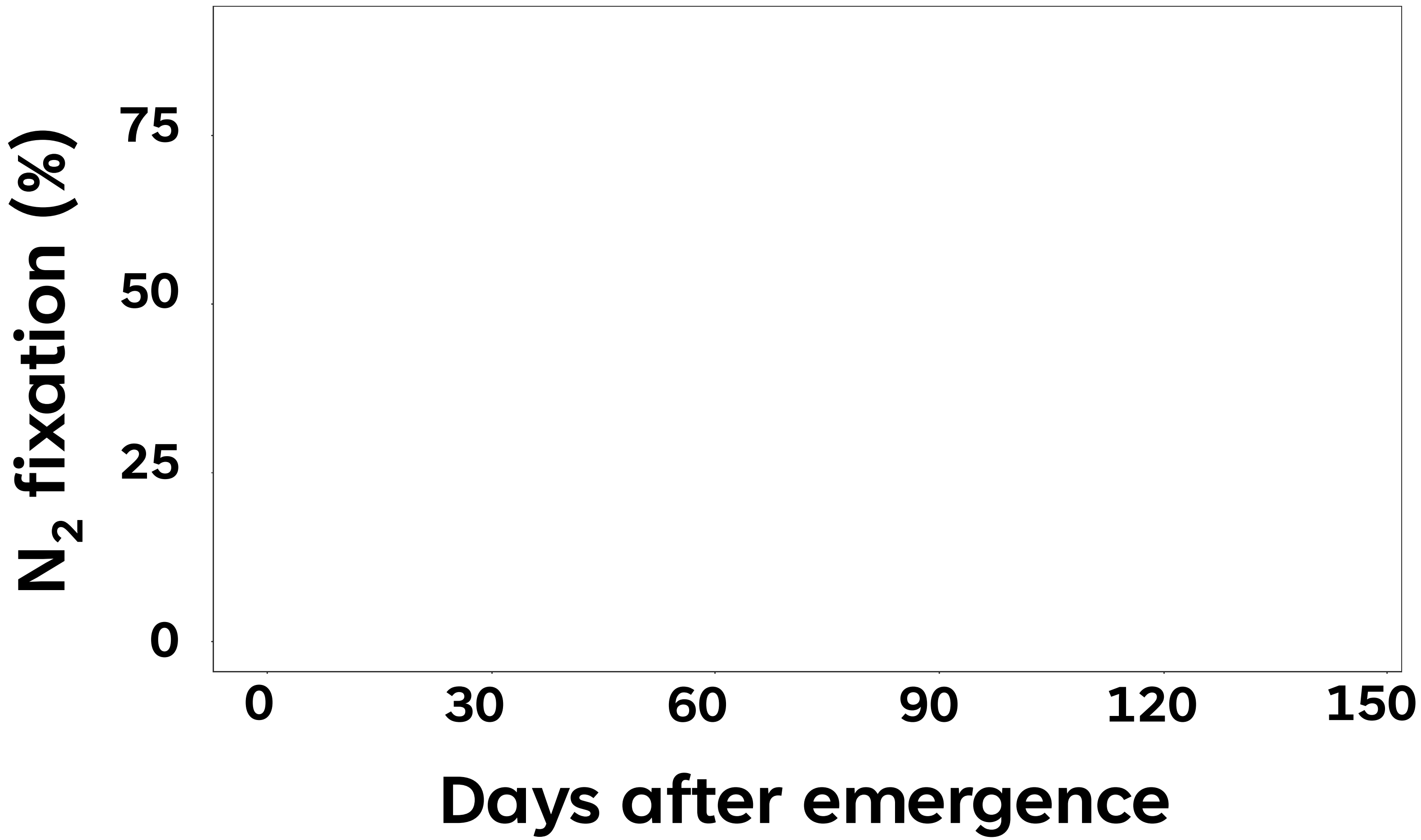


Seasonal N fixation increases with the growing season with a peak during the seed filling period, declining rate after R5.5 stage

Moro Rosso, Ciampitti, et al. 2022 (Europ J Agron)

Seasonal N Fixation Patterns (38 sites, data from 2021 and 2022)

Ciampitti Lab



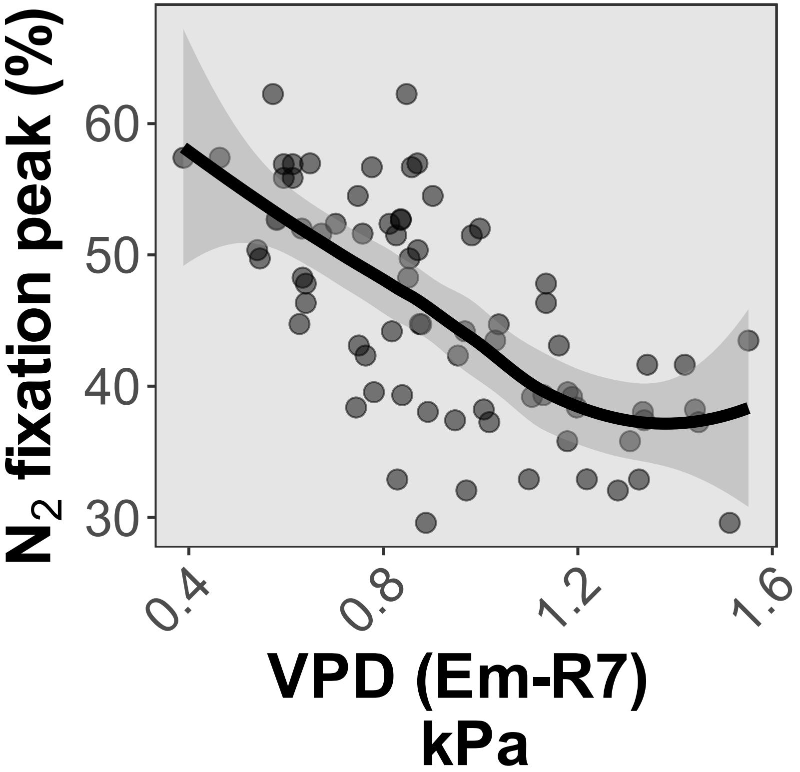
N fixation peaked from 7 to 90%, occurring more often between full-pod and full-seed

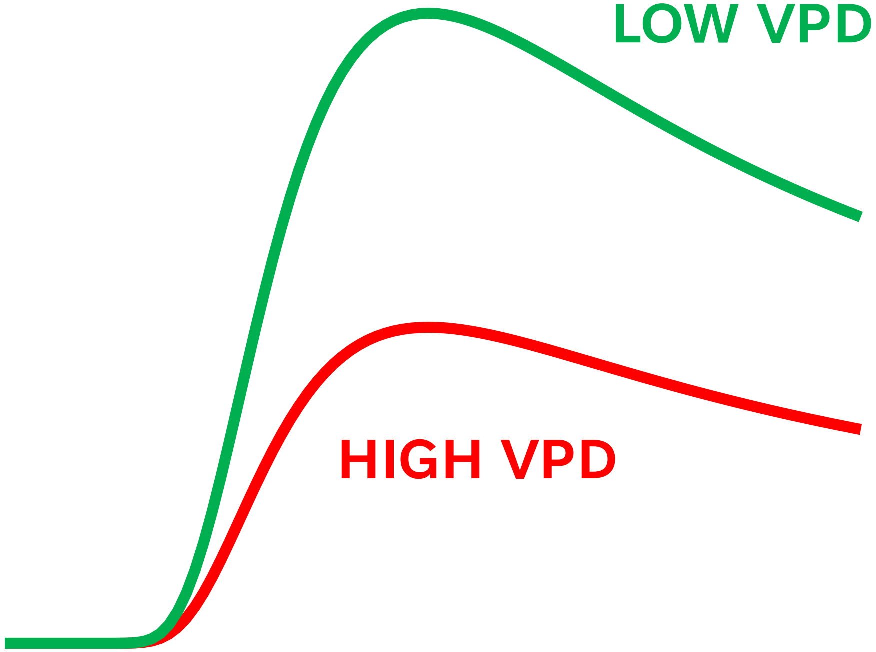
Under review - Almeida, Ciampitti et al. (2024)

Seasonal N Fixation Drivers

Ciampitti Lab

N_2 FIXATION PEAK

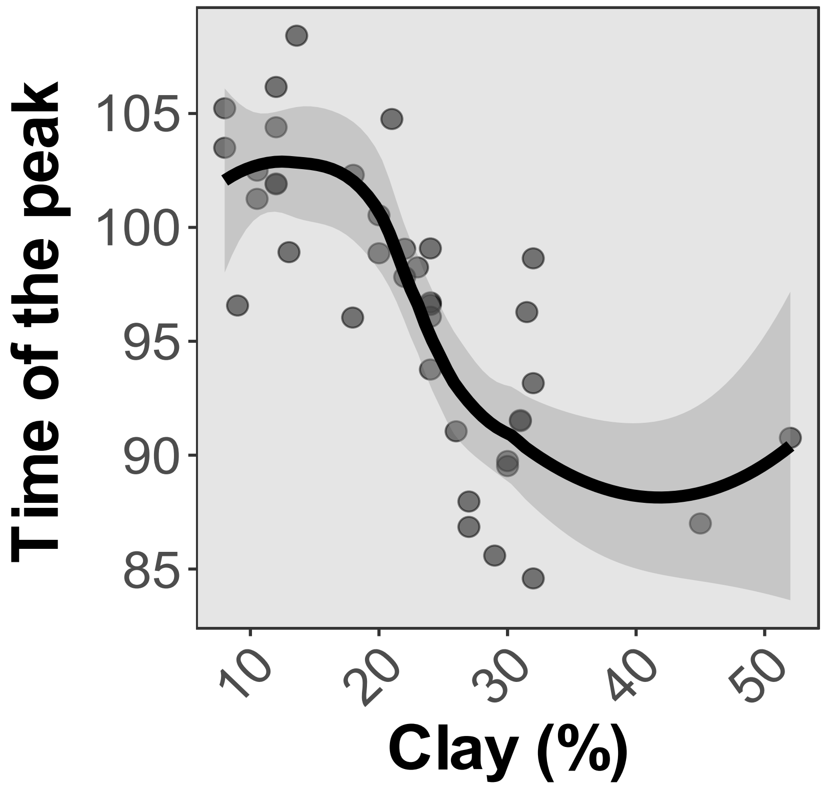


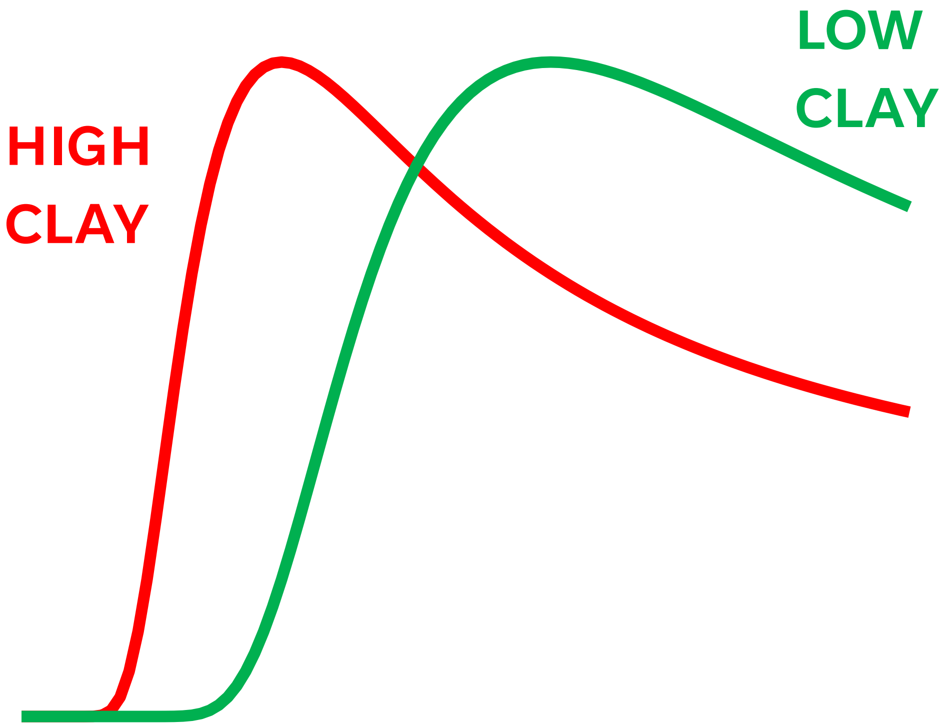


Under review - Almeida, Ciampitti et al. (2024)

Seasonal N Fixation Drivers

TIME OF THE PEAK





Under review - Almeida, Ciampitti et al. (2024)

Importance on quantifying N fixation, non-nod

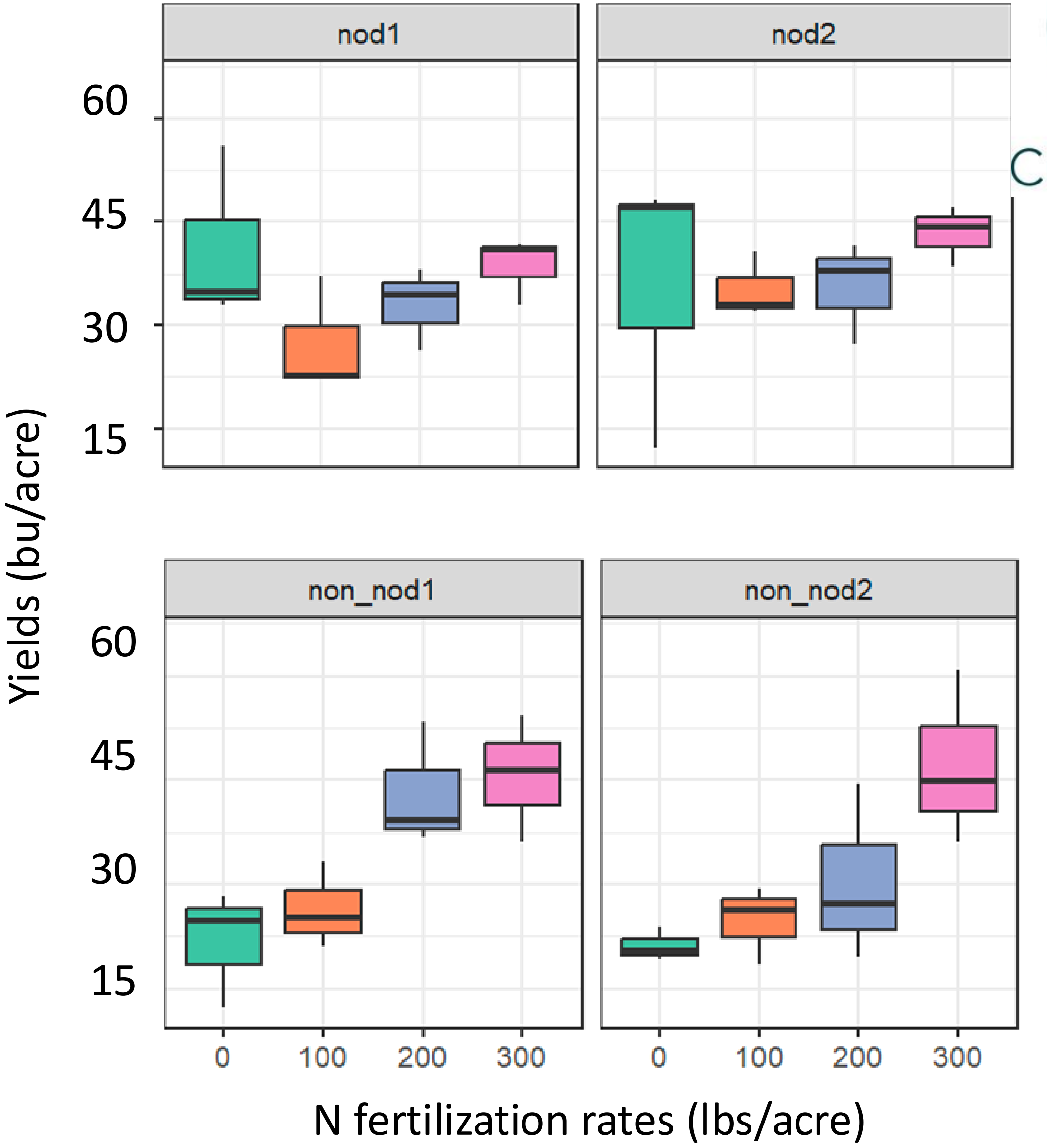
Ciampitti Lab

Responsiveness to N fixation (non-nod on the left, and nod on the right)

Heinz , Ciampitti et al. (2024)

Non-nod varieties

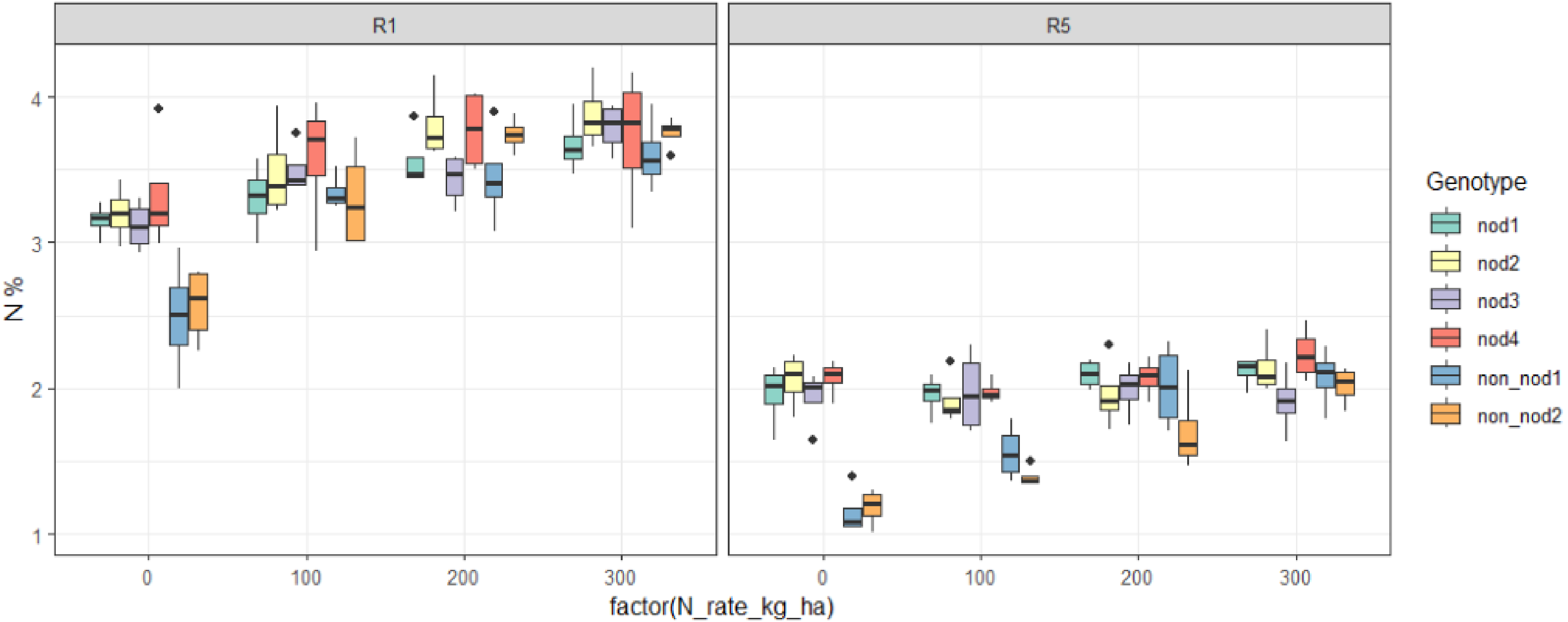
Non-nod, poorly nodulated with at least 20 bushels less, only compensated by applying 200 lbs N/acre.



Heinz , Ciampitti et al. (2024)

Non-nod varieties

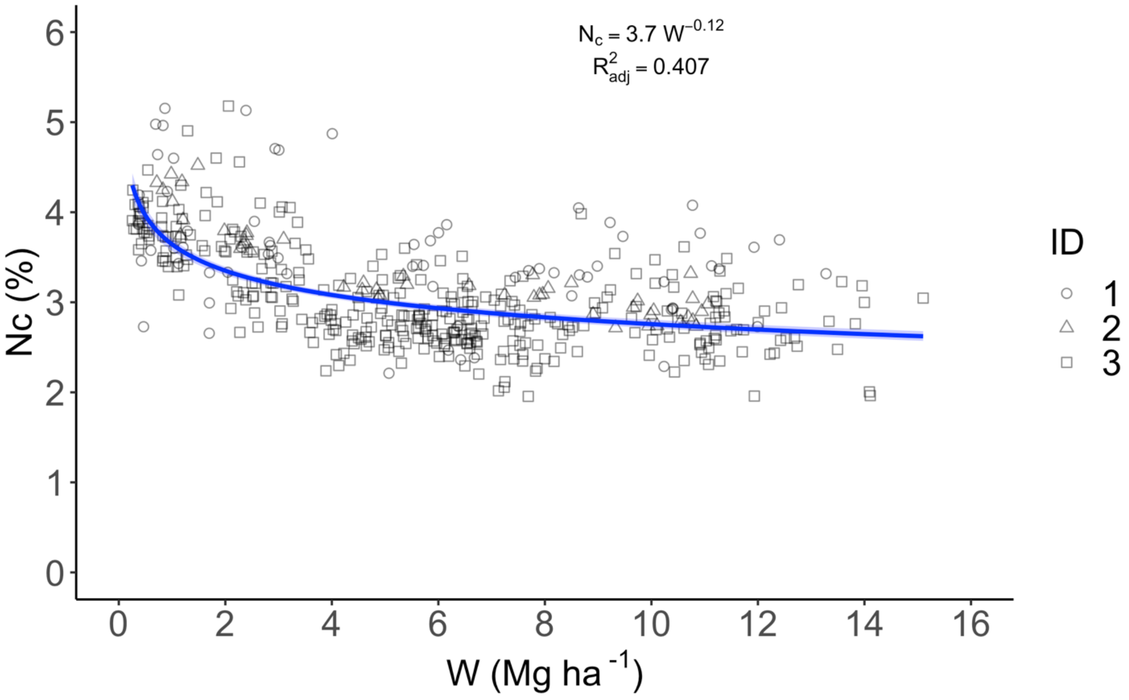
Heinz , Ciampitti et al. (2024)



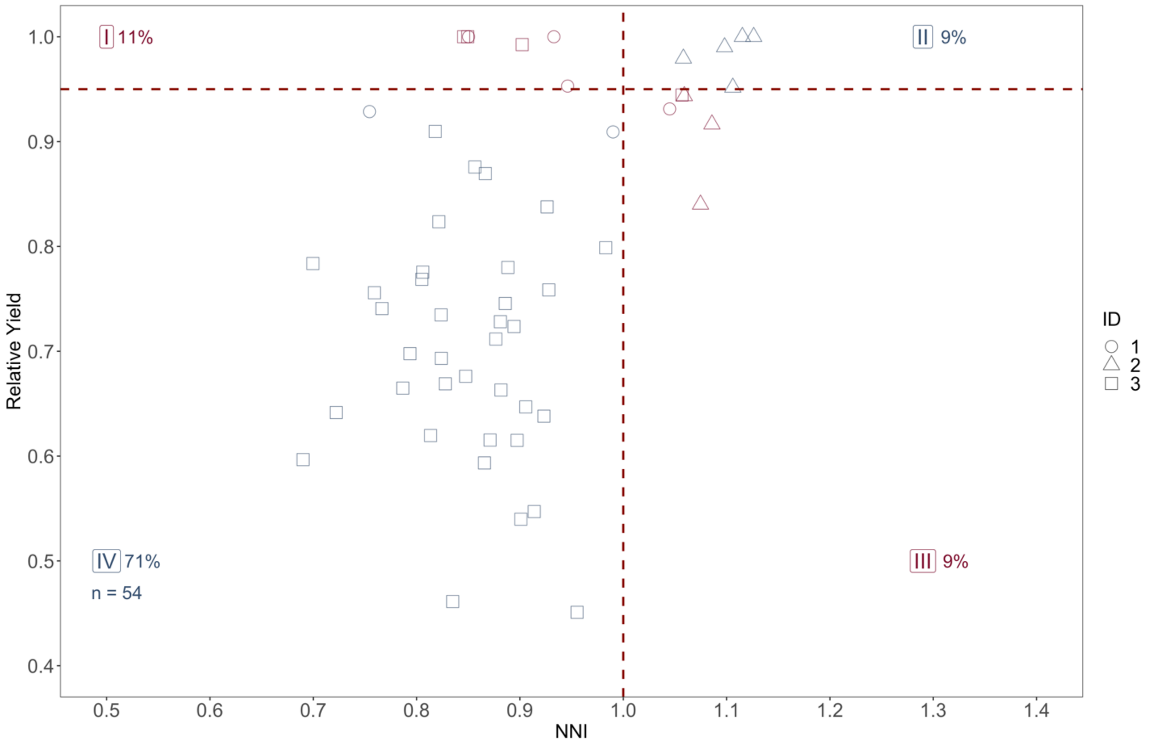
Non-nod starts showing less plant %N around flowering time (R1) and not even 100 lbs/acre helps to offset this N deficiency by seed filling stage (R5).

In-season screening for N deficiency?

Ciampitti Lab



Establishing a plant %N dilution curve, points above this blue line present "luxury N uptake" and below the line are "N deficient".

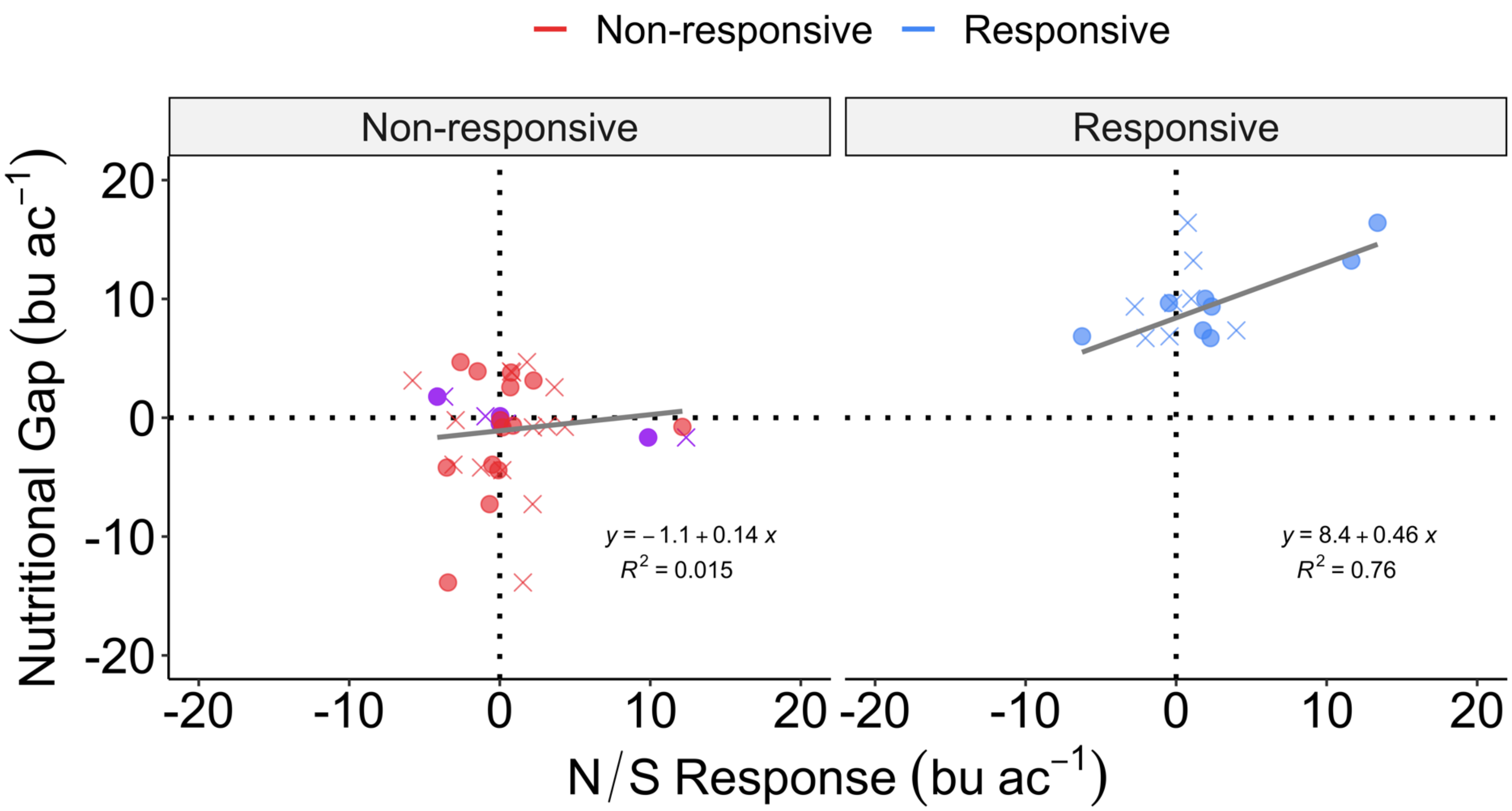


Around full flowering (R2 growth stage), values of Nitrogen Nutrition Index (NNI) below 1 refer to potential situations with yield responses.

Bosche, Ciampitti et al. 2024

In-season screening for N deficiency?
Is this a nitrogen problem?

Responses to fertilization are more correlated to S than to N.



●— Sulfur

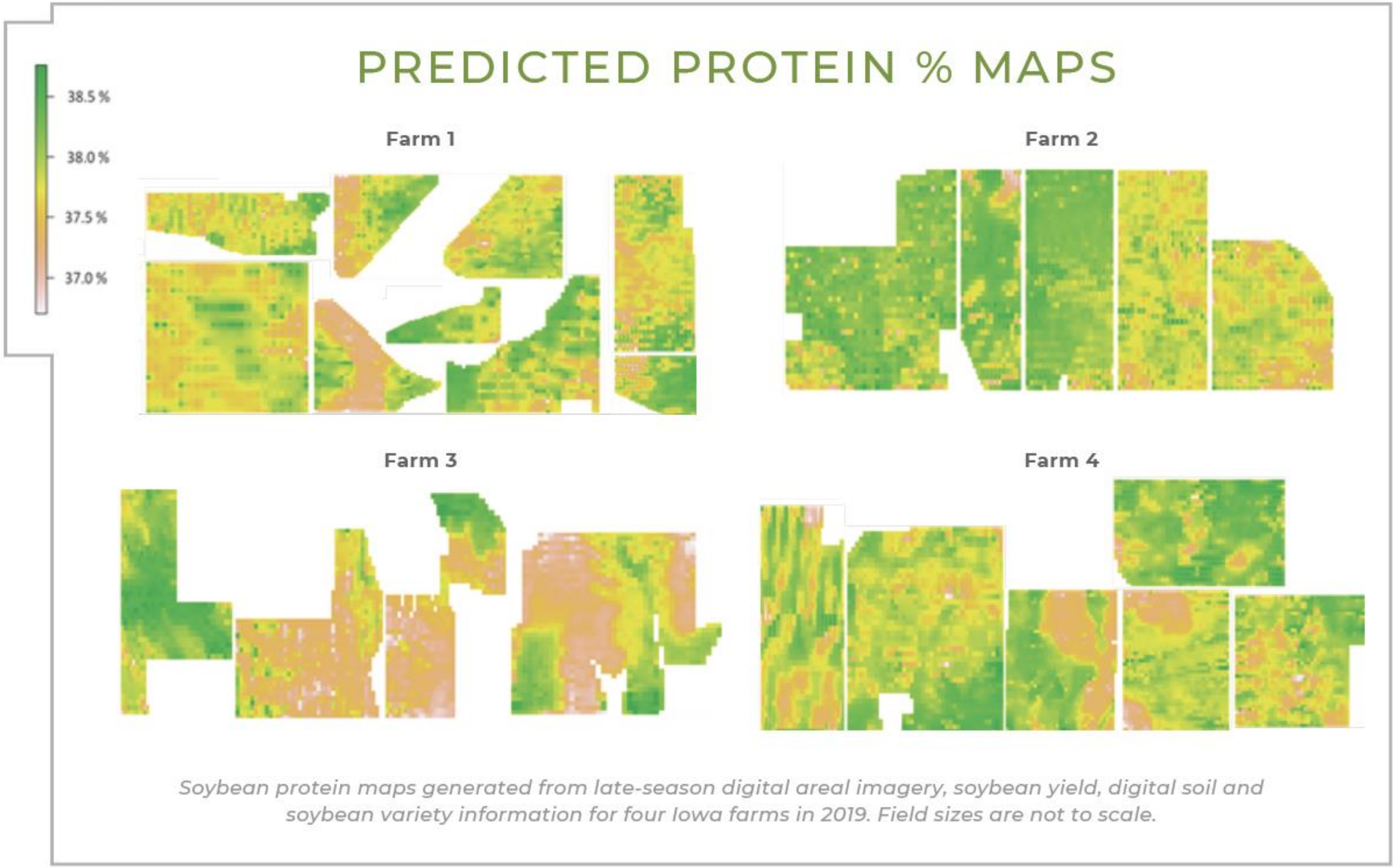
× Nitrogen

● KS Sites

Bosche, Ciampitti et al. 2024

Opportunity for segregation of soybean seed quality within a field

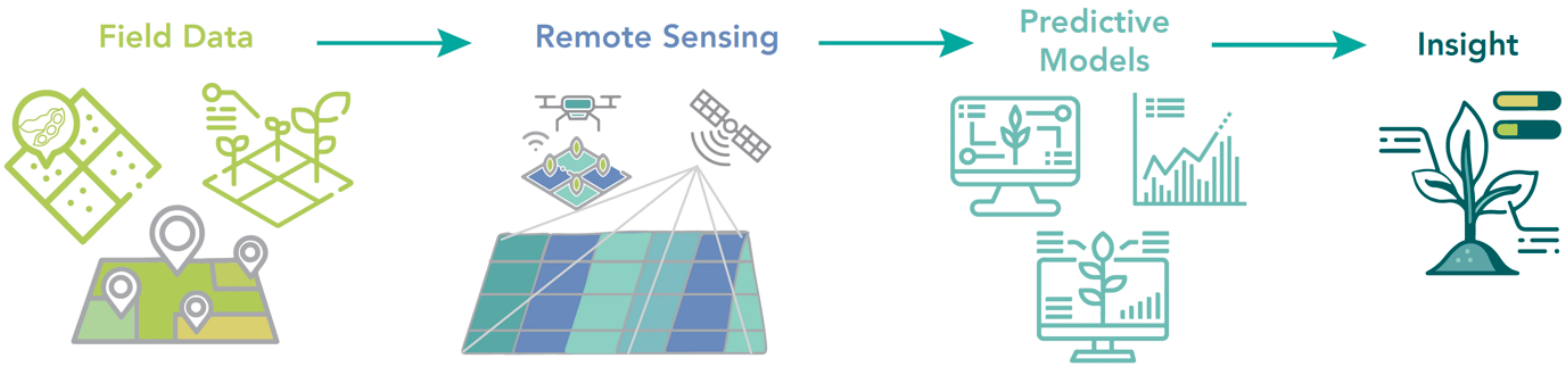
Changes within-the field of protein in farmer fields at harvest time, using combine protein sensor.



Soybean protein maps generated from late-season digital areal imagery, soybean yield, digital soil and soybean variety information for four Iowa farms in 2019. Field sizes are not to scale.

Objectives

- Determine the best crop season period to perform the estimation
- Develop a soybean seed protein and oil concentration predictive model
- Determine the best model and best satellite variables to perform the estimation



Framework of data processing and development of soybean mapping quality tool

Ciampitti Lab

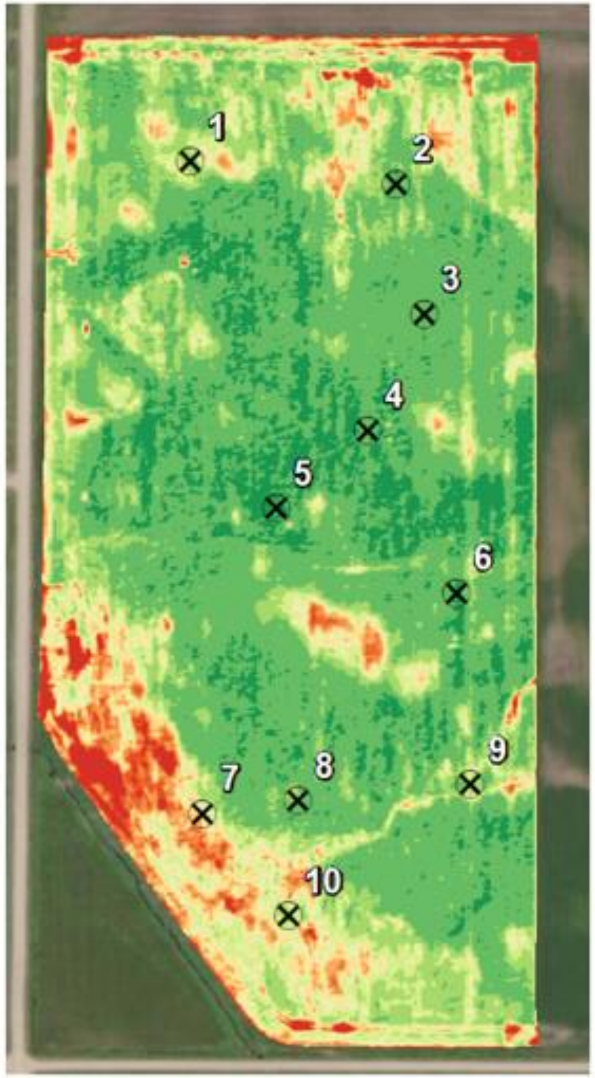
- Objectives
- Materials and methods
- Results
- Take home messages
- Next Steps

KANSAS STATE UNIVERSITY

Development of a multi-state database

Development of a farmer field-scale protocol for sampling soybean seed quality.

A protocol was established using available satellite data from past years and for defining zones within a field with different productivity to "direct" the sampling for seed quality.



Field sampling protocol based on satellite imagery of soybean canopy and soil type.

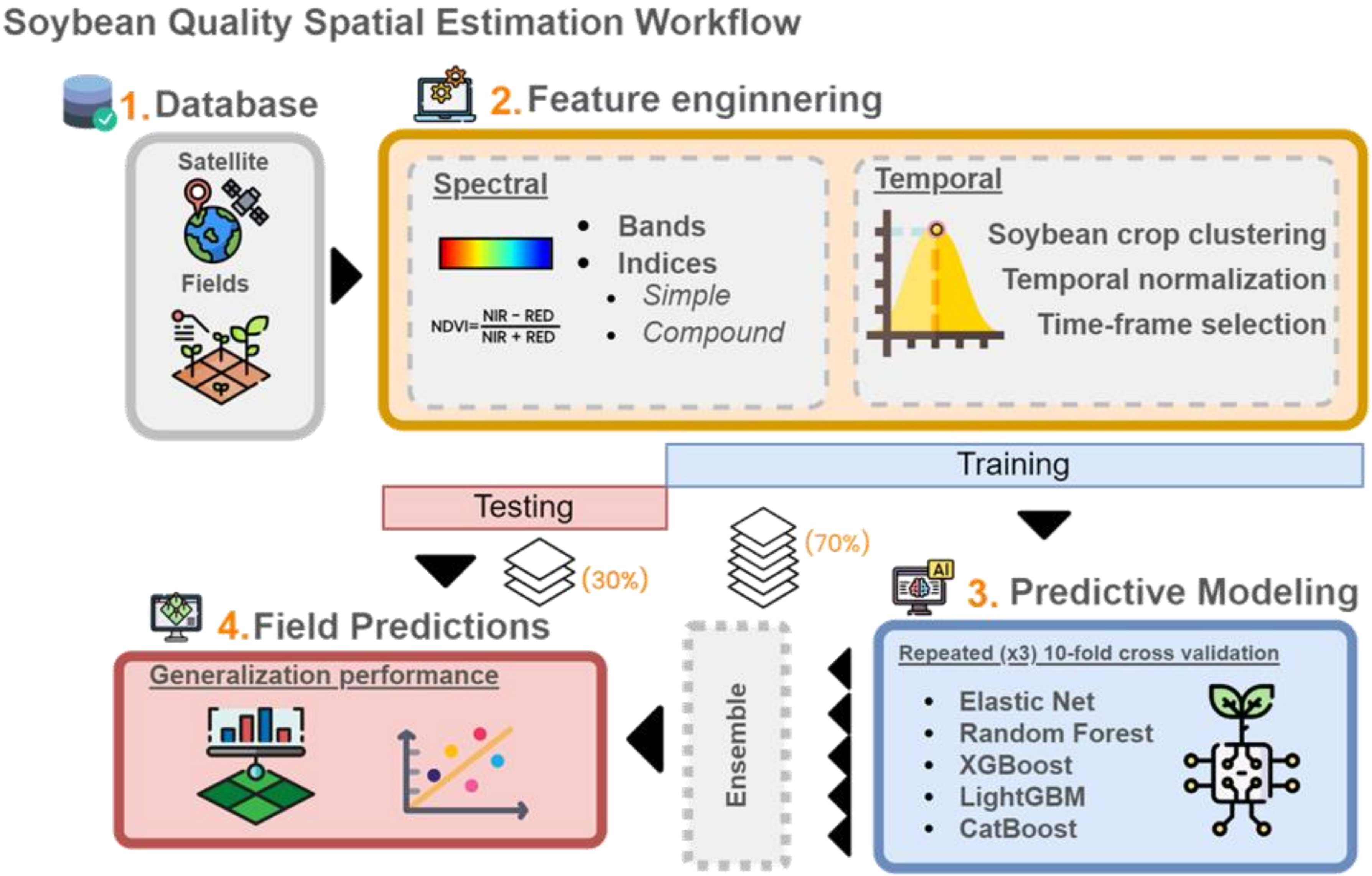
Hernandez and Ciampitti, 2024

Ciampitti Lab

- Objectives
- Materials and methods
- Results
- Take home messages
- Next Steps

KANSAS STATE UNIVERSITY

From digital tools to solving complex problems around the globe



Soybean Quality Spatial Estimation Workflow

1. Database

Satellite

Fields

2. Feature enginnering

Spectral

- Bands
- Indices
  - *Simple*
  - *Compound*

$NDVI = \frac{NIR - RED}{NIR + RED}$

Temporal

Soybean crop clustering

Temporal normalization

Time-frame selection

Training

Testing

(70%)

(30%)

3. Predictive Modeling

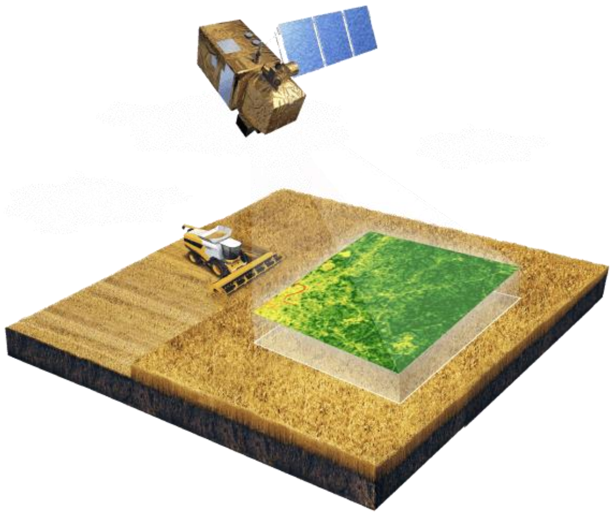
Repeated (x3) 10-fold cross validation

- Elastic Net
- Random Forest
- XGBoost
- LightGBM
- CatBoost

Ensemble

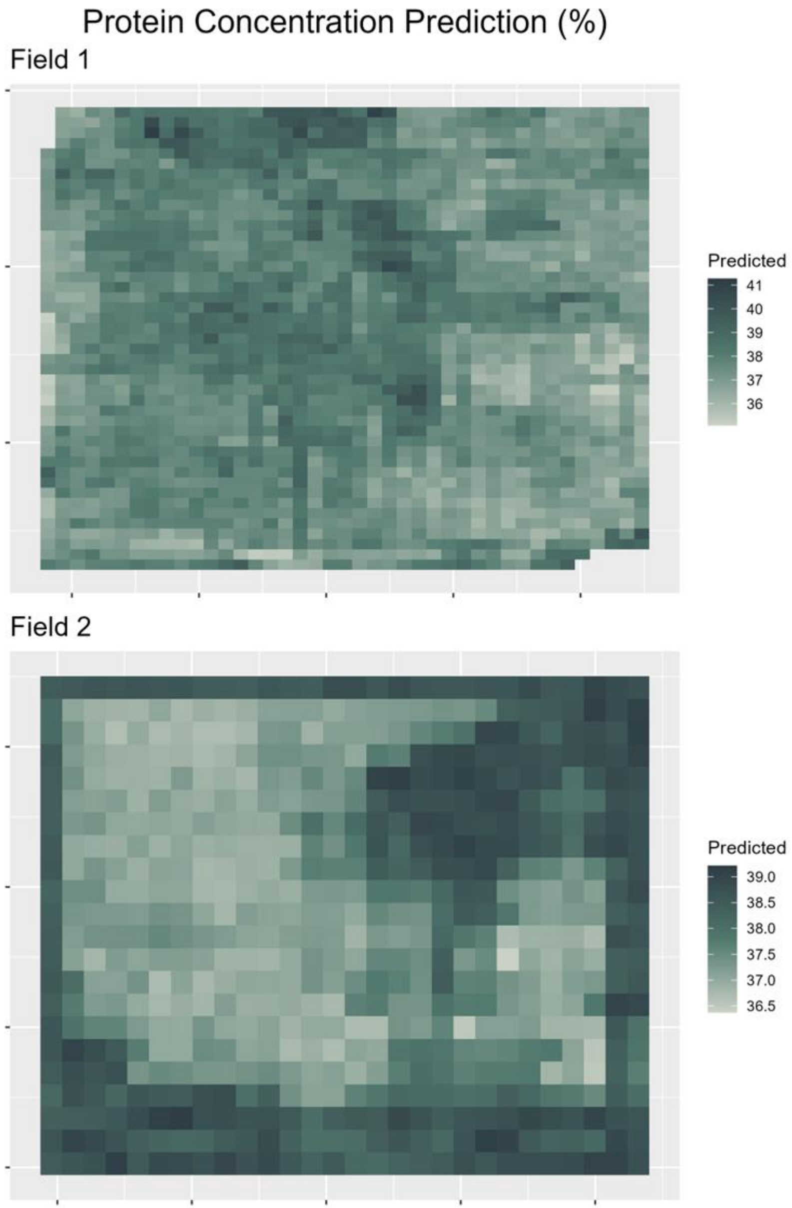
4. Field Predictions

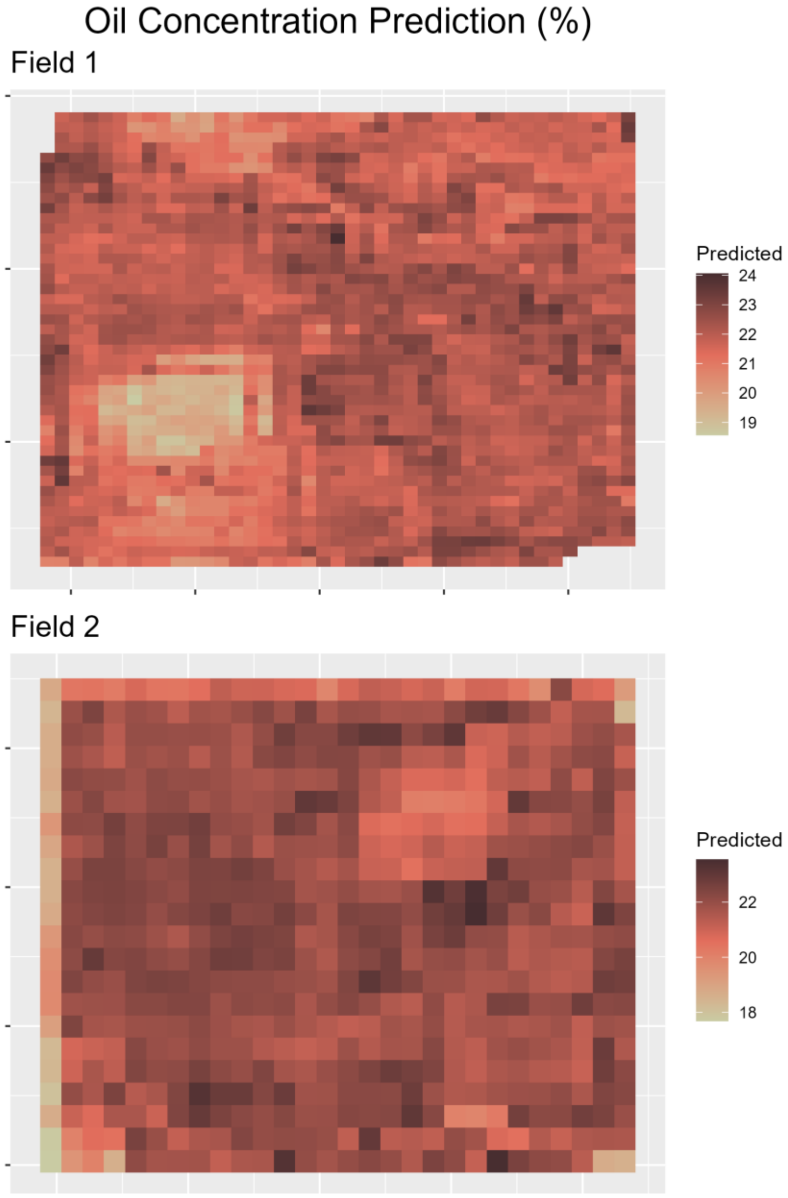
Generalization performance



Hernandez and Ciampitti, 2024

From digital tools to solving complex problems around the globe

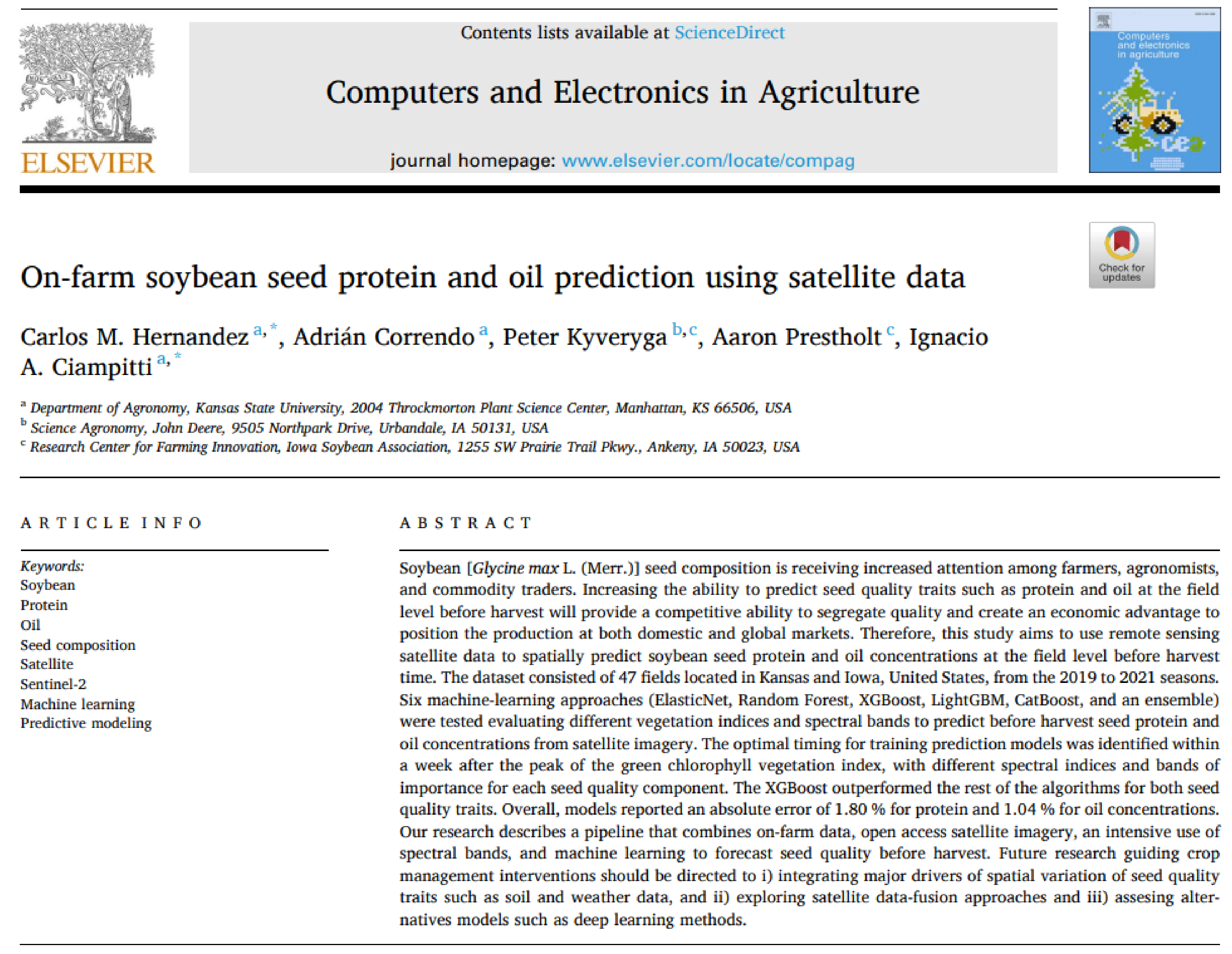




- The optimal timing for making predictions was identified around a week after the peak of green chlorophyll vegetation index.
- The XGBoost was identified as the best predictive model for both quality traits.
- Overall, models reported an absolute error of 1.7% for protein and 1.1% for oil concentrations.

Hernandez and Ciampitti, 2024

Check all the details of this project in our publication.



Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

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

On-farm soybean seed protein and oil prediction using satellite data

Carlos M. Hernandez [a,*], Adrián Correndo [a], Peter Kyveryga [b,c], Aaron Prestholt [c], Ignacio A. Ciampitti [a,*]

[a] *Department of Agronomy, Kansas State University, 2004 Throckmorton Plant Science Center, Manhattan, KS 66506, USA*
[b] *Science Agronomy, John Deere, 9505 Northpark Drive, Urbandale, IA 50131, USA*
[c] *Research Center for Farming Innovation, Iowa Soybean Association, 1255 SW Prairie Trail Pkwy., Ankeny, IA 50023, USA*

ARTICLE INFO

Keywords:
Soybean
Protein
Oil
Seed composition
Satellite
Sentinel-2
Machine learning
Predictive modeling

ABSTRACT

Soybean [*Glycine max* L. (Merr.)] seed composition is receiving increased attention among farmers, agronomists, and commodity traders. Increasing the ability to predict seed quality traits such as protein and oil at the field level before harvest will provide a competitive ability to segregate quality and create an economic advantage to position the production at both domestic and global markets. Therefore, this study aims to use remote sensing satellite data to spatially predict soybean seed protein and oil concentrations at the field level before harvest time. The dataset consisted of 47 fields located in Kansas and Iowa, United States, from the 2019 to 2021 seasons. Six machine-learning approaches (ElasticNet, Random Forest, XGBoost, LightGBM, CatBoost, and an ensemble) were tested evaluating different vegetation indices and spectral bands to predict before harvest seed protein and oil concentrations from satellite imagery. The optimal timing for training prediction models was identified within a week after the peak of the green chlorophyll vegetation index, with different spectral indices and bands of importance for each seed quality component. The XGBoost outperformed the rest of the algorithms for both seed quality traits. Overall, models reported an absolute error of 1.80 % for protein and 1.04 % for oil concentrations. Our research describes a pipeline that combines on-farm data, open access satellite imagery, an intensive use of spectral bands, and machine learning to forecast seed quality before harvest. Future research guiding crop management interventions should be directed to i) integrating major drivers of spatial variation of seed quality traits such as soil and weather data, and ii) exploring satellite data-fusion approaches and iii) assesing alternatives models such as deep learning methods.

Thanks for your time

Ignacio Ciampitti

Professor, Digital Ag Farming Systems

Director, Digital Ag Institute

Ciampitti Lab

https://ciampittilab.wixsite.com/ciampitti-lab