Study of plant and soil factors affecting seasonal nitrogen fixation, yield formation and seed composition in soybeans

by

Luiz Moro Rosso

B.S., Federal University of Santa Maria, 2018

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Agronomy
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:

Major Professor
Dr. Ignacio A. Ciampitti

Copyright

Abstract

Soybean [*Glycine max* (L.) Merr.] production currently faces several challenges linked to global food security (both quantity and quality) raised by an overgrowing human population, limited cropland, and diversified dietary in developed regions. To sustain seed yield and high protein levels, soybeans depend on large nitrogen (N) uptake, which is mostly attained by the symbiotic N fixation (SNF) process. Although SNF has been extensively investigated with single assessments during the season, just a few recent reports looked at the temporality of N sources (soil and SNF) while taking into consideration seasonal dry matter accumulation and soil nitrate (NO_3) and ammonium (NH_4) availability. Furthermore, it is still unclear how the overall changes in N uptake dynamics supports yield formation and seed components among canopy portions, especially considering the branches as potential contributors for high yield in modern genotypes.

Following this rationale, this project presents two overall objectives: 1) to identify the impact of soil NO_3 and NH_4 temporal availability on seasonal SNF [N derived from the atmosphere (Ndfa)], N uptake, and dry matter accumulation (herein called study 1); and 2) to characterize seed yield, protein, oil, amino acids (AA), and fatty acids (FA) from the main stem and branches (herein called study 2) for different commercial soybean varieties. To address the first objective, four genotypes were field grown at Manhattan (Kansas, US) during 2019 and 2020 growing seasons. Dry matter, N concentration, N uptake, Ndfa, fixed N, soil NO_3, and NH_4 (60-cm depth) were measured at six phenological stages, along with seed yield, protein, and oil concentration at harvest time. Seasonal exposure to NH_4 (area under the curve) showed a stronger suppression of Ndfa at the end of the season than NO_3. However, a mid-season NO_3 peak delayed uptake from soil and SNF, but only decreased maximum uptake rates from SNF. Additionally, dry matter was used as a seasonal linear predictor of fixed N to simplify the

process model. However, this relationship was deeply affected by soil N availability across environments (boundary functions) and also by a potential dry matter threshold around 5 Mg ha^{-1} across genotypes and site-years. For the second objective, another four genotypes were field-grown during the 2018 and 2019 growing seasons at Ashland Bottoms and Rossville (Kansas, US), respectively. At harvest time, seeds from the upper, middle, lower main stem, and branch nodes were manually separated and assessed for yield, seed size, protein, and oil (seed content and concentration), abundance of limiting AA within protein, and FA ratio (oleic / linoleic + linolenic). The accumulation of protein was more responsive to node position than oil, determining high protein concentration in the upper main stem and high oil concentration in the lower main stem nodes. However, the protein quality (limiting AA) was higher in the lower main stem, while the FA ratio (oil quality) was greater in the upper section of the plant. Branches presented the less-rich seed composition relative to the main stem, but their contribution to yield was positively associated with oil and limiting AA abundance across genotypes.

In summary, the main outcomes of the present thesis are related to 1) the importance of soil NO_3 and NH_4 to regulate Ndfa during the season, 2) the timing of Ndfa assessment or prediction for an accurate fixed N calculation throughout the season, and 3) the underlaying effect of branch yield allocation on the seed composition of the whole soybean plant, plausibly moderating changes across genotypes, environments, and management practices. A better understanding of soybean N acquisition and allocation for yield and quality formation within the plant is important to sustain the yield increase, offset protein decay, and assure cropping systems sustainability and food security in a long-term standpoint.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Dedication

The completion of this goal is dedicated to my parents. You have always been an example of courage, integrity, and transparency that shapes my character and decisions.

Chapter 1 - General Introduction

Soybean [*Glycine max* (L.) Merr.] global production doubled over the past 20 years, reaching ca. 334 Tg in 2019, after increasing harvested area and seed yield mainly in Brazil, USA, and Argentina (FAO, 2021). After the 1980s, the US yield increase was approximately 30 kg ha^{-1} yr^{-1}, which was mainly attained by the release of modern genotypes (Specht et al., 2015). High yields are accompanied by a large N requirement (80 kg Mg^{-1} of seeds), which is partially attended (50-60%) by symbiotic N fixation (SNF) and complemented by the indigenous soil N supply (Ciampitti and Salvagiotti, 2018). During the seed filling period, remobilization rates to the seed exceed the plant ability to uptake N and significant senescence of vegetative organs takes place (Sinclair and De Wit, 1975). This behavior indicates a tight relationship between the N sources and sink in order to sustain seed yield and protein levels. Therefore, modern genotypes under high-yielding conditions are responsive to greater N uptake, especially towards the end of the season (Gaspar et al., 2017). However, increasing total N uptake in soybeans has been proven difficult mainly for two reasons: 1) increasing the soil N supply via N fertilization inhibits SNF (Salvagiotti et al., 2009) showing minor effects on yield (Mourtzinis et al., 2018b); and 2) management practices to improve SNF itself are not consistently efficient in the US (Carciochi et al., 2019). Although a greater availability of soil N has similar effect as N fertilization, just a few reports have investigated the temporal relationship between soil nitrate (NO$_3$) and ammonium (NH$_4$) availability with the uptake from SNF and soil N supply. The natural variation in soil N availability could improve soybean N uptake, compensating shortages on fixed N when SNF is being established or shutting down towards the end of the growing season.

Remarkably, a single pre-sowing assessment of soil N has weaker relationship with SNF than N fertilization, highlighting the importance of a temporal characterization of N availability

(de Borja Reis et al., 2021). In addition, plant breeding has not altered the Ndfa levels of modern genotypes (Donahue et al., 2020), leaving the control of SNF contribution up to the environment and management. As a moderator of changes in environment and management, dry matter has been proposed as a linear predictor of fixed-N throughout the season (Córdova et al., 2019). However, different from the prediction of N uptake, which follows a relatively stable dilution process, fixed N is governed by the Ndfa levels, which are variable during the season and across site-years. Despite an expected weaknesses of using dry matter to predict fixed N, a simplification of the seasonal model of SNF can benefit future research investigating interactions with soil N dynamics. Furthermore, a deeper understanding on the SNF evolution could lead to improved N uptake to support high yield without protein decay, sustain positive or neutral N balances, and assuring the sustainability of soybean production systems.

Further from improving N uptake and SNF, it is crucial to unravel the N allocation within the plant canopy (node position), defining the concentration of different components and attaining modern industrial requirements. Despite the unintended consequences of plant breeding, such as a decay on seed protein concentration (de Borja Reis et al., 2020), modern genotypes are more competitive under high plant densities and also compensate the absence of plants with branching (Suhre et al., 2014). However, the branching process is usually neglected, especially considering the overall effect of this physiological change on soybean seed composition (Huber et al., 2016). Furthermore, besides the accumulation of protein and oil, the amino acids (AA) and fatty acids (FA) profiles are critical to determine monogastric weight gain from soybean rations (Mourtzinis et al., 2018a), and oil quality for specific industrial objectives (Gao et al., 2009). For instance, changes on yield allocation within the plant architecture caused by genotype, management, or environment, could moderate overall soybean seed composition.

The increasing soybean demand have been raising commodity prices and it is possible premium payments or preferential markets are established in the near future for farmers producing high quality seeds (Brumm and Hurburgh, 2006). Although significant research has been recently executed to uncover external factors of soybean seed composition, just a few reports have looked at variations within the plant canopy and how producers could take advantage of that. Finally, understanding the internal allocation of seed components could bring insights into soybean nitrogen (N) nutrition and the balance between N sources to attend total uptake.

The present study was design to develop two aspects of soybean production with emphasis on N uptake sources and seed component allocation within the plant considering main stem and branches. The main objectives of Chapter 2 were to 1) describe seasonal changes on soil-plant N dynamics among genotypes with contrasting genetic background, 2) asses the influence of soil NO_3 and NH_4 on the temporal SNF changes, and 3) quantify the variation on N uptake and fixed N as a function of seasonal dry matter accumulation. The main objectives of Chapter 3 were to 1) characterize seed yield and composition (protein, oil, AA, and FA) at three segments of the main stem (vertical canopy profile) and stem branches; and 2) quantify the impact of canopy yield allocation on seed composition, focusing on branches as potential contributor for high yields in modern genotypes.

References

Brumm, T.J., Hurburgh, C.R., 2006. Changes in long-term soybean compositional patterns. JAOCS, J. Am. Oil Chem. Soc. 83, 981–983. https://doi.org/10.1007/s11746-006-5056-4

Carciochi, W.D., Rosso, L.H.M., Secchi, M.A., Torres, A.R., Naeve, S., Casteel, S.N., Kovács, P., Davidson, D., Purcell, L.C., Archontoulis, S., Ciampitti, I.A., 2019. Soybean yield, biological N_2 fixation and seed composition responses to additional inoculation in the United States. Sci. Rep. 9, 19908. https://doi.org/10.1038/s41598-019-56465-0

Ciampitti, I.A., Salvagiotti, F., 2018. New insights into soybean biological nitrogen fixation. Agron. J. 110, 1185–1196. https://doi.org/10.2134/agronj2017.06.0348

Córdova, S.C., Castellano, M.J., Dietzel, R., Licht, M.A., Togliatti, K., Martinez-Feria, R., Archontoulis, S. V., 2019. Soybean nitrogen fixation dynamics in Iowa, USA. F. Crop. Res. 236, 165–176. https://doi.org/10.1016/j.fcr.2019.03.018

de Borja Reis, A.F., Moro Rosso, L., Purcell, L.C., Naeve, S., Casteel, S.N., Kovács, P., Archontoulis, S., Davidson, D., Ciampitti, I.A., 2021. Environmental Factors Associated With Nitrogen Fixation Prediction in Soybean. Front. Plant Sci. 12. https://doi.org/10.3389/fpls.2021.675410

de Borja Reis, A.F., Tamagno, S., Moro Rosso, L.H., Ortez, O.A., Naeve, S., Ciampitti, I.A., 2020. Historical trend on seed amino acid concentration does not follow protein changes in soybeans. Sci. Rep. 10, 17707. https://doi.org/10.1038/s41598-020-74734-1

Donahue, J.M., Bai, H., Almtarfi, H., Zakeri, H., Fritschi, F.B., 2020. The quantity of nitrogen derived from symbiotic N fixation but not the relative contribution of N fixation to total N uptake increased with breeding for greater soybean yields. F. Crop. Res. 259, 107945. https://doi.org/10.1016/j.fcr.2020.107945

FAO, 2021. FAOSTAT Online Database [WWW Document]. URL http://faostat.fao.org/

Gao, J., Hao, X., Thelen, K.D., Robertson, G.P., 2009. Agronomic management system and precipitation effects on soybean oil and fatty acid profiles. Crop Sci. 49, 1049–1057. https://doi.org/10.2135/cropsci2008.08.0497

Gaspar, A.P., Laboski, C.A.M., Naeve, S.L., Conley, S.P., 2017. Dry matter and nitrogen uptake, partitioning, and removal across a wide range of soybean seed yield levels. Crop Sci. 57, 2170–2182. https://doi.org/10.2135/cropsci2016.05.0322

Huber, S.C., Li, K., Nelson, R., Ulanov, A., DeMuro, C.M., Baxter, I., 2016. Canopy position has a profound effect on soybean seed composition. PeerJ 2016, e2452. https://doi.org/10.7717/peerj.2452

Mourtzinis, S., Borg, B.S., Naeve, S.L., Osthus, J., Conley, S.P., 2018a. Characterizing soybean meal value variation across the United States: A swine case study. Agron. J. 110, 2343–2349. https://doi.org/10.2134/agronj2017.11.0624

Mourtzinis, S., Kaur, G., Orlowski, J.M., Shapiro, C.A., Lee, C.D., Wortmann, C., Holshouser, D., Nafziger, E.D., Kandel, H., Niekamp, J., Ross, W.J., Lofton, J., Vonk, J., Roozeboom, K.L., Thelen, K.D., Lindsey, L.E., Staton, M., Naeve, S.L., Casteel, S.N., Wiebold, W.J., Conley, S.P., 2018b. Soybean response to nitrogen application across the United States: A synthesis-analysis. F. Crop. Res. 215, 74–82. https://doi.org/10.1016/j.fcr.2017.09.035

Salvagiotti, F., Specht, J.E., Cassman, K.G., Walters, D.T., Weiss, A., Dobermann, A., 2009. Growth and nitrogen fixation in high-yielding soybean: Impact of nitrogen fertilization. Agron. J. 101, 958–970. https://doi.org/10.2134/agronj2008.0173x

Sinclair, T.R., De Wit, C.T., 1975. Photosynthate and nitrogen requirements for seed production by various crops. Science (80-.). 189, 565–567. https://doi.org/10.1126/science.189.4202.565

Specht, J.E., Diers, B.W., Nelson, R.L., de Toledo, J.F.F., Torrion, J.A., Grassini, P., 2015. Soybean. pp. 311–355. https://doi.org/10.2135/cssaspecpub33.c12

Suhre, J.J., Weidenbenner, N.H., Rowntree, S.C., Wilson, E.W., Naeve, S.L., Conley, S.P., Casteel, S.N., Diers, B.W., Esker, P.D., Specht, J.E., Esker, P.D., 2014. Soybean yield partitioning changes revealed by genetic gain and seeding rate interactions. Agron. J. 106, 1631–1642. https://doi.org/10.2134/agronj14.0003

Chapter 2 - Temporal Variation of Soil N Availability Defines N Fixation in Soybeans

Abstract

Soybean [*Glycine max* (L.) Merr.] plays a critical role in global food security and agriculture sustainability, acquiring most of the nitrogen (N) required for growth and high-protein seeds from symbiotic N fixation (SNF). However, there is scarce information on how soil N supply interacts with the proportion of N derived from the atmosphere (Ndfa, %) and fixed N (kg ha^{-1}) throughout the growing season. This study aims to 1) describe seasonal changes on soil-plant N dynamics among genotypes with contrasting genetic background, 2) asses the influence of soil nitrate (NO_3) and ammonium (NH_4) on the temporal SNF changes, and 3) quantify the variation on N uptake and fixed N as a function of seasonal dry matter accumulation. Four genotypes were field-grown during two growing seasons. Bayesian non-linear models described seasonal change on dry matter, N uptake, Ndfa, soil NO_3, and NH_4 (0-60 cm depth). Seasonal exposure to NH_4 (given by the area under the curve, AUC) suppressed end-season Ndfa four-fold more than NO_3. A mid-season NO_3 peak was associated with a 20-day delay in maximum N fixation rates compared to an early-season NO_3 peak. In addition, the mid-season NO_3 peak did not increase uptake rates from the soil supply but suppressed maximum uptake from SNF (fixed N). Using dry matter as a predictor of fixed N simplifies the seasonal model to a slope-only equation (N requirement). However, in-season Ndfa variations controlled by soil N availability challenge the effectiveness of dry matter predicting fixed N across environments. Furthermore, a slope change around 5 Mg ha^{-1} (beginning of seed filling, R5 growth stage) points out the need for two Ndfa assessments during the season. Future research must explore high-yielding and

broader SNF conditions for understanding to what extent soil N can suppress SNF without decreasing seed yield in soybeans

Introduction

Soybean [*Glycine max* (L.) Merr.] production is essential for global nourishment and agriculture sustainability due to a high seed protein and oil concentration and low reliance on nitrogen (N) fertilization [less than 30% of the US cropped area receives about 20 kg N ha^{-1} (USDA, 2018)]. The symbiotic association with Bradyrhizobium spp. contributes to 50-60% of the soybean N uptake (Salvagiotti et al., 2008), with the remaining demand mainly attained by the soil N supply. However, a substantial yield increase over the last decades raised concerns on symbiotic N2 fixation (SNF) effectiveness to sustain yields while maintain protein levels (de Borja Reis et al., 2020). Unfortunately, the well-known antagonism of soybean N sources is a challenge to increase N uptake, particularly to provide an adequate level of N supply during the seed filling (Sinclair and De Wit, 1975). Increasing soil N supply via N fertilization downregulates the proportion of N derived from the atmosphere (Ndfa), jeopardizing potential benefits in field conditions (Mourtzinis et al., 2018). Recently explored strategies to enhance SNF [e.g., in-season re-inoculation (Carciochi et al., 2019), and co-inoculation with Azospirillum spp. (de Borja Reis, unpublished)] seem ineffective in producing yield benefits in the US. Although sparse in the current literature, a seasonal characterization of soybean N sources (Ciampitti et al., 2021) could shed light on the plant strategy of N acquisition and provide new insights for designing soybean systems with improved N uptake, yield, and seed composition.

While there is evidence that plant breeding did not affect soybean Ndfa levels (Donahue et al., 2020), soil and weather conditions (mediated or not by plant growth) seem to control SNF across environments (de Borja Reis et al., 2021). The higher yields in modern genotypes depend on greater remobilization and N uptake, especially during the seed filling, with the latter also increasing protein concentration (Gaspar et al., 2017). Thus, if genetics have not improved Ndfa, genotype choice could be less relevant for overcoming a negative partial N balance in soybean (Donahue et al., 2020). The soil naturalization of highly efficient Bradyrhizobium spp. strains have also limited potential benefits of inoculation under field conditions (Ambrosini et al., 2019; McLoughlin et al., 1990), leaving up to environmental factors the regulation of Ndfa. Schipanski et al. (2010) found soil texture explaining 20% of Ndfa variation across fields in the US. Drought and flooding are known to strongly suppress SNF during the vegetative or reproductive growth (Pasley et al., 2020; Purcell et al., 2004; Santachiara et al., 2019). Soil temperature, fallow precipitation, and air temperature during the seed filling also impact SNF contributions to N uptake (Collino et al., 2015; Lindemann and Ham, 1979).

Many recent reports have in common is a strong SNF suppression by N fertilization aiming to recreate a condition of high soil mineral N supply (Moro Rosso et al., 2021; Santachiara et al., 2018; Tamagno et al., 2018). Latimore et al. (1977) reported both NO_3 and NH_4 reduced nodule energetic supply, but NO_3 had a greater impact on nitrogenase activity. Also under greenhouse conditions, a better uptake mechanism of NO_3 than NH_4 could be responsible for stronger SNF inhibition by NO_3 in peas (*Pisum sativum*) and alfalfa (*Medicago sativa*) (Houwaard, 1980; Richardson et al., 1957). However, the metabolism of NO_3 is slower and energetically more expensive than NH_4 due to the need of reductive amination, unless NH_4 is supplied in luxurious amounts (Street and Sheat, 1958). Under field conditions, seed yield

response to starter N fertilizer is small (5% or 0.15 Mg ha^{-1}) or inconsistent across environments (Osborne and Riedell, 2006; Starling et al., 1998) and growth stages (Wesley et al., 1998; Wood et al., 1993). Only a temporal assessment and modeling (statistically or mechanistically) of NO$_3$ and NH$_4$ can identify potential interactions with SNF and N uptake. For instance, de Borja Reis et al. (2021) found no relationship between SNF and soil NO$_3$ or NH$_4$ at sowing when N fertilization still suppressed the relative abundance of ureides (RAU). It is possible the temporality of soil N availability has a strong control over SNF and fertilization responses.

Dry matter accumulation regulates N uptake due to the role N plays in plant growth (Novoa and Loomis, 1981). Investments in structural tissue (low N concentration), leaf expansion (metabolic tissue with high N concentration), and finally reproductive organs decrease the ratio between metabolic and structural compartments (Greenwood et al., 1990). Therefore, shoot N concentration follows a dilution process as the plant develops and accumulates dry matter (Lemaire and Gastal, 1997). Although after pod setting the N dilution curve is attenuated in soybeans (Divito et al., 2016), crop phenology itself is not a good predictor of N concentration changes. Following these lines, Ratjen et al. (2018) has shown that dry matter is a better predictor of plant N status than crop phenology, for capturing its allometric relationship with the leaf area index for both maize (*Zea mays*) and wheat (*Triticum aestivum*). Furthermore, Córdova et al. (2019) proposed dry matter throughout the growing season as the best predictor of fixed N, potentially moderating the effect of other factors on N fixation. A slope-only equation describing fixed N as a function of dry matter (N requirement) could provide quantitative data on the Ndfa variation at similar dry matter levels during the season.

This study aims to 1) describe seasonal changes on soil-plant N dynamics among genotypes with contrasting genetic background, 2) assess the influence of soil nitrate (NO$_3$) and

ammonium (NH_4) on the temporal SNF changes, and 3) quantify the variation on N uptake and fixed N as a function of seasonal dry matter accumulation.

Materials and methods

Experimental design and growing conditions

Field experiments were performed during the 2019 and 2020 growing seasons, located at 39.2° North, 96.6° West, 330 m elevation, in Manhattan, Kansas, US. Both experimental areas have been under a soybean-corn (Zea mays) rotation with conventional tillage performed prior to sowing. Both soils were from the Wymore series, with Fine, Aquertic Argiudolls taxonomic classification. Initial soil characterization (texture and chemical properties) was obtained 30 days before the onset of the experiment (Table 2.1). Climate was classified as Cfa (humid subtropical) presenting evenly distributed precipitation throughout the year (Köppen, 2011). Weather data was retrieved from a weather station located 500 m from the sites (Kansas Mesonet, 2017). Weather variables (Table 2.1) were summarized for different periods: 30 days prior emergence [VE stage (Fehr and Caviness, 1977)], from VE to full flowering (R2), from R2 to beginning seed filling (R5), and from R5 to physiological maturity (R7). The phenological stages were weekly recorded from 10 targeted plants in each experimental unit.

Sowing dates were June 4 in 2019 and June 8 in 2020, with crop emergence 5 days after sowing, on June 9 and June 13, respectively. Soybean genotypes composed the single treatment factor, with four levels: Williams 82 [maturity group (MG) 3.8, released in 1981], P34T43R2 (MG 3.4, 2014), P35T75X (MG 3.5, 2019), and P37T51PR (MG 3.7, 2019) (Corteva Agriscience, Johnston, IA, US). Genotype selection aimed contrasting genetic backgrounds, with variability on year of release (Williams 82) and seed composition (high-oleic trait in the

P37T51PR). However, all genotypes had similar MG and presented minimum differences on phenology throughout the season. Experimental design was a randomized complete block with four repetitions in both site-years, herein called 2019 and 2020. Experimental units were composed by four rows spaced 0.75 m apart with total area of 25 m^2. Seeding rate was 300,000 seeds ha^{-1} and ca. 240,000 plants ha^{-1} composed final stands. Seed inoculation was done before sowing with Vault HP Rhizobia Inoculant (BASF, Ludwigshafen, Germany), containing 3.0 × 10^9 colony forming unit mL^{-1} of Bradyrhizobium japonicum. Management of weeds, insects, and diseases followed best agronomic practices (Ciampitti et al., 2016).

Samplings and laboratory analysis

Six sampling stages were targeted: late vegetative (seventh leaf, V7); R2; pod setting (R4); R5; full seed (R6); and R7 (Fehr and Caviness, 1977). For 2019, samplings corresponded to 30, 46, 57, 68, 86, and 99 days after emergence (DAE); while for 2020 they corresponded to 24, 38, 52, 62, 79, and 97 DAE. The DAE were chosen to describe seasonal time due to 1) simple interpretation in non-linear model parameters; 2) to fit soil processes which are less dependent on soybean growth and development; and 3) similar temperature, sowing dates, and phenology between the two site-years. At each targeted stage, soil and plant samples were collected from all experimental units. Soil samples were collected from a 60-cm depth layer (six cores per plot) and stored in an icebox until further laboratory analysis. The NO$_3$ and NH$_4$ were extracted from 2 g of dry grounded soil (2 mm sieve) with a potassium chloride solution (KCl, 1 mol L^{-1}) and quantified (mg dm^{-3}) by colorimetric procedures in a flow analyzer (Brown, 1998). Shoot samples were collected from a 0.5 m^2 area, excluding border rows, and dried in an air-forced oven (65 °C) until constant weight for estimating aboveground dry matter (Mg ha^{-1}). A

sub-sample of five plants was ground in a micro-mill (0.1 mm particle size) and subjected to

chemical analysis. The Ndfa (%), a time-integrated measurement of the proportion of

atmospheric N within the plant tissue, was estimated using the natural abundance method

according to the following equation (Unkovich et al., 2008):

$$Ndfa(\%) = \frac{\delta^{15}No freferenceplant - \delta^{15}Nof soybeans}{\delta^{15}No freferenceplant - Bvalue} \times 100 \,, \tag{1}$$

in which δ^{15}N is the natural excess of the ^{15}N isotope in the plant tissue. Two options of

reference plants were sampled: 1) a non-nodulating genotype (Lee, MG 5.6) randomized within

each block; and 2) an unfertilized corn strip at the side of the experiment. Unfertilized corn and

non-nodulating Lee provided similar Ndfa values despite potential phenology differences (Sup.

Fig. S1). In addition, the non-nodulating soybean roots were checked to assure the absence of

nodules and used for final Ndfa calculations. The B-value of -2.54 was chosen as the reported

median of previous literature (Balboa and Ciampitti, 2020). Along with δ^{15}N, total N

concentration was estimated using an isotope ratio mass spectrometer, allowing for calculation of

N uptake (kg ha^{-1}), fixed N, and N uptake from the soil (as the remaining uptake). At harvest

time, two central adjacent rows covering 2 m^2 were manually collected from each plot and

machine threshed for obtaining seed samples and estimating seed yield (Mg ha^{-1}) and seed

weight (mg seed^{-1}, herein called seed size), both adjusted to 130 g kg^{-1} moisture basis.

Statistical analyses

Hierarchical Bayesian models were employed for all response variables. In this

framework, observations (y) were assumed to come from a normal distribution centered in the

true state (z) (data model), described by a process model, and including parameter models (distribution of process model parameters) (Wikle et al., 2019). Linear mixed models were used to describe the end-season data (seed yield and size). Site-year and genotype were included as fixed effect factors and generated eight treatment means for each parameter. The blocks were assigned with a random intercept, accounting for a portion of the variance as random effect factors. Linear model assumptions were checked with posterior predictive and residual plots. The same uninformative prior distribution was set to all model parameters (treatment means), according to the response variable. Posterior distributions were compared based on high-density continuous intervals (95% probability) using pairwise hypothesis testing. Finally, a Tukey adjust compared genotypes within site-years and site-years within genotypes, simplifying the presentation and interpretation of posterior distributions contrast analysis.

Non-linear equations were adopted as process models for all in-season response variables, using DAE to account for time (x, independent variable) and assigning a random intercept for blocks. Besides primary model parameters, secondary parameters (derived quantities) were calculated based on all posterior draws for each site-year and genotype interaction. Comparison of non-linear parameter posterior distributions was done in similar fashion as for linear mixed models aforementioned. For dry matter and N uptake, the following logistic equation described the temporal process (process model):

$$z = \frac{m}{1+e^{-k(x-g)}},$$
(2)

in which *m* is the asymptote or maximum predicted value; *k* controls the growth rate; and *g* controls the timing of maximum growth rate during the season. Two secondary parameters were calculated: maximum growth rate (*r*); and predicted value at R7 (*f*). Uninformative truncated

normal distributions were set as priors to all primary parameters, restricting posterior samples to positive values. The Ndfa was assigned with a gaussian growth equation, as follows:

$$z = me^{\frac{-(x-t)^2}{xc^2}},$$ (3)

in which m also represents the maximum predicted value (peak), t is the timing of the peak, and c controls the growth and decay rates. Four secondary parameters were calculated: timing of maximum growth (g), maximum growth rate (r), predicted value at R7 (f), and area under the curve (AUC). Truncated normal distributions were used as priors and shown in Sup. Fig. S2. Soil NO_3 and NH_4 were modeled with the gaussian peak function, intending to capture an in-season increase from the baseline condition, described by the following equation:

$$z = b + me^{\frac{-(x-t)^2}{w^2}},$$ (4)

in which b is the baseline concentration, and w controls the growth and decay rates (width). Only the AUC was calculated as secondary parameter, representing the overall crop exposure to soil N supply during the growing season. All prior distributions were uninformative and bounded by zero (truncated normal), assuming a potential peak during the season.

      The shoot N concentration (total and fixed) was described as a function of dry matter using the dilution curve (Greenwood et al., 1986), an exponential equation from Lemaire et al. (2008). In addition, N uptake (total and from SNF) was predicted from dry matter using the slope-only function from Córdova et al. (2019), simplifying the seasonal process model. The slope-only model was also employed over the 10th and 90th percentiles to explore potential

variation on N uptake and fixed N across dry matter levels. The slope estimates are expressed as the amount of N per unit of dry matter, herein called N requirement (kg Mg^{-1}). The slopes ratio (fixed N over N uptake) was considered as an estimation of overall Ndfa throughout the season and compared to 1) observed Ndfa at R7, 2) ratio of fixed N and N uptake predicted at R7 by the logistic models, and 3) end-season gaussian growth model prediction.

All linear and non-linear models were evaluated in terms of R^2, and root mean squared error (RMSE). Markov chain Monte Carlo (MCMC) was used to obtain posterior draws using the brms package (Bürkner, 2018, 2017) in R software (R Core Team, 2021). For all model parameters, two-thousand posterior samples were generated before and six-thousand after algorithm warm-up, discarding four every five samples (thinning) to improve chain mixing (four total) and reduce correlation of consecutive samples (Hooten and Hefley, 2019).

Results

Seed yield, dry matter, and N uptake

With a relatively late sowing in both site-years, seed yields averaged 3.2 Mg ha^{-1} and did not differ among any treatment levels (Table 2.2). Seed size did not differ among genotypes within site-years, but it was overall 10% greater in 2020 (140 mg seed^{-1}) than 2019 (125 mg seed^{-1}). Dry matter accumulation was described ($R^2 = 0.90$) by the logistic model (Fig. 2.1A-B), with no difference in the main primary parameters [asymptote (*m*), and timing of maximum growth (*g*)] and maximum growth rate (*r*) coming from either genotype or site-year (Supp. Table S1). However, the level of dry matter predicted at R7 (*f*) was ca. 25% greater in 2020 (8.9 Mg ha^{-1}) than 2019 (7 Mg ha^{-1}) for all genotypes. Total N uptake at R7 was overall 15% greater in 2020 than 2019 (Fig. 2.1C-D), but only significant for P37T51PR (with 215 and 165 kg ha^{-1},

respectively). Although genotypes did not differ within site-years, Williams 82 had smaller N

uptake from the soil supply at R7 in 2020 than 2019, with 65 and 95 kg ha^{-1}, respectively (Fig.

2.1E-F). Temporal changes on fixed N were remarkably different between site-years, with 2020

reaching R7 values of ca. 120 kg ha^{-1} and 2019 about 90 kg ha^{-1} (35% difference), not significant

for the P35T75X genotype (Fig. 2.1G-H). For 2019, genotypes also differed on R7 fixed N, with

P35T75X presenting the greatest (110 kg ha^{-1}) and P37T51PR the smallest value (70 kg ha^{-1}). All

parameter medians followed by the Tukey results for site-years and genotypes total N update,

uptake from soil and SNF are shown in Supp. Table S2.

Ndfa and soil mineral N

Besides modeling N uptake and fixed N, Ndfa seasonal changes were described (R^2 of

0.76) with the gaussian growth function (Fig. 2.2A-D). Prediction credible intervals shows large

variability on the posterior samples, making most of the primary and secondary parameters

statistically similar among genotypes within site-years (RMSE of ca. 25%). The AUC was

almost two-fold in 2020 compared to 2019, with the former presenting a peak within the season

(ca. 85 DAE) and the latter peaking after R7. The Ndfa maximum rate (r) was ca. 65% higher in

2020 (1.65 kg ha^{-1} day^{-1}) than 2019 (1 kg ha^{-1} day^{-1}) (not significant for the P35T75X genotype).

Remarkably, Ndfa values at R7 were similar between site-years (ca. 55%), except for Williams

82, which reached ca. 20% less Ndfa in 2019 (44%) than 2020 (64%). Overall, Ndfa at R7 did

not differ across genotype and site-year combinations (Supp. Table S3).

Both NO_3 and NH_4 presented contrasting site-year seasonal availability, but genotypes

only differed on the NO_3 peak (m) in 2019, with Williams 82 reaching ca. 25 mg dm^{-3} and

P34T43R2 ca. 15 mg dm^{-3} (Fig. 2.2E-H). The site-year 2019 presented a NO_3 peak around 50

DAE (R3 stage), while the NH_4 peak was estimated after R7 (Fig. 2.2J-M). On the other hand, 2020 showed a NO_3 peak around 7 DAE, with NH_4 peaking approximately 30 DAE (R1 stage). In terms of NH_4, site-years differed on the baseline values, with 2019 showing ca. 6 mg dm^{-3} and 2020 ca. 4.3 mg dm^{-3}. Although 2020 reached an NH_4 peak during the season, the AUC was larger for 2019. The AUC, as an indicator of soybean exposure to mineral N, had a negative relationship with Ndfa at R7 (f) for both NO_3 (Fig. 2.2I) and NH_4 (Fig. 2.2N). In 87% of the posterior draws, slope estimates were negative for NO_3 AUC and Ndfa, with a 3.1% Ndfa decay every 100 mg dm^{-3} day. The relationship was stronger for NH_4, showing a 11.7% decay every 100 mg dm^{-3} day (negative at 98% of the times). All primary and secondary parameter medians followed by the CLD (Tukey test at $p < 0.05$) are presented in Supp. Table S4.

Seasonality of uptake rates

Maximum N uptake from the soil supply (Fig. 2.3A-D) occurred before the peak of N fixation (Fig. 2.3F-I), almost 30 days between peaks in 2019 (mid-season NO_3 peak, R3 stage) and about 20 days in 2020 (early season NO_3 peak, 7 DAE). In 2019, maximum soil and SNF uptake rates were delayed, occurring ca. 50 (R3 stage) and 80 DAE (R6 stage), respectively. For 2020, maximum soil and SNF uptake rates happened around 40 (V7 stage) and 60 DAE (R4 stage). Besides temporal differences, maximum uptake rates of fixed N were almost three-fold in 2020 (7.5 kg ha^{-1} day^{-1}) compared to 2019 (2.5 kg ha^{-1} day^{-1}), while uptake rates from the soil were similar across genotypes and site-years (ca. 3 kg ha^{-1} day^{-1}). For 2020, uptake rates of fixed N were greater for P34T43R2 and smaller for the P37T51PR genotype, 8.2 and 3.6 kg ha^{-1} day^{-1}, respectively. Although soil NO_3 peak and timing of maximum uptake from soil coincided in 2019 (ca. 50 DAE at the R3 stage), maximum rates were not significantly greater than 2020 (Fig.

2.3E), with positive slope estimates only 74% of the times (median of 0.013 kg ha^{-1} day^{-1}). On the other hand, when the NO_3 peak was delayed a day, uptake rates from fixation decreased 0.08 kg ha^{-1} day^{-1} across genotypes and site-year combinations (negative slope at 100% of the times) (Fig. 2.3J). Inference on the temporal effect of NH_4 peak was not explored due to a relatively flat availability in 2019 (peak likely happened after R7) (Fig. 2.2J-M).

Dry matter as predictor for N uptake

Instead of using time (DAE) as a seasonal descriptor of N uptake and fixed N, dry matter accumulation simplifies the process model to a slope-only equation. In addition to modeling the true state (z), boundary functions over percentiles are useful to define potential scenarios of maximum and minimum variation. Across genotypes and site-years, N uptake had a potential variation of ca. 60% while fixed N varied almost 400% throughout the season (Fig. 2.4A). For instance, at 3 Mg ha^{-1} dry matter (roughly the R2 stage) fixed N could vary 40 kg ha^{-1}, while at 6 Mg ha^{-1} (R5 stage), it could vary 90 kg ha^{-1}. Because N dilution curves are similar across genotypes (Supp. Fig. S3), plant N concentration produced relatively low variability in total N uptake. On the other hand, Ndfa introduced large variation into the fixed N, because of its sensitivity to soil N availability (Figs. 2.2-2.3). Based on the boundaries, seasonal Ndfa could range from 15 to 90% across dry matter levels. Furthermore, fixed N observations below 5 Mg ha^{-1} (R4-R5) produce a smaller slope than observations up to maturity (R7) or after 5 Mg ha^{-1}, possibly related to the SNF establishment (Fig. 2.4B). This breaking point also challenges dry matter as a linear predictor of fixed N, indicating a need for Ndfa assessments.

The slope-only model was able to predict N uptake ($R^2 = 0.95$) and fixed N ($R^2 = 0.84$), even though the latter displayed wider credible intervals (Fig. 2.4C-D). While a linear regression

might oversimplify plant N dynamics early in the season (high N concentration and neglectable SNF), the slopes ratio (fixed N over N uptake) produces an overall Ndfa estimate. Remarkably, the slopes ratio was comparable (95% credible intervals) to Ndfa observations and predictions at the end of the season (Fig. 2.4E), likely due to a smaller impact of low dry matter observations (below 5 Mg ha^{-1}). Finally, slope estimates can be translated into N requirement (kg N Mg^{-1} of dry matter). The site-year 2020 had an overall lower N requirement to produce dry matter (26 kg Mg^{-1}) but presented more fixed N per unit of dry matter (15 kg Mg^{-1}). For 2019, the genotype P37T51PR had the smallest requirement of total N (26 kg Mg^{-1}) and fixed N (9 kg Mg^{-1}), while the other genotypes required ca. 29 kg Mg^{-1} and 13 kg Mg^{-1}, respectively.

Discussion

The present study expands previous findings on seasonal SNF characterization by modeling NO_3 and NH_4 availability, which had significant impact on the N uptake sources. In addition, we exposed potential limitations of dry matter as a linear predictor of fixed N throughout the season, proposing a Bayesian modeling framework for future studies.

Our results opposed those of a previous report on soybean under greenhouse conditions, which shows greater SNF suppression from NO_3 than NH_4 (Latimore et al., 1977). Here, the seasonal exposure to NH_4 (AUC) suppressed Ndfa four-fold more than NO_3, possibly explained by a lower energetic cost of NH_4 assimilation under continuous supply compared to NO_3 (Streeter, 1972). Remarkably, field studies controlling NO_3 and NH_4 supply as a treatment factor are scarce due to inherent N dynamics in the soil. Our results with contrasting NO_3 and NH_4 availability between site-years allowed novel inference about their relationship with soybeans SNF. Besides a mid-season NO_3 peak delaying N uptake from SNF, the temporal match of N

uptake from soil supply and high NO_3 availability did not increase uptake rates from the soil but strongly suppressed the ones from SNF. Although Cafaro La Menza et al. (2020) reported temporal dynamics for N fixation and soil N supply, it was unclear how temporal soil N availability could interact with uptake timings and rates. Furthermore, inconsistent yield responses to N fertilization (Mourtzinis et al., 2018) could be connected to seasonal N availability, but also explained by the compensation of these N sources over time.

Despite the complex relationship between soil water dynamics and NO_3 availability, excessive precipitation could have delayed the mineralization process in 2019 (Linn and Doran, 1984). However, those differences in early-season precipitation might have also shifted the NO_3 peak from early to mid-season and impaired SNF. Interestingly, an NH_4 peak was not observed before the NO_3, indicating a quick immobilization into the soil microbiota (released after crop residue decomposition) or a prompt nitrification process. Furthermore, it is important to acknowledge the indissociable character of N uptake and growth (Briat et al., 2020), and how environmental factors such as precipitation could interact with either of these factors.

Simplifications on modeling SNF seasonal changes are desirable and a slope-only equation can be useful in conditions of stable Ndfa (Córdova et al., 2019). However, over simplifying SNF prediction overlooks that fixed N variation increases with plant biomass, substantiating the end-season observations from Ciampitti and Salvagiotti (2018) (Supp. Fig. S4). In addition, the linear relationship had two different phases, before and after the 5 Mg ha^{-1} threshold (early seed filling), possibly related with an attenuation of fixed N concentration within the plant, as observed for total N after pod setting (Divito et al., 2016). It is worth noting that N uptake prediction with mechanistic models (Archontoulis et al., 2020) is reasonably straightforward compared to fixed N because N uptake is not as much affected by SNF

(Santachiara et al., 2018) as Ndfa by environmental factors. Therefore, the approach from Fig. 2.4 could couple N uptake simulations with Ndfa assessments to obtain accurate fixed N estimations across environments, relaying on fewer (potentially two) Ndfa assessments. Finally, the N uptake and fixed N slopes ratio (overall Ndfa) matches the end-season observations and predictions by other non-linear models, indicating a timing for single Ndfa assessment.

We acknowledge the small number of genotypes as a potential limitation to characterize contrasting yield levels and variations in soybean seed composition. In addition, most of the environmental differences (soil mineral N) originated from site-years and not from a controlled treatment. Due to the weather conditions, both experiments reached similar and intermediate yield levels, leaving a knowledge gap under higher N uptake (and yield) scenarios. While accounting for the limitations of the present study, future research should focus on 1) describing in more details (statistically or mechanistically) processes involved with soil N availability (e.g., mineralization, denitrification, immobilization) as the timing of NO_3 peak could eventually impair SNF and reduce N uptake; 2) identifying environmental factors controlling Ndfa that are not directly translated into dry matter changes, and; 3) validating the proposed dry matter threshold for predicting fixed N throughout the season with a small number of Ndfa assessments (or predictions), highly relevant for rapid phenotyping. Although challenging, SNF enhancement must be a priority for the overall sustainability of soybean production systems, from both standpoints of food security (yield) and biofortification (seed quality).

Conclusions

The availability of soil N seems to be the primary driver of seasonal Ndfa changes and temporal compensation between N sources (soil supply and SNF). Furthermore, soybeans

exposure to NH_4 and NO_3 significantly suppress Ndfa throughout the growing season. Remarkably, a late NO_3 peak (mid-season) delayed soil N uptake and SNF, but only reduced the latter. Although dry matter accumulation simplified the process to model N uptake and fixed N (slope-only equation), Ndfa variations across dry matter levels represent a challenge for the seasonal prediction of fixed N across environments. Future studies under a wide seed yield gradient are necessary to undercover the actual compensation mechanisms of N sources in satisfying the plant demand under contrasting soil N availability and SNF status, quantifying potential impacts on yield and nutritional quality.

Acknowledgements

References

Ambrosini, V.G., Fontoura, S.M. V., de Moraes, R.P., Tamagno, S., Ciampitti, I.A., Bayer, C., 2019. Soybean yield response to *Bradyrhizobium* strains in fields with inoculation history in Southern Brazil. J. Plant Nutr. 42, 1941–1951. https://doi.org/10.1080/01904167.2019.1648680

Archontoulis, S. V., Castellano, M.J., Licht, M.A., Nichols, V., Baum, M., Huber, I., Martinez-Feria, R., Puntel, L., Ordóñez, R.A., Iqbal, J., Wright, E.E., Dietzel, R.N., Helmers, M., Vanloocke, A., Liebman, M., Hatfield, J.L., Herzmann, D., Córdova, S.C., Edmonds, P., Togliatti, K., Kessler, A., Danalatos, G., Pasley, H., Pederson, C., Lamkey, K.R., 2020. Predicting crop yields and soil-plant nitrogen dynamics in the US Corn Belt. Crop Sci. 60, 721–738. https://doi.org/10.1002/csc2.20039

Balboa, G.R., Ciampitti, I.A., 2020. Estimating biological nitrogen fixation in field-grown soybeans: impact of B value. Plant Soil 446, 195–210. https://doi.org/10.1007/s11104-019-04317-1

Briat, J.-F., Gojon, A., Plassard, C., Rouached, H., Lemaire, G., 2020. Reappraisal of the central role of soil nutrient availability in nutrient management in light of recent advances in plant nutrition at crop and molecular levels. Eur. J. Agron. 116, 126069. https://doi.org/10.1016/j.eja.2020.126069

Brown, J.R., 1998. Recommended chemical soil test procedures for the North Central Region, 221st ed. Missouri Agricultural Experiment Station, University of Missouri--Columbia.

Bürkner, P.-C., 2018. Advanced Bayesian Multilevel Modeling with the R Package *brms*. R J. 10, 395. https://doi.org/10.32614/RJ-2018-017

Bürkner, P.-C., 2017. brms: An R Package for Bayesian Multilevel Models Using Stan. J. Stat. Softw. 80, 1–28. https://doi.org/10.18637/jss.v080.i01

Cafaro La Menza, N., Monzon, J.P., Lindquist, J.L., Arkebauer, T.J., Knops, J.M.H., Unkovich, M., Specht, J.E., Grassini, P., 2020. Insufficient nitrogen supply from symbiotic fixation reduces seasonal crop growth and nitrogen mobilization to seed in highly productive soybean crops. Plant Cell Environ. 43, 1958–1972. https://doi.org/10.1111/pce.13804

Carciochi, W.D., Rosso, L.H.M., Secchi, M.A., Torres, A.R., Naeve, S., Casteel, S.N., Kovács, P., Davidson, D., Purcell, L.C., Archontoulis, S., Ciampitti, I.A., 2019. Soybean yield, biological N_2 fixation and seed composition responses to additional inoculation in the United States. Sci. Rep. 9, 19908. https://doi.org/10.1038/s41598-019-56465-0

Ciampitti, I., Schapaugh, W.T., Shoup, D., Duncan, S., Diaz, D.R., Peterson, D., Rogers, D.H., Whitworth, J., Schwarting, H., Jardine, D., others, 2016. Soybean Production Handbook. Manhattan, Kansas.

Ciampitti, I.A., De Borja Reis, A.F., Córdova, S.C., Castellano, M.J., Archontoulis, S., Correndo, A.A., Almeida, L.F.A., Moro Rosso, L.H., 2021. Revisiting biological nitrogen fixation dynamics in soybeans. Front. Plant Sci. https://doi.org/10.3389/fpls.2021.727021

Ciampitti, I.A., Salvagiotti, F., 2018. New insights into soybean biological nitrogen fixation. Agron. J. 110, 1185–1196. https://doi.org/10.2134/agronj2017.06.0348

Collino, D.J., Salvagiotti, F., Perticari, A., Piccinetti, C., Ovando, G., Urquiaga, S., Racca, R.W., 2015. Biological nitrogen fixation in soybean in Argentina: relationships with crop, soil, and meteorological factors. Plant Soil 392, 239–252. https://doi.org/10.1007/s11104-015-2459-8

Córdova, S.C., Castellano, M.J., Dietzel, R., Licht, M.A., Togliatti, K., Martinez-Feria, R., Archontoulis, S. V., 2019. Soybean nitrogen fixation dynamics in Iowa, USA. F. Crop. Res. 236, 165–176. https://doi.org/10.1016/j.fcr.2019.03.018

de Borja Reis, A.F., Moro Rosso, L., Purcell, L.C., Naeve, S., Casteel, S.N., Kovács, P., Archontoulis, S., Davidson, D., Ciampitti, I.A., 2021. Environmental Factors Associated With Nitrogen Fixation Prediction in Soybean. Front. Plant Sci. 12. https://doi.org/10.3389/fpls.2021.675410

de Borja Reis, A.F., Tamagno, S., Moro Rosso, L.H., Ortez, O.A., Naeve, S., Ciampitti, I.A., 2020. Historical trend on seed amino acid concentration does not follow protein changes in soybeans. Sci. Rep. 10, 17707. https://doi.org/10.1038/s41598-020-74734-1

Divito, G.A., Echeverría, H.E., Andrade, F.H., Sadras, V.O., 2016. Soybean shows an attenuated nitrogen dilution curve irrespective of maturity group and sowing date. F. Crop. Res. 186, 1–9. https://doi.org/10.1016/j.fcr.2015.11.004

Donahue, J.M., Bai, H., Almtarfi, H., Zakeri, H., Fritschi, F.B., 2020. The quantity of nitrogen derived from symbiotic N fixation but not the relative contribution of N fixation to total N uptake increased with breeding for greater soybean yields. F. Crop. Res. 259, 107945. https://doi.org/10.1016/j.fcr.2020.107945

Fehr, W.R., Caviness, C.E., 1977. Stages of Soybean Development. Spec. Rep. 80, 11.

Gaspar, A.P., Laboski, C.A.M., Naeve, S.L., Conley, S.P., 2017. Dry matter and nitrogen uptake, partitioning, and removal across a wide range of soybean seed yield levels. Crop Sci. 57, 2170–2182. https://doi.org/10.2135/cropsci2016.05.0322

Greenwood, D.J., Lemaire, G., Gosse, G., Cruz, P., Draycott, A., Neeteson, J.J., 1990. Decline in percentage N of C3 and C4 crops with increasing plant mass. Ann. Bot. 66, 425–436. https://doi.org/10.1093/oxfordjournals.aob.a088044

Greenwood, D.J., Neeteson, J.J., Draycott, A., 1986. Quantitative relationships for the dependence of growth rate of arable crops on their nitrogen content, dry weight and aerial environment. Plant Soil 91, 281–301. https://doi.org/10.1007/BF02198111

Hooten, M.B., Hefley, T.J., 2019. Bringing Bayesian Models to Life. CRC Press, Boca Raton, FL : CRC Press, Taylor & Francis Group, 2019. https://doi.org/10.1201/9780429243653

Houwaard, F., 1980. Influence of ammonium and nitrate nitrogen on nitrogenase activity of pea plants as affected by light intensity and sugar addition. Plant Soil 54, 271–282. https://doi.org/10.1007/BF02181853

Kansas Mesonet, 2017. Kansas Mesonet Historical Data [WWW Document]. URL http://mesonet.k-state.edu/weather/historical (accessed 6.30.21).

Köppen, W., 2011. The thermal zones of the Earth according to the duration of hot, moderate and cold periods and to the impact of heat on the organic world. Meteorol. Zeitschrift 20, 351–360. https://doi.org/10.1127/0941-2948/2011/105

Latimore, M., Giddens, J., Ashley, D.A., 1977. Effect of Ammonium and Nitrate Nitrogen upon Photosynthate Supply and Nitrogen Fixation by Soybeans 1 . Crop Sci. 17, 399–404. https://doi.org/10.2135/cropsci1977.0011183x001700030015x

Lemaire, G., Gastal, F., 1997. N Uptake and Distribution in Plant Canopies, in: Diagnosis of the Nitrogen Status in Crops. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3–43. https://doi.org/10.1007/978-3-642-60684-7_1

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

Lindemann, W.C., Ham, G.E., 1979. Soybean Plant Growth, Nodulation, and Nitrogen Fixation as Affected by Root Temperature. Soil Sci. Soc. Am. J. 43, 1134–1137. https://doi.org/10.2136/sssaj1979.03615995004300060014x

Linn, D.M., Doran, J.W., 1984. Effect of Water-Filled Pore Space on Carbon Dioxide and Nitrous Oxide Production in Tilled and Nontilled Soils. Soil Sci. Soc. Am. J. 48, 1267–1272. https://doi.org/10.2136/sssaj1984.03615995004800060013x

McLoughlin, T.J., Hearn, S., Alt, S.G., 1990. Competition for nodule occupancy of introduced Bradyrhizobium japonicum strains in a Wisconsin soil with a low indigenous bradyrhizobia population. Can. J. Microbiol. 36, 839–845. https://doi.org/10.1139/m90-145

Moro Rosso, L.H., Tamagno, S., da Silva, A.L., Torres, A.R., Schwalbert, R.A., Ciampitti, I.A., 2021. Relative abundance of ureides differs among plant fractions in soybean. Eur. J. Agron. 122, 126175. https://doi.org/10.1016/j.eja.2020.126175

Mourtzinis, S., Kaur, G., Orlowski, J.M., Shapiro, C.A., Lee, C.D., Wortmann, C., Holshouser, D., Nafziger, E.D., Kandel, H., Niekamp, J., Ross, W.J., Lofton, J., Vonk, J., Roozeboom, K.L., Thelen, K.D., Lindsey, L.E., Staton, M., Naeve, S.L., Casteel, S.N., Wiebold, W.J., Conley, S.P., 2018. Soybean response to nitrogen application across the United States: A synthesis-analysis. F. Crop. Res. 215, 74–82. https://doi.org/10.1016/j.fcr.2017.09.035

Novoa, R., Loomis, R.S., 1981. Nitrogen and plant production. Plant Soil 58, 177–204. https://doi.org/10.1007/BF02180053

Osborne, S.L., Riedell, W.E., 2006. Starter Nitrogen Fertilizer Impact on Soybean Yield and Quality in the Northern Great Plains. Agron. J. 98, 1569–1574. https://doi.org/10.2134/agronj2006.0089

Pasley, H.R., Huber, I., Castellano, M.J., Archontoulis, S. V., 2020. Modeling Flood-Induced Stress in Soybeans. Front. Plant Sci. 11. https://doi.org/10.3389/fpls.2020.00062

Purcell, L.C., Serraj, R., Sinclair, T.R., De, A., 2004. Soybean N 2 Fixation Estimates, Ureide Concentration, and Yield Responses to Drought. Crop Sci. 44, 484–492. https://doi.org/10.2135/cropsci2004.4840

R Core Team, 2021. R: A Language and Environment for Statistical Computing.

Ratjen, A.M., Lemaire, G., Kage, H., Plénet, D., Justes, E., 2018. Key variables for simulating leaf area and N status: Biomass based relations versus phenology driven approaches. Eur. J. Agron. 100, 110–117. https://doi.org/10.1016/j.eja.2018.04.008

Richardson, D.A., Jordan, D.C., Garrard, E.H., 1957. The Influence of Combined Nitrogen on Nodulation and Nitrogen Fixation By Rhizobium Meliloti Dangeard. Can. J. Plant Sci. 37, 205–214. https://doi.org/10.4141/cjps57-025

Salvagiotti, F., Cassman, K.G., Specht, J.E., Walters, D.T., Weiss, A., Dobermann, A., 2008. Nitrogen uptake, fixation and response to fertilizer N in soybeans: A review. F. Crop. Res. 108, 1–13. https://doi.org/10.1016/j.fcr.2008.03.001

Santachiara, G., Salvagiotti, F., Gerde, J.A., Rotundo, J.L., 2018. Does biological nitrogen fixation modify soybean nitrogen dilution curves? F. Crop. Res. 223, 171–178. https://doi.org/10.1016/j.fcr.2018.04.001

Santachiara, G., Salvagiotti, F., Rotundo, J.L., 2019. Nutritional and environmental effects on biological nitrogen fixation in soybean: A meta-analysis. F. Crop. Res. 240, 106–115. https://doi.org/10.1016/j.fcr.2019.05.006

Schipanski, M.E., Drinkwater, L.E., Russelle, M.P., 2010. Understanding the variability in soybean nitrogen fixation across agroecosystems. Plant Soil 329, 379–397. https://doi.org/10.1007/s11104-009-0165-0

Sinclair, T.R., De Wit, C.T., 1975. Photosynthate and nitrogen requirements for seed production by various crops. Science (80-.). 189, 565–567. https://doi.org/10.1126/science.189.4202.565

Starling, M.E., Wood, C.W., Weaver, D.B., 1998. Starter Nitrogen and Growth Habit Effects on Late-Planted Soybean. Agron. J. 90, 658–662. https://doi.org/10.2134/agronj1998.00021962009000050015x

Street, H.E., Sheat, D.E.G., 1958. The absorption and availability of nitrate and ammonia, in: Der Stickstoffumsatz / Nitrogen Metabolism. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 150–165. https://doi.org/10.1007/978-3-642-94733-9_6

Streeter, J.G., 1972. Nitrogen Nutrition of Field-grown Soybean Plants: I. Seasonal Variations in Soil Nitrogen and Nitrogen Composition of Stem Exudate 1. Agron. J. 64, 311–314. https://doi.org/10.2134/agronj1972.00021962006400030016x

Tamagno, S., Sadras, V.O., Haegele, J.W., Armstrong, P.R., Ciampitti, I.A., 2018. Interplay between nitrogen fertilizer and biological nitrogen fixation in soybean: implications on seed yield and biomass allocation. Sci. Rep. 8, 17502. https://doi.org/10.1038/s41598-018-35672-1

Unkovich, M., Herridge, D., Peoples, M., Cadisch, G., Boddey, B., Giller, K., Alves, B., Chalk, P., 2008. Measuring plant-associated nitrogen fixation in agricultural systems, Monograph. ed. Australian Centre for International Agricultural Research (ACIAR), Canberra, Australia.

USDA, 2018. Agricultural Chemical Usage.

Wesley, T.L., Lamond, R.E., Martin, V.L., Duncan, S.R., 1998. Effects of late-season nitrogen fertilizer on irrigated soybean yield and composition. J. Prod. Agric. 11, 331–336. https://doi.org/10.2134/jpa1998.0331

Wikle, C.K., Zammit-Mangion, A., Cressie, N., 2019. Spatio-Temporal Statistics with R. Chapman and Hall/CRC, Boca Raton, Florida : CRC Press, [2019]. https://doi.org/10.1201/9781351769723

Wood, C.W., Torbert, H.A., Weaver, D.B., 1993. Nitrogen Fertilizer Effects on Soybean Growth, Yield, and Seed Composition. J. Prod. Agric. 6, 354–360. https://doi.org/10.2134/jpa1993.0354
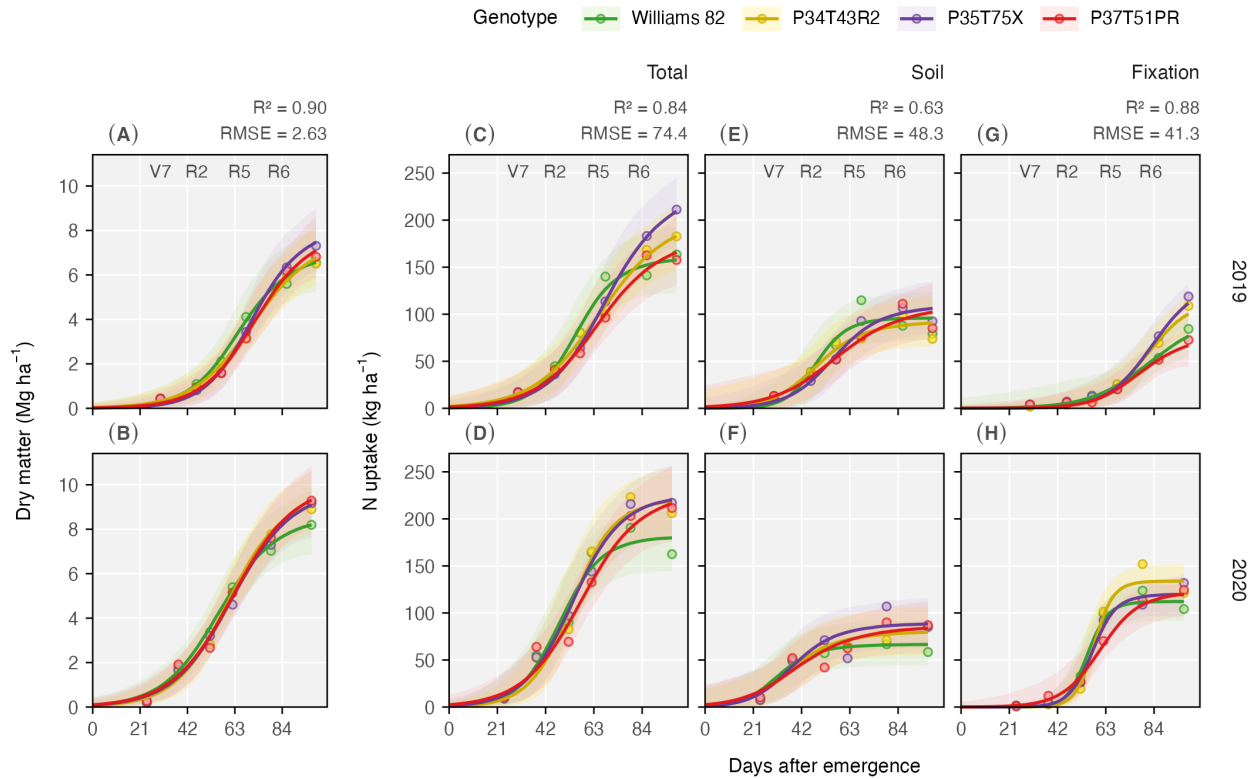
Figure 2.1. Dry matter accumulation for the site-year in 2019 (A) and 2020 (B), total N uptake (C and D), uptake from the soil N supply (E and F), and from the N fixation process (G and H). The logistic model described seasonal changes for all response variables and genotypes. Points represent observation means across the six sampling stages. The stages of seventh leaf (V7), full flowering (R2), beginning of seed filling (R5), and full seed (R6) are shown within the panels. Solid lines represent the posterior prediction median and shaded area represents the 95% credible interval. The coefficient of determination (R^2) and root mean squared error (RMSE) are presented as a measure of goodness-of-fit for each response variable.
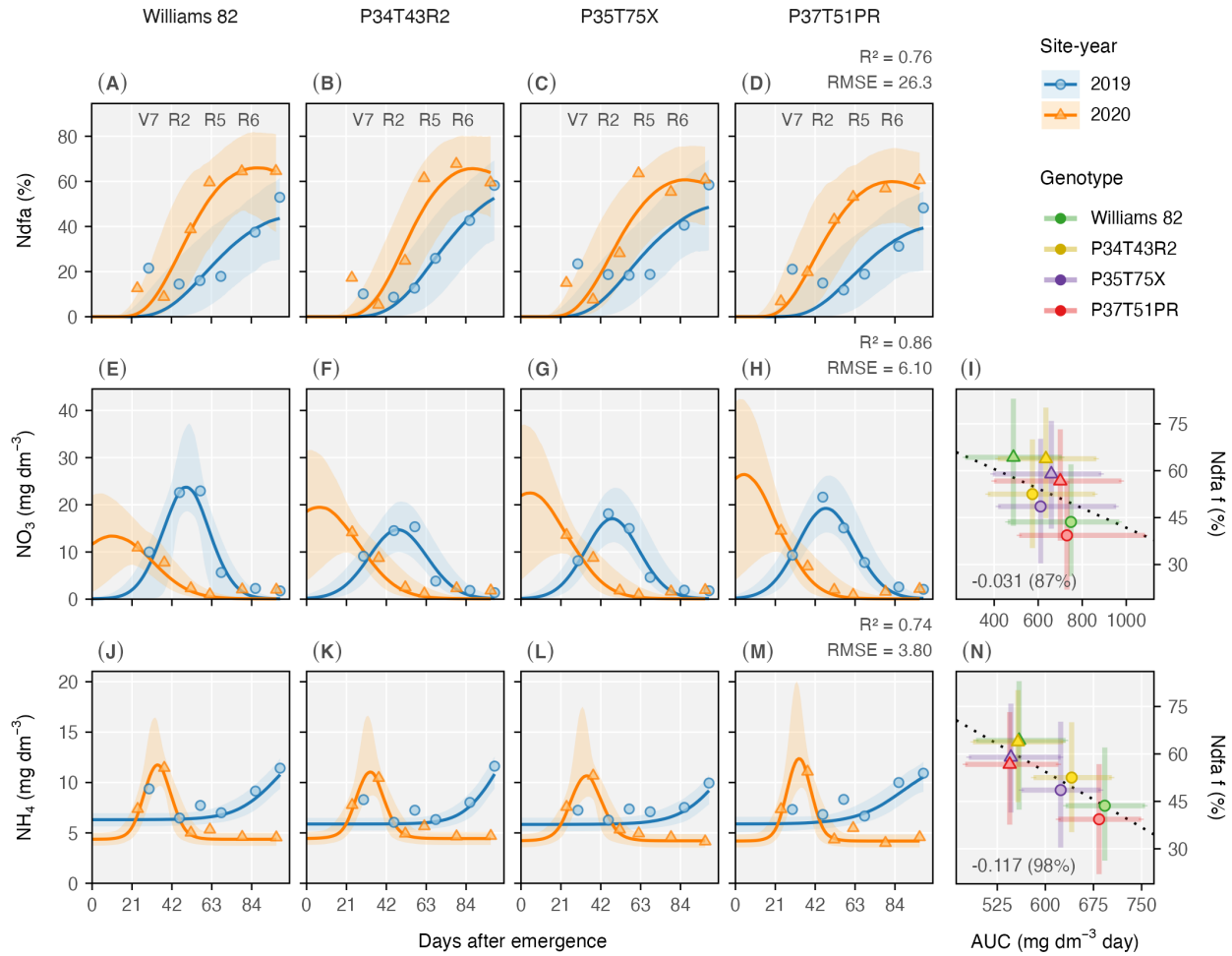
Figure 2.2. Seasonal changes on N derived from the atmosphere (Ndfa, A-D), soil nitrate (NO_3, E-H) and ammonium (NH_4, J-M) at a 60-cm depth layer across site-years (2019 and 2020) and genotypes. The coefficient of determination (R^2) and root mean squared error (RMSE) were presented for the gaussian growth (Ndfa) and gaussian peak models (NO_3 and NH_4). Points represent observation means (six sampling stages), solid lines represent posterior prediction medians, and shaded areas the 95% credible intervals. The stages of seventh leaf (V7), full flowering (R2), beginning of seed filling (R5), and full seed (R6) are shown within the panels. The Ndfa prediction at physiological maturity (R7) (f) was regressed over NO_3 (I) and NH_4 (N) area under the curve (AUC). Median of slope estimates are followed by the probability of negative values considering regressions over all posterior draws.
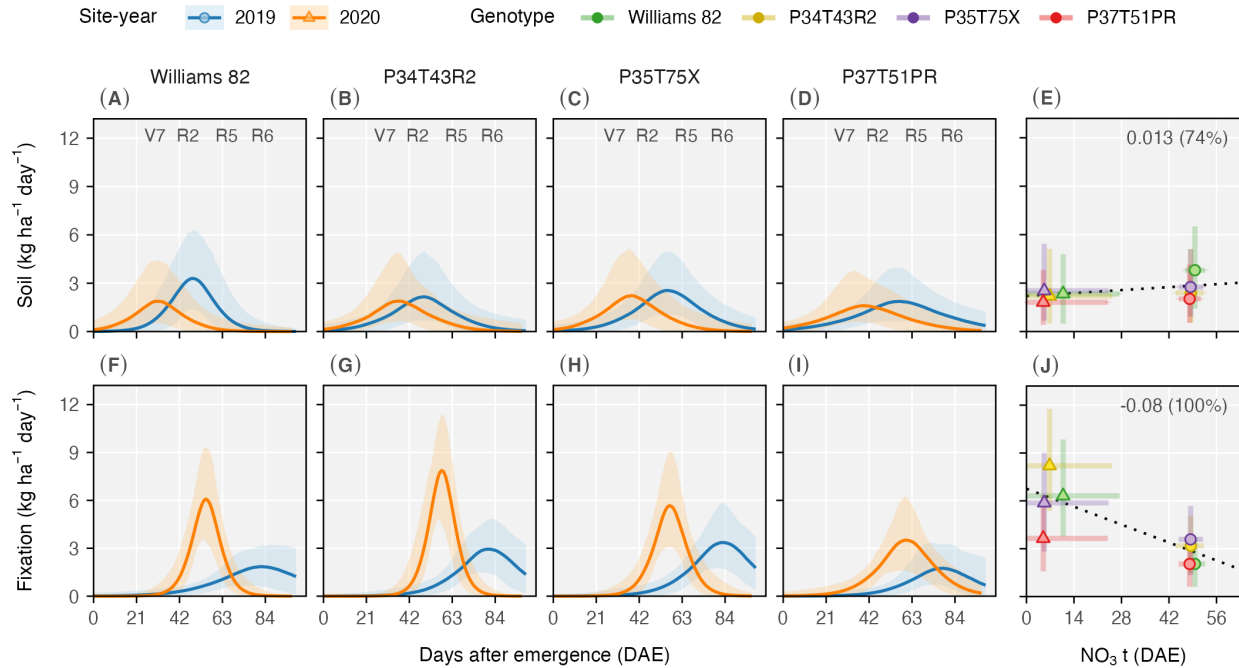
Figure 2.3. Nitrogen uptake rates from the soil (A-D) and N fixation supply (F-I) throughout the 2019 and 2020 growing seasons (site-years) for each genotype. Solid lines represent posterior prediction medians, and shaded areas represent the 95% credible intervals from the logistic non-linear models. The maximum uptake rate from the soil N supply (E) and N fixation (J) was regressed over the timing of soil nitrate (NO_3) peak. Median of slope estimates are followed by the probability of either positive (E) or negative (J) values considering regressions over all the posterior draws. The stages of seventh leaf (V7), full flowering (R2), beginning of seed filling (R5), and full seed (R6) are shown within the upper panels.
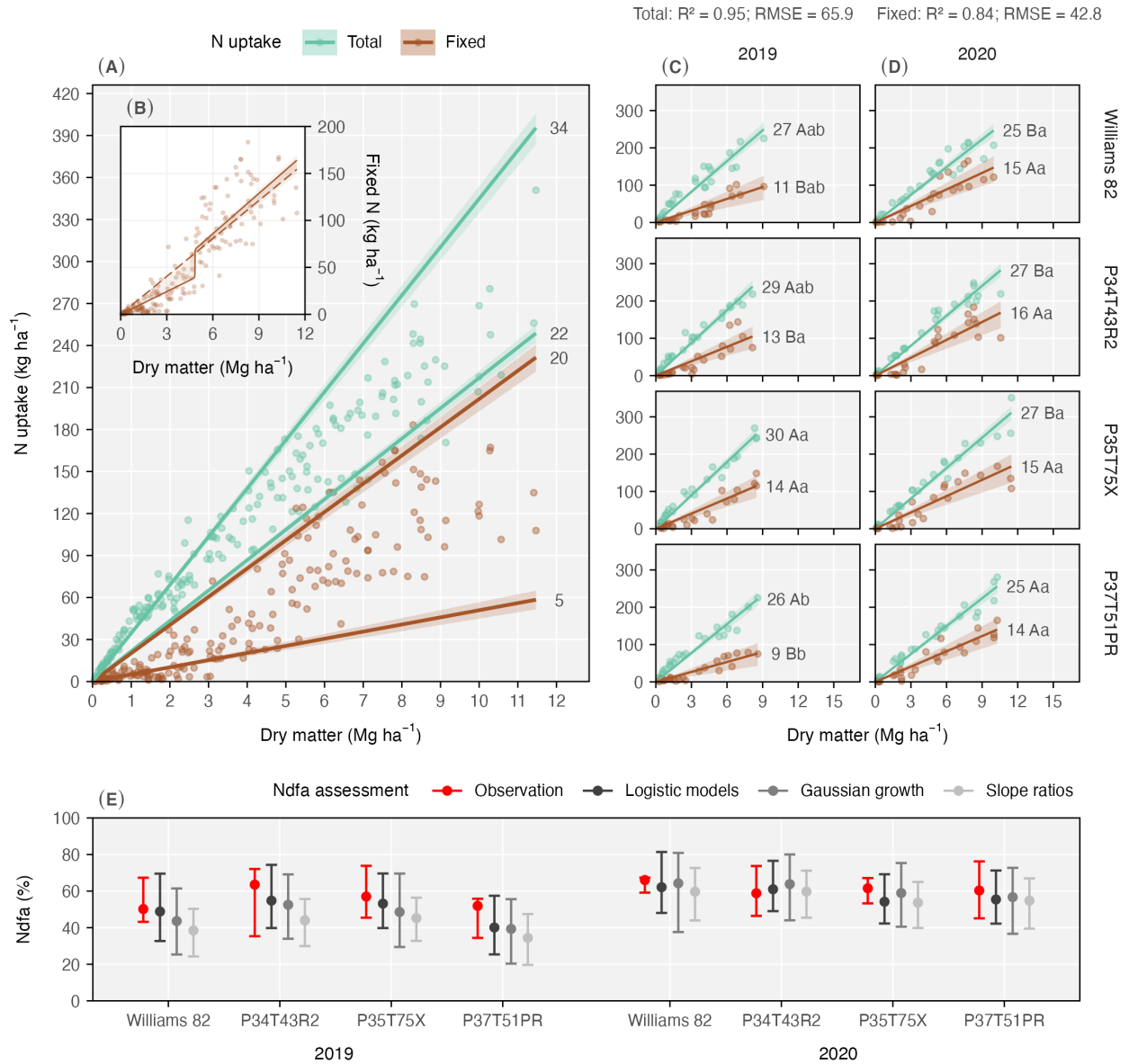
Figure 2.4. Relationship between seasonal dry matter and nitrogen (N) uptake with boundary functions over the 10th and 90th percentiles (A). Breaking point of fixed N requirement (solid line) over dry matter as compared to the overall seasonal requirement (dashed line) (B). Total N uptake and fixed N for the 2019 (C) and 2020 (D) site-years across genotypes. Solid lines represent posterior prediction medians and shaded areas the 95% credible intervals (CI). Slope medians of posterior distributions are presented at the end of the lines and followed by letters from the Tukey test (p < 0.05). Uppercase letters compare site-years within genotypes, and lowercase letters compare genotypes within site-years. Posterior medians and 95% CI of N derived from the atmosphere (Ndfa, %) predicted by different models (E).

31

Tables

Table 2.1. Soil and weather conditions for Manhattan (KS) during the 2019 and 2020 soybean growing seasons (site-years). Composite soil samples were collected approximately 30 days before sowing from a 0.15 m depth layer, except for nitrate (NO_3), ammonium (NH_4), and sulfate (SO_4), measured from a 0.60 m depth layer. Weather variables were summarized for different periods of the crop development, considering crop emergence (VE), full flowering (R2), beginning of seed filling (R5), and physiological maturity (R7).

Variable	2019	2020
Pre-sowing soil		
pH[a]	6.3	7.1
P[b], mg dm^{-3}	18	16
SOM[c], g kg^{-1}	32	26
CEC[d], cmol$_c$ dm^{-3}	30	21
Clay, g kg^{-1}	300	290
Sand, g kg^{-1}	130	120
Silt, g kg^{-1}	570	590
NO_3, mg dm^{-3}	3.2	4.3
NH_4, mg dm^{-3}	8.0	8.9
SO_4, mg dm^{-3}	3.9	2.7
Weather variables		
Precipitation, mm		
30 days prior VE	168	128
VE to R2	297	170
R2 to R5	138	65
R5 to R7	146	68
Soil mean temperature, °C		
30 days prior VE	20	21
VE to R2	25	26
R2 to R5	26	25
R5 to R7	25	22
Air mean temperature, °C		
30 days prior VE	20	22
VE to R2	25	26
R2 to R5	26	25
R5 to R7	25	22
Solar radiation, MJ m^{-2}		
30 days prior VE	343	393
VE to R2	656	555
R2 to R5	250	298
R5 to R7	282	373

[a] Water based; [b] Phosphorus extracted by Mehlich-3; [c] Soil organic matter via loss on ignition (LOI); [d] Cation exchange capacity.

Table 2.2. Soybean seed yield, seed size, seed number, protein, and oil concentration for each genotype during the 2019 and 2020 growing seasons (site-years). Values represent the median of estimated marginal means (emmeans) posterior distributions, and letters represent Tukey test results ($p < 0.05$) within the Bayesian statistical framework. Uppercase letters compare site-years within genotypes, while lowercase letters compare the genotypes within site-years. Letters were omitted when treatment levels did not differ.

Genotype	Yield	Seed size [a]	Seed number [b]	Protein	Oil
	Mg ha^{-1}	mg seed^{-1}	seed m^{-2}	g kg^{-1}	g kg^{-1}
2019					
Williams 82	2.90	122 B	2360	435 Aa	206 B
P34T43R2	2.99	132 A	2250	432 Aab	199 B
P35T75X	3.11	124 B	2500	420 Abc	201 B
P37T51PR	2.95	125 B	2320	406 Ac	202 B
2020					
Williams 82	3.24	137 A	2370	410 Ba	228 A
P34T43R2	3.52	142 A	2460	400 Bab	223 A
P35T75X	3.43	137 A	2500	399 Bab	220 A
P37T51PR	3.63	144 A	2500	392 Bb	222 A

[a] Soybean seed weight; [b] Seed number was calculated based on seed yield and seed size.

Chapter 3 - Vertical Canopy Profile and the Impact of Branches on Soybean Seed Composition

Abstract

Soybean [*Glycine max* (L.) Merr.] seeds are of global importance for human and animal nutrition due to their high protein and oil concentrations, and their complete amino acid (AA) and fatty acid (FA) profiles. However, a detailed description of seed composition at different canopy portions (i.e., main stem and branch nodes) is currently lacking in scientific literature. This study aims to (1) characterize seed yield and composition (protein, oil, AA, and FA) at the main stem (exploring a vertical canopy profile) and stem branches and (2) quantify the impact of canopy yield allocation on seed composition, focusing on branches as a potential contributor for higher yields. Four genotypes were field-grown during the 2018 and 2019 seasons, with seeds manually harvested from all the branches and three main stem segments (lower, middle, and upper). Seed samples were analyzed for seed yield (Mg/ha), seed size (mg/seed), protein and oil content (mg/seed) and their respective concentrations (g/kg), and AA and FA concentrations within protein and oil (g/100 g), herein called abundance. The upper main stem produced greater protein (25%) and oil (15%) content relative to the lower section; however, oil concentration increased from top to bottom while protein concentration followed the opposite vertical gradient. Limiting AAs (lysine, cysteine, methionine, threonine, and tryptophan) were more abundant in the lower main stem, while the oleic/(linoleic + linolenic) ratio was greater in the upper segment. Overall, branches produced seeds with inferior nutritional quality than the main stem. However, the contribution of branches to yield (%) was positively related to limiting AA abundance and oil concentration across soybean genotypes. Future research studies should consider the

morphological process of stem branching as a critical factor intimately involved with soybean seed composition across environments, genotypes, and management practices.

Introduction

High concentrations of seed protein and oil have expanded soybean [*Glycine max* (L.) Merr.] production worldwide. In 2018, 345 Tg of soybean seeds were produced (FAO, 2021). Considering a safe protein intake of ~60 g/adult/day (WHO/FAO/UNU, 2007), soybeans alone can supply roughly 75% of the global protein need and contribute to 30% of global vegetable oil production (FAO, 2021). In the United States (US), dry basis protein and oil concentration are about 400 and 215 g/kg, respectively (Rotundo et al., 2016). Environmental conditions are known to modify protein and oil concentrations by roughly 20% (Rotundo and Westgate, 2009), with these factors dominating the variation on soybean seed composition (Assefa et al., 2019). However, changes in seed composition within the plant canopy (Collins and Cartter, 1956) have received less attention, especially considering seeds from the branches.

Modern soybean genotypes not only produce high yields under high plant density but also compensate for the absence of plants with enhanced branching (Suhre et al., 2014). This flexibility might favor yield stability (Agudamu and Shiraiwa, 2016), especially under adverse conditions of stand establishment in northern latitudes (Lamichhane et al., 2020). Along with yield increase, soybeans have been reported to have decreased protein and increased oil concentration (Rincker et al., 2014). However, it is unclear that how branches contributed to the yield-protein-oil relationship in soybeans. Furthermore, the concentration of a seed component is a consequence of its content, which depends on assimilate supply (Rotundo et al., 2009) and varies within the canopy. Despite genotype leaf characteristics, greater net radiation is found in

the upper section of the canopy, and greater temperature, smaller vapor-pressure deficit (VPD), and smaller carbon dioxide (CO_2) concentration, all relative to lower canopy (Baldocchi et al., 1983, 1985).

Genetics and the environment have a great influence on stem branching in soybeans (Shim et al., 2017, 2019). Remarkably, low plant density enhances branching (Carpenter and Board, 1997), possibly associated with radiation quality (e.g., red to far-red ratio) at the ground level (Toyota et al., 2017). Branch leaves unroll about 30 days after sowing and develop under continuous shading, presenting thinner leaves compared to the main stem (Koller, 1972). In addition, branch nodes have late flowering and pod set but similar physiological maturity compared to the main stem (Munier-Jolain et al., 1994). This internal ontogenesis variation relates to reproductive abortion, seed-filling rate and duration, and yield (Egli and Bruening, 2006a,b). During the seed filling, 30% of a leaf carbon (C) assimilate is remobilized to pods on the same node and the other 30–40% to the four neighbor nodes (Stephenson and Wilson, 1977). Therefore, differences in source for C assimilation might prevail on the seed composition of a stem segment, mainly for protein due to the strong nitrogen (N) remobilization process (Sinclair and De Wit, 1975).

At physiological maturity, upper main stem nodes have greater protein and lower oil concentration than the lower main stem (Sharma et al., 2013), but little is known about the concentration of seed components in the branches. Under low plant densities, this vertical gradient for seed composition is reduced but still maintained (Huber et al., 2016). Protein and oil vertical profiles are consistent regardless of soybean growth type (determinate or indeterminate) and genotype protein level (Escalante and Wilcox, 1993a,b). A few reports have explored the concentration of amino acids (AAs) and fatty acids (FAs) as a measure of the soybean nutritional

value. Bennett et al. (2003) found a greater concentration of sulfur-containing AAs in the lower main stem seeds, while seeds in the upper main stem nodes presented higher oleic acid. Greater oleic concentration in the top main stem nodes was confirmed by Bellaloui and Gillen (2010), contributing to heat stability and shelf life for food preparation and biodiesel industry (Carrera and Dardanelli, 2017). Sulfur AAs (cysteine and methionine) are among the five limiting AAs (with lysine, threonine, and tryptophan) often supplemented in monogastric dietary (Thakur and Hurburgh, 2007).

Although a majority of the soybean industry does not reward superior nutritional quality (protein, oil, AA, and FA) (Brumm and Hurburgh, 2006), the increasing demand for sustainable food production could disseminate premium payments, promote seed-quality segregation at the field level, and enhance competitiveness, and marketability. Differences in seed yield and composition from the main stem and branches should be explored to understand potential unintended changes in these critical plant traits at the whole plant level. Following this rationale, the aims of this study were to (1) characterize the seed yield and composition (protein, oil, AA, and FA) at three segments of the main stem (vertical canopy profile) and stem branches; and (2) quantify the impact of canopy yield allocation on seed composition, focusing on branches as a potential contributor for high yields in modern genotypes.

Materials and Methods

Experimental Design and Growing Conditions

Field experiments were performed during the 2018 and 2019 growing seasons, at the Ashland Bottoms Agronomy Farm (39.14° North, 96.64° West, 315 m elevation, in Manhattan, Kansas, US) and Kansas River Valley Experimental Field (39.12° North, 95.92° West, 280 m

elevation, in Rossville, Kansas, US), respectively. Conventional tillage was performed before

sowing, and composite soil samples were collected to characterize texture and initial chemical

properties (Table 3.1). Both fields have been under soybean-corn (*Zea mays*) rotation, and

irrigation was not adopted during the growing season. Climate is classified as Cfa (humid

subtropical) with evenly distributed precipitation throughout the year (Köppen, 2011).

Four genotypes with contrasting branching potential were selected: P31T11R [maturity

group (MG) 3.1, released in 2014]; P34T43R2 (3.4, 2014); P35T58R (3.5, 2013); and P39T67R

(3.9, 2013) (Corteva Agriscience, Johnston, Iowa, US). Experimental design followed a

randomized complete block with four repetitions in both site-years. Treatment factors were (1)

genotype (four levels) and (2) canopy portion from which seeds were produced (lower, middle,

and upper main stem segments and the branches altogether). Herein, we use the term branches to

represent a morphological structure accounting for a fraction of the seed yield, not the process of

branching itself. Experimental plots consisted of six rows spaced 0.75 m and a plot size of 60 m^2.

The sowing dates were April 29, 2018, and June 9, 2019. All genotypes were sown at 300,000

seeds/ha that resulted in approximately 240,000 plants/ha at harvest time. Soybean seeds were

inoculated before sowing with Vault HP Rhizobia Inoculant (BASF, Ludwigshafen, Germany)

containing at least 3.0×10^9 colony-forming unit/ml of Bradyrhizobium japonicum. Weeds,

insects, and diseases were managed according to the best agronomic practices.

Weather variables were retrieved from the DAYMET database (Thornton et al., 2020)

and summarized from soybean emergence (VE) to physiological maturity (R7 stage, one pod in

the main stem had reached mature pod color) (Fehr and Caviness, 1977) according to Correndo

et al. (2021) (Table 3.1). The soybean cycle reached ~120 days in 2018 and ~105 days in 2019.

Differences in MG across the tested genotypes introduced an overall season-length variation of

less than a week. The sowing date affected the seasonal weather conditions, with Rossville 2019 (late sowing) presenting lower temperatures and solar radiation. Ashland 2018 was less humid and had greater VPD and lower precipitation compared to Rossville 2019.

Measurements and Laboratory Analysis

At harvest time (R8 stage), three central adjacent rows covering ~3.4 m^2 were manually harvested from each plot. Main stems were divided into three segments (lower, middle, and upper), with five to six nodes in each segment. Cotyledonary and unifoliolate nodes were not considered due to the absence of pods. Branches were collected as a unique segment, adding up to the four canopy portions evaluated. Samples were machine threshed and taken to the laboratory for determining seed yield (Mg/ha) and seed weight (mg/seed, herein called seed size), both adjusted to 130 g/kg moisture basis. Seed number (1,000 seeds/m^2) was calculated based on yield and seed size. Finally, seed samples (~500 g) were oven-dried (65°C) until constant weight is obtained and ground to 0.1 mm particle size. Protein, oil, AA, and FA concentrations were determined via near-infrared spectroscopy (NIR) using the Perten DA7200 Feed Analyzer (Perten Instruments, Stockholm, Sweden). The ground material was scanned between 1,000 and 2,500 nm wavelength, and normalized reflectance readings were used to estimate each seed component. The calibration method was based on Honigs et al. (1985) and evaluated with cross-validation using the coefficient of determination (r^2). After accounting for protein and oil, the remaining seed size was classified as residue fraction, mostly carbohydrates.

Besides dry basis concentration (g/kg), seed components were expressed in content (mg/seed), a consequence of the seed-filling process, speaking to industry and agronomists. The AA and FA concentrations were expressed within protein (g/100 g protein) and oil (g/100 g oil),

respectively, as a measure of abundance within each component (Gerde and White, 2008). Because NIR does not differentiate asparagine and aspartate, or glutamine and glutamate, these AAs were expressed as aspartic and glutamic acid, respectively. Therefore, the 20 primary AAs were analyzed as a total of 18 types and then added within three groups: (1) non-essential (alanine, arginine, aspartic acid, glutamic acid, glycine, proline, serine, and tyrosine); (2) essential non-limiting (isoleucine, leucine, histidine, phenylalanine, and valine); and (3) essential limiting (lysine, cysteine, methionine, threonine, and tryptophan) following Pfarr et al. (2018). The abundance of those five limiting AAs (LAAs, g/100 g protein) was considered the main descriptor of protein quality hereafter. Five FAs were determined: linoleic, oleic, palmitic, linolenic, and stearic. However, the oleic/(linoleic + linolenic) ratio was the main descriptor of oil quality (Gao et al., 2009). Whole plant and main stem data were calculated as the weighted average of the containing portions, considering respective yields.

Statistical Analysis

Linear mixed models related dependent (e.g., yield and protein content) with independent variables (i.e., genotype and canopy portion). Those variables were tested using three models, considering different levels of the plant canopy: (1) whole plant (only testing the effect of genotype); (2) two canopy portions (also comparing main stem and branches); and (3) four canopy portions (lower, middle, and upper main stem segments and the branches). For the first model, genotype was the only fixed effect (four levels), with a random intercept for site-year, block, and block within site-year. For the other two models, fixed effects were canopy portion, genotype, and their interaction, with random effects also including genotype nested in block × site-year, because canopy portions were observed on the same plant sample. To investigate the

relationship between seed yield of the whole plant or entire main stem and the contribution of branches to yield, a regression model was proposed across all soybean genotypes. In this case, seed yield (Mg/ha) was the dependent, and branch-yield contribution (%) was the independent variable, both continuous and with a random intercept for site-year.

Finally, and related to the second objective, seed yield of the whole plant and branch-yield contribution were tested as fixed effects (independent variables) describing protein and oil concentrations, LAA abundance, and oleic/(linoleic + linolenic) ratio (four dependent variables). A random intercept for site-year was also included, and dependent variables were considered at the whole plant level or by canopy portion (checking for interactions). This model intends to dissect the effect of branches from the effect of yield variation on soybean seed composition. The independent variables were centered (subtracted by the mean) and scaled (divided by the standard deviation) before model fitting and ANOVA. The center-scale transformation was performed due to contrasting variable magnitude, which could impair the hypothesis testing.

Statistical analysis was performed in the R software (R Core Team, 2020). For all variables, normality and homogeneity of variance were checked using Shapiro–Wilk and Bartlett's test, and data transformations were not employed. The *lme4* package (Bates et al., 2015) was used to fit the models, and the *car* package (Fox and Weisberg, 2019) was used to perform type III ANOVA. Significant effects represented p value < 0.05 (F-test). A protected Fisher's least significant difference (LSD) test was adopted for means comparison using the *multcomp* package (Hothorn et al., 2008). The least-square means (LSMEANS) were computed for all the treatment combinations. The figures presented in this manuscript were generated using the *ggplot2* package (Wickham, 2016).

Results

Seed Yield and Macrocomponent Content

Results from the three linear mixed models, comparing genotypes and canopy portions, are displayed using bar charts that resemble a soybean plant (Figure 3.1A). Remarkably, across the tested soybean genotypes greater yields were attainable as the branch contribution to seed yield increased (Figure 3.1B), at the expense of seed yield derived from the main stem. At the whole plant level, seed yield differed among genotypes, ranging from 3.7 to 4.7 Mg/ha (Figures 3.2A–D), mainly driven by seed number rather than seed size (Supplementary Figures 3.1A–D). Although the main stem produced most of the yield (~70%), significant yield differences among genotypes were only captured in the branches, ranging from 0.6 to 1.9 Mg/ha. The high-yielding genotypes (P35T58R and P39T67R) produced greater seed yield coming from the branches (Figures 3.2C,D). Seed size in the main stem was also similar across genotypes but was variable in the branches (Figures 3.2F–I). The lower main stem yielded overall 25% less than the middle and upper (Figure 3.2E) segments, while branches yielded more than the main stem segments in the P35T58R and P39T67R (genotypes with high branch yield, ~37% contribution to yield), and the same or less in the P31T11R and P34T43R2 (genotypes with low branch yield, ~24% contribution to yield). Yield variations within the main stem were not connected to seed number, but to seed size, decreasing ~17% from top to bottom nodes (Figure 3.2J). On the other hand, the yield from branches was more proportional to changes in seed number, as seed size was similar to the entire main stem. The ANOVA coefficients for the mixed models testing genotype and canopy portion are shown in Supplementary Table 1.

Seeds from the upper main stem accumulated 25% more protein (Figure 3.2O) and 15% more oil (Figure 3.2T) contents than the lower main stem. However, the vertical gradient of

protein content was attenuated for the soybean genotypes presenting high branch yield. Across

the tested genotypes, protein content within the branches was slightly greater than the main stem

(without separation by the means comparison test) and usually similar to the middle main stem

(Figures 3.2F–J). Branch oil content was the same as that of the main stem, with values ranging

between the lower and middle segments (Figure 3.2T). Although protein and oil content did not

differ at the whole plant level, segment means (main stem and branches) differed among

genotypes, with an evident trade-off for the genotype P34T43R2 (Figures 3.2L,Q). The residue

content followed a similar pattern as the protein and oil, with lower values in the lower section of

the main stem and greater content in the upper main stem segment (Supplementary Figure 1J).

Remarkably, protein was the seed component with the greatest content variation among canopy

portions, relative to oil and residue.

Nutritional Quality of Soybean Seeds (Concentration)

In the soybean industry, nutritional quality is evaluated in a unit of mass (concentration),

not in terms of content per seed. In this scenario, a high concentration can be achieved with

increased content of a given component or with decreased content of the other components. For

protein, the upper main stem concentration was ~9% greater than the lower main stem, with the

middle stem and branches reaching similar values (Figure 3.3E). Despite genotype interactions,

oil concentration was almost 3% greater in the lower main stem than in the upper main stem

(Figure 3.3J). Comparing the entire main stem and branches, protein concentration was similar

for those fractions; however, oil decreased from 233 to 227 g/kg, respectively. Genotypes

differed on the portion means (main stem and branches) for both protein and oil concentration,

with the genotype with the smallest yield (P34T43R2) presenting high protein (Figure 3.3B) and

low oil (Figure 3.3G) concentrations. Across genotypes, oil concentration ranged from 219 to 240 g/kg and protein concentration ranged from 368 to 386 g/kg. The residue concentration (Supplementary Figure 1O) followed the oil trend, highlighting protein as the seed component with the greatest canopy variation.

Besides protein and oil concentration, soybean nutritional value is determined by protein and oil quality. Here, the protein quality is expressed as the abundance of LAA within a protein and oil as the oleic/(linoleic + linolenic) ratio. Both variables differed across genotypes, with LAA abundance ranging from 15.0 to 15.3 g/100 g protein and the FA ratio ranging from 0.27 to 0.34. The low-yielding genotype (P34T43R2) had the smallest LAA abundance (Figure 3.3L), while the lower FA ratio was found in the P39T67R (Figure 3.3S) genotype. Genotypes did not interact with the canopy portion for either LAA or FA. The LAA abundance in the lower stem section was ~3% greater than the upper stem section, while branches were ~0.5% lower than the entire main stem. The oleic/(linoleic + linolenic) ratio was slightly lower in the branches (without separation on the means comparison test), but the upper main stem surpassed the lower segment by ~30%. Because protein decreased from top to bottom main stem segments while LAA increased, the vertical gradient of protein concentration and protein quality was opposite, similar to the oil concentration and quality (measured as the FA ratio).

Branch-Yield Contribution Affects Seed Composition

The whole plant yield was positively associated with the branch-yield contribution (Figure 3.1B), increasing about 30 kg/ha when branch contribution increased by 1%. Due to the yield-branch significant relationship, a simple linear regression exploring the effect of branches on seed components would likely confound the two factors. Therefore, protein, oil, LAA

concentration, and the oleic/(linoleic + linolenic) ratio were modeled as a function of both variables. Branches had lower oil concentration, LAA abundance, and FA ratio than the main stem (Figure 3.3). However, when there was more yield coming from branches, the FA ratio was not affected, while oil concentration and LAA abundance increased at the whole plant level (Table 3.2). For protein, neither yield nor branches had significant slopes. Only yield was found to have a negative relationship with the FA ratio, with more yield meaning poor oil quality. However, for oil concentration and LAA abundance, branch-yield contribution alone was related to greater values, with positive slopes of 0.59 and 0.01, respectively. Considering branch contribution to yield ranged from ~10 to 50%, oil concentration was predicted to increase 23.6 g/kg from low to high branch-yielding conditions. Although the LAA had a smaller rate of change, it represented 0.4 g/100 g protein, matching the overall genotype range presented in Figure 3.2. The ANOVA coefficients are shown in Supplementary Table 2.

Discussion

Results from this study showed the importance of branch-yield contribution for soybean seed composition, expanding previous findings on protein and oil vertical gradient in the main stem. Additionally, this study provides novel analysis, including a characterization of the abundance of limiting AA and oleic/(linoleic + linolenic) ratio, as parameters of protein and oil quality. We acknowledge some limitations of this study, such as the limited number of tested genotypes and collinearity between yield and branch-yield contribution. However, manipulating branches under field conditions while attaining comparable seed yield is difficult and highly sensitive to genotype, environment, and management (G × E × M) interactions.

A vertical gradient of seed protein and oil concentration was expected. Escalante and Wilcox (1993a) found more protein in the upper main stem (~40 g/kg) than in the lower main stem. Additionally, the same authors documented genotypes with contrasting protein concentration increased ~7 g/kg/node from lower to upper main stem (Escalante and Wilcox, 1993b). Bellaloui and Gillen (2010) found more protein and less oil concentrations in the upper nodes, with differences attributed to changes in genotype and light distribution within the canopy. Sharma et al. (2013) observed the same protein-oil vertical gradient going beyond physiological maturity and affecting composition during storage. Our results confirm the main stem vertical gradient widely reported, with protein concentration decreasing (~9%) and oil concentration increasing (~3%) both from top to bottom main stem nodes. However, greater oil concentration in the lower nodes was not associated with greater seed oil content but with a proportionally greater reduction in protein content than both oil and residue compounds. Protein was the seed component with the greatest variation among canopy portions, pointing to protein accumulation as a critical process determining the concentration of others seed components within the plant.

Assimilate supply and ontogenesis temporal variation might be the complementary factors defining protein and oil content among nodes. Our genotypes presented slightly greater protein content in the branches than the main stem and greater seed size in the upper nodes than the lower nodes. Indeterminate soybean genotypes start setting pods from lower to upper nodes (Egli and Bruening, 2006a) and later in branches than the main stem (Munier-Jolain et al., 1994). However, a possibly shorter seed-filling period might be compensated by a greater accumulation rate (Egli et al., 1978), since seeds from upper nodes are not necessarily smaller (Parvej et al., 2016), and there is no clear association between timing of fruit initiation and seed size (Egli,

2012). Greater seed size might be connected to CO_2 assimilation and N concentration in upper leaves (Boon et al., 1983), as the protein content in the branches could be driven by greater light exposure relative to the lower main stem. On the other hand, considering protein is accumulated before oil (Poeta et al., 2014), a shorter seed-filling duration in branches and upper nodes could limit the oil deposition, affecting its overall final accumulation (evidenced by a relative stability of oil compared to protein).

Greater protein quality (LAA abundance) was found in the lower main stem, while greater oil quality (FA ratio) was found in the upper main stem. These differences might be related to microclimatic canopy changes, especially temperature and light (both are greater in the upper canopy). Higher temperatures promote the synthesis of oleic acid at the cost of linoleic and linolenic (Wolf et al., 1982). Because LAAs are less dependent on carbon supply, their abundance is maintained or increased under shading conditions (Pfarr et al., 2018). Within the plant, low mobility of sulfur could enhance the differences on LAA along the main stem (Sexton et al., 2002), since two out of the five LAAs are rich in sulfur (cysteine and methionine). Regarding branches, our results confirm the expected FA trend, with a lower oleic/(linoleic + linolenic) ratio, possibly due to reduced temperature compared to the upper main stem. However, branches presented a smaller LAA abundance relative to the entire main stem, possibly indicating greater light exposure than the lower main stem nodes, decreasing the abundance of LAA (Pfarr et al., 2018).

Our results point to branch yield as an underlying factor of seed composition. It is possible that a greater branch yield favors the accumulation of oil and LAA at the whole plant level, even though branch seeds have a lower concentration of those components. The importance of branch yield for seed yield formation and stability has been highlighted by

47

Carpenter and Board (1997) and Suhre et al. (2014), but information on seed quality changes was lacking. From our study, we demonstrated that more branch yield drove the whole plant seed composition toward the lower main stem characteristics (high oil concentration and LAA abundance); however, the overall branch seed composition is not similar to the lower main stem. Changes in plant density and arrangement could affect the branch length and microclimate within the canopy, making branch seeds similar to either lower or middle-upper main stem segments. However, under a reduced row spacing (0.45 m) branch seed composition was consistent with our results (Werner et al., 2021).

Changing the yield allocation among canopy portions could be manipulated by plant breeding and management practices. For instance, enhanced branching in modern genotypes could be related to stable LAA concentration despite protein decay over the last 40 years (de Borja Reis et al., 2020). Future research must consider to study the branching process (initiation, progress rate and duration, dry matter, and harvest index) as a potential moderator of soybean seed composition, plausibly benefitting human nutrition with improved AA profile and oil concentration. Findings of this study denote two relevant points from a breeding standpoint: (i) expanding assessment on seed quality to include not only protein and oil but also AA and FA and (ii) acknowledging the contribution of the breeding process on branching, exploiting genetic variation on this trait to assess potential improvements in seed quality.

Conclusions

In conclusion, soybean seeds from the upper main stem segment accumulated more protein and oil than the lower main stem segment. However, the upper main stem section presented greater protein and lower oil concentrations relative to the lower main stem section of

this segment. Across genotypes with contrasting branch yield, seeds from the branches presented similar protein concentration as the main stem but lower oil concentration, oleic/(linoleic + linolenic) ratio, and LAA abundance. However, branch-yield contribution was related to greater oil concentration and LAA abundance across genotypes. This study highlights the importance of improving our knowledge on yield contribution and seed composition from different canopy portions for benefitting food production and soybean markets.

Acknowledgments

References

Agudamu, Yoshihira, T., Shiraiwa, T., 2016. Branch development responses to planting density and yield stability in soybean cultivars. Plant Prod. Sci. 19, 331–339. https://doi.org/10.1080/1343943X.2016.1157443

Assefa, Y., Purcell, L.C., Salmeron, M., Naeve, S., Casteel, S.N., Kovács, P., Archontoulis, S., Licht, M., Below, F., Kandel, H., Lindsey, L.E., Gaska, J., Conley, S., Shapiro, C., Orlowski, J.M., Golden, B.R., Kaur, G., Singh, M., Thelen, K., Laurenz, R., Davidson, D., Ciampitti, I.A., 2019. Assessing variation in us soybean seed composition (protein and oil). Front. Plant Sci. 10. https://doi.org/10.3389/fpls.2019.00298

Baldocchi, D.D., Verma, S.B., Rosenberg, N.J., 1983. Microclimate in the soybean canopy. Agric. Meteorol. 28, 321–337. https://doi.org/10.1016/0002-1571(83)90009-2

Baldocchi, D.D., Verma, S.B., Rosenberg, N.J., Blad, B.L., Specht, J.E., 1985. Microclimate-plant architectural interactions: Influence of leaf width on the mass and energy exchange of a soybean canopy. Agric. For. Meteorol. 35, 1–20. https://doi.org/10.1016/0168-1923(85)90070-X

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects Models Using {lme4}. J. Stat. Softw. 67, 1–48. https://doi.org/10.18637/jss.v067.i01

Bellaloui, N., Gillen, A.M., 2010. Soybean seed protein, oil, fatty acids, N, and S partitioning as affected by node position and cultivar differences. Agric. Sci. 01, 110–118. https://doi.org/10.4236/as.2010.13014

Bennett, J.O., Krishnan, A.H., Wiebold, W.J., Krishnan, H.B., 2003. Positional Effect on Protein and Oil Content and Composition of Soybeans. J. Agric. Food Chem. 51, 6882–6886. https://doi.org/10.1021/jf0343711

Boon-Long, P., Egli, D.B., Leggett, J.E., 1983. Leaf N and Photosynthesis during Reproductive Growth in Soybeans 1. Crop Sci. 23, 617–620. https://doi.org/10.2135/cropsci1983.0011183x002300040005x

Bronikowski, A., Webb, C., 1996. Appendix: A critical examination of rainfall variability measures used in behavioral ecology studies. Behav. Ecol. Sociobiol. 39, 27–30. https://doi.org/10.1007/s002650050263

Brumm, T.J., Hurburgh, C.R., 2006. Changes in long-term soybean compositional patterns. JAOCS, J. Am. Oil Chem. Soc. 83, 981–983. https://doi.org/10.1007/s11746-006-5056-4

Carpenter, A.C., Board, J.E., 1997. Branch yield components controlling soybean yield stability across plant populations. Crop Sci. 37, 885–891. https://doi.org/10.2135/cropsci1997.0011183X003700030031x

Carrera, C.S., Dardanelli, J.L., 2017. Water deficit modulates the relationship between temperature and unsaturated fatty acid profile in soybean seed oil. Crop Sci. 57, 3179–3189. https://doi.org/10.2135/cropsci2017.04.0214

Collins, F.I., Cartter, J.L., 1956. Variability in Chemical Composition of Seed From Different Portions of the Soybean Plant 1. Agron. J. 48, 216–219. https://doi.org/10.2134/agronj1956.00021962004800050006x

de Borja Reis, A.F., Tamagno, S., Moro Rosso, L.H., Ortez, O.A., Naeve, S., Ciampitti, I.A., 2020. Historical trend on seed amino acid concentration does not follow protein changes in soybeans. Sci. Rep. 10, 17707. https://doi.org/10.1038/s41598-020-74734-1

Egli, D.B., 2012. Timing of Fruit Initiation and Seed Size in Soybean. J. Crop Improv. 26, 751–766. https://doi.org/10.1080/15427528.2012.666784

Egli, D.B., Bruening, W.P., 2006a. Fruit development and reproductive survival in soybean: Position and age effects. F. Crop. Res. 98, 195–202. https://doi.org/10.1016/j.fcr.2006.01.005

Egli, D.B., Bruening, W.P., 2006b. Temporal profiles of pod production and pod set in soybean. Eur. J. Agron. 24, 11–18. https://doi.org/10.1016/j.eja.2005.04.006

Egli, D.B., Leggett, J.E., Wood, J.M., 1978. Influence of Soybean Seed Size and Position on the Rate and Duration of Filling. Agron. J. 70, 127–130. https://doi.org/10.2134/agronj1978.00021962007000010029x

Escalante, E.E., Wilcox, J.R., 1993a. Variation in seed protein among nodes of determinate and indeterminate soybean near-isolines. Crop Sci. 33, 1166–1168. https://doi.org/10.2135/cropsci1993.0011183X003300060012x

Escalante, E.E., Wilcox, J.R., 1993b. Variation in seed protein among nodes of normal- and high-protein soybean genotypes. Crop Sci. 33, 1164–1166. https://doi.org/10.2135/cropsci1993.0011183X003300060011x

FAO, 2021. FAOSTAT Online Database [WWW Document]. URL http://faostat.fao.org/

Fehr, W.R., Caviness, C.E., 1977. Stages of Soybean Development. Spec. Rep. 80, 11.

Fox, J., Weisberg, S., 2019. An *R* Companion to Applied Regression, Third. ed. Sage, Thousand Oaks {CA}.

Gao, J., Hao, X., Thelen, K.D., Robertson, G.P., 2009. Agronomic management system and precipitation effects on soybean oil and fatty acid profiles. Crop Sci. 49, 1049–1057. https://doi.org/10.2135/cropsci2008.08.0497

Gerde, J.A., White, P.J., 2008. Lipids, in: Johnson, L.A., White, P.J., Galloway, R. (Eds.), Soybeans: Chemistry, Production, Processing, and Utilization. AOCS Press, Ames, IA, pp. 193–227. https://doi.org/10.1016/B978-1-893997-64-6.50010-X

Honigs, D.E., Hieftje, G.M., Mark, H.L., Hirschfeld, T.B., 1985. Unique-Sample Selection via Near-Infrared Spectral Subtraction. Anal. Chem. 57, 2299–2303. https://doi.org/10.1021/ac00289a029

Hothorn, T., Bretz, F., Westfall, P., 2008. Simultaneous Inference in General Parametric Models. Biometrical J. 50, 346–363.

Huber, S.C., Li, K., Nelson, R., Ulanov, A., DeMuro, C.M., Baxter, I., 2016. Canopy position has a profound effect on soybean seed composition. PeerJ 2016, e2452. https://doi.org/10.7717/peerj.2452

Joint FAO/WHO/UNU Expert Consultation on Protein and Amino Acid Requirements in Human Nutrition (2002 : Geneva Switzerland) Food and Agriculture Organization of the United Nations World Health Organization & United Nations University, 2007. Protein and amino acid requirements in human nutrition : report of a joint FAO/WHO/UNU expert consultation. WHO technical report series ; no. 935.

Koller, H.R., 1972. Leaf Area-Leaf Weight Relationships in the Soybean Canopy 1. Crop Sci. 12, 180–183. https://doi.org/10.2135/cropsci1972.0011183x001200020007x

Köppen, W., 2011. The thermal zones of the Earth according to the duration of hot, moderate and cold periods and to the impact of heat on the organic world. Meteorol. Zeitschrift 20, 351–360. https://doi.org/10.1127/0941-2948/2011/105

Lamichhane, J.R., Constantin, J., Schoving, C., Maury, P., Debaeke, P., Aubertot, J.N., Dürr, C., 2020. Analysis of soybean germination, emergence, and prediction of a possible northward establishment of the crop under climate change. Eur. J. Agron. 113, 125972. https://doi.org/10.1016/j.eja.2019.125972

Munier-Jolain, N.G., Ney, B., Duthion, C., 1994. Reproductive development of an indeterminate soybean as affected by morphological position. Crop Sci. 34, 1009–1013. https://doi.org/10.2135/cropsci1994.0011183X003400040033x

Parvej, M.R., Slaton, N.A., Purcell, L.C., Roberts, T.L., 2016. Soybean yield components and seed potassium concentration responses among nodes to potassium fertility. Agron. J. 108, 854–863. https://doi.org/10.2134/agronj2015.0353

Pfarr, M.D., Kazula, M.J., Miller-Garvin, J.E., Naeve, S.L., 2018. Amino acid balance is affected by protein concentration in soybean. Crop Sci. 58, 2050–2062. https://doi.org/10.2135/cropsci2017.11.0703

Poeta, F.B., Rotundo, J.L., Borrás, L., Westgate, M.E., 2014. Seed water concentration and accumulation of protein and oil in soybean seeds. Crop Sci. 54, 2752–2759. https://doi.org/10.2135/cropsci2014.03.0204

R Core Team, 2020. R: A Language and Environment for Statistical Computing.

Rincker, K., Nelson, R., Specht, J., Sleper, D., Cary, T., Cianzio, S.R., Casteel, S., Conley, S., Chen, P., Davis, V., Fox, C., Graef, G., Godsey, C., Holshouser, D., Jiang, G.-L., Kantartzi, S.K., Kenworthy, W., Lee, C., Mian, R., McHale, L., Naeve, S., Orf, J., Poysa, V., Schapaugh, W., Shannon, G., Uniatowski, R., Wang, D., Diers, B., 2014. Genetic Improvement of U.S. Soybean in Maturity Groups II, III, and IV. Crop Sci. 54, 1419–1432. https://doi.org/10.2135/cropsci2013.10.0665

Rotundo, J.L., Borrás, L., Westgate, M.E., Orf, J.H., 2009. Relationship between assimilate supply per seed during seed filling and soybean seed composition. F. Crop. Res. 112, 90–96. https://doi.org/10.1016/j.fcr.2009.02.004

Rotundo, J.L., Miller-Garvin, J.E., Naeve, S.L., 2016. Regional and temporal variation in soybean seed protein and oil across the United States. Crop Sci. 56, 797–808. https://doi.org/10.2135/cropsci2015.06.0394

Rotundo, J.L., Westgate, M.E., 2009. Meta-analysis of environmental effects on soybean seed composition. F. Crop. Res. 110, 147–156. https://doi.org/10.1016/j.fcr.2008.07.012

Sexton, P.J., Paek, N.C., Naeve, S.L., Shibles, R.M., 2002. Sulfur metabolism and protein quality of soybean. J. Crop Prod. 5, 285–308. https://doi.org/10.1300/J144v05n01_12

Sharma, S., Kaur, A., Bansal, A., Gill, B.S., 2013. Positional effects on soybean seed composition during storage. J. Food Sci. Technol. 50, 353–359. https://doi.org/10.1007/s13197-011-0341-0

Shim, S., Ha, J., Kim, M.Y., Choi, M.S., Kang, S.T., Jeong, S.C., Moon, J.K., Lee, S.H., 2019. GmBRC1 is a candidate gene for branching in soybean (*Glycine max* (L.) Merrill). Int. J. Mol. Sci. 20, 135. https://doi.org/10.3390/ijms20010135

Shim, S., Kim, M.Y., Ha, J., Lee, Y.H., Lee, S.H., 2017. Identification of QTLs for branching in soybean (*Glycine max* (L.) Merrill). Euphytica 213, 225. https://doi.org/10.1007/s10681-017-2016-z

Sinclair, T.R., De Wit, C.T., 1975. Photosynthate and nitrogen requirements for seed production by various crops. Science (80-.). 189, 565–567. https://doi.org/10.1126/science.189.4202.565

Stephenson, R.A., Wilson, G.L., 1977. Patterns of assimilate distribution in soybeans at maturity. I. The influence of reproductive development stage and leaf position. Aust. J. Agric. Res. 28, 203–209. https://doi.org/10.1071/AR9770203

Suhre, J.J., Weidenbenner, N.H., Rowntree, S.C., Wilson, E.W., Naeve, S.L., Conley, S.P., Casteel, S.N., Diers, B.W., Esker, P.D., Specht, J.E., Esker, P.D., 2014. Soybean yield partitioning changes revealed by genetic gain and seeding rate interactions. Agron. J. 106, 1631–1642. https://doi.org/10.2134/agronj14.0003

Thakur, M., Hurburgh, C.R., 2007. Quality of US soybean meal compared to the quality of soybean meal from other origins. JAOCS, J. Am. Oil Chem. Soc. 84, 835–843. https://doi.org/10.1007/s11746-007-1107-8

Thornton, M.M., Thornton, P.E., Wei, Y., Mayer, B.W., Cook, R.B., Vose, R.S., 2017. Daymet: Monthly Climate Summaries on a 1-km Grid for North America, Version 3. Oak Ridge Natl. Lab. https://doi.org/10.3334/ORNLDAAC/1855

Toyota, M., Maitree, L., Chomsang, K., 2017. Changes in radiation interception and R:FR over time and with canopy depth of two soybean cultivars with different branching characteristics. Plant Prod. Sci. 20, 205–214. https://doi.org/10.1080/1343943X.2017.1294464

Werner, F., de Aguiar e Silva, M.A., Ferreira, A.S., Zucareli, C., Balbinot, A.A., 2021. Grain, oil, and protein production on soybean stems and branches under reduced densities. Rev. Bras. Ciencias Agrar. 16, 1–9. https://doi.org/10.5039/AGRARIA.V16I1A7439

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wolf, R.B., Cavins, J.F., Kleiman, R., Black, L.T., 1982. Effect of temperature on soybean seed constituents: Oil, protein, moisture, fatty acids, amino acids and sugars. J. Am. Oil Chem. Soc. 59, 230–232. https://doi.org/10.1007/BF02582182
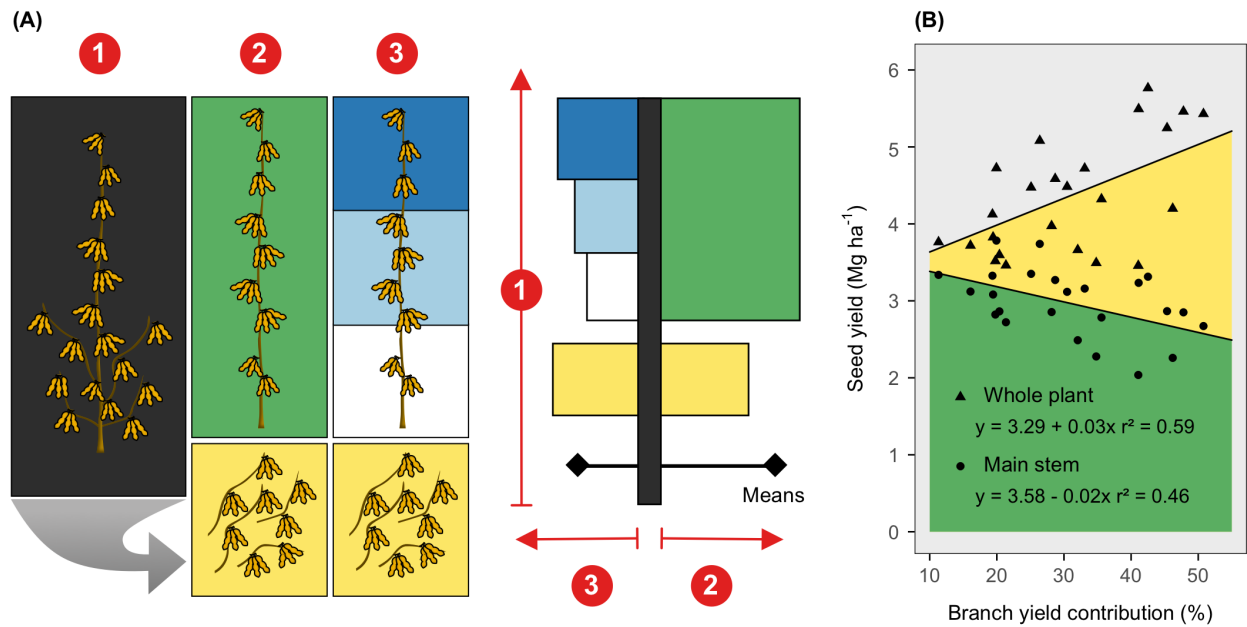
Figure 3.1. (A) Soybean seed harvesting and bar chart representation of observed variables at three canopy levels: (1) whole plant, (2) entire main stem (green color) and stem branches (yellow color), and (3) upper (dark blue), middle (light blue), lower (white color) main stem segments and the branches. The bar chart representation was meant to resemble a soybean plant, concisely depicting the linear mixed models testing genotype and canopy portion. (B) Seed yield from the main stem (green) and branches (yellow color) relative to the branch-yield contribution (%). Regression lines were fit across genotypes and considered a random intercept for site-year.
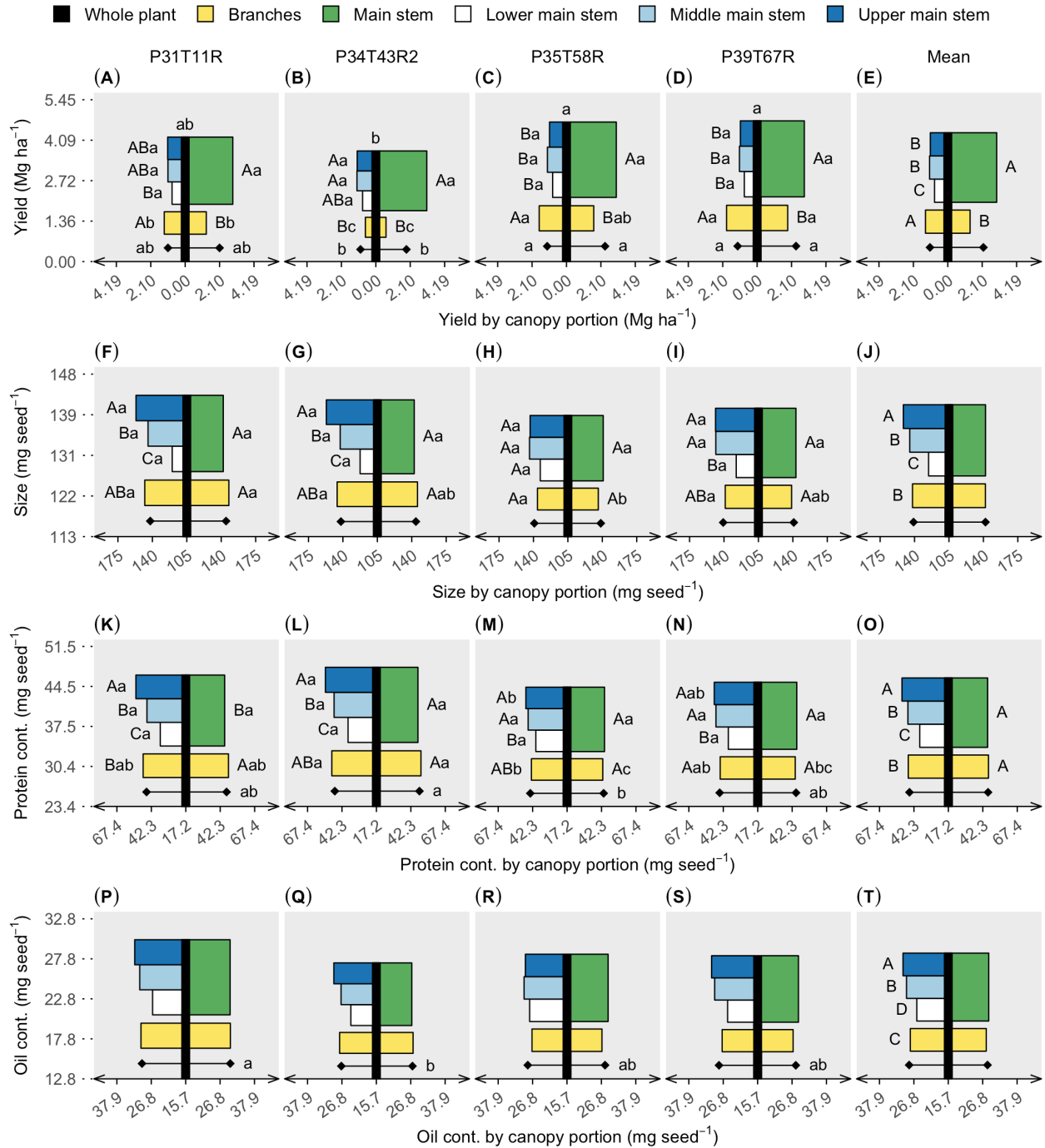
Figure 3.2. Soybean seed protein (A–E) and oil concentration (F–J), limiting amino acids (LAAs) abundance (K–O), and oleic/(linoleic + linolenic) ratio (P–T). Vertical black bars refer to the whole plant data in the y axis, with lowercase letters on top comparing genotypes (model 1). Horizontal bars are centered on the black bar (x axis), referring to two canopy portions on the right side (main stem and branches) (model 2) and four canopy portions on the left side (lower, middle, and upper main stem segments and branches) (model 3). Diamonds represent the genotype mean of all canopy portions for models 2 and 3. Uppercase letters compare stem segments within genotype (interaction) or on the overall mean (portion effect). Lowercase letters

compare genotypes within canopy portions (interaction) or on the genotype mean, diamonds (genotype effect). Each panel row portraits three linear models of a response variable, for the whole plant, two and four canopy portions. Absence of letters represents no significant difference ($p < 0.05$).
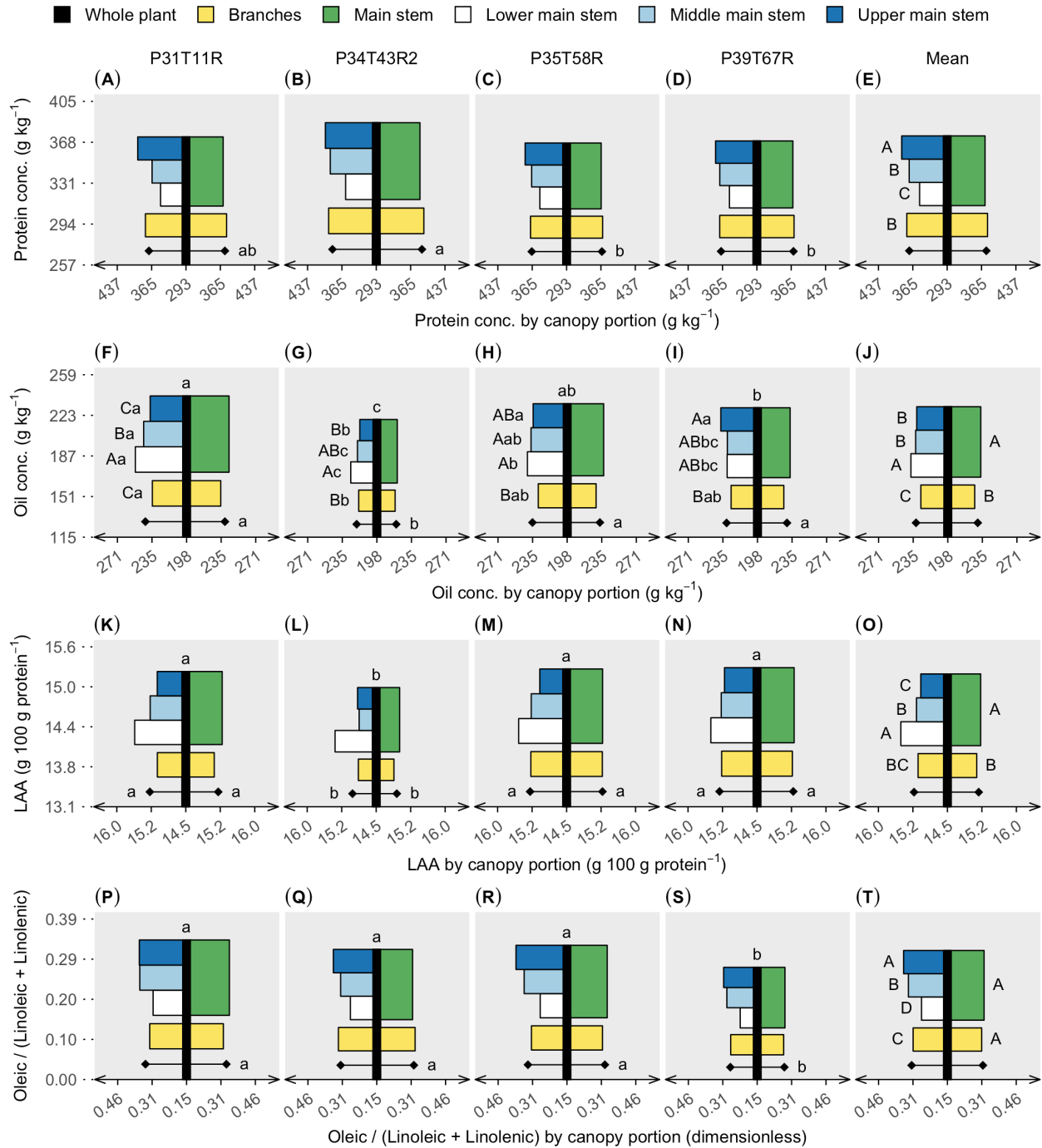
Figure 3.3. Soybean seed protein (A–E) and oil concentration (F–J), limiting amino acids (LAAs) abundance (K–O), and oleic/(linoleic + linolenic) ratio (P–T). Vertical black bars refer to the whole plant data, in the y axis, with lowercase letters comparing genotypes. Horizontal bars are centered on the black bar (x axis), referring to two canopy portions on the right side (main stem and branches) and four canopy portions on the left side (lower, middle, and upper main stem segments and branches). Diamonds represent canopy portion means on each side. Uppercase letters compare canopy portions within genotype (significant interaction) or on the overall mean (canopy portions single effect). Lowercase letters compare genotypes within

58

canopy portions (interaction) or on the genotype mean, diamonds (genotype single effect). The absence of letters represents no significant difference in the analysis of variance (p < 0.05).

Tables

Table 3.1. Soil and weather variables characterizing Ashland Bottoms 2018 and Rossville 2019 experimental sites.

Variable	Ashland 2018	Rossville 2019
Soil variables		
Soil texture, g kg^{-1}		
Clay	18.0	17.3
Sand	28.0	30.0
Silt	54.0	52.7
Water pH	7.6	7.0
SOM, g kg^{-1} [a]	21.0	15.0
NO$_3$, mg dm^{-3}	4.0	2.7
SO$_4$, mg dm^{-3}	1.0	2.3
P, mg dm^{-3} [b]	90.2	43.0
Weather variables		
Mean temperature, °C	25.7	24.3
Maximum temperature, °C	32.2	29.7
Minimum temperature, °C	19.1	18.9
Solar radiation, MJ m^{-2} day^{-1}	1687	1296
Evapotranspiration, mm [c]	754	474
Rainfall precipitation, mm	338	518
Precipitation SDI [d]	0.59	0.65
Relative humidity, %	59.7	77.5
Mean VPD, kPa [e]	1.65	0.73

[a] Soil organic matter via loss on ignition (LOI); [b] Phosphorus extracted by Mehlich-3; [c] Reference evapotranspiration; [d] Shannon-diversity index (Bronikowski and Webb, 1996); [e] Vapor pressure deficit.

Table 3.2. Whole plant soybean seed protein, oil, and limiting amino acids (LAAs) concentration, and oleic/(linoleic + linolenic) ratio as a function of whole plant yield (Mg/ha) and branch contribution to the whole plant yield (%).

Parameter	Estimate	Standard error	p-value	Back transformed slope estimate
Protein, g kg^{-1}				
Intercept	373.6	11.65	2.21e-02 *	
Whole plant yield	-0.02	3.53	9.96e-01	-0.02
Branch contribution	-5.74	3.39	1.06e-01	-0.51
Oil, g kg^{-1}				
Intercept	231.7	5.17	1.95e-02 *	
Whole plant yield	-3.79	2.58	1.56e-01	-5.10
Branch contribution	6.63	2.49	1.48e-02 *	0.59
LAA, g 100 g protein^{-1}				
Intercept	15.21	0.35	1.50e-02 *	
Whole plant yield	0.02	0.05	6.98e-01	0.02
Branch contribution	0.10	0.04	3.77e-02 *	0.01
Oleic / (Linoleic + Linolenic)				
Intercept	0.31	0.02	5.62e-02	
Whole plant yield	-0.02	0.01	2.55e-02 *	-0.03
Branch contribution	0.01	0.01	4.23e-01	0.00

* Estimate is significant at the 0.05 probability level ($p < 0.05$).

Chapter 4 - Conclusions and Future Implications

The present study expanded the current understanding of the process of allocation of seed quality components within the soybean canopy and how yield and protein formation (SNF) is temporally affected by changes in SNF and soil N supply in the forms of NO_3 and NH_4. Regarding the temporal dynamics of N sources, the use of non-linear models to describe NO_3 and NH_4 availability along with Ndfa was valuable to undercover their overall relationship. Remarkably, even though NH_4 levels were smaller than NO_3, the total NH_4 exposure was more suppressive of Ndfa at physiological maturity than NO_3. However, because NH_4 did not present a strong and clear peak during the season in both site-years, we could only explore the temporal effect of NO_3 on N uptake and SNF. A mid-season NO_3 peak matched the timing of maximum N uptake from the soil supply, which did not benefit N uptake rates from soil but delayed and suppressed the fixed N by approximately 20 days. In an attempt to simplify the seasonal modeling of SNF, dry matter throughout the season was adopted as a predictor of fixed N instead of days after emergence (DAE). Although this approach resonates with the broadly characterized N dilution process, it overlooks seasonal variation on Ndfa and exposes difficulties for predicting fixed N across environments. For instance, our slope-only model (N requirement) had a change around 5 Mg ha^{-1} of dry matter in the pooled data for all genotypes and site-years. However, despite of the limitations, the approach presented here is valuable to guide future studies looking at temporal variations of Ndfa. Considering the rationale of crop growth models that follows the hierarchy of 1) environmental and management conditions modulating changes on dry matter, and 2) dry matter moderating N uptake, it would be reasonable to propose a third step, focusing on N uptake regulating fixed N. Therefore, the challenge becomes to predict or assess an Ndfa value that could be applicable throughout the entire season. According to our results, at least two

assessments would be needed, before and after 5 Mg ha^{-1} dry matter, which might need to be revisited under high yield conditions (yield levels > 4 Mg ha^{-1}).

Regarding the internal variation on soybean seed composition, a previously reported vertical profile was confirmed for protein and oil concentration. However, protein accumulation was more responsive to the vertical position than oil and residuals, causing variations in concentration in those three components. For instance, the upper canopy accumulates more protein in the seeds and has a large protein concentration than the bottom, which has less accumulation of protein and therefore a higher oil concentration. Furthermore, high concentration of either protein or oil is accompanied by a poor quality of the respective components inside of the plant. Seeds with high protein concentration in the upper nodes display a smaller abundance of limiting AA, because the higher carbon supply (light) from the upper leaves favors the production of non-limiting AA such as glutamic and aspartic acids. On the other hand, seeds in the bottom of the canopy produce less oleic acid compared to linoleic and linolenic, which is not desirable in the industry considering shelf life and heat stability. Most importantly, this is one of the few reports considering the seed yield from branches, which was found with lower oil concentration, low oleic / (linoleic + linolenic) ratio, and low limiting AA abundance. However, across the tested genotypes, a greater branch yield contribution was positively related with limiting AA and oil concentration at the whole canopy level.

From the seed composition standpoint, the present research has implications for researchers, breeders, producers, and consumers. Considering plant architecture and possible yield allocation within the canopy as moderator of seed composition, future research should explore the impact of branching on the overall effect of other management practices on protein, oil, AA, and FA. Furthermore, soybean breeders must be aware of potential unintended effects

on seed composition coming from genotypes more prone to branching. Because canopy portion separation during harvest is currently unpractical with modern machinery, it is unlikely farmers and industry could take advantage of the internal variation directly, but new technologies will be able to segregate seed composition at the field level (by regions). From the SNF standpoint, this study highlights the importance of Ndfa assessment during the season to undercover a strong effect of soil N availability on N uptake. In addition, a simplified approach for predicting seasonal fixed-N should be further explored to indicate the advantages but also the limitations of predictions across environments, which must be sustained by at least two Ndfa assessments during the crop growing season.