UAV-based proximal remote sensing for nitrogen stress detection and grain yield and protein prediction in winter wheat

by

Kaden Spencer

B.S., Brigham Young University–Idaho, 2024

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

2026

Approved by:

Major Professor
Deepak R. Joshi

Copyright

© Kaden Spencer 2026.

Abstract

Nitrogen (N) management influences both grain yield and grain protein concentration in hard red winter wheat, but crop response to fertilizer N varies across environments. Unmanned aerial vehicle (UAV) multispectral imagery can monitor crop status during the growing season. Its usefulness for N management and pre-harvest prediction, however, depends on the environmental context and on whether models transfer to new locations. This thesis evaluated UAV multispectral sensing of winter wheat across seven Kansas environments during the 2024–2025 growing season. At each location, eight cultivars were grown under seven N rates (0 to 202 kg N ha^{-1}) in a split-plot design. Imagery was collected at up to six growth stages and summarized into eight vegetation indices across three developmental stages: tillering (Feekes 3–5), stem elongation (Feekes 6–9), and grain formation (Feekes 10–11.4).

The first study evaluated whether grain yield and protein could be predicted at locations not used for model training, using nested leave-one-location-out validation. Partial least squares regression (PLSR) outperformed random forest and XGBoost. GNDVI at stem elongation and N rate were the dominant yield predictors. MTCI at grain formation, N rate, and plant height were the most important protein predictors. The final PLSR models achieved $R^2 = 0.55$, concordance correlation coefficient (CCC) = 0.72, and RMSE = 516.4 kg ha^{-1} for grain yield, and $R^2 = 0.47$, CCC = 0.67, and RMSE = 1.09 percentage points for grain protein. Removing cultivar information reduced protein R^2 from 0.47 to 0.14.

The second study evaluated UAV detectability of fertilizer-N response across environments with contrasting background soil N supply. Profile soil NO_3-N ranged from 29 to 270 kg N ha^{-1}, and grain-yield response to fertilizer N ranged from −0.7% to 105%. N

detectability increased with crop development. Red-edge-based indices were more sensitive than NDVI during stem elongation (Feekes 6–9). Greater soil nitrate tended to correspond with weaker yield response and weaker red-edge detectability ($\rho = -0.68$, $P = 0.094$). Red-edge detectability ranked environments in the same order as relative yield response ($\rho = 1.00$). Normalizing vegetation indices to a local N-rich reference strengthened their relationship with yield response; for NDVI, $|\rho|$ increased from 0.36–0.79 to 0.89–0.96.

Together, these results show that UAV multispectral imagery provides useful information for winter wheat N management and pre-harvest yield and protein prediction. That information was most reliable when interpreted alongside crop developmental stage, local soil N supply, an N-rich reference, and known management and cultivar information.

Keywords: precision nitrogen management, leave-one-location-out validation, nitrogen rich sufficiency index, grain protein concentration, hard red winter wheat.

Table of Contents

List of Figures viii
List of Tables xi
Acknowledgements xii
Chapter 1 - Introduction 1
Background 1
UAV Multispectral Sensing of Winter Wheat 2
Interpreting UAV Data Across Environments 4
Research Objectives 5
Thesis Organization 6
References 6
Chapter 2 - Transferable Prediction of Winter Wheat Grain Yield and Protein Using UAV Multispectral Imagery and Agronomic Variables 11
Abstract 11
1 Introduction 12
2 Materials and Methods 15
2.1 Study Sites and Environmental Context 15
2.2 Experimental Design and Treatments 16
2.3 Plant Height, Biomass, Grain Yield and Protein Measurement 18
2.4 UAV Imagery Acquisition 18
2.5 Radiometric Calibration and Image Processing 19
2.6 Soil Background Removal and Plot-Level Reflectance Extraction 20
2.7 Vegetation Index Calculation 20
2.7.1 Crop Growth Stage Classification 22
2.7.2 Exploratory Statistical Analysis 23
2.8 Machine Learning Model Development and Evaluation 24
2.8.1 Response Variables, Predictors, and Preprocessing 24
2.8.2 Nested Leave-One-Location-Out Validation and Predictor Selection 25
2.8.3 Machine-Learning Algorithms and Hyperparameter Tuning 26
2.8.4 Model Evaluation Metrics 27

2.9 Predictor Importance and Genotype Representation 29
3 Results 31
3.1 Variation in Grain Yield and Protein Across Locations 31
3.2 Growth-Stage-Specific Spectral Relationships 32
3.3 Relationships with Management and Crop Structural Traits 33
3.4 Selection of Transferable UAV and Genotype Predictors 34
3.5 Cross-Location Prediction and Algorithm Comparison 35
3.6 Predictor Importance and Genotype Representation 36
4 Discussion 39
5 Limitations and Future Research 45
6 Conclusion 46
References 47
Chapter 3 - UAV-Based Nitrogen Stress Detectability in Winter Wheat Across Contrasting Soil Nitrogen Environments 52
Abstract 52
1 Introduction 53
2 Materials and Methods 56
2.1 Study Sites and Experimental Design 56
2.2 Soil N Supply and Agronomic N Response 58
2.3 UAV Multispectral Imagery and Vegetation Indices 59
2.4 Crop Growth Stage Classification 61
2.5 Standardized Spectral N Detectability 62
2.6 Comparison of Vegetation Index Sensitivity 63
2.7 N-Rich Sufficiency Index 64
2.8 Cross-Environment Relationships and Sensitivity Analyses 66
3 Results 68
3.1 Background Soil N Supply and Grain Yield and Protein Responses to Fertilizer Nitrogen 68
3.2 Vegetation Index Sensitivity to Fertilizer N Across Crop Developmental Stages 69
3.3 Background Soil N, Agronomic Response, and Cross-Location Variation in Spectral N Detectability 72

3.4 Agronomic Relevance of Spectral N Detectability and N-Rich Normalization 74
4 Discussion 78
4.1 Stage- and Index-Specific Sensitivity to Fertilizer N 78
4.2 Soil N Supply and Agronomic Responsiveness Associated with UAV-Based N Detectability 80
4.3 Local N-Rich Normalization and Implications for Precision N Management 81
5 Limitations and Future Research 81
6 Conclusion 83
References 83
Chapter 4 - Conclusions 89
Summary of Findings 89
Synthesis Across Chapters 90
Implications for Nitrogen Management 91
Limitations and Future Research 91
Appendix A - Supplementary Figures for Chapter 3 93

List of Figures

Figure 2.1. Distribution of (A) grain yield and (B) grain protein across seven Kansas winter wheat environments, and (C) the pooled relationship between grain yield and grain protein. In A and B, half-violin distributions show plot-level observations, boxplots show the median and interquartile range, and black diamonds indicate location means. In C, point color represents N rate and the solid line is the fitted linear relationship. 31

Figure 2.2. Pooled Pearson correlations between stage-averaged UAV vegetation indices and (A) grain yield and (B) grain protein concentration across the tillering, stem elongation, and grain formation stages. Correlations are descriptive and based on pooled plot-level observations. * FDR-adjusted $q < 0.05$; ** $q < 0.01$. 33

Figure 2.3. Pooled Pearson correlations of N rate, plant height, and anthesis biomass with grain yield and grain protein concentration. * FDR-adjusted $q < 0.05$; ** $q < 0.01$. 34

Figure 2.4. Observed versus leave-one-location-out predicted values for the final PLSR models: (A) grain yield and (B) grain protein. Each observation was predicted when its corresponding location was completely excluded from model training. The dashed line represents the 1:1 relationship. 36

Figure 2.5. Predictor importance and genotype representation in the final partial least squares regression (PLSR) models for grain yield and grain protein. Mean variable importance in projection (VIP) scores are shown for (A) grain yield and (B) grain protein; error bars represent the standard deviation across the seven outer leave-one-location-out (LOLO) folds, and the dashed vertical line indicates VIP = 1. Panels (C) and (D) compare LOLO prediction R^2 among models containing no genotype information, agronomic cultivar traits (yield plasticity, PP; grain-protein tendency, GPC; and maturity), or cultivar identity for grain yield and grain protein, respectively. For grain yield, the trait-based and cultivar-identity models produced similar predictive performance, whereas cultivar identity provided greater predictive value than the three agronomic traits for grain protein. 38

Figure 3.1. Field experiments were conducted during the 2024–2025 winter wheat growing season at seven Kansas locations: Manhattan, Ashland Bottoms, Belleville, Russell, Hutchinson, Hays, and Hoisington. 57

Figure 3.2. Background soil nitrogen supply and grain yield and protein responses to fertilizer N across seven winter wheat environments. (A) Profile soil NO_3-N (0–61 cm) at each location, ordered from lowest to highest background nitrate supply. (B) Mean grain yield at 0 and 202 kg N ha^{-1}, with labels indicating the relative grain-yield response to fertilizer N. (C) Mean grain protein concentration at 0 and 202 kg N ha^{-1}, with labels indicating the absolute grain-protein response expressed in percentage points (pp). 69
Figure 3.3. Vegetation-index sufficiency index (SI) of the 0-N treatment relative to the local N-rich reference (202 kg N ha^{-1}) across seven Kansas winter wheat environments and three developmental stages. 71
Figure 3.4. Mean standardized fertilizer-N detectability of eight UAV-derived vegetation indices across broad winter wheat developmental stages. Tillering, stem elongation, and grain formation correspond to Feekes 3–5, 6–9, and 10–11.4, respectively. Cell values represent the mean standardized N detectability across the seven environments, excluding pre-fertilizer UAV observations. 72
Figure 3.5. Cross-environment relationships among background soil nitrate supply, agronomic response to fertilizer N, and UAV-derived red-edge N detectability across seven winter wheat environments. (A) Relationship between profile soil NO3-N (0–61 cm) and relative grain-yield response to fertilizer N. (B) Relationship between profile soil NO3-N and overall red-edge N detectability, calculated by first summarizing the five red-edge indices (NDRE, CIRE, REDVI, MTCI, and DATT) within each developmental stage and then averaging across the three developmental stages with equal stage weighting. (C) Relationship between fertilizer-N yield response and overall red-edge N detectability. (D) Relationship between grain-protein response and overall detectability of red-edge N. 73
Figure 3.6. Comparison of the relationships between fertilizer-N yield response and absolute vegetation-index values versus N-rich sufficiency-index values for (A) NDVI and (B) NDRE across the tillering, stem elongation, and grain formation stages of winter wheat. Points represent Spearman's rank correlation coefficients across seven environments. More negative values indicate a stronger inverse association with fertilizer-N yield response. 76
Figure 3.7. Association between vegetation-index sufficiency indices and grain-yield response across the full fertilizer-N gradient and winter wheat developmental stages. Cell values represent the median within-location Spearman rank correlation (ρ) between the vegetation-

index sufficiency index and the yield ratio to the 202 kg N ha^{-1} reference across fertilizer-N rates from 0 to 168 kg N ha^{-1}. Values below each correlation coefficient indicate the number of environments with a positive association out of seven total environments. 77

Appendix Figure A.1. Field layout and split-plot experimental design used for the winter wheat nitrogen study. The upper panel shows a representative UAV image of the experimental field, and the lower panel illustrates the treatment arrangement within one replicate block. 93

Appendix Figure A.2. Cultivar modification of UAV spectral response to fertilizer N across winter wheat developmental stages. (A) Percentage and number of environments exhibiting a significant N rate × cultivar interaction for each vegetation index. Values within cells indicate the percentage of the seven environments with a significant interaction; the corresponding number of environments is shown below. (B) Overall percentage of significant N rate × cultivar interaction tests across eight vegetation indices and seven environments within each developmental stage. Spectral response to fertilizer N was most consistent among cultivars during stem elongation. 94

List of Tables

Table 2.1. Location, soil, and weather characteristics of the seven Kansas winter wheat environments during the 2024–2025 growing season. 16
Table 2.2. Company, year of release, and contrasting traits of the eight hard red winter wheat cultivars evaluated. 17
Table 2.3. Sowing, fertilizer-N application, and UAV acquisition timing across seven Kansas winter wheat environments during the 2024–2025 growing season. N application and flight timing are expressed as days after sowing (DAS); Feekes growth stage at each event is shown in parentheses. 21
Table 2.4. Vegetation indices derived from UAV multispectral reflectance..... 21
Table 2.5. Hyperparameter tuning used for the final cross-location model comparison. 27
Table 2.6. Nested leave-one-location-out predictor selection and performance of the final parsimonious PLSR models for winter wheat grain yield and protein. 35
Table 3.1. Sowing, fertilizer-N application, and UAV acquisition timing across seven Kansas winter wheat environments during the 2024–2025 growing season. Values in parentheses are the Feekes growth stage on each date (e.g., F11.2 = Feekes 11.2). * Flight collected before fertilizer-N application. † Multispectral imagery from this flight was not usable. NA, no flight. 60
Table 3.2. Vegetation indices derived from UAV multispectral reflectance..... 60

Acknowledgements

I would like to express my sincere gratitude to my major professor, Dr. Deepak Joshi, for the opportunity to become his first graduate student and for his continued support, guidance, and dedication to my academic and professional development. I am also grateful to my committee members, Dr. Romulo Lollato and Dr. Andres Patrignani. I thank Dr. Lollato for the opportunity to collaborate with him and his team on their wheat research trials, and Dr. Patrignani for the knowledge and insight he shared through both formal coursework and countless informal discussions. I am also thankful to my lab colleagues, Sangam Bishwakarma and Matias Lucero, for their assistance, collaboration, and support throughout my research.

I would also like to extend my appreciation to Dr. Nels Hansen and Dr. David Burchfield at Brigham Young University–Idaho for their encouragement and mentorship. Their support motivated me to pursue graduate education, and their excellent teaching helped develop my curiosity and ability to think critically and scientifically.

Finally, I owe my deepest gratitude to my wife and son, as well as my parents and siblings, for their unconditional love, patience, and continued encouragement throughout this journey. Their support has meant more to me than I can adequately express.

Chapter 1 - Introduction

Background

Winter wheat (*Triticum aestivum* L.) is an important cereal crop worldwide and a major source of food, energy, and protein for the human population (Curtis and Halford, 2014). In the United States, Kansas is the leading winter wheat-producing state, accounting for approximately 27% of the total U.S. wheat-producing area and 23% of total U.S. production (Jaenisch et al., 2026). For hard red winter wheat, the dominant class grown in the region, both grain yield and grain protein concentration determine the value of the crop. Grain yield is the primary measure of productivity, while grain protein concentration is a key quality trait that influences end-use value and, in turn, the price growers receive.

Nitrogen (N) management strongly influences both traits. Nitrogen deficiency limits crop growth by reducing chlorophyll concentration, leaf area development, and biomass accumulation (Hirel et al., 2007; Xu et al., 2012), whereas excessive N application can lower N-use efficiency and increase the risk of environmental N losses (Raun and Johnson, 1999; Hawkesford, 2014). Grain yield and protein also respond to N in different ways. Grain protein concentration depends on cultivar, N rate, and the timing of N application (Lollato et al., 2021), and late-season N application has been shown to increase grain protein concentration while having little overall effect on grain yield (Giordano et al., 2023). Because soil N supply, crop N demand, and growing conditions vary across fields and seasons, timely information on crop N status and expected crop performance is needed to guide fertilizer, harvest, and marketing decisions.

Precision agriculture emerged in the late 1980s and 1990s, when yield monitors, satellite positioning, geographic information systems, and variable-rate equipment made it possible to measure and manage variability within fields rather than treating each field as uniform (Mulla,

2013). Remote sensing became central to this approach because canopy reflectance provides spatially continuous information on crop status, but satellite and manned-aircraft imagery were often limited by revisit time, cloud cover, spatial resolution, and cost (Zhang & Kovacs, 2012; Mulla, 2013). Unmanned aircraft first entered agriculture as application platforms. In Japan, remotely piloted helicopters such as the Yamaha R-50, completed in 1987, and its successor, the RMAX, were developed for aerial spraying of rice and other crops on small fields (Sato, 2003). Their use as sensing platforms followed in the 2000s, when researchers mounted consumer digital and color-infrared cameras on model aircraft to estimate crop biomass and N status (Hunt et al., 2005) and used light UAVs to monitor wheat canopy variables in small research plots (Lelong et al., 2008). Lighter airframes, miniaturized multispectral and thermal sensors, and photogrammetric processing software subsequently made UAV imagery practical for routine field research (Berni et al., 2009; Colomina & Molina, 2014). In the United States, the small unmanned aircraft rule (14 CFR Part 107), effective in 2016, provided a framework for routine commercial operation (Federal Aviation Administration, 2016). UAVs are now used in precision agriculture for crop scouting, nutrient and water-stress assessment, weed mapping, yield estimation, and high-throughput phenotyping (Maes & Steppe, 2019; Tsouros et al., 2019).

UAV Multispectral Sensing of Winter Wheat

Proximal remote sensing provides a non-destructive means of monitoring these crop responses during the growing season rather than waiting until harvest. Changes in chlorophyll concentration, green leaf area, and canopy structure alter canopy reflectance in the visible, red-edge, and near-infrared regions, allowing remote sensing to track crop development and N status over time (Joshi et al., 2024; Singh et al., 2025). Vegetation indices derived from these spectral regions have been used extensively for in-season assessment of wheat N status, fertilizer response,

and N recommendations (Raun et al., 2005). Unmanned aerial vehicles (UAVs) are well suited to this work because they provide non-destructive, scalable, high-spatial-resolution imagery and can be flown repeatedly at specific crop growth stages (Sharma et al., 2022). Previous studies have shown that UAV-derived vegetation indices can be useful for predicting both wheat grain yield and grain protein (Zhou et al., 2021; Walsh et al., 2022).

The information provided by a vegetation index, however, depends on its spectral formulation and on when it is measured. The normalized difference vegetation index (NDVI) is widely used because of its simplicity and robustness, but its sensitivity can decline under high-biomass conditions due to spectral saturation (Li et al., 2010; Cao et al., 2015). Red-edge-based indices are generally less susceptible to saturation and remain more responsive to chlorophyll and crop N status as the canopy develops (Cao et al., 2015; Prey and Schmidhalter, 2019). The choice of vegetation index and sensing window is therefore central to the use of UAV imagery for both N management and crop prediction.

Much of the foundation for using NDVI in wheat N management and yield prediction was developed at Oklahoma State University. Stone et al. (1996) used canopy spectral radiance to detect and correct in-season N deficiencies in winter wheat, and Raun et al. (2001) showed that an in-season yield estimate, calculated from two postdormancy NDVI readings divided by the growing degree days between them, explained 83% of the variation in winter wheat grain yield across nine locations over two years. This work became the basis of optical sensor-based N fertilization algorithms that combine predicted yield potential with an in-season response index measured against an N-rich reference strip, using the GreenSeeker active sensor (Raun et al., 2002, 2005; Mullen et al., 2003), and the approach was later extended to corn (Teal et al., 2006). The same group showed that NDVI measured late in the season could predict both wheat grain yield

and grain protein (Freeman et al., 2003), and that simple red–green–blue (RGB) digital images could estimate wheat vegetation cover, which was closely related to canopy NDVI (Lukina et al., 1999). UAV multispectral imagery extends these approaches by providing similar spectral information at high spatial resolution across entire fields or experiments.

Interpreting UAV Data Across Environments

Canopy reflectance is also shaped by the environment in which the crop is growing. Background soil N supply, water status, and other site conditions can alter canopy development and spectral response independently of the fertilizer rate applied (Bausch and Delgado, 2005; Clay et al., 2012; Wang et al., 2025). Consequently, the same fertilizer-N gradient may produce strong spectral separation in one environment and little separation in another, depending on how much N the soil already supplies. One way to account for this is to express canopy measurements relative to an N-rich reference grown under the same conditions, commonly as a sufficiency index (Raun et al., 2002; Flowers et al., 2003; Holland and Schepers, 2010). Clay et al. (2012), for example, found that the N sufficiency index of winter wheat was more closely related to N-induced yield response when under-fertilized plants were referenced to well-fertilized plants grown in the same water environment. Relatively few multi-environment UAV studies, however, have tested whether differences in background soil N supply explain differences in both spectral N detectability and agronomic N responsiveness across wheat environments.

Environmental variation similarly complicates the prediction of grain yield and protein. Prediction models are often trained and tested on randomly divided observations from the same fields, which evaluates performance only within environments already represented during model development. When entire field-years are withheld, wheat yield and protein prediction accuracy has been shown to decline (Tilse et al., 2025), and UAV-based wheat yield models have not

consistently generalized to a new location when no data from that location were included in model training (Kaushal et al., 2026). In addition, spectral information alone may not capture all of the variation in final yield and protein, because these traits are also influenced by management, genetic, and environmental factors (Lollato et al., 2021; Maatougui et al., 2026). Evaluating whether UAV-based models transfer to locations not used for model development is therefore necessary before such models can be relied upon in practice.

Research Objectives

This thesis addresses these challenges using a multi-environment field experiment conducted at seven Kansas locations (Manhattan, Ashland Bottoms, Belleville, Russell, Hutchinson, Hays, and Hoisington) during the 2024–2025 growing season. At each location, eight hard red winter wheat cultivars were grown under seven N rates ranging from 0 to 202 kg N ha^{-1} in a split-plot design, and UAV multispectral imagery was collected at up to six growth stages from late tillering (Feekes 3) through physiological maturity (Feekes 11.4). The overall goal of this research was to evaluate how UAV multispectral sensing can support N management, particularly by identifying where in a field N management may need attention, and pre-harvest prediction of grain yield and protein in Kansas winter wheat across contrasting environments.

The specific objectives were to (i) evaluate the transferability of grain yield and protein predictions from multitemporal UAV vegetation indices and agronomic variables to locations not used for model training, using nested leave-one-location-out validation, and identify the spectral and agronomic predictors that support transferable prediction (Chapter 2); and (ii) evaluate the sensitivity of non-red-edge and red-edge vegetation indices to fertilizer-N response across crop developmental stages, determine whether variation in spectral N detectability among environments is associated with background soil N supply and agronomic N responsiveness, and test whether

normalization to a local N-rich reference strengthens the relationship between spectral and agronomic N response (Chapter 3).

Thesis Organization

This thesis is organized into four chapters. Chapter 1 provides a general introduction and the objectives of the research. Chapter 2 evaluates the cross-location prediction of winter wheat grain yield and protein from UAV multispectral imagery and agronomic variables, comparing partial least squares regression, random forest, and extreme gradient boosting under nested leave-one-location-out validation. Chapter 3 evaluates UAV-based detectability of fertilizer-N response across seven environments with contrasting background soil N supply and the value of local N-rich normalization. Chapter 4 summarizes the main conclusions of the thesis, their implications for N management, and needs for future research. Chapters 2 and 3 are written as standalone manuscripts, each with its own abstract, introduction, materials and methods, results, discussion, and references; as a result, some overlap exists in the descriptions of the study sites, experimental design, and UAV data collection.

References

Bausch, W. C., & Delgado, J. A. (2005). Impact of residual soil nitrate on in-season nitrogen applications to irrigated corn based on remotely sensed assessments of crop nitrogen status. *Precision Agriculture, 6*(6), 509–519. https://doi.org/10.1007/s11119-005-5641-9

Berni, J. A. J., Zarco-Tejada, P. J., Suárez, L., & Fereres, E. (2009). Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. IEEE Transactions on Geoscience and Remote Sensing, 47(3), 722–738. https://doi.org/10.1109/TGRS.2008.2010457

Cao, Q., Miao, Y., Feng, G., Gao, X., Li, F., Liu, B., Yue, S., Cheng, S., Ustin, S. L., & Khosla, R. (2015). Active canopy sensing of winter wheat nitrogen status: An evaluation of two sensor systems. Computers and Electronics in Agriculture, 112, 54–67. https://doi.org/10.1016/j.compag.2014.08.012

Clay, D. E., Kharel, T. P., Reese, C., Beck, D., Carlson, C. G., Clay, S. A., & Reicks, G. (2012). Winter wheat crop reflectance and nitrogen sufficiency index values are influenced by

nitrogen and water stress. Agronomy Journal, 104(6), 1612–1617. https://doi.org/10.2134/agronj2012.0216

Colomina, I., & Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 92, 79–97. https://doi.org/10.1016/j.isprsjprs.2014.02.013

Curtis, T., & Halford, N. G. (2014). Food security: The challenge of increasing wheat yield and the importance of not compromising food safety. Annals of Applied Biology, 164(3), 354–372. https://doi.org/10.1111/aab.12108

Federal Aviation Administration. (2016). Operation and certification of small unmanned aircraft systems (Final rule). Federal Register, 81(124), 42064–42214.

Flowers, M., Weisz, R., & Heiniger, R. (2003). Quantitative approaches for using color infrared photography for assessing in-season nitrogen status in winter wheat. Agronomy Journal, 95, 1189–1200. https://doi.org/10.2134/agronj2003.1189

Freeman, K. W., Raun, W. R., Johnson, G. V., Mullen, R. W., Stone, M. L., & Solie, J. B. (2003). Late-season prediction of wheat grain yield and grain protein. Communications in Soil Science and Plant Analysis, 34, 1837–1852.

Giordano, N., Sadras, V. O., & Lollato, R. P. (2023). Late-season nitrogen application increases grain protein concentration and is neutral for yield in wheat: A global meta-analysis. Field Crops Research, 290, 108740. https://doi.org/10.1016/j.fcr.2022.108740

Hawkesford, M. J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. Journal of Cereal Science, 59(3), 276–283. https://doi.org/10.1016/j.jcs.2013.12.001

Hirel, B., Le Gouis, J., Ney, B., & Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: Towards a more central role for genetic variability and quantitative genetics within integrated approaches. Journal of Experimental Botany, 58(9), 2369–2387. https://doi.org/10.1093/jxb/erm097

Holland, K. H., & Schepers, J. S. (2010). Derivation of a variable rate nitrogen application model for in-season fertilization of corn. Agronomy Journal, 102, 1415–1424. https://doi.org/10.2134/agronj2010.0015

Hunt, E. R., Jr., Cavigelli, M., Daughtry, C. S. T., McMurtrey, J. E., III, & Walthall, C. L. (2005). Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. Precision Agriculture, 6(4), 359–378. https://doi.org/10.1007/s11119-005-2324-5

Jaenisch, B. R., Giordano, N., Sadras, V. O., Nelson, N. O., Guttieri, M. J., Fritz, A. K., Ruiz Diaz, D., & Lollato, R. P. (2026). Nitrogen and sulfur co-limitation and stoichiometry modulate yield and nutrient use efficiency of bread wheat. Field Crops Research, 344, 110539. https://doi.org/10.1016/j.fcr.2026.110539

Joshi, D. R., Clay, S. A., Sharma, P., Rekabdarkolaee, H. M., Kharel, T., Rizzo, D. M., Thapa, R., & Clay, D. E. (2024). Artificial intelligence and satellite-based remote sensing can be used to predict soybean (Glycine max) yield. Agronomy Journal, 116(3), 917–930. https://doi.org/10.1002/agj2.21473

Kaushal, S., Billah, M. M., Singh, M., Koupal, D., Sitaula, P. R., Halder, J., Sidhu, J., Maimaitijiang, M., & Sehgal, S. K. (2026). UAV-based deep transfer learning to improve grain yield prediction in winter wheat across temporal and spatial variability. The Plant Phenome Journal, 9(1), e70085. https://doi.org/10.1002/ppj2.70085

Lelong, C. C. D., Burger, P., Jubelin, G., Roux, B., Labbé, S., & Baret, F. (2008). Assessment of unmanned aerial vehicles imagery for quantitative monitoring of wheat crop in small plots. Sensors, 8(5), 3557–3585. https://doi.org/10.3390/s8053557

Li, F., Zhang, H., Jia, L., Bareth, G., Miao, Y., & Chen, X. (2010). Estimating winter wheat biomass and nitrogen status using an active crop sensor. Intelligent automation and soft computing, 16(6), 1221-1230.

Lollato, R. P., Jaenisch, B. R., & Silva, S. R. (2021). Genotype-specific nitrogen uptake dynamics and fertilizer management explain contrasting wheat protein concentration. Crop Science, 61(3), 2048–2066. https://doi.org/10.1002/csc2.20442

Lukina, E. V., Stone, M. L., & Raun, W. R. (1999). Estimating vegetation coverage in wheat using digital images. Journal of Plant Nutrition, 22(2), 341–350. https://doi.org/10.1080/01904169909365631

Maatougui, M., Phillips, S., Abiola, S. O., & Arnall, D. B. (2026). Wheat yield prediction using machine learning, proximal sensing, nitrogen and climate data for Oklahoma. Precision Agriculture, 27, 148. https://doi.org/10.1007/s11119-026-10462-2

Maes, W. H., & Steppe, K. (2019). Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture. Trends in Plant Science, 24(2), 152–164. https://doi.org/10.1016/j.tplants.2018.11.007

Mulla, D. J. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosystems Engineering, 114(4), 358–371. https://doi.org/10.1016/j.biosystemseng.2012.08.009

Mullen, R. W., Freeman, K. W., Raun, W. R., Johnson, G. V., Stone, M. L., & Solie, J. B. (2003). Identifying an in-season response index and the potential to increase wheat yield with nitrogen. Agronomy Journal, 95(2), 347–351. https://doi.org/10.2134/agronj2003.3470

Prey, L., & Schmidhalter, U. (2019). Sensitivity of vegetation indices for estimating vegetative N status in winter wheat. Sensors, 19(17), 3712. https://doi.org/10.3390/s19173712

Raun, W. R., & Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. Agronomy Journal, 91(3), 357–363. https://doi.org/10.2134/agronj1999.00021962009100030001x

Raun, W. R., Johnson, G. V., Stone, M. L., Solie, J. B., Lukina, E. V., Thomason, W. E., & Schepers, J. S. (2001). In-season prediction of potential grain yield in winter wheat using canopy reflectance. Agronomy Journal, 93(1), 131–138.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K. W., Thomason, W. E., & Lukina, E. V. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agronomy Journal, 94(4), 815–820. https://doi.org/10.2134/agronj2002.8150

Raun, W. R., Solie, J. B., Stone, M. L., Martin, K. L., Freeman, K. W., Mullen, R. W., Zhang, H., Schepers, J. S., & Johnson, G. V. (2005). Optical sensor-based algorithm for crop nitrogen fertilization. Communications in Soil Science and Plant Analysis, 36(19), 2759–2781. https://doi.org/10.1080/00103620500303988

Sato, A. (2003). The RMAX helicopter UAV (Report No. ADA427393). Yamaha Motor Co., Ltd.; Defense Technical Information Center.

Sharma, P., Leigh, L., Chang, J., Maimaitijiang, M., & Caffé, M. (2022). Above-ground biomass estimation in oats using UAV remote sensing and machine learning. Sensors, 22(2), 601. https://doi.org/10.3390/s22020601

Singh, R., Kaur, S., Joshi, D. R., Iboyi, J., Dar, E., Sharma, L. K., & Singh, H. (2025). Estimating cotton biomass and nitrogen content using high-resolution satellite and UAV data fusion with machine learning. Smart Agricultural Technology, 12, 101191.

Stone, M. L., Solie, J. B., Raun, W. R., Whitney, R. W., Taylor, S. L., & Ringer, J. D. (1996). Use of spectral radiance for correcting in-season fertilizer nitrogen deficiencies in winter wheat. Transactions of the ASAE, 39(5), 1623–1631.

Teal, R. K., Tubana, B., Girma, K., Freeman, K. W., Arnall, D. B., Walsh, O., & Raun, W. R. (2006). In-season prediction of corn grain yield potential using normalized difference vegetation index. Agronomy Journal, 98(6), 1488–1494.

Tilse, M. J., Bishop, T. F. A., & Filippi, P. (2025). Predicting within-field grain protein content at scale using agronomic and remote sensing variables, and machine learning. Precision Agriculture, 26, 78. https://doi.org/10.1007/s11119-025-10267-9

Tsouros, D. C., Bibi, S., & Sarigiannidis, P. G. (2019). A review on UAV-based applications for precision agriculture. Information, 10(11), 349. https://doi.org/10.3390/info10110349

Walsh, O. S., Marshall, J., Jackson, C., Nambi, E., Shafian, S., Jayawardena, D. M., Lamichhane, R., Owusu Ansah, E., & McClintick-Chess, J. R. (2022). Wheat yield and protein estimation with handheld- and UAV-based reflectance measurements.

Agrosystems, Geosciences & Environment, 5(4), e20309. https://doi.org/10.1002/agg2.20309

Wang, J., Meyer, S. T., Xu, X., Weisser, W. W., & Yu, K. (2025). Drone multispectral imaging captures the effects of soil mineral nitrogen on canopy structure and nitrogen use efficiency in wheat. Computers and Electronics in Agriculture, 235, 110342. https://doi.org/10.1016/j.compag.2025.110342

Xu, G., Fan, X., & Miller, A. J. (2012). Plant nitrogen assimilation and use efficiency. Annual Review of Plant Biology, 63, 153–182. https://doi.org/10.1146/annurev-arplant-042811-105532

Zhang, C., & Kovacs, J. M. (2012). The application of small unmanned aerial systems for precision agriculture: A review. Precision Agriculture, 13(6), 693–712. https://doi.org/10.1007/s11119-012-9274-5

Zhou, X., Kono, Y., Win, A., Matsui, T., & Tanaka, T. S. T. (2021). Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. Plant Production Science, 24(2), 137–151. https://doi.org/10.1080/1343943X.2020.1819165

Chapter 2 - Transferable Prediction of Winter Wheat Grain Yield and Protein Using UAV Multispectral Imagery and Agronomic Variables

Abstract

Unmanned aerial vehicle (UAV)-based pre-harvest estimates of winter wheat grain yield and protein concentration provide a non-destructive, scalable method of prediction that can support growers in crop management and grain quality decisions. However, prediction models trained at one location may not transfer well to new locations. Therefore, the main objective of this study was to use an innovative leave-one-location-out (LOLO) validation method to evaluate the transferability of grain yield and protein predictions from machine learning (ML) models trained on multitemporal UAV multispectral imagery and agronomic variables. For that, a field experiment was conducted across seven Kansas environments during the 2024–2025 growing season, evaluating eight hard red winter wheat cultivars across seven nitrogen (N) fertilizer rates. UAV images were collected during tillering, stem elongation, and grain formation to derive eight red-edge- and non-red-edge-based vegetation indices, as well as agronomic variables such as plant biomass and height. Across the 1,344 plots, grain yield ranged from 1,497 to 6,237 kg ha^{-1} and grain protein concentration from 8.8% to 17.5%, providing a wide range of conditions for testing model transferability. Partial least squares regression (PLSR), random forest (RF), and XGBoost were evaluated via nested LOLO validation with VI and agronomic variables. PLSR provided better overall cross-location performance than the nonlinear models for both grain yield and protein. The final PLSR models achieved $R^2 = 0.55$, concordance correlation coefficient (CCC) = 0.72, and RMSE = 516.4 kg ha^{-1} for yield, and $R^2 = 0.47$, CCC = 0.67, and RMSE = 1.09

percentage points for protein. Predictor-importance analysis identified the green normalized difference vegetation index (GNDVI) during stem elongation and the N fertilizer rate as the main yield predictors, whereas the MERIS terrestrial chlorophyll index (MTCI) during grain formation, the N rate, and plant height were most important for protein. These results showed that combining UAV-based spectral features with agronomic information (N rate, cultivar, and plant height) improved prediction across unseen locations.

1 Introduction

Wheat (*Triticum aestivum* L.) supplies a large share of the energy and protein in the human diet (Curtis and Halford, 2014), and Kansas, in the southern Great Plains, accounts for about 27% of the U.S. wheat area and 23% of U.S. wheat production (Jaenisch et al., 2026). Due to the versatile uses of wheat, total yield and production quality, such as protein content, are very important to end users, consequently impacting growers in this region. Therefore, accurate pre-harvest yield estimation can help growers make decisions regarding input application, marketing, and storage. Similarly, pre-harvest protein content estimation can help identify high-protein areas in the field, supporting selective harvesting and grain segregation to achieve better market prices, as proposed as a strategy in other important wheat growing regions (Larroque et al., 2022).

With advancements in precision agriculture, unmanned aerial vehicle (UAV) remote sensing provides a non-destructive, scalable way to monitor these crop responses during the growing season rather than waiting until harvest (Sharma et al., 2022). As crops progress through developmental stages, changes in canopy characteristics and plant greenness are reflected in spectral response, allowing remote sensing to track crop development over time (Joshi et al., 2024). Unmanned aerial vehicles (UAVs) are well-suited to this type of work because they provide high spatial resolution and can be flown repeatedly at specific crop growth stages. Various previous

studies have shown that UAV-based vegetation indices (VI) can be useful for predicting both wheat yield and grain protein (Walsh et al., 2022; Zhou et al., 2021; Fu et al., 2022; Billah et al., 2026). For example, Zhou et al. (2021) used various UAV-derived VI's, including normalized difference vegetation index (NDVI), green normalized difference vegetation index (GNDVI), two-band enhanced vegetation index (EVI2), green chlorophyll index (CIgreen), red-edge chlorophyll index (CIred-edge), and normalized difference red-edge index (NDRE), and showed that these indices had useful relationships with both winter wheat grain yield and protein content for prediction.

Spectral information derived from UAV imagery provides useful insights into in-season crop growth conditions and development at the time of sensing, but spectral features alone may not capture all the variation in final grain yield and protein concentration. Final crop performance is also influenced by agronomic and environmental factors, such as N fertilizer rate, cultivar, and growing conditions, as well as crop structures such as biomass and plant height, which can affect yield and protein. For example, Lollato et al. (2021), based on a study conducted at six different environmental locations, reported wheat yields ranging from 2,853 to 8,023 kg ha^{-1} and grain protein concentrations from 88 to 152 g kg^{-1}. Moreover, they demonstrated that protein content is determined by cultivar selection, N fertilizer rate, and timing of fertilizer application, indicating that genotype and N dynamics can contribute substantially to grain protein content. Similarly, Giordano et al. (2024) demonstrated that cultivar yield and protein plasticity (i.e., environmental responsiveness) modulated yield and protein response to applied nitrogen, especially in high-expression environments when nitrogen was not limiting. Meanwhile, Giordano et al. (2023), in a global meta-analysis of 38 published studies, found that increasing N fertilizer rate and late-season N application (i.e., after the boot stage, Feekes 10) increased grain protein concentration by

approximately 4% while having little overall effect on grain yield. Similar work in other cropping systems has shown that crop yield results from several interacting management and agronomic variables and that the most important influencing factors can differ among crops and environments (Lucero et al., 2026). Together, these studies show that UAV-derived spectral information is important, but prediction of wheat yield and protein is likely to benefit from combining spectral features with management and genetic information.

Combining UAV-based spectral features across different wheat growth stages with management and cultivar information can provide more robust predictions of yield and protein content. For example, in a recent study conducted in Oklahoma at 57 sites, combining NDVI, N management, and weather- and soil-related variables explained up to 81% of yield variability and resulted in a prediction error (RMSE) of 0.62 t ha^{-1} (Maatougui et al., 2026). Likewise, Maranhão et al. (2025) demonstrated that adding weather and management variables to vegetation indices improved the field-level yield prediction power of machine learning models. However, a main concern about the prediction model is that when it is trained and tested on data from the same field, it may not perform as well when applied to an unseen field not used for training. Tilse et al. (2025), for example, reported lower wheat yield and protein prediction accuracy when entire field-years were withheld from model development. Similarly, Kaushal et al. (2026) reported that wheat yield models developed from UAV imagery did not consistently generalize to a new location when no data from that location were included during model development.

Therefore, this study evaluated the transferability of predicting winter wheat grain yield and protein using multitemporal UAV multispectral imagery together with agronomic variables. The main objectives of this study were (i) to use leave-one-location-out (LOLO) validation to test yield and protein prediction accuracy at unseen locations, with each of the seven locations withheld

in turn as the test set while models were trained on the remaining six; (ii) to evaluate the contribution of N rate, cultivar, crop canopy features, and UAV spectral information to variation in yield and protein prediction; (iii) to identify a small set of spectral predictors for grain yield and protein using nested leave-one-location-out validation; and (iv) to compare the cross-location prediction performance of PLSR, random forest, and extreme gradient boosting.

Our emphasis was not only on achieving high prediction accuracy but also on determining whether the selected predictors continued to perform when the model was applied to a location not represented during model development.

2 Materials and Methods

2.1 Study Sites and Environmental Context

Field trials were conducted during the 2024–2025 winter wheat growing season at seven locations distributed across the state of Kansas, representing the breadth of agronomic environments and climatic conditions characteristic of the state's hard red winter wheat production region. Trial sites were established at Manhattan (Kansas State University North Farm), Ashland Bottoms, Hoisington, Russell, Hays, Belleville, and Hutchinson. The region experiences a temperate continental climate, with 30-year mean annual precipitation ranging from approximately 600 mm at the westernmost site (Hays) to 870 mm at the easternmost site (Manhattan) and mean monthly temperatures varying from approximately −6°C in winter to 33°C in summer. This west-to-east precipitation gradient, combined with variation in soil series and texture across sites, provided a diverse set of growing conditions for evaluating the effects of cultivar and nitrogen rate on grain yield and protein concentration. Each location's weather conditions, including average growing-season temperature (Tmean), precipitation (P), and evapotranspiration (ET), together

with soil properties such as soil texture (sand, silt, and clay), organic matter (OM), and pH, are summarized in Table 2.1.

Table 2.1. Location, soil, and weather characteristics of the seven Kansas winter wheat environments during the 2024–2025 growing season.

Location	Coordinates		SSURGO soil series	Soil properties (0–15 cm)					Weather, sowing to harvest		
	Lat. (°N)	Long. (°W)		Sand (%)	Silt (%)	Clay (%)	OM (%)	pH	P (mm)	Tmean (°C)	ETo (mm)
Ashland	39.1432	96.6312	Belvue silt loam	26	62	12	1.8	6.4	540	9.0	663
Belleville	39.8148	97.6717	Crete silt loam	14	62	24	2.6	5.3	462	8.5	675
Hays	38.8540	99.3411	Harney silt loam	16	52	32	2.5	6.7	339	9.1	758
Hoisington	38.6334	98.8680	Harney silt loam	16	57	27	2.6	6.0	318	9.5	757
Hutchinson	37.9278	98.0263	Ost loam	39	43	18	1.7	5.9	484	9.1	609
Manhattan	39.2173	96.5893	Kahola silt loam	14	60	26	3.1	6.8	483	9.6	682
Russell†	38.8941	98.8476	Crete silt loam	14	62	24	2.8	6.4	339	9.1	758

Notes: **Coordinates** are decimal degrees (WGS 84). **SSURGO soil series** is the soil map unit at each site. **Soil properties** are from soil samples collected at 0–15 cm at each site; **OM**, organic matter. **Weather** values are totals or means of daily data from the nearest Kansas Mesonet station between sowing and harvest (226 to 268 days): **P**, total precipitation; **Tmean**, mean of the daily maximum and minimum air temperatures; **ETo**, cumulative grass-reference evapotranspiration. † No weather station at the Russell site; values are from the Hays station (about 43 km west), as in the analysis dataset.

2.2 Experimental Design and Treatments

Treatments were arranged in a split-plot design, with nitrogen (N) rate assigned as the whole-plot factor and winter wheat cultivar as the subplot factor within each replicate block. Seven N rates were evaluated: 0, 34, 67, 101, 135, 168, and 202 kg N ha^{-1}. Eight commercially available hard red winter wheat cultivars were included: Bob Dole (developed by Kansas State University and marketed by AgriPro/Syngenta), LCS Chrome and LCS Julep (Limagrain Cereal Seeds), SY Monument and SY Wolverine (AgriPro/Syngenta), and WB4269, WB4401, and WB4699 (WestBred/Bayer). These cultivars were selected to represent contrasting phenotypic plasticity of yield, grain protein concentration (GPC), and maturity, with each of the eight combinations of high or low yield plasticity, high or low GPC, and early or late maturity represented by one cultivar

(Table 2.2). Based on previous work by Giordano et al. (2024) and Cossani and Sadras (2021), we hypothesized that these traits would result in large variability in cultivar-specific response to nitrogen, increasing the range in measured values. Each replicate block therefore contained 56 treatment combinations (seven N rates × eight cultivars). The experiment was replicated three times at Manhattan, Ashland Bottoms, Belleville, and Hoisington and four times at Hutchinson, Russell, and Hays. Individual plots were seven rows wide, with 19.1-cm row spacing, and approximately 9.1 m long. Plots were sown using a Great Plains 606 no-till drill. Other agronomic practices, including sowing date, seeding rate, and pest and disease management, followed locally recommended practices for each location so that they were non-limiting to grain yield and were applied consistently within each site. Nitrogen fertilizer was applied according to the assigned treatment rates at each location.

Table 2.2. Company, year of release, and contrasting traits of the eight hard red winter wheat cultivars evaluated.

Variety	Company	Year of release	Phenotypic plasticity of yield	GPC	Maturity
SY Wolverine	AgriPro (Syngenta)	2019[b]	High	High	Early
WB4401	WestBred (Bayer)	2020	High	Low	Early
WB4699	WestBred (Bayer)	2018[c]	High	Low	Late
LCS Julep	Limagrain Cereal Seeds	2020	High	High	Late
WB4269	WestBred (Bayer)	2017	Low	Low	Early
SY Monument	AgriPro (Syngenta)	2014	Low	Low	Late
LCS Chrome	Limagrain Cereal Seeds	2016	Low	High	Late
Bob Dole	AgriPro (Syngenta)[a]	2018	Low	High	Early

Notes: **GPC**, grain protein concentration. Trait classes are relative among the eight cultivars. Release years are from state variety-testing reports and company announcements. [a] Developed by Kansas State University and marketed by AgriPro (Syngenta). [b] Certified seed first available for fall 2020 planting. [c] Reported as 2018 by the University of Nebraska–Lincoln and as 2019 by Texas A&M AgriLife.

2.3 Plant Height, Biomass, Grain Yield and Protein Measurement

Plant height and aboveground biomass measurements were collected at anthesis (Feekes 10.5) to characterize crop development and vegetative growth. Plant height was measured from the soil surface to the tip of the tallest part of the plant (excluding awns), with measurements taken from five representative plants within each plot. Measurements were recorded in centimeters (cm) and averaged to obtain a plot-level estimate.

Aboveground biomass was collected by harvesting plant material from a designated sampling area within each plot. Samples were dried at 60-65 °C for approximately 48 hours until constant weight and weighed to determine dry biomass. Biomass measurements were expressed as g m^{-2}.

At physiological maturity (Feekes 11.4), grain yield was measured from each plot using a Massey Ferguson 8XP small-plot combine. Grain yield was adjusted to 13.5% moisture content and expressed in kg ha^{-1}. Following harvest, grain subsamples were cleaned and analyzed for protein concentration using near-infrared (NIR) reflectance spectroscopy, which estimates protein content based on absorbance characteristics associated with nitrogen-containing compounds in the grain. Protein concentration values were expressed on a dry-weight basis as a percentage of total grain mass.

2.4 UAV Imagery Acquisition

Multispectral aerial imagery was collected approximately every two weeks throughout the growing season, with six total flights conducted between late March and early July 2025. Flights coincided with key phenological stages spanning late tillering (Feekes 3) through physiological maturity (Feekes 11.4), capturing the predominant stages when large biomass accumulation occurs (Lollato and Edwards, 2015). At Hoisington and Russell, the first flight preceded fertilizer-N

application (Table 2.3), capturing initial plot conditions before any change driven by the N fertilizer response. Data collection was suspended or rescheduled when wind speeds exceeded safe operating limits or when cloud cover prevented the clear-to-near-clear sky conditions required for radiometric calibration.

Imagery was acquired using a senseFly eBee X fixed-wing unmanned aerial vehicle (senseFly, Cheseaux-sur-Lausanne, Switzerland) equipped with a Parrot Sequoia Duet M multispectral camera. The Sequoia Duet M simultaneously captures surface reflectance in four narrow spectral bands: green (550 nm), red (660 nm), red-edge (735 nm), and near-infrared (790 nm), along with a co-registered RGB image. An integrated irradiance sensor mounted on the camera body continuously logged incoming solar radiation throughout each flight to support radiometric calibration.

All flights were conducted at a fixed altitude of 50 m above ground level, yielding a multispectral ground sampling distance (GSD) of 4.8 cm. Forward and lateral image overlap were both set to 75% to ensure sufficient image redundancy for accurate orthomosaic reconstruction. Flights were planned and executed using eMotion flight-planning software and conducted near solar noon, between 11:00 and 14:00 local time, under clear-to-near-clear sky conditions.

2.5 Radiometric Calibration and Image Processing

Prior to each flight, a reflectance calibration panel with known spectral properties was photographed under ambient illumination. The panel reflectance values, combined with irradiance data continuously logged by the integrated sensor during flight, were used to convert raw digital number values to calibrated surface reflectance. Radiometric calibration and image stitching were performed in Pix4Dfields (Pix4D SA, Prilly, Switzerland), which automatically applies panel-

based radiometric correction during orthomosaic generation, producing georeferenced, surface-reflectance orthomosaics for each spectral band at every flight date.

2.6 Soil Background Removal and Plot-Level Reflectance Extraction

Following orthomosaic generation, soil background pixels were removed from each image prior to plot-level data extraction using a custom Python pipeline. Plot boundaries were delineated by digitizing polygons over individual plots within Pix4Dfields. Mean reflectance values for each spectral band were then extracted within each plot polygon, yielding a single plot-level mean reflectance value per band per flight date.

2.7 Vegetation Index Calculation

Eight vegetation indices were calculated from the plot-level green, red, red-edge, and near-infrared reflectance values generated from each usable UAV acquisition (Table 2.4). The indices were selected to represent complementary spectral responses to canopy greenness, chlorophyll status, biomass, and red-edge behavior. NDVI, SAVI, and GNDVI were treated as non-red-edge indices, whereas NDRE, CI_RE, REDVI, MTCI, and DATT were treated as red-edge-based indices. The same index formulas were applied to all locations and acquisition dates before growth-stage aggregation. Belleville and Hays did not have usable multispectral imagery for Flight 2; all other usable acquisitions were retained for stage-based summarization.

The eight indices were chosen to cover the main spectral formulations used for in-season assessment of wheat N status that can be calculated from the four bands of the sensor. NDVI was included as the most widely used benchmark of canopy greenness and biomass and as the basis of established sensor-based N algorithms for winter wheat (Rouse et al., 1974; Raun et al., 2002), and SAVI was included to reduce soil-background effects early in the season, when canopy cover was incomplete (Huete, 1988). GNDVI replaces the red band with the green band, which remains

more sensitive to chlorophyll concentration at moderate to high leaf area (Gitelson et al., 1996). The five red-edge indices were included because reflectance in the red-edge region is closely related to chlorophyll and N status and is less prone to saturation in dense canopies than red-based indices (Cao et al., 2015; Prey and Schmidhalter, 2019). These indices use the red-edge band in different ways: as a normalized difference with NIR (NDRE; Barnes et al., 2000), as a ratio (CIRE; Gitelson et al., 2005), as a simple difference (REDVI; Cao et al., 2013), and with the red band as an additional reference (MTCI, Dash & Curran, 2004; DATT, Datt, 1999). Including several formulations made it possible to test whether red-edge information in general, rather than one particular index, provided the most transferable predictors.

Table 2.3. Sowing, fertilizer-N application, and UAV acquisition timing across seven Kansas winter wheat environments during the 2024–2025 growing season. N application and flight timing are expressed as days after sowing (DAS); Feekes growth stage at each event is shown in parentheses.

Location	Sowing date	N application	Flight 1	Flight 2	Flight 3	Flight 4	Flight 5	Flight 6
Ashland	Oct. 24, 2024	153 (F3)	161 (F4)	172 (F5)	181 (F8)	196 (F10)	216 (F11.2)	NA
Belleville	Oct. 15, 2024	164 (F3)	175 (F3)	185 (F6)	196 (F9)	206 (F10)	220 (F10.5)	239 (F11.4)
Hays	Oct. 8, 2024	171 (F3)	181 (F3)	185 (F4)	195 (F6)	206 (F10)	218 (F10.5)	234 (F11)
Hoisington	Sep. 27, 2024	195 (F3)	185 (F3)	196 (F3)	206 (F6)	217 (F9)	229 (F10.5)	245 (F11.2)
Hutchinson	Nov. 16, 2024	137 (F3)	142 (F3)	150 (F3)	160 (F5)	170 (F7)	187 (F10)	205 (F11.4)
Manhattan	Oct. 21, 2024	156 (F3)	164 (F4)	175 (F6)	184 (F7)	199 (F9)	219 (F11.2)	235 (F11.4)
Russell	Oct. 14, 2024	171 (F3)	168 (F3)	179 (F4)	189 (F6)	200 (F10)	212 (F10.5)	228 (F11.2)

Table 2.4. Vegetation indices derived from UAV multispectral reflectance

Vegetation index	Formula	Reference
NDVI	(NIR - Red) / (NIR + Red)	Rouse et al. (1974)
GNDVI	(NIR - Green) / (NIR + Green)	Gitelson et al. (1996)
NDRE	(NIR - RedEdge) / (NIR + RedEdge)	Barnes et al. (2000)
CI_{RE}	(NIR / RedEdge) - 1	Gitelson et al. (2005)
SAVI	1.5 × (NIR - Red) / (NIR + Red + 0.5)	Huete (1988)
REDVI	NIR - RedEdge	Cao et al. (2013)
MTCI	(NIR - RedEdge) / (RedEdge - Red)	Dash & Curran (2004)
DATT	(NIR - RedEdge) / (NIR - Red)	Datt (1999)

Abbreviations: **NDVI**, normalized difference vegetation index; **GNDVI**, green normalized difference vegetation index; **NDRE**, normalized difference red-edge index; **CI_{RE}**, red-edge chlorophyll index; **SAVI**, soil-adjusted vegetation index; **REDVI**, red-edge difference vegetation index; **MTCI**, MERIS terrestrial chlorophyll index; **DATT**, Datt chlorophyll index; **NIR**, near-infrared reflectance; **RedEdge**, red-edge reflectance; **Red**, red reflectance; **Green**, green reflectance.

2.7.1 Crop Growth Stage Classification

Because UAV image acquisition dates were not evenly spaced across locations, flight number was treated only as a sampling identifier and not as a common biological time scale. Instead, UAV observations were classified according to the Feekes growth stage (Large, 1954) recorded at each location and flight.

For cross-environment comparisons, observations were grouped into three broad developmental stages:

- **Tillering:** Feekes 3–5
- **Stem elongation:** Feekes 6–9
- **Grain formation (heading/grain fill):** Feekes 10–11.4

These groups were selected to provide comparable phenological windows across the seven environments while retaining sufficient observations within each stage for statistical comparison. The three windows also correspond to distinct physiological phases: tiller production and early

canopy expansion (tillering); rapid biomass and N accumulation and the critical period for grain-number determination (stem elongation); and spike emergence, anthesis, and grain filling, when N is remobilized to the grain and grain protein is determined (grain formation; Fischer, 1985; Kichey et al., 2007; Slafer et al., 2014).

Because the interval between fertilizer application and UAV acquisition varied across locations, developmental-stage comparisons were interpreted as phenological rather than as estimates of the exact timing or onset of fertilizer-N response. Thus, differences among the three developmental stages reflect broad changes in canopy development and spectral sensitivity across crop growth rather than standardized days after N application.

2.7.2 Exploratory Statistical Analysis

Before model development, grain yield and grain protein distributions were summarized by location to describe environmental variation in the two response traits. Pearson correlation coefficients were then used to quantify (i) the pooled relationship between grain yield and grain protein, (ii) relationships between each stage-averaged vegetation index and each response trait, and (iii) relationships of N rate, plant height, and anthesis biomass with grain yield and protein. For the vegetation-index analysis, false discovery rate (FDR) control was applied separately for grain yield and grain protein across the 24 index-by-stage correlations (8 indices x 3 developmental stages) using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995). For management and crop-structural analyses, FDR adjustment was applied across the six correlations (3 predictors × 2 response traits). Adjusted q values below 0.05 were considered significant. These pooled correlations were used as descriptive analyses of trait and spectral behavior and were not interpreted as independent treatment-effect tests because observations were clustered within

locations and replicate blocks. Cross-environment predictive performance was evaluated separately using the location-level validation framework described below.

2.8 Machine Learning Model Development and Evaluation

2.8.1 Response Variables, Predictors, and Preprocessing

Grain yield and grain protein concentration were predicted separately. The final modeling datasets contained 1,344 observations for grain yield and 1,343 observations for grain protein. Candidate predictors represented management, genotype, crop structure, and UAV spectral information. N rate was entered as a continuous numeric predictor. Genotype was represented in two alternative ways: cultivar identity, encoded as eight binary indicator variables, or three cultivar-level agronomic traits available in the dataset—yield plasticity (PP; LOW/HIGH), grain-protein tendency (GPC; LOW/HIGH), and maturity class (EARLY/LATE). For modeling, PP and GPC were encoded as 0 = LOW and 1 = HIGH, and maturity was encoded as 0 = EARLY and 1 = LATE. Cultivar identity and the three agronomic traits were evaluated as alternative genotype representations rather than entered simultaneously. Plant height was used as the crop-structural predictor in the final protein candidate set, while anthesis biomass was retained for the exploratory correlation analysis described above. UAV predictors consisted of the 24 stage-averaged vegetation-index variables (8 indices × 3 crop growth stages). Location was used only to define cross-validation groups and was not included as a predictor, thereby preventing the models from using site identity directly.

Rows with a missing response value were removed separately for the two traits. Within each outer validation fold, missing predictor values were imputed using the median estimated from the training locations only. For PLSR, the imputed predictors were standardized to zero mean and unit variance using parameters estimated from the same training data, then PLSRegression was

performed with internal scaling disabled (scale=False). RF and XGBoost used the same training-fold median imputation but did not require standardization. The fitted imputer and, where applicable, scaler were then applied unchanged to the held-out location. Thus, imputation, scaling, predictor selection, and hyperparameter tuning were completed without information from the outer test location.

2.8.2 Nested Leave-One-Location-Out Validation and Predictor Selection

Cross-environment performance was evaluated using an outer leave-one-location-out (LOLO) design. In each of seven outer folds, all observations from one location were withheld as an independent test set, and the remaining six locations were used for model development. The procedure was repeated until every location had served once as the test environment. Predictions from all seven held-out locations were combined into a single out-of-fold prediction set, which was used to calculate the overall model performance metrics. No observations from the outer test location were used for preprocessing, predictor selection, component selection, or hyperparameter tuning.

Predictor selection was performed on the outer training data using location-grouped inner cross-validation. For grain yield, each of the 24 stage-specific UAV variables (8 vegetation indices × tillering, stem elongation, and grain formation) was evaluated individually after being added to a baseline that included N rate, PP, GPC, and maturity. After identifying the most stable yield spectral predictor, all eight possible subsets of PP, GPC, and maturity, including a no-trait model, were evaluated within the same training-only nested framework, while N rate and the selected spectral predictor were retained. For grain protein, the descriptive stage-wise analysis showed that useful spectral relationships were concentrated during grain formation; therefore, the targeted protein search compared the eight Grain-formation vegetation indices after addition to a baseline

containing N rate, cultivar indicators, and plant height. Within each outer fold, candidate models were ranked by inner-validation RMSE, and the number of PLS components was tuned within the same training-only procedure. Selected predictors were recorded for each outer fold to assess selection stability. This design provided leakage-controlled predictor selection and enabled evaluation of prediction at a location that had not contributed to model fitting.

After nested predictor selection, the most stable predictors were used to define common parsimonious models applicable across locations. The final grain-yield model included N rate, PP, GPC, maturity, and GNDVI at stem elongation. The final grain-protein model included N rate, cultivar indicators, plant height, and MTCI at grain formation. Results from the fully nested selection analyses were retained alongside the fixed-predictor models to assess the effect of predictor-selection uncertainty rather than relying solely on the final fixed specifications.

2.8.3 Machine-Learning Algorithms and Hyperparameter Tuning

Partial least squares regression (PLSR) was used as the primary predictive model because it is well-suited to correlated predictors and provides a compact latent-variable representation that can be interpreted using VIP scores and standardized coefficients (Wold et al., 2001). Random forest (RF; Breiman, 2001) and extreme gradient boosting (XGBoost; Chen & Guestrin, 2016) were included as nonlinear benchmarks. To ensure a fair comparison of algorithms, RF and XGBoost were evaluated with the same final predictor sets and the same outer LOLO test folds as PLSR; they were not given an additional feature-selection step.

For PLSR, the number of latent components was tuned by exhaustive grid search from 1 to min(10, p), where p was the number of predictors in the candidate model. Inner validation used location-grouped cross-validation with up to five folds and negative RMSE as the scoring function. For RF, exhaustive grid search evaluated n_estimators = {300, 600}, max_depth = {3, 6, None},

and min_samples_leaf = {3, 8}. The RF search used leave-one-location-out validation across the six outer-training locations, negative RMSE scoring, and random_state = 42; other RF settings remained at their scikit-learn defaults.

XGBoost tuning used RandomizedSearchCV with 15 sampled parameter combinations in each outer fold. The search space was n_estimators = {100, 200, 300, 500}, max_depth = {2, 3, 4, 5}, learning_rate = {0.01, 0.03, 0.05, 0.10}, subsample = {0.6, 0.8, 1.0}, colsample_bytree = {0.6, 0.8, 1.0}, min_child_weight = {1, 3, 5}, reg_alpha = {0, 0.1, 0.5, 1.0}, and reg_lambda = {1, 2, 5, 10}. Inner validation used five-fold GroupKFold, grouped by location, with negative RMSE scoring. XGBoost used objective = reg:squarederror, tree_method = hist, and random_state = 42. For each algorithm, the hyperparameter setting with the lowest mean inner-validation RMSE was refit using all six outer-training locations before predicting at the held-out location. The search spaces and inner tuning procedures for all three algorithms are summarized in Table 2.5.

Table 2.5. Hyperparameter tuning used for the final cross-location model comparison.

Model	Inner tuning procedure	Hyperparameters/search space
PLSR	GridSearchCV; location-grouped CV; negative RMSE	n_components = 1, ..., min(10, p); median imputation + StandardScaler; PLSRegression(scale=False)
Random forest	GridSearchCV; leave-one-location-out within the six outer-training locations; negative RMSE	n_estimators = 300 or 600; max_depth = 3, 6, or None; min_samples_leaf = 3 or 8; random_state = 42
XGBoost	RandomizedSearchCV; 15 draws; 5-fold GroupKFold by location; negative RMSE	n_estimators = 100, 200, 300, 500; max_depth = 2, 3, 4, 5; learning_rate = 0.01, 0.03, 0.05, 0.10; subsample = 0.6, 0.8, 1.0; colsample_bytree = 0.6, 0.8, 1.0; min_child_weight = 1, 3, 5; reg_alpha = 0, 0.1, 0.5, 1.0; reg_lambda = 1, 2, 5, 10; objective = reg:squarederror; tree_method = hist; random_state = 42

2.8.4 Model Evaluation Metrics

Final performance was calculated from the pooled out-of-fold predictions generated for the seven held-out locations rather than from model fit to the training data or from a simple average

of fold-level statistics. The primary metrics were the coefficient of determination (R^2), Lin's concordance correlation coefficient (CCC; Lin, 1989), root mean square error (RMSE), mean absolute error (MAE), and mean bias error (Bias). Mean absolute percentage error (MAPE) was also calculated during predictor-set screening. Let y_i denote the observed value for plot i, $\hat{y}_i$ the corresponding prediction, $\bar{y}$ the mean observed value, and n the number of held-out observations. Metrics were calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (1)$$

$$\mathrm{CCC} = \frac{2\sigma_{y\hat{y}}}{\sigma_y^2 + \sigma_{\hat{y}}^2 + (\mu_y - \mu_{\hat{y}})^2} \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (3)$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad (4)$$

$$\mathrm{Bias} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \quad (5)$$

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (6)$$

For CCC, μ_γ and $\mu\hat{y}$ are the means of the observed and predicted values, σ_γ^2 and $\sigma\hat{y}^2$ are their variances, and $\sigma_\gamma\hat{y}$ is their covariance. Positive bias indicates overall overprediction and negative bias indicates underprediction. RMSE and MAE were reported in kg ha^{-1} for grain yield and percentage points for grain protein. Location-specific R^2, RMSE, MAE, and bias were also calculated to examine variation in transferability among environments. Model ranking was based on held-out-location performance.

All analyses were conducted in Python. PLSR and RF were implemented with scikit-learn (Pedregosa et al., 2011), and XGBoost was implemented with the XGBoost package (Chen & Guestrin, 2016; version 3.2.0). A fixed random seed (random_state = 42) was used for stochastic tuning procedures to support reproducibility.

2.9 Predictor Importance and Genotype Representation

Predictor importance was evaluated from the final PLSR models because PLSR provided the strongest overall cross-location performance and permits direct examination of predictor contributions. Importance statistics were calculated separately within each outer LOLO fold using only the six training locations. For each fold, the number of PLS components was re-tuned using the same inner location-grouped procedure described above. Variable importance in projection (VIP) was then calculated from the fitted PLS scores, X-weights, and Y-loadings. For predictor j, VIP was calculated as:

$$\mathrm{VIP}_j = \sqrt{p \, \frac{\sum_{a=1}^{A} \mathrm{SSY}_a \left(w_{ja}^2 / \sum_{k=1}^{p} w_{ka}^2 \right)}{\sum_{a=1}^{A} \mathrm{SSY}_a}} \tag{7}$$

where p is the number of predictors, a indexes the retained PLS components, w_{ja} is the X-weight for predictor j on component a, and SSY_a is the amount of response variation represented by component a, calculated from the component scores and Y-loadings. VIP values were averaged across the seven outer folds and their standard deviations were calculated. A mean VIP greater than 1 was used as a conventional indicator of relatively strong model contribution, but it was treated as a descriptive threshold rather than a hard rule for retaining or discarding predictors.

PLSR coefficients were extracted from the same fold-specific fitted models. Because the predictors were standardized before fitting while the response remained in its original units, these coefficients are X-standardized and describe the change in the modeled response associated with a one-standard-deviation change in the predictor, conditional on the other variables in the fitted PLSR model. Coefficient signs were recorded across all seven outer folds to assess directional stability. VIP values and standardized coefficients were interpreted as model-based predictive associations rather than causal effects. Coefficients for binary genotype variables, including

cultivar indicators and the binary agronomic trait variables, were used mainly to assess direction and relative contribution rather than as raw treatment effects.

Genotype representation was compared under the same outer LOLO framework using three alternatives: (i) no genotype information, (ii) the three agronomic cultivar traits PP, GPC, and maturity, and (iii) cultivar identity represented by eight binary indicators. Non-genotype predictors were held constant within each response trait.

All grain-yield models included N rate and GNDVI at stem elongation, whereas all grain-protein models included N rate, plant height, and MTCI at grain formation. The trait models additionally included PP, GPC, and maturity, and the cultivar-identity models additionally included the eight cultivar indicators.

Each model was refitted independently within every outer LOLO fold, including training-fold imputation, scaling, and PLS-component tuning. Pooled held-out predictions were compared using R^2, CCC, RMSE, MAE, and bias.

3 Results

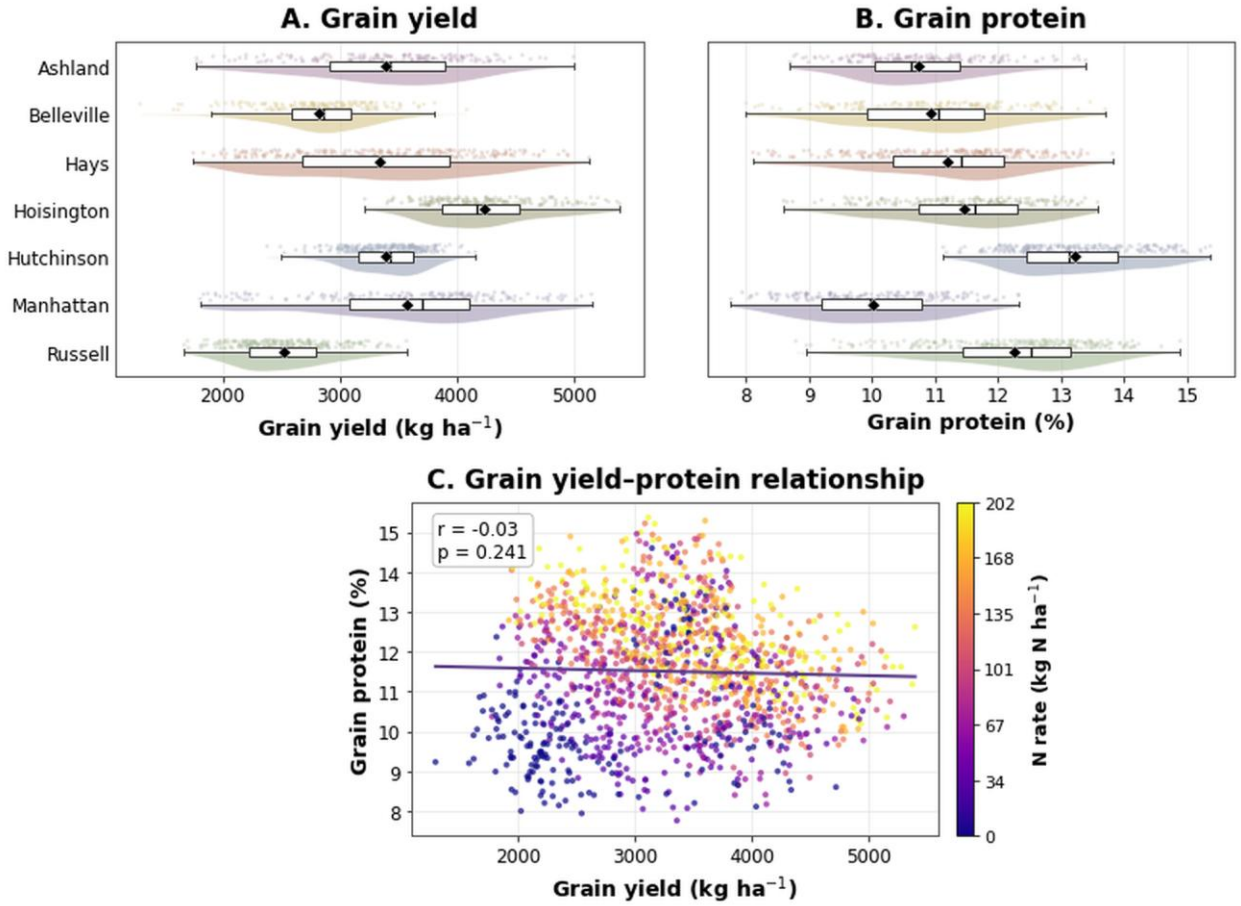
3.1 Variation in Grain Yield and Protein Across Locations

Grain yield ranged from 1,497 to 6,237 kg ha^{-1} and grain protein varied from 8.8 to 17.5% across the seven locations (Figure 2.1A, B). Hoisington had the highest grain-yield distribution, whereas Russell was among the lowest-yielding environments. Grain protein showed a different location pattern, with Hutchinson having comparatively high protein concentrations and Manhattan generally having lower values. These differences were retained in the analysis rather than removing locations with weaker or different spectral responses.

When all plot observations were pooled, grain yield and protein were not significantly related ($r = -0.03$, $P = 0.24$; Figure 2.1C). The scatterplot nevertheless showed a clear nitrogen-rate gradient in grain protein, with higher N rates generally associated with higher protein values. The absence of a corresponding overall yield-protein relationship supported treating grain yield and protein as separate prediction targets.

Figure 2.1. Distribution of (A) grain yield and (B) grain protein across seven Kansas winter wheat environments, and (C) the pooled relationship between grain yield and grain protein. In A and B, half-violin distributions show plot-level observations, boxplots show the median and interquartile

range, and black diamonds indicate location means. In C, point color represents N rate and the solid line is the fitted linear relationship.

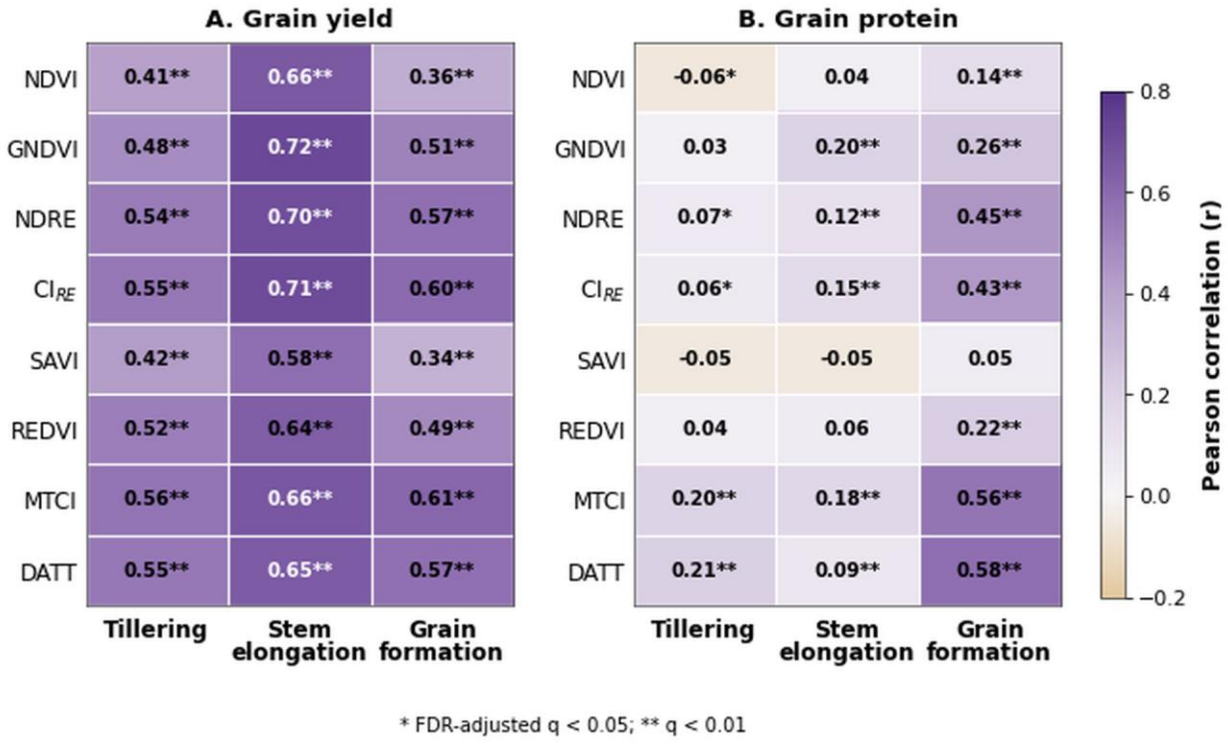


3.2 Growth-Stage-Specific Spectral Relationships

Relationships between UAV vegetation indices and final grain yield and protein changed with crop development (Figure 2.2 A & B). Across the eight indices, correlations with yield were generally moderate during tillering, strongest during stem elongation, and weaker again during grain formation. GNDVI during stem elongation had the strongest pooled correlation with grain yield (r = 0.72), followed by CI_RE (r = 0.71) and NDRE (r = 0.70). MTCI, NDVI, DATT, and REDVI also showed moderate relationships with yield during stem elongation (r = 0.64–0.66). During the later growing season at grain formation, the strongest correlations were observed for MTCI (r = 0.61) and CI_RE (r = 0.60), while NDVI and SAVI declined to r = 0.36 and 0.34, respectively.

Grain protein showed a different temporal pattern (Figure 2.2B). Most vegetation indices had weak relationships with final protein concentration during tillering and stem elongation. The spectral relationships strengthened during grain formation, particularly for the red-edge-based indices. DATT and MTCI had the strongest correlations with grain protein during grain formation (r = 0.58 and 0.56, respectively), followed by NDRE (r = 0.45) and CI_RE (r = 0.43). In contrast, NDVI and SAVI remained weakly related to protein during grain formation (r = 0.14 and 0.05, respectively).

Figure 2.2. Pooled Pearson correlations between stage-averaged UAV vegetation indices and (A) grain yield and (B) grain protein concentration across the tillering, stem elongation, and grain formation stages. Correlations are descriptive and based on pooled plot-level observations. * FDR-adjusted q < 0.05; ** q < 0.01.

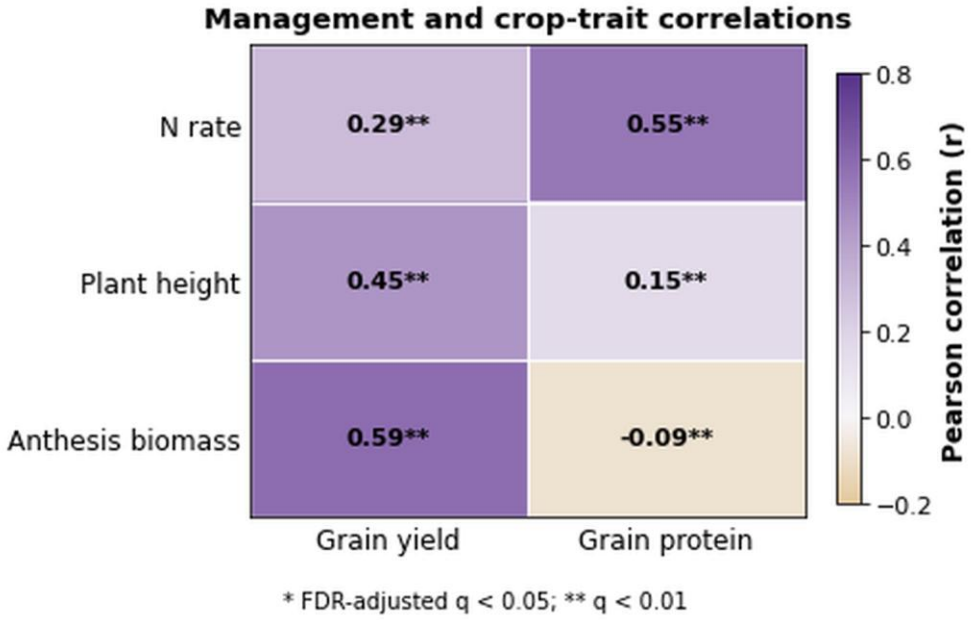


3.3 Relationships with Management and Crop Structural Traits

The management and crop structural variables were also related differently to grain yield and protein (Figure 2.3). Grain yield was positively correlated with N rate (r = 0.29), plant height

(r = 0.45), and anthesis biomass (r = 0.59). For grain protein, N rate had the strongest positive relationship (r = 0.55), while plant height had a much smaller positive correlation (r = 0.15). Anthesis biomass was weakly and negatively correlated with grain protein (r = -0.09). All six pooled correlations remained significant after false discovery rate adjustment; however, the biomass-protein relationship was weak in magnitude.

Figure 2.3. Pooled Pearson correlations of N rate, plant height, and anthesis biomass with grain yield and grain protein concentration. * FDR-adjusted q < 0.05; ** q < 0.01.



3.4 Selection of Transferable UAV and Genotype Predictors

Nested leave-one-location-out selection consistently identified GNDVI during stem elongation as the most transferable UAV predictor of grain yield. GNDVI at stem elongation was selected in all seven outer LOLO folds when evaluated together with N rate and the cultivar-level traits yield plasticity (PP), grain-protein tendency (GPC), and maturity. The resulting trait-based PLSR model achieved an overall out-of-fold R^2 of 0.545, CCC of 0.722, RMSE of 516.4 kg ha^{-1}, MAE of 401.5 kg ha^{-1}, and bias of 9.3 kg ha^{-1}. Thus, GNDVI during stem elongation was retained as the common spectral predictor in the final grain-yield model.

For grain protein, useful spectral information was concentrated during grain formation. MTCI at grain formation was selected in five of the seven outer LOLO folds, whereas DATT was selected in the remaining two. The fully nested single-index procedure produced an R^2 of 0.45 and RMSE of 1.11 percentage points. Using MTCI at grain formation as the common spectral predictor together with N rate, cultivar identity, and plant height resulted in an R^2 of 0.470, CCC of 0.668, RMSE of 1.089 percentage points, MAE of 0.861 percentage points, and bias of −0.051 percentage points. Although adding DATT provided a small numerical improvement, MTCI and DATT were highly collinear; therefore, MTCI alone was retained in the final parsimonious protein model. A summary of nested predictor selection and final parsimonious model performance is provided in Table 2.6.

Table 2.6. Nested leave-one-location-out predictor selection and performance of the final parsimonious PLSR models for winter wheat grain yield and protein.

Trait	Baseline predictors	Nested UAV selection	Final fixed model	Final LOLO performance
Grain yield	N rate + PP + GPC + maturity	GNDVI at stem elongation (7/7 folds)	N rate + PP + GPC + maturity + GNDVI at stem elongation	R^2 = 0.545; CCC = 0.722; RMSE = 516.4 kg ha^{-1}; MAE = 401.5 kg ha^{-1}
Grain protein	N rate + cultivar identity + plant height	MTCI at grain formation (5/7 folds); DATT (2/7 folds)	N rate + cultivar identity + plant height + MTCI at grain formation	R^2 = 0.470; CCC = 0.668; RMSE = 1.089 percentage points; MAE = 0.861 percentage points

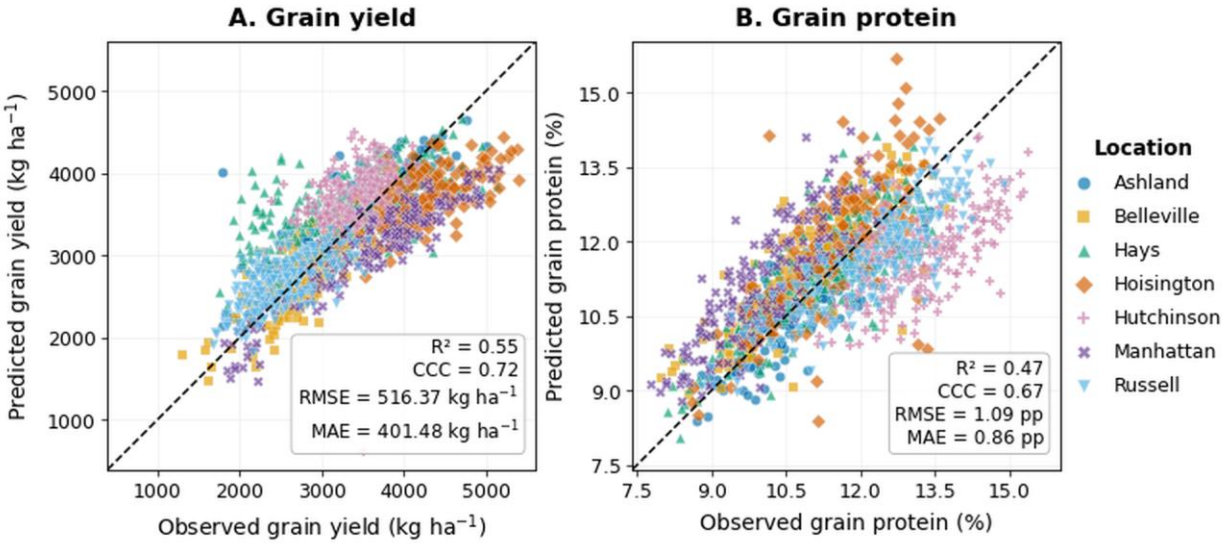
3.5 Cross-Location Prediction and Algorithm Comparison

Across the three evaluated algorithms, PLSR provided the strongest overall cross-location prediction for both grain yield and grain protein. For grain yield, the final trait-based PLSR model achieved R^2 = 0.55, CCC = 0.72, RMSE = 516.4 kg ha^{-1}, and MAE = 401.5 kg ha^{-1}. XGBoost produced slightly lower performance, with R^2 = 0.51, CCC = 0.69, RMSE = 536.8 kg ha^{-1}, and MAE = 422.4 kg ha^{-1}. Random forest had the lowest overall yield performance, with R^2 = 0.447, CCC = 0.66, RMSE = 569.2 kg ha^{-1}, and MAE = 450.0 kg ha^{-1}.

The same overall ranking was observed for grain protein. PLSR achieved $R^2 = 0.47$, CCC = 0.67, RMSE = 1.09 percentage points, and MAE = 0.86 percentage points. XGBoost achieved $R^2 = 0.33$, CCC = 0.57, RMSE = 1.23 percentage points, and MAE = 0.98 percentage points, whereas random forest achieved $R^2 = 0.31$, CCC = 0.56, RMSE = 1.25 percentage points, and MAE = 0.99 percentage points. Thus, the nonlinear tree-based models did not improve overall prediction when an entire location was excluded from model development.

Observed-versus-predicted relationships from the final PLSR models showed closer agreement with the 1:1 line for grain yield than for grain protein, consistent with the higher concordance observed for yield. Overall prediction bias was small for both response traits, indicating little systematic over- or underprediction across the combined held-out locations.

Figure 2.4. Observed versus leave-one-location-out predicted values for the final PLSR models: (A) grain yield and (B) grain protein. Each observation was predicted when its corresponding location was completely excluded from model training. The dashed line represents the 1:1 relationship.



3.6 Predictor Importance and Genotype Representation

PLSR variable-importance results were consistent across the seven outer LOLO folds (Figure 2.5 A, B). For grain yield, GNDVI at stem elongation was the dominant predictor, with a

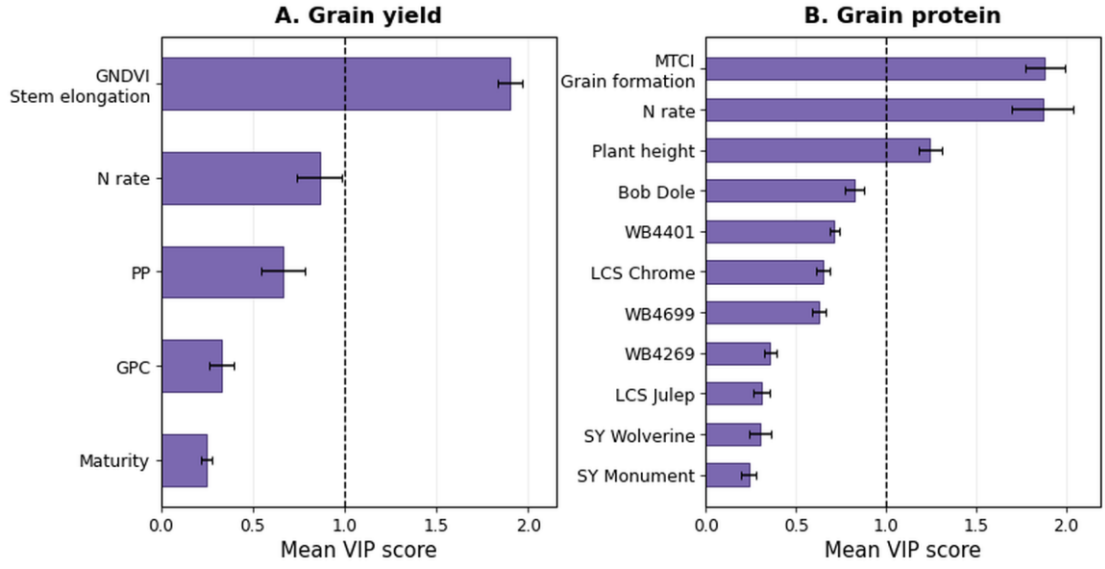
mean VIP of 1.90 ± 0.07. N rate had the next highest VIP (0.86 ± 0.12), followed by yield plasticity (PP; 0.66 ± 0.12), grain-protein tendency (GPC; 0.33 ± 0.07), and maturity class (0.25 ± 0.03). Although only GNDVI exceeded the conventional VIP = 1 reference threshold, all three cultivar traits were retained together in all seven nested LOLO folds. The X-standardized coefficients were directionally stable across folds: GNDVI at stem elongation, N rate, and PP were positive in all seven folds, whereas GPC and maturity were negative in all seven folds. Mean coefficients were 540.59 ± 52.81 kg ha^{-1} for GNDVI, 105.85 ± 16.59 kg ha^{-1} for N rate, 215.01 ± 32.32 kg ha^{-1} for PP, −116.55 ± 13.89 kg ha^{-1} for GPC, and −77.92 ± 7.98 kg ha^{-1} for maturity per one-standard-deviation increase in the corresponding predictor.

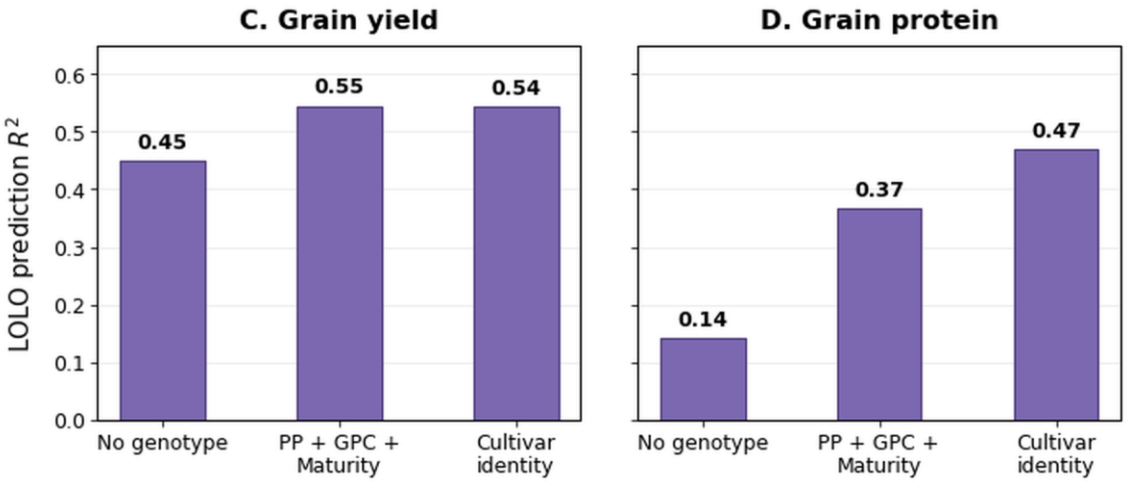
For grain protein, MTCI at grain formation (VIP = 1.88 ± 0.11), N rate (VIP = 1.87 ± 0.17), and plant height (VIP = 1.24 ± 0.06) exceeded the VIP = 1 reference threshold. MTCI and N rate had positive X-standardized coefficients in all seven folds (1.02 ± 0.16 and 0.40 ± 0.10 percentage points, respectively), whereas plant height had a negative coefficient in all seven folds (−0.73 ± 0.08 percentage points). Individual cultivar indicators had VIP values below 1, although several showed consistent coefficient directions across folds.

Genotype representation affected grain yield and grain protein predictions differently (Figure 2.5C, D). For grain yield, the model without genotype information achieved LOLO R^2 = 0.45, CCC = 0.64, RMSE = 568.05 kg ha^{-1}, and MAE = 437.77 kg ha^{-1}. Adding PP, GPC, and maturity increased R^2 to 0.55 and CCC to 0.72, while reducing RMSE to 516.37 kg ha^{-1} and MAE to 401.48 kg ha^{-1}. This performance was essentially equivalent to the cultivar-identity model (R^2 = 0.54, CCC = 0.72, RMSE = 517.71 kg ha^{-1}, MAE = 403.94 kg ha^{-1}), indicating that the three agronomic traits captured most of the genotype-related information relevant to grain-yield prediction.

For grain protein, genotype representation had a larger effect. The no-genotype model achieved $R^2 = 0.14$, CCC = 0.40, RMSE = 1.39 percentage points, and MAE = 1.15 percentage points. Adding PP, GPC, and maturity improved prediction to $R^2 = 0.37$, CCC = 0.59, RMSE = 1.19 percentage points, and MAE = 0.97 percentage points. However, retaining cultivar identity produced the strongest protein prediction, with $R^2 = 0.47$, CCC = 0.67, RMSE = 1.09 percentage points, and MAE = 0.86 percentage points. Thus, broad agronomic cultivar traits were sufficient to represent genotype effects for grain yield but did not fully capture cultivar-specific information relevant to grain protein concentration.

Figure 2.5. Predictor importance and genotype representation in the final partial least squares regression (PLSR) models for grain yield and grain protein. Mean variable importance in projection (VIP) scores are shown for (A) grain yield and (B) grain protein; error bars represent the standard deviation across the seven outer leave-one-location-out (LOLO) folds, and the dashed vertical line indicates VIP = 1. Panels (C) and (D) compare LOLO prediction R^2 among models containing no genotype information, agronomic cultivar traits (yield plasticity, PP; grain-protein tendency, GPC; and maturity), or cultivar identity for grain yield and grain protein, respectively. For grain yield, the trait-based and cultivar-identity models produced similar predictive performance, whereas cultivar identity provided greater predictive value than the three agronomic traits for grain protein.





4 Discussion

The main objective of this study was to determine how well UAV-based spectral and agronomic predictors of winter wheat grain yield and protein concentration transfer to a new location that was not represented during model training. Under the leave-one-location-out (LOLO) framework, the final PLSR model achieved an R^2 of 0.55, CCC of 0.72, and RMSE of 516.40 kg ha^{-1} for grain yield, while the grain-protein model achieved an R^2 of 0.47, CCC of 0.67, and RMSE of 1.09 percentage points. These accuracies are somewhat lower than those reported in several previous UAV-based wheat studies (Walsh et al., 2022; Zhou et al., 2021; Bian et al., 2022), but those studies generally used less restrictive validation approaches. For example, Billah et al. (2026) reported R^2 values of 0.65 for grain yield and 0.61 for grain protein using multitemporal UAV imagery and deep learning across seven South Dakota wheat sites. However, their data were randomly split into 70% for training and 30% for testing, allowing observations from the same experimental sites to appear in both sets. Similarly, Zhou et al. (2021) reported validation R^2 values of approximately 0.62 for yield and 0.63 for protein, but their observations were also randomly divided into training and testing sets. Such approaches primarily evaluate prediction within environments already represented during model development and may therefore overestimate performance when models are applied to a completely unseen location.

A recent study conducted in Oklahoma across 57 site-years achieved substantially higher accuracy in predicting wheat yield with CatBoost ($R^2 = 0.81$, RMSE = 0.62 t ha^{-1}) (Maatougui et al., 2026). Their evaluation was more rigorous than a random plot-level split because complete site-years were assigned to either training or testing groups. However, several differences likely contributed to their higher accuracy. Their dataset contained 57 site-years, with 45 available for model training and 12 held out for independent testing, providing considerably greater environmental diversity than the seven locations from a single growing season evaluated in the present study. They also included a broader range of predictors, including early-season NDVI, pre-plant N rate, soil texture, temperature, precipitation, evapotranspiration, growing degree days, and water balance. In addition, because validation was conducted by site-year rather than by completely withholding observations from a given physical location, observations from another year at the same location could remain in the training dataset. In contrast, the LOLO framework used here excluded all observations from the target location during preprocessing, predictor selection, model tuning, and prediction. Therefore, the lower predictive accuracy observed in the present study should be interpreted in the context of a more stringent spatial-transfer problem rather than simply as poorer model performance.

Weather and soil variables were also evaluated as additional predictors during model development but were not retained in the final parsimonious models because they did not improve cross-location prediction under LOLO validation. Their limited contribution may partly reflect the small number of independent site-level environments, because these variables varied primarily among the seven locations rather than among individual plots. Under LOLO validation, relationships involving these variables were estimated from only six training environments and then applied to the held-out location, which may have limited their stability and transferability.

Thus, adding site-level environmental variables increased model complexity without providing sufficient improvement in prediction of the unseen location. Future studies incorporating additional site-years should further evaluate whether environmental information available before or at the time of sensing, such as cumulative growing degree days, precipitation before stem elongation, and soil nitrate at sowing, can improve prediction across environments. Previous studies have shown that combining remotely sensed vegetation indices with weather and management information can improve wheat-yield prediction, suggesting that environmental variables may become more useful when substantially greater environmental diversity is available for model development.

Studies using location- or field-level holdout validation further illustrate the difficulty of developing models that generalize across environments. Kaushal et al. (2026) evaluated UAV-based winter wheat yield prediction using a LOLO approach in which models were trained on two South Dakota locations and tested at a third. Without information from the target location, cross-location R^2 values ranged from -15.30 to -0.40, demonstrating the difficulty of spatial transferability. Prediction accuracy increased to approximately 0.29-0.69 only after 40% of observations from the target location were used for fine-tuning. Tilse et al. (2025) reached a similar conclusion for wheat yield and grain protein, with prediction accuracy declining when an entire field-year was excluded compared with situations in which part of the same field was available for model calibration. In this context, the current LOLO prediction accuracies for grain yield (R^2 = 0.55) and grain protein (R^2 = 0.47) are encouraging, as they were achieved without using any observations from the target location during model development. This distinction is particularly important for practical deployment, where predictions may be required in fields for which no local calibration data are available.

The timing of the most informative UAV predictors differed markedly between grain yield and grain protein concentration. GNDVI during stem elongation was the most stable spectral predictor of grain yield and was selected in all seven outer LOLO folds. The pooled correlation analysis similarly showed that GNDVI during stem elongation had the strongest relationship with final grain yield ($r = 0.72$). In contrast, grain-protein relationships with UAV indices were relatively weak during tillering and stem elongation but strengthened during grain formation, particularly for red-edge-based indices. DATT and MTCI during grain formation had the strongest pooled relationships with final protein concentration ($r = 0.58$ and 0.56, respectively), and MTCI provided the most stable spectral contribution to the final protein model. These results indicate that grain yield and grain protein concentration are associated with canopy information expressed at different developmental stages and support the use of phenology-specific rather than single-date remote-sensing predictors.

The physiological basis for these stage-specific relationships is also consistent with wheat development. Much of the variation in final grain yield is associated with grain number, which is largely determined during stem elongation, as spike growth, floret development, and floret survival respond to crop growth, radiation capture, and N availability before anthesis (Fischer, 1985; Slafer et al., 2014). GNDVI is sensitive to canopy greenness and chlorophyll status and therefore integrates aspects of crop vigor and N status during the period when yield potential is being established. This provides a physiological basis for the strong predictive contribution of GNDVI during stem elongation. Grain protein concentration, in contrast, is influenced by the amount of N accumulated and subsequently remobilized to the grain relative to carbohydrate deposition during grain filling. A substantial proportion of grain N originates from remobilization of N stored in vegetative tissues before anthesis, with post-anthesis N uptake contributing the remainder (Kichey

et al., 2007; Bogard et al., 2010). Consequently, red-edge indices measured during grain formation may provide useful information about canopy chlorophyll status, N retention, and the progression of senescence during the period when N is being transferred to developing grain. The importance of MTCI during grain formation is therefore consistent with its sensitivity to red-edge spectral behavior and canopy chlorophyll status. This interpretation also agrees with evidence that late-season N management has a substantially greater influence on grain protein concentration than on grain yield (Giordano et al., 2023).

The genotype-representation analysis provided additional insight into how cultivar information contributed to transferable prediction. For grain yield, the model without genotype information achieved an R^2 of 0.45, whereas inclusion of the cultivar-level traits yield plasticity (PP), grain-protein tendency (GPC), and maturity increased R^2 to 0.55. This performance was essentially equivalent to that obtained when cultivar identity was represented directly using cultivar indicators ($R^2 = 0.54$). Thus, for grain yield, the measured cultivar traits captured nearly all of the predictive information associated with cultivar identity. This finding is important because trait-based genotype representation is more biologically interpretable than categorical cultivar labels and may offer greater potential for application to cultivars not represented in the original model-development dataset. Rather than requiring a model to recognize a specific cultivar name, a trait-based approach may allow cultivars with comparable agronomic characteristics to be represented by common biological descriptors.

For grain protein, however, cultivar identity retained substantially greater predictive information than the three cultivar traits evaluated. The protein model without genotype information achieved an R^2 of only 0.14. Adding PP, GPC, and maturity increased R^2 to 0.37, indicating that these traits captured a substantial portion of the genotype-related variation in protein

concentration. However, representing genotype using cultivar identity further increased R^2 to 0.47. Therefore, the available cultivar traits did not completely explain the cultivar-specific information associated with grain protein. This result agrees with previous research in Kansas demonstrating that genotype-specific differences in N uptake, N remobilization, and N partitioning can strongly influence final grain protein concentration and end-use quality (Raj et al., 2023; Giordano et al., 2024; Lollato et al., 2021). Protein concentration is a complex genotype-dependent trait, and physiological characteristics associated with N uptake, post-anthesis N acquisition, remobilization efficiency, senescence, and grain filling may not be completely represented by the relatively broad PP, GPC, and maturity classes evaluated here.

The contrast between yield and protein suggests that genotype information need not be represented in the same way across different prediction targets. For grain yield, biologically interpretable cultivar traits were sufficient to reproduce essentially the same predictive performance as cultivar identity. For grain protein, however, cultivar-specific information beyond the three evaluated traits remained important. This distinction has implications for developing more generalizable prediction models. Trait-based representation may offer greater potential for extending yield models to newly released cultivars, whereas protein models may require a richer set of physiological, phenological, or breeding-related genotype descriptors. Future studies should therefore evaluate whether quantitative cultivar characteristics, such as heading date, plant-height potential, grain-protein deviation, N uptake efficiency, remobilization efficiency, or stay-green behavior, can explain additional genotype-specific variation currently represented indirectly by cultivar identity.

The superior overall performance of PLSR relative to RF and XGBoost further suggests that greater algorithmic complexity did not improve spatial transferability in this dataset. The final

models contained a relatively small number of biologically related predictors, several of which represented correlated spectral, management, and crop-development processes. Under these conditions, PLSR may have benefited from summarizing correlated information via latent components while avoiding the additional flexibility of tree-based models, which can capture location-specific nonlinear relationships that do not generalize well to unseen environments. Thus, model generalizability depended not only on the choice of algorithm but also on the selection of biologically meaningful predictors and on evaluation within a validation framework that realistically represented the intended application. Overall, combining phenology-specific UAV spectral information with N management, crop structural measurements, and appropriately represented genotype information provided useful prediction of winter wheat grain yield and protein concentration across previously unseen Kansas environments.

5 Limitations and Future Research

This study was conducted during a single growing season, which limits direct evaluation of year-to-year transferability. Weather is an important source of interannual variation in winter wheat yield and grain protein concentration; however, the seven study locations represented a broad range of environmental conditions within the 2024–2025 growing season. Cumulative growing-season precipitation ranged from approximately 318 mm at Hoisington to 540 mm at Ashland Bottoms. Mean maximum air temperature during the warmest period ranged from approximately 26.2 to 27.7 °C among locations, while mean minimum air temperature during December–February ranged from approximately −9.0 to −6.8 °C. The locations also differed substantially in soil texture, organic matter, pH, and background soil N supply. Thus, although year-to-year variability was not directly represented, the spatial design exposed the models to considerable variation in growing conditions across environments. Because each LOLO model had

to predict an environment it had not seen, much of the environmental variation that would normally be sampled across years was represented across locations. Temporal transferability was not tested directly, however, and seasons with conditions outside the range observed here could reduce prediction accuracy. The experiments were also conducted in replicated research plots with imposed N rates and cultivars, but their applicability in real farmers' fields still needs to be tested. Therefore, multi-year testing in commercial fields with field-scale soil and weather information is needed.

The models developed here were designed to evaluate transferability among locations represented within the same regional production system. Their performance in substantially different soil, climatic, and management conditions remains unknown. Future research should therefore test these models across broader geographic regions and investigate whether locally calibrated or transfer-learning approaches can improve prediction when applied to environments that differ substantially from those used for model development.

6 Conclusion

This study evaluated whether UAV multispectral imagery combined with agronomic and genotype-related information could support cross-location prediction of winter wheat grain yield and grain protein across seven Kansas environments. The results showed clear differences in both the timing and type of information associated with the two response traits. GNDVI during stem elongation was the most stable spectral predictor of grain yield and was selected in all seven outer LOLO folds, whereas MTCI during grain formation provided the most consistent spectral contribution to grain protein when combined with N rate, cultivar information, and plant height.

The final PLSR models achieved $R^2 = 0.55$ and CCC = 0.72 for grain yield and $R^2 = 0.47$ and CCC = 0.67 for grain protein. PLSR outperformed RF and XGBoost under the same held-out-

location framework, indicating that greater model complexity did not improve spatial transferability in this dataset.

Genotype representation also differed in importance between grain yield and protein. For grain yield, cultivar-level agronomic traits provided essentially the same predictive performance as cultivar identity, indicating that biologically interpretable traits can effectively represent much of the genotype-related information needed for yield prediction. For grain protein, however, cultivar identity retained additional predictive information beyond the evaluated cultivar traits, suggesting that protein prediction depends on genotype-specific characteristics not fully captured by yield plasticity, grain-protein tendency, and maturity.

Taken together, these findings demonstrate that UAV-based spectral information is most useful when combined with complementary agronomic and genotype-related predictors, and that the optimal set of predictors depends on the target trait. The LOLO framework provided evidence of spatial transferability among the seven Kansas environments represented in this study; however, validation across additional years, commercial fields, cultivars, and environmental conditions will be required before the models can be considered broadly transferable.

References

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R. J., Li, H., & Moran, M. S. (2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In P. C. Robert, R. H. Rust, & W. E. Larson (Eds.), Proceedings of the Fifth International Conference on Precision Agriculture. ASA, CSSA, and SSSA, Madison, WI.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B (Methodological), 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bian, C., Shi, H., Wu, S., Zhang, K., Wei, M., Zhao, Y., Sun, Y., Zhuang, H., Zhang, X., & Chen, S. (2022). Prediction of field-scale wheat yield using machine learning method and

multi-spectral UAV data. Remote Sensing, 14(6), 1474. https://doi.org/10.3390/rs14061474

Billah, M. M., Maimaitijiang, M., Kaushal, S., Millett, B., Khan, S. N., Halder, J., Kleinjan, J., & Sehgal, S. K. (2026). Predicting wheat yield and grain quality with UAV multispectral imagery and deep learning. Frontiers in Plant Science, 17, 1812052. https://doi.org/10.3389/fpls.2026.1812052

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J.-M., Jeuffroy, M.-H., Gate, P., Martre, P., & Le Gouis, J. (2010). Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. Journal of Experimental Botany, 61(15), 4303–4312. https://doi.org/10.1093/jxb/erq238

Breiman, L. (2001). Random forests. Machine Learning, 45, 5–32. https://doi.org/10.1023/A:1010933404324

Cao, Q., Miao, Y., Feng, G., Gao, X., Li, F., Liu, B., Yue, S., Cheng, S., Ustin, S. L., & Khosla, R. (2015). Active canopy sensing of winter wheat nitrogen status: An evaluation of two sensor systems. Computers and Electronics in Agriculture, 112, 54–67. https://doi.org/10.1016/j.compag.2014.08.012

Cao, Q., Miao, Y., Wang, H., Huang, S., Cheng, S., Khosla, R., & Jiang, R. (2013). Non-destructive estimation of rice plant nitrogen status with Crop Circle multispectral active canopy sensor. Field Crops Research, 154, 133–144. https://doi.org/10.1016/j.fcr.2013.08.005

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). https://doi.org/10.1145/2939672.2939785

Cossani, C. M., & Sadras, V. O. (2021). Nitrogen and water supply modulate the effect of elevated temperature on wheat yield. European Journal of Agronomy, 124, 126227.

Curtis, T., & Halford, N. G. (2014). Food security: The challenge of increasing wheat yield and the importance of not compromising food safety. Annals of Applied Biology, 164(3), 354–372. https://doi.org/10.1111/aab.12108

Dash, J., & Curran, P. J. (2004). The MERIS terrestrial chlorophyll index. International Journal of Remote Sensing, 25(23), 5403–5413. https://doi.org/10.1080/0143116042000274015

Datt, B. (1999). A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using Eucalyptus leaves. Journal of Plant Physiology, 154(1), 30–36. https://doi.org/10.1016/S0176-1617(99)80314-9

Fischer, R. A. (1985). Number of kernels in wheat crops and the influence of solar radiation and temperature. Journal of Agricultural Science, 105(2), 447–461. https://doi.org/10.1017/S0021859600056495

Fu, Z., Yu, S., Zhang, J., Xi, H., Gao, Y., Lu, R., Zheng, H., Zhu, Y., Cao, W., & Liu, X. (2022). Combining UAV multispectral imagery and ecological factors to estimate leaf nitrogen and grain protein content of wheat. European Journal of Agronomy, 132, 126405. https://doi.org/10.1016/j.eja.2021.126405

Giordano, N., Sadras, V. O., & Lollato, R. P. (2023). Late-season nitrogen application increases grain protein concentration and is neutral for yield in wheat: A global meta-analysis. Field Crops Research, 290, 108740. https://doi.org/10.1016/j.fcr.2022.108740

Giordano, N., Sadras, V. O., Correndo, A. A., & Lollato, R. P. (2024). Cultivar-specific phenotypic plasticity of yield and grain protein concentration in response to nitrogen in winter wheat. Field Crops Research, 306, 109202.

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, 58(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. Geophysical Research Letters, 32, L08403. https://doi.org/10.1029/2005GL022688

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment, 25(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Jaenisch, B. R., Giordano, N., Sadras, V. O., Nelson, N. O., Guttieri, M. J., Fritz, A. K., Ruiz Diaz, D., & Lollato, R. P. (2026). Nitrogen and sulfur co-limitation and stoichiometry modulate yield and nutrient use efficiency of bread wheat. Field Crops Research, 344, 110539. https://doi.org/10.1016/j.fcr.2026.110539

Joshi, D. R., Clay, S. A., Sharma, P., Rekabdarkolaee, H. M., Kharel, T., Rizzo, D. M., Thapa, R., & Clay, D. E. (2024). Artificial intelligence and satellite-based remote sensing can be used to predict soybean (Glycine max) yield. Agronomy Journal, 116(3), 917–930. https://doi.org/10.1002/agj2.21473

Kaushal, S., Billah, M. M., Singh, M., Koupal, D., Sitaula, P. R., Halder, J., Sidhu, J., Maimaitijiang, M., & Sehgal, S. K. (2026). UAV-based deep transfer learning to improve grain yield prediction in winter wheat across temporal and spatial variability. The Plant Phenome Journal, 9(1), e70085. https://doi.org/10.1002/ppj2.70085

Kichey, T., Hirel, B., Heumez, E., Dubois, F., & Le Gouis, J. (2007). In winter wheat (Triticum aestivum L.), post-anthesis nitrogen uptake and remobilisation to the grain correlates with agronomic traits and nitrogen physiological markers. Field Crops Research, 102(1), 22–32. https://doi.org/10.1016/j.fcr.2007.01.002

Large, E. C. (1954). Growth stages in cereals: Illustration of the Feekes scale. Plant Pathology, 3(4), 128–129.

Larroque, O. R., Calderini, D. F., & Angus, J. F. (2022). Managing dryland wheat to produce high-quality grain. Field Crops Research, 280, 108473.

Lin, L. I. (1989). A concordance correlation coefficient to evaluate reproducibility. Biometrics, 45(1), 255–268. https://doi.org/10.2307/2532051

Lollato, R. P., & Edwards, J. T. (2015). Maximum attainable wheat yield and resource-use efficiency in the southern Great Plains. Crop Science, 55(6), 2863–2876.

Lollato, R. P., Jaenisch, B. R., & Silva, S. R. (2021). Genotype-specific nitrogen uptake dynamics and fertilizer management explain contrasting wheat protein concentration. Crop Science, 61(3), 2048–2066. https://doi.org/10.1002/csc2.20442

Lucero, M., Adee, E., Ruiz Diaz, D., Sullivan, T., Jha, G., Simon, L., & Joshi, D. R. (2026). Machine learning insights into long-term soil fertility legacy and seasonal weather effects in corn–soybean rotation. Soil Science Society of America Journal, 90(5), e70334. https://doi.org/10.1002/saj2.70334

Maatougui, M., Phillips, S., Abiola, S. O., & Arnall, D. B. (2026). Wheat yield prediction using machine learning, proximal sensing, nitrogen and climate data for Oklahoma. Precision Agriculture, 27, 148. https://doi.org/10.1007/s11119-026-10462-2

Maranhão, R. L. A., Caldas, M. M., Kastens, J., Watson, J., & Lollato, R. P. (2025). Assessing NDVI, climate, and management to predict winter wheat yields at field scale in Kansas, USA. Remote Sensing, 17(20), 3500.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

Prey, L., & Schmidhalter, U. (2019). Sensitivity of vegetation indices for estimating vegetative N status in winter wheat. Sensors, 19(17), 3712. https://doi.org/10.3390/s19173712

Raj, A. S., Siliveru, K., McLean, R., Prasad, P. V. V., & Lollato, R. P. (2023). Intensive management simultaneously reduces yield gaps and improves milling and baking properties of bread wheat. Crop Science, 63(2), 936–955.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K. W., Thomason, W. E., & Lukina, E. V. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agronomy Journal, 94(4), 815–820. https://doi.org/10.2134/agronj2002.8150

Rouse, J. W., Jr., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In Third Earth Resources Technology Satellite-1 Symposium (Vol. 1, pp. 309–317). NASA SP-351.

Sharma, P., Leigh, L., Chang, J., Maimaitijiang, M., & Caffé, M. (2022). Above-ground biomass estimation in oats using UAV remote sensing and machine learning. Sensors, 22(2), 601. https://doi.org/10.3390/s22020601

Slafer, G. A., Savin, R., & Sadras, V. O. (2014). Coarse and fine regulation of wheat yield components in response to genotype and environment. Field Crops Research, 157, 71–83. https://doi.org/10.1016/j.fcr.2013.12.004

Tilse, M. J., Bishop, T. F. A., & Filippi, P. (2025). Predicting within-field grain protein content at scale using agronomic and remote sensing variables, and machine learning. Precision Agriculture, 26, 78. https://doi.org/10.1007/s11119-025-10267-9

Walsh, O. S., Marshall, J., Jackson, C., Nambi, E., Shafian, S., Jayawardena, D. M., Lamichhane, R., Owusu Ansah, E., & McClintick-Chess, J. R. (2022). Wheat yield and protein estimation with handheld- and UAV-based reflectance measurements. Agrosystems, Geosciences & Environment, 5(4), e20309. https://doi.org/10.1002/agg2.20309

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. Chemometrics and Intelligent Laboratory Systems, 58(2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1

Zhou, X., Kono, Y., Win, A., Matsui, T., & Tanaka, T. S. T. (2021). Predicting within-field variability in grain yield and protein content of winter wheat using UAV-based multispectral imagery and machine learning approaches. Plant Production Science, 24(2), 137–151. https://doi.org/10.1080/1343943X.2020.1819165

Chapter 3 - UAV-Based Nitrogen Stress Detectability in Winter Wheat Across Contrasting Soil Nitrogen Environments

Abstract

Precision nitrogen (N) management requires reliable detection of crop response to fertilizer N across environments with contrasting background soil N supply. This study evaluated UAV multispectral detection of fertilizer-N response in winter wheat across seven Kansas environments during the 2024–2025 growing season. Eight hard red winter wheat cultivars (Bob Dole, LCS Chrome, LCS Julep, SY Monument, SY Wolverine, WB4269, WB4401, and WB4699) received seven N rates (0, 34, 67, 101, 135, 168, and 202 kg N ha^{-1}) in a split-plot design. High-resolution UAV-based multispectral imagery was acquired during the six different growth stages at each location, and eight vegetation indices, including non-red-edge indices (NDVI, SAVI, GNDVI) and red-edge-based indices (NDRE, CI_RE, REDVI, MTCI, DATT), were summarized into three growth stages of winter wheat: tillering (Feekes 3–5), stem elongation (Feekes 6–9), and grain formation (Feekes 10–11.4). Spectral N detectability was quantified using a standardized N-rate response from cluster-robust regression. Profile soil NO_3-N ranged from 29 to 270 kg N ha^{-1}, while grain-yield response to fertilizer N ranged from −0.7% to 105%. Detectability increased with crop development, with the clearest advantage of red-edge-based VIs occurring during stem elongation, when NDRE, CI_{RE}, MTCI, and DATT were more detectable than NDVI. Greater background soil NO_3-N tended to be associated with weaker yield response and weaker red-edge detectability (ρ = −0.68, P = 0.094). Overall red-edge detectability closely reflected agronomic responsiveness, with identical rank ordering among environments and relative yield response (ρ = 1.00, P < 0.001). These findings indicate that red-edge-based UAV VIs provide sensitive measures of fertilizer-N response, but their interpretation depends on local soil N supply and agronomic responsiveness.

This is particularly relevant for on-farm N management, where spatial variability in soil N supply can cause different areas of the same field to respond differently to fertilizer N. Integrating soil N information with red-edge UAV sensing and a local N-rich reference may therefore help distinguish N-responsive areas from those where additional fertilizer is less likely to provide agronomic benefit.

1 Introduction

Nitrogen (N) fertilizer management strongly influences winter wheat canopy development, grain yield, and grain protein. Both insufficient and excessive N application can reduce the sustainability and profitability of wheat production. For instance, nitrogen deficiency limits crop growth by reducing chlorophyll concentration, leaf area development, and biomass accumulation (Hirel et al. (2007) and Xu et al. (2012)), whereas excessive N application can lower N-use efficiency and increase the risk of environmental N losses (Raun and Johnson, 1999; Hawkesford, 2014). Therefore, using a precision N management approach in wheat production is important, potentially utilizing various sensing methods for accurate, timely detection of N stress to guide fertilizer application at the appropriate rate and time. Such an assessment can support in-season fertilizer decisions and improve N-use efficiency by accounting for the spatial and temporal variability in crop N status. Multiple approaches, including the application of remote sensing using either satellite or unmanned aerial vehicle (UAV) platforms, provide a practical means of monitoring crop N responses, as changes in chlorophyll concentration, green leaf area, and canopy structure alter reflectance in the visible, red-edge, and near-infrared regions (Singh et al., 2025; Joshi et al., 2024; Kharel et al., 2025). Vegetation indices (VIs) derived from these spectral regions have consequently been used extensively for in-season assessment of wheat N status, fertilizer response, and recommendation decision (Raun et al., 2005; Li et al., 2014).

Many vegetation indices have been used for crop N sensing, but their sensitivity depends on spectral formulation, crop developmental stage, biomass, canopy density, and crop N status. Sensor-based N decision systems also tend to perform best when canopy measurements are interpreted together with other information on soil, crop, and environment rather than in isolation (Colaço and Bramley, 2018; Colaço et al., 2021). Red- and red-edge-based indices are among the most used for evaluating crop N status. The normalized difference vegetation index (NDVI), which uses red and near-infrared reflectance, is widely used because of its simplicity and robustness. Its sensitivity, however, can decline at high wheat biomass production conditions due to spectral saturation (Li et al., 2010; Cao et al., 2015). Therefore, many studies have overcome this saturation limitation by using Red-edge-based VIs, which are generally less susceptible to saturation and remain responsive to changes in chlorophyll and crop N status under dense canopy conditions (Cao et al., 2015; Prey and Schmidhalter, 2019). For example, Cao et al. (2015) reported that NDRE/GOSAVI were more sensitive than conventional NDVI before aboveground biomass reached approximately 5,000 kg ha^{-1}, whereas the red-edge chlorophyll index CI_{RE} retained greater sensitivity after biomass exceeded approximately 5,000 kg ha^{-1}. Their results, therefore, demonstrated the advantage of red-edge information as wheat canopy biomass increased. Similarly, Prey and Schmidhalter (2019) found that red-edge-based indices were particularly advantageous under higher N-status conditions; for plants with nitrogen nutrition index (NNI > 0.8), their best red-edge index achieved mean R^2 values of approximately 0.46 for NNI and 0.48 for N uptake, supporting the greater usefulness of red-edge information when conventional red–NIR indices become less sensitive. Most previous comparisons among vegetation indices, however, have focused on correlations or prediction accuracy for plant N concentration, N uptake,

biomass, or N nutrition index. Less attention has been given to how strongly individual indices distinguish an imposed fertilizer-N response across contrasting field environments.

Another approach to access crop response to fertilizer N is to use a sufficiency index (SI), commonly calculated as the ratio of the VI measured in the under-fertilized treatment to that measured in the N-rich reference (Raun et al., 2002; Flowers et al., 2003; Holland and Schepers, 2010). When SI approaches 1, it indicates that the spectral condition of the under-fertilized crop is similar to that of the N-rich reference or N sufficiency condition, whereas lower values indicate increasing spectral difference and potentially indicate N deficiency in the crop (Guillard et al., 2021). Because the sensitivity of UAV-based VIs to crop N status is influenced by the background environment in which the crop is growing, spectral responses can vary with soil conditions, crop growth, and other environmental factors (Clay et al., 2006, 2012; Li et al., 2024). Therefore, using SI can reduce these background environmental effects and provide a more locally relevant indicator of crop response to N. For example, the importance of the background environment was demonstrated by Clay et al. (2012) in winter wheat under contrasting soil residual N and water conditions. Both N and water stress affected canopy reflectance, and the N sufficiency index was more closely related to N-induced yield response when under-fertilized wheat was referenced to well-fertilized plants grown under the same water environment. This showed that environmental background can alter the spectral interpretation of crop N status and that a locally appropriate reference can reduce this effect and improve the remote sensing-based N stress detection. Similar concerns apply to background soil N supply. Bausch and Delgado (2005) showed that residual soil nitrate influenced remotely sensed assessment of crop N status in corn, while Argento et al. (2021) and Wang et al. (2025) demonstrated that soil mineral N affects wheat N uptake, canopy development, and spectral response. Because crop response depends on both fertilizer N and

residual soil N, fields with high residual N may show little canopy or yield response to fertilizer, whereas the same N gradient in low-N soils may produce strong spectral separation (Lucero et al., 2026). Relatively few multi-environment UAV studies have directly tested whether variation in background soil N supply explains differences in both spectral N detectability and agronomic N responsiveness on yield and protein content across wheat environments.

Therefore, the overall objectives of this study were –

1. To compare the sensitivity of non-red-edge and red-edge-based vegetation indices for detecting fertilizer-N response across crop developmental stages and determine whether spectral N response is consistent among winter wheat cultivars.
2. To quantify cross-location variation in UAV-based spectral N detectability and determine whether that variation is associated with residual soil N supply.
3. To determine whether spectral N detectability corresponds with agronomic responsiveness to fertilizer N, as measured by grain yield and protein response across locations.
4. To evaluate whether expressing canopy spectral measurements relative to a local N-rich reference strengthens their relationship with agronomic N responsiveness and whether spectral sufficiency tracks grain-yield and protein response across different N fertilizer rates.

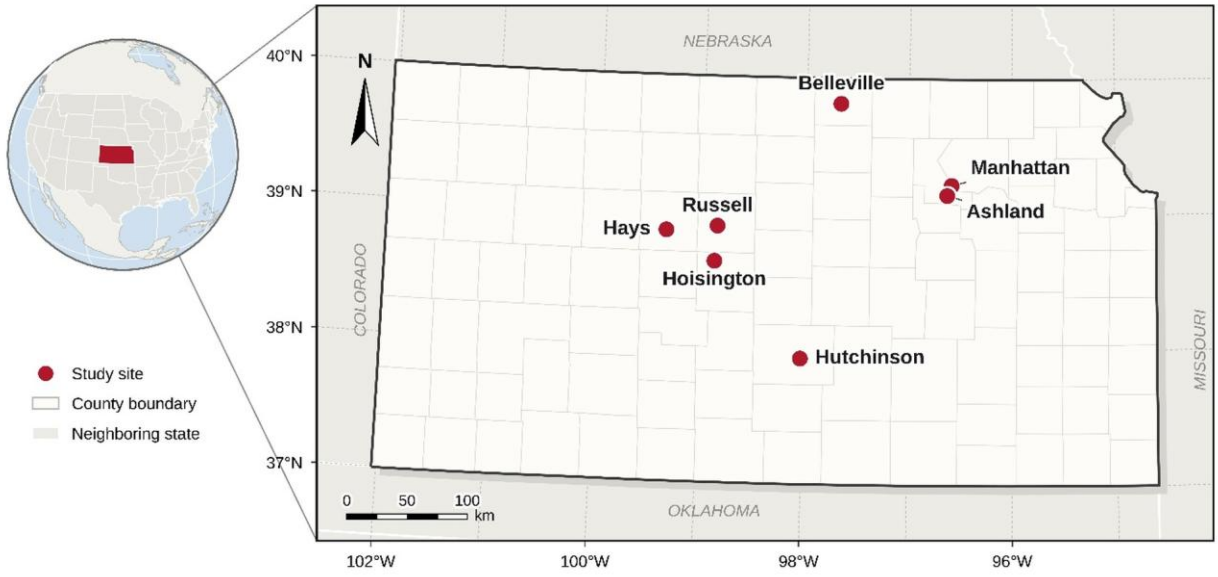
2 Materials and Methods

2.1 Study Sites and Experimental Design

The experiment was conducted at seven Kansas locations during the 2024–2025 season: Manhattan, Ashland Bottoms, Hoisington, Russell, Hays, Belleville, and Hutchinson (Figure 3.1). The sites span the main hard red winter wheat production region of the state along a west-to-east precipitation gradient, with 30-year mean annual precipitation increasing from approximately 600

mm at Hays to 870 mm at Manhattan, and they differ in soil series and texture. Together, they provided a range of background soil N supply and crop N responsiveness against which the same fertilizer-N treatments could be compared. Site coordinates, soil series and textural classes, and sowing-to-harvest weather for each location are summarized in Table 2.1 (Chapter 2).

Figure 3.1. Field experiments were conducted during the 2024–2025 winter wheat growing season at seven Kansas locations: Manhattan, Ashland Bottoms, Belleville, Russell, Hutchinson, Hays, and Hoisington.



The eight cultivars were Bob Dole (developed by Kansas State University and marketed by AgriPro/Syngenta), LCS Chrome and LCS Julep (Limagrain Cereal Seeds), SY Monument and SY Wolverine (AgriPro/Syngenta), and WB4269, WB4401, and WB4699 (WestBred/Bayer); their release years and contrasting traits are listed in Table 2.2 (Chapter 2). Granular urea fertilizer (46-00-00) was applied at 0, 34, 67, 101, 135, 168, and 202 kg N ha^{-1}. Treatments were arranged as a split-plot design, with N rate assigned as the whole-plot factor and cultivar as the subplot factor within replicate blocks (Appendix Figure A.1). Each complete block, therefore, contained 56 plots (seven N rates × eight cultivars). Manhattan, Ashland Bottoms, Belleville, and Hoisington contained three complete replicate blocks, whereas Russell, Hutchinson, and Hays contained four

blocks. Within each block, each N-rate whole plot consisted of eight cultivar subplots. Individual plots were seven rows wide, with 19.1-cm row spacing, and approximately 9.1 m long, resulting in plot dimensions of approximately 1.34 m × 9.1 m. The combined experiment included 1,344 plots representing a common set of eight hard red winter wheat cultivars and seven spring fertilizer-N rates across contrasting soil-N environments.

Grain yield was measured at maturity using a plot combine and expressed as kg ha^{-1}. For the cross-environment analysis, site-level agronomic response to fertilizer N was quantified by comparing mean grain yield at 0 and 202 kg N ha^{-1}.

2.2 Soil N Supply and Agronomic N Response

Soil nitrate-N was measured separately for the 0–15 and 15–61 cm depth increments at each location. Nitrate concentrations reported in ppm were converted to kg ha^{-1} according to Ruiz Diaz et al. (2023). This conversion follows the standard furrow-slice approximation and therefore does not incorporate site-specific measured bulk density. Soil N variables were treated as site-level characteristics because soil measurements were not available for individual experimental plots. An additional estimate of indigenous N supply was calculated by combining profile nitrate-N with a surface-soil organic-matter N credit. Following the structure used in Kansas wheat N recommendations (Ruiz Diaz et al., 2023), 11.2 kg N ha^{-1} was credited for each percentage point of organic matter in the 0–15 cm soil layer:

$$\text{Indigenous N supply } (\text{kg ha}^{-1}) = \text{Profile } NO_3\text{-N} + 11.2 \times \text{SOM (\%)} \quad (1)$$

Profile nitrate-N was used as the primary indicator of background soil N in the cross-environment analyses, while indigenous N supply was used as a complementary site-level descriptor.

Agronomic responsiveness to fertilizer N was quantified from the grain-yield contrast between the 0 and 202 kg N ha^{-1} treatments within each location. Absolute yield response was calculated as:

$$\text{Yield response}_{kg\ ha^{-1}} = Y_{202N} - Y_{0N} \quad (2)$$

and the relative yield response was calculated as:

$$\text{Yield response (\%)} = \frac{Y_{202N} - Y_{0N}}{Y_{0N}} \times 100 \quad (3)$$

where Y_0N and $Y_{202}N$ represent the mean grain yields of the 0 and 202 kg N ha^{-1} treatments, respectively. Relative yield response was used as the primary site-level measure of agronomic N responsiveness, while absolute yield gain was retained for sensitivity analyses.

Similarly, grain protein response to fertilizer N was evaluated as a secondary agronomic endpoint. For each environment, the protein response was calculated as the difference in mean grain protein concentration between the 202 and 0 kg N ha^{-1} treatments, expressed as percentage points. This metric was used to characterize the magnitude of fertilizer-N effects on grain protein independently of the relative grain-yield response.

2.3 UAV Multispectral Imagery and Vegetation Indices

Up to six UAV multispectral flights were conducted at each location during the 2024–2025 winter wheat growing season (Table 3.1). Multispectral imagery was collected using a Parrot Sequoia sensor with green, red, red-edge, and near-infrared (NIR) spectral bands centered approximately at 550, 660, 735, and 790 nm, respectively. The sensor was mounted on a senseFly eBee X fixed-wing UAV (senseFly, Cheseaux-sur-Lausanne, Switzerland) flown at 50 m above ground level, yielding a multispectral ground sampling distance (GSD) of 4.8 cm. Images were radiometrically calibrated using reflectance-panel and irradiance measurements and processed to generate georeferenced reflectance products. Image stitching and orthomosaic generation were

performed in Pix4Dfields (version 2.13.4; Pix4D SA, Prilly, Switzerland). Mean band reflectance was subsequently summarized at the individual-plot level using custom Python workflows (Python version 3.13; Python Software Foundation, 2024). Belleville and Hays did not have usable multispectral imagery for Flight 2. Eight vegetation indices representing conventional red/NIR, green/NIR, and red-edge spectral responses were calculated from plot-level reflectance (Table 3.2). NDVI, SAVI, and GNDVI were classified as non-red-edge indices, whereas NDRE, CI_{RE}, REDVI, MTCI, and DATT were classified as red-edge-based indices for spectral-family comparisons.

Table 3.1. Sowing, fertilizer-N application, and UAV acquisition timing across seven Kansas winter wheat environments during the 2024–2025 growing season. Values in parentheses are the Feekes growth stage on each date (e.g., F11.2 = Feekes 11.2). * Flight collected before fertilizer-N application. † Multispectral imagery from this flight was not usable. NA, no flight.

Location	Sowing date	N application	Flight 1	Flight 2	Flight 3	Flight 4	Flight 5	Flight 6
Ashland	Oct. 24, 2024	Mar. 26, 2025 (F3)	Apr. 3 (F4)	Apr. 14 (F5)	Apr. 23 (F8)	May 8 (F10)	May 28 (F11.2)	NA
Belleville	Oct. 15, 2024	Mar. 28, 2025 (F3)	Apr. 8 (F3)	Apr. 18 (F6)†	Apr. 29 (F9)	May 9 (F10)	May 23 (F10.5)	Jun. 11 (F11.4)
Hays	Oct. 8, 2024	Mar. 28, 2025 (F3)	Apr. 7 (F3)	Apr. 11 (F4)†	Apr. 21 (F6)	May 2 (F10)	May 14 (F10.5)	May 30 (F11)
Hoisington	Sep. 27, 2024	Apr. 10, 2025 (F3)	Mar. 31 (F3)*	Apr. 11 (F3)	Apr. 21 (F6)	May 2 (F9)	May 14 (F10.5)	May 30 (F11.2)
Hutchinson	Nov. 16, 2024	Apr. 2, 2025 (F3)	Apr. 7 (F3)	Apr. 15 (F3)	Apr. 25 (F5)	May 5 (F7)	May 22 (F10)	Jun. 9 (F11.4)
Manhattan	Oct. 21, 2024	Mar. 26, 2025 (F3)	Apr. 3 (F4)	Apr. 14 (F6)	Apr. 23 (F7)	May 8 (F9)	May 28 (F11.2)	Jun. 13 (F11.4)
Russell	Oct. 14, 2024	Apr. 3, 2025 (F3)	Mar. 31 (F3)*	Apr. 11 (F4)	Apr. 21 (F6)	May 2 (F10)	May 14 (F10.5)	May 30 (F11.2)

Table 3.2. Vegetation indices derived from UAV multispectral reflectance

Vegetation index	Formula	Reference
NDVI	*(NIR - Red) / (NIR + Red)*	Rouse et al. (1974)
GNDVI	*(NIR - Green) / (NIR + Green)*	Gitelson et al. (1996)
NDRE	*(NIR - RedEdge) / (NIR + RedEdge)*	Barnes et al. (2000)
CI_{RE}	*(NIR / RedEdge) - 1*	Gitelson et al. (2005)
SAVI	*1.5 × (NIR - Red) / (NIR + Red + 0.5)*	Huete (1988)
REDVI	*NIR - RedEdge*	Cao et al. (2013)
MTCI	*(NIR - RedEdge) / (RedEdge - Red)*	Dash & Curran (2004)
DATT	*(NIR - RedEdge) / (NIR - Red)*	Datt (1999)

Abbreviations: **NDVI**, normalized difference vegetation index; **GNDVI**, green normalized difference vegetation index; **NDRE**, normalized difference red-edge index; **CI_{RE}**, red-edge chlorophyll index; **SAVI**, soil-adjusted vegetation index; **REDVI**, red-edge difference vegetation index; **MTCI**, MERIS terrestrial chlorophyll index; **DATT**, Datt chlorophyll index; **NIR**, near-infrared reflectance; **RedEdge**, red-edge reflectance; **Red**, red reflectance; **Green**, green reflectance.

2.4 Crop Growth Stage Classification

Flight dates varied by location, so a given flight number did not represent the same point in crop development at every site. Each UAV observation was therefore assigned the Feekes growth stage recorded at that location and date (Large, 1954), and observations were pooled into three developmental windows for comparisons among environments:

- **Tillering:** Feekes 3–5
- **Stem elongation:** Feekes 6–9
- **Grain formation (heading/grain fill):** Feekes 10–11.4

Grouping observations by developmental window rather than by date kept the stages comparable across environments while providing enough observations in each window for statistical comparison. The windows also separate the main phases of crop N demand: tiller production and early canopy expansion, rapid biomass and N accumulation during stem elongation, and spike development, anthesis, and grain filling.

The first flights at Hoisington and Russell were collected 10 and three days, respectively, before fertilizer N was applied. Because these observations could not reflect a fertilizer response, they were excluded from all stage-level summaries of fertilizer-N detectability.

The time between N application and each flight also differed among sites (Table 3.1). Stage comparisons were therefore interpreted as differences in crop development and spectral sensitivity, not as estimates of how many days after application a fertilizer response first became detectable.

2.5 Standardized Spectral N Detectability

Spectral N detectability was quantified separately for each Location × UAV flight × vegetation-index combination. Because fertilizer N was applied as the whole-plot treatment and the eight cultivars were subplots within each N-rate whole plot, the analysis accounted for the shared experimental structure among cultivar subplots.

For each combination, vegetation-index values were modeled as:

$$VI = \beta_0 + \beta_1(\text{N rate}) + \text{Cultivar} + \text{Block} + \varepsilon$$

where N rate was treated as a continuous variable (0–202 kg N ha^{-1}), and cultivar and block were included as fixed effects. Standard errors were estimated using cluster-robust covariance at the whole-plot level, where each whole plot was defined by the unique Location × Block × N-rate combination. This approach accounted for the non-independence of the eight cultivar subplots within each fertilizer-N whole plot.

The coefficient for N rate (β_1) represented the change in a vegetation index across the imposed fertilizer-N gradient after accounting for cultivar and block effects. Because the eight vegetation indices differed substantially in their numerical ranges and residual variability, raw N-rate slopes were standardized by the residual standard deviation of each fitted model and expressed per 100 kg N ha^{-1}:

$$\text{Standardized N detectability} = (\beta_1 \times 100) / \sigma_{\text{residual}}$$

where σ_{residual} is the residual standard deviation from the corresponding Location × Flight × Index model.

A standardized detectability value of 1.0, therefore, indicates that a 100 kg N ha^{-1} increase in fertilizer rate was associated with a one-residual-standard-deviation increase in the vegetation index. Positive values indicate increasing vegetation-index values with increasing N rate, whereas values near zero indicate little spectral differentiation across the fertilizer-N gradient.

Standardized N detectability was used as the common response metric for comparing vegetation indices, crop developmental stages, and environments. True pre-fertilizer observations at Hoisington and Russell were excluded from post-fertilizer developmental-stage summaries but retained separately for quality control.

2.6 Comparison of Vegetation Index Sensitivity

Vegetation-index sensitivity to fertilizer N was compared using the standardized N detectability metric described above. To avoid giving greater weight to locations with more UAV observations within a developmental stage, standardized detectability was first summarized at the flight level and then averaged within each Location × Stage × Index combination. The resulting site-level stage summaries, therefore, contributed equally to cross-environment comparisons.

Differences among the eight vegetation indices were evaluated separately for each developmental stage. Within each stage, an overall Friedman repeated-measures test was used to determine whether standardized N detectability differed among indices, with location treated as the repeated observational unit. When the Friedman test was significant, paired Wilcoxon signed-rank tests were used for pairwise comparisons among vegetation indices. Pairwise P-values were adjusted within each developmental stage using the Benjamini-Hochberg false discovery rate (FDR) procedure.

For spectral-family comparisons, NDVI, SAVI, and GNDVI were grouped as non-red-edge indices, whereas NDRE, CIRE, REDVI, MTCI, and DATT were grouped as red-edge-based indices. Within each Location × Stage combination, the median standardized detectability was calculated across indices within each spectral family. Paired comparisons between the non-red-edge and red-edge families were then performed across locations using Wilcoxon signed-rank tests.

Pre-fertilizer observations from Hoisington and Russell were excluded from these developmental-stage comparisons.

To determine whether cultivar modified spectral response to fertilizer N, an additional analysis was conducted for each Location × developmental stage × vegetation-index combination. Stage-level vegetation-index values were modeled as a function of continuous N rate, cultivar, the N rate × cultivar interaction, and block. Standard errors were estimated using cluster-robust covariance at the N-rate whole-plot level. The joint significance of all N rate × cultivar interaction terms was evaluated using a finite-sample-corrected Wald F-test, and P-values were adjusted within each developmental stage using the Benjamini–Hochberg false-discovery-rate procedure.

Statistical significance was evaluated at $P < 0.05$ unless otherwise stated.

2.7 N-Rich Sufficiency Index

To reduce cross-environment differences in absolute vegetation-index values, spectral measurements were also expressed relative to a locally N-rich reference. For each Location × UAV flight × Cultivar × Vegetation index combination, the N-rich reference was defined as the mean vegetation-index value of plots receiving 202 kg N ha^{-1}. This within-location, within-flight, and within-cultivar reference was used to account for environmental, developmental, and cultivar-related differences in canopy reflectance.

The N-rich sufficiency index (SI) for each plot was calculated as:

$$SI = \frac{VI_{plot}}{\overline{VI}_{202N}}$$

where VIplot is the vegetation-index value for an individual plot and $\bar{V}I_{202}N$ is the mean vegetation-index value of the 202 kg N ha^{-1} treatment for the corresponding Location × Flight × Cultivar × Index combination.

An SI value approaching 1 indicates that the spectral response of a plot is similar to that of the local N-rich reference, whereas values below 1 indicate increasing spectral divergence from the N-rich treatment. As a quality-control check, the mean SI of the 202 kg N ha^{-1} treatment equaled 1.0 for each vegetation index.

For cross-environment comparisons, the analysis focused on the 0-N treatment, which represented the greatest fertilizer-N contrast relative to the N-rich reference. Mean 0-N sufficiency index values were summarized for each Location × developmental stage × vegetation index combination after exclusion of the true pre-fertilizer UAV observations at Hoisington and Russell.

To evaluate whether N-rich normalization improved agronomic interpretation across environments, relationships between yield response to fertilizer N and absolute vegetation-index values were compared with the corresponding relationships using the 0-N sufficiency index. These comparisons were conducted separately for NDVI and NDRE within each developmental stage.

To determine whether the relationship between spectral and agronomic response extended across the complete fertilizer-N gradient, sufficiency indices were additionally summarized for all N rates. Grain yield at each N rate was expressed as a ratio to mean yield at the 202 kg N ha^{-1} reference within each environment. Within each Location × developmental stage × vegetation-index combination, Spearman correlations between spectral sufficiency index and yield ratio were calculated across the 0–168 kg N ha^{-1} treatments. The 202 kg N ha^{-1} treatment was excluded from correlation analysis because both the spectral sufficiency index and yield ratio equal 1 by definition at the reference treatment. Correlations were summarized across environments using the median coefficient and number of locations showing positive associations.

2.8 Cross-Environment Relationships and Sensitivity Analyses

Cross-environment relationships were evaluated at the location level because profile soil N measurements were available at the site level, not the plot level. Profile soil NO3-N was used as the primary indicator of background soil N supply. Agronomic responsiveness was represented primarily by relative grain-yield response to fertilizer N, calculated from the 0- and 202-kg N ha^{-1} treatment means, with absolute yield gain (kg ha−1) retained for sensitivity analysis.

For developmental-stage summaries, standardized N detectability was first averaged across available post-fertilizer UAV observations within each Location × Stage × Index combination. This approach prevented environments with more UAV observations within a developmental stage from receiving greater weight. True pre-fertilizer observations at Hoisington and Russell were excluded from these post-fertilizer summaries.

To summarize red-edge response across indices, the five red-edge-based indices (NDRE, CI_{RE}, REDVI, MTCI, and DATT) were first combined within each Location × developmental stage using the median standardized N detectability. The resulting stage-level values were then averaged across the three developmental stages with equal weighting to obtain an overall red-edge N-detectability value for each location. This two-step procedure prevented any single index or developmental stage from dominating the site-level summary.

Associations among profile soil NO_3-N, agronomic N responsiveness, and spectral N detectability were evaluated using both Spearman rank correlation and Pearson correlation across the seven environments. Spearman's correlation was emphasized for site-level inference due to the small number of independent environments and the primary focus on monotonic relationships. Pearson correlations were reported as complementary measures of linear association. Statistical

significance was assessed at $P < 0.05$, but interpretation of site-level relationships emphasized effect size, direction, and consistency, given that only seven independent locations were available.

The robustness of the relationship between spectral N detectability and agronomic N responsiveness was evaluated in three ways. First, the analysis was repeated separately for NDRE, CI_RE, REDVI, MTCI, and DATT rather than relying only on the combined red-edge summary. Second, yield response was expressed both as relative yield response (%) and absolute yield gain (kg ha^{-1}). Third, a leave-one-location-out sensitivity analysis was conducted by sequentially removing each environment and recalculating the correlation between overall red-edge detectability and yield response using the remaining six locations. These analyses were used to determine whether the cross-environment relationship depended strongly on a particular vegetation index, yield-response definition, or individual location.

To evaluate the effect of local N-rich normalization, Spearman correlations between yield response and absolute vegetation-index values were compared descriptively with the corresponding correlations obtained using the 0-N sufficiency index. These comparisons were conducted separately for NDVI and NDRE within each developmental stage. Because the correlations were derived from the same seven environments and were not formally compared using a test for dependent correlations, differences were interpreted as strengthening or weakening of the observed association rather than as formal evidence that one correlation statistically exceeded another. As an exploratory secondary analysis, grain-protein response was also related to the profile of soil NO_3-N and overall red-edge N detectability using Spearman rank correlation, with Pearson correlation reported as a complementary measure of linear association.

All statistical analyses were conducted in Python using pandas for data processing, SciPy for correlation and nonparametric tests, and statsmodels for regression models, cluster-robust covariance estimation, and false-discovery-rate adjustment.

3 Results

3.1 Background Soil N Supply and Grain Yield and Protein Responses to Fertilizer Nitrogen

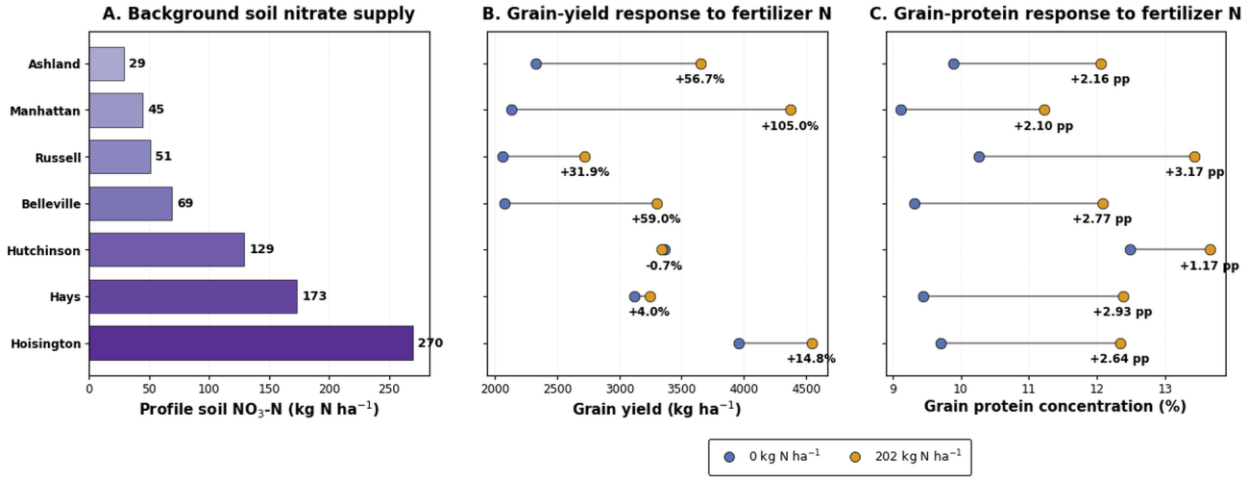
Background soil N supply varied substantially among the seven environments (Figure 3.2A). Profile soil NO3-N ranged from 29.4 kg N ha^{-1} at Ashland to 269.7 kg N ha^{-1} at Hoisington. Estimated indigenous N supply showed a similar gradient, ranging from 49.5 kg N ha^{-1} at Ashland to 298.8 kg N ha^{-1} at Hoisington. Manhattan, Russell, and Belleville also had relatively low profile NO3-N concentrations of 44.6, 50.7, and 69.0 kg N ha^{-1}, respectively, whereas Hutchinson and Hays had higher background nitrate levels of 129.3 and 173.1 kg N ha^{-1}, respectively.

Grain-yield response to fertilizer N also differed markedly among environments (Figure 3.2B). Manhattan had the greatest response, with yield increasing from 2,465 kg ha^{-1} at 0 N to 5,053 kg ha^{-1} at 202 kg N ha^{-1}, corresponding to a 105.0% increase. Belleville and Ashland showed yield increases of 59.0% and 56.7%, respectively, while Russell showed an intermediate response of 31.9%. In contrast, yield responses were relatively small at Hoisington (14.8%) and Hays (4.0%), and no positive response was observed at Hutchinson (−0.7%). Across environments, yield response tended to decrease with increasing profile soil NO3-N (Spearman ρ=-0.679, P=0.094; Pearson r = -0.657, P=0.109), although the relationship was not statistically significant at P<0.05.

Grain protein concentration also responded to fertilizer N, but the magnitude of response followed a different pattern from grain yield (Figure 3.2C). Protein concentration increased from the 0- to 202-kg N ha^{-1} treatment at all seven environments, with increases ranging from 1.17

percentage points at Hutchinson to 3.17 percentage points at Russell. Hays and Belleville also showed relatively large protein responses of 2.93 and 2.77 percentage points, respectively, despite contrasting grain-yield responses among environments.

Figure 3.2. Background soil nitrogen supply and grain yield and protein responses to fertilizer N across seven winter wheat environments. (A) Profile soil NO_3-N (0–61 cm) at each location, ordered from lowest to highest background nitrate supply. (B) Mean grain yield at 0 and 202 kg N ha^{-1}, with labels indicating the relative grain-yield response to fertilizer N. (C) Mean grain protein concentration at 0 and 202 kg N ha^{-1}, with labels indicating the absolute grain-protein response expressed in percentage points (pp).



3.2 Vegetation Index Sensitivity to Fertilizer N Across Crop Developmental Stages

N-rich sufficiency indices revealed substantial variation in the magnitude and temporal development of spectral N stress among environments (Figure 3.3). During tillering, SI values were generally close to 1 across most vegetation indices and environments, indicating limited spectral separation between the 0-N treatment and the local N-rich reference. As crop development progressed, SI declined substantially at several environments, particularly Ashland and Manhattan, indicating increasing spectral expression of N limitation. The decline was generally greater for red-edge-based indices such as NDRE, CI_RE, REDVI, and MTCI than for conventional greenness indices. In contrast, environments such as Hutchinson and Hays maintained SI values

relatively close to 1 across developmental stages, indicating comparatively weak spectral response to fertilizer N.

These environment- and stage-dependent patterns were further quantified using standardized fertilizer-N detectability. Standardized fertilizer-N detectability varied among vegetation indices and across broad crop developmental stages (Figure 3.4). Detectability was generally low during tillering (Feekes 3–5), increased during stem elongation (Feekes 6–9), and was greatest during grain formation (Feekes 10–11.4). Mean standardized N detectability during tillering ranged from 0.11 for SAVI to 0.32 for DATT. During stem elongation, values increased to 0.74–1.16, with MTCI (1.16), DATT (1.14), CI_{RE} (1.12), and NDRE (1.11) showing the greatest mean responses. During grain formation, detectability ranged from 1.57 for SAVI to 2.00 for CI_{RE}, with NDRE (1.99) and GNDVI (1.86) also showing strong responses.

Vegetation indices differed significantly in standardized N detectability across developmental stages. Friedman tests indicated significant differences among the indices during tillering ($\chi2 = 24.14$, $P = 0.0011$), stem elongation ($\chi2 = 38.42$, $P < 0.001$), and grain formation ($\chi2 = 17.38$, $P = 0.015$). The clearest separation among individual indices occurred during stem elongation. Following false-discovery-rate correction, NDRE, CI_{RE}, MTCI, and DATT showed significantly greater detectability than NDVI ($P_{FDR}=0.031$). These red-edge-based indices also generally showed greater detectability than SAVI and GNDVI during stem elongation.

When indices were grouped by spectral formulation, red-edge-based indices generally exhibited greater standardized N detectability than non-red-edge indices across the three developmental stages, with the largest contrast occurring during stem elongation. However, although the overall among-index tests were significant during tillering and grain formation, individual pairwise differences were less consistently significant after multiple-comparison

correction. Thus, the advantage of red-edge-based indices was most clearly expressed during stem elongation.

Cultivar modification of spectral N response also differed strongly among developmental stages (Appendix Figure A.2). After FDR adjustment, significant N rate × cultivar interactions occurred in 19 of 56 Location × vegetation-index combinations (33.9%) during tillering and 30 of 56 combinations (53.6%) during grain formation. In contrast, only one of 56 tests (1.8%) was significant during stem elongation. Among individual indices, cultivar interactions at grain formation were most frequent for MTCI, occurring in five of seven environments. These results indicate that spectral response to fertilizer N was comparatively consistent among the eight cultivars during stem elongation.

Figure 3.3. Vegetation-index sufficiency index (SI) of the 0-N treatment relative to the local N-rich reference (202 kg N ha^{-1}) across seven Kansas winter wheat environments and three developmental stages.

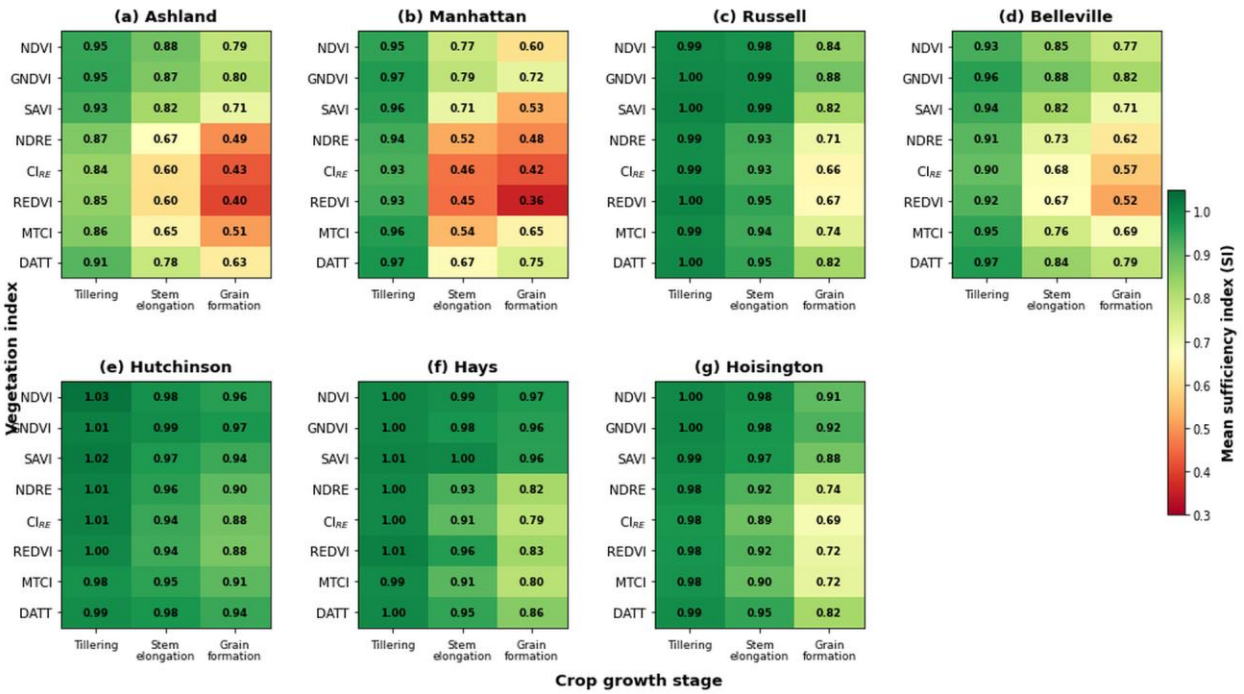
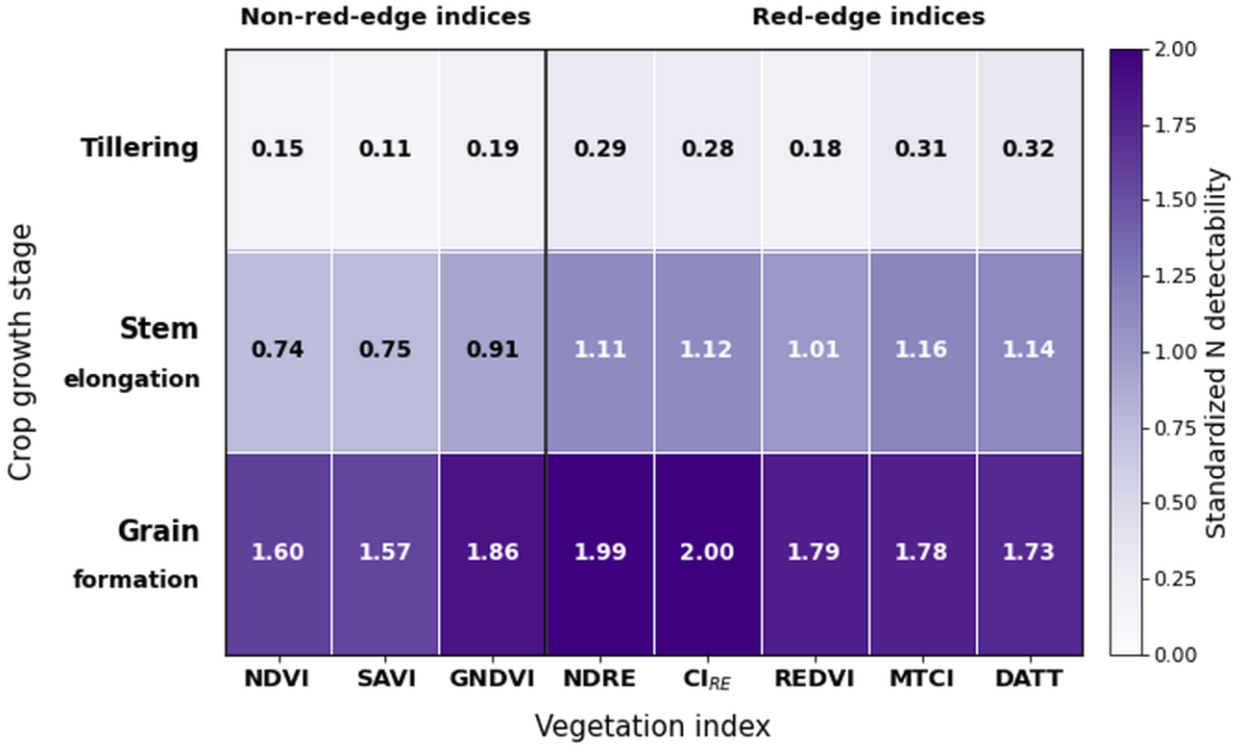


Figure 3.4. Mean standardized fertilizer-N detectability of eight UAV-derived vegetation indices across broad winter wheat developmental stages. Tillering, stem elongation, and grain formation correspond to Feekes 3–5, 6–9, and 10–11.4, respectively. Cell values represent the mean standardized N detectability across the seven environments, excluding pre-fertilizer UAV observations.



3.3 Background Soil N, Agronomic Response, and Cross-Location Variation in Spectral N Detectability

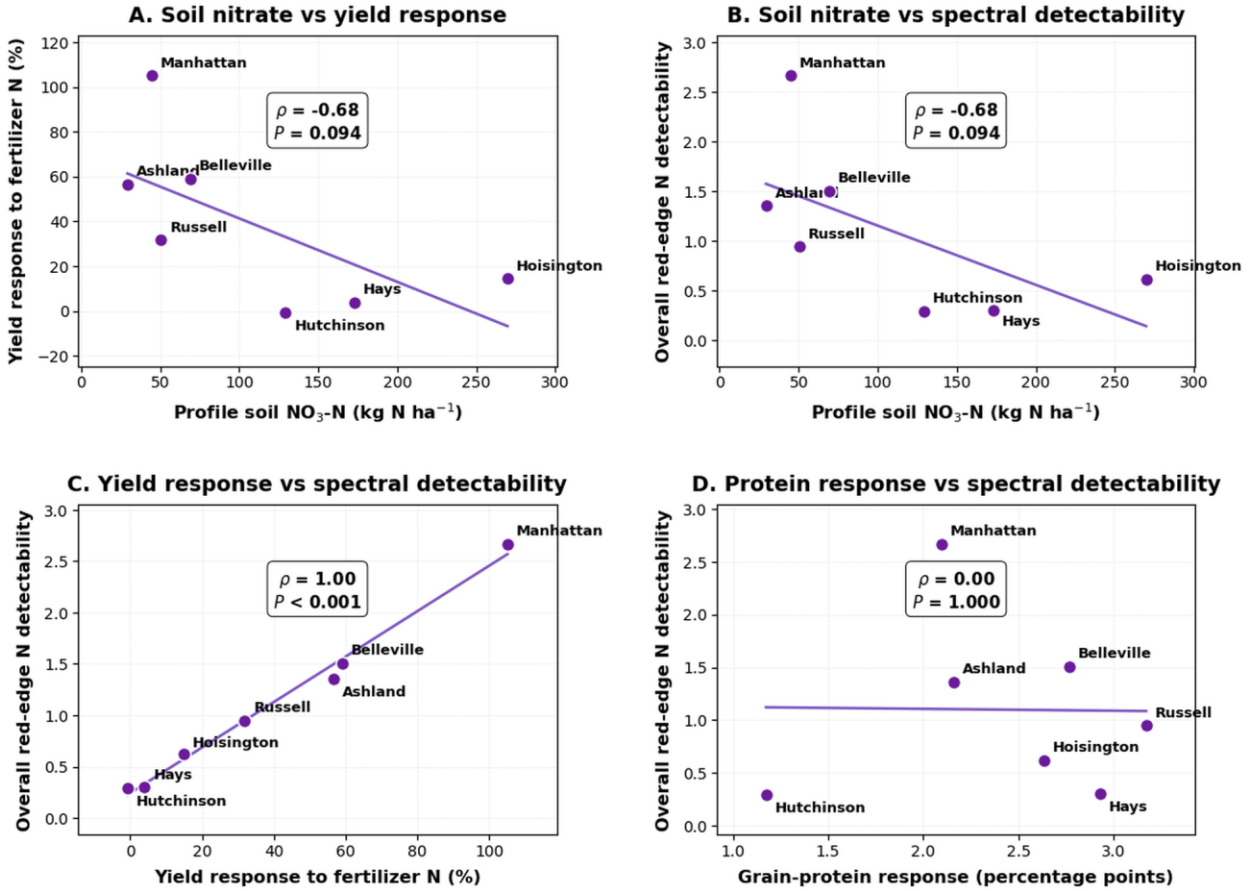
Background soil N supply was negatively associated with both agronomic and spectral responsiveness to fertilizer N across the seven environments (Figure 3.5A–B). Grain-yield response tended to decline as profile soil NO_3-N increased, with a Spearman correlation of ($\rho = -0.68$, P=0.094; Figure 3.5A). A similar negative relationship was observed between profile soil NO_3-N and seasonally summarized red-edge N detectability (ρ= -0.68, P=0.094; Figure 3.5B). Although neither relationship reached P<0.05, the direction of association was consistent with weaker fertilizer-N responsiveness at environments with greater background nitrate supply.

The site-level pattern was also evident in the magnitude of spectral response. Locations with relatively low-profile soil NO_3-N, including Manhattan, Belleville, and Ashland, showed comparatively strong red-edge N detectability, whereas Hutchinson and Hays, which had greater background nitrate supply and little agronomic response to fertilizer N, showed much weaker spectral separation. Hoisington had the highest-profile soil NO3-N concentration but retained a modest spectral and yield response, indicating that background nitrate alone did not fully explain the among-environment variation in fertilizer-N responsiveness.

When relationships were examined by developmental stage and spectral family, associations between profile soil NO_3-N and standardized N detectability were consistently negative. The strongest relationship was observed for non-red-edge indices during grain formation (ρ= -0.82, P=0.023), while red-edge detectability during grain formation also showed a strong negative association (ρ=-0.75, P=0.052). Relationships at earlier stages were generally weaker, although their direction remained negative. Given the limited number of environments (n=7), these stage-specific relationships were interpreted primarily by the magnitude and consistency of the associations rather than by statistical significance alone.

Figure 3.5. Cross-environment relationships among background soil nitrate supply, agronomic response to fertilizer N, and UAV-derived red-edge N detectability across seven winter wheat environments. (A) Relationship between profile soil NO3-N (0–61 cm) and relative grain-yield response to fertilizer N. (B) Relationship between profile soil NO_3-N and overall red-edge N detectability, calculated by first summarizing the five red-edge indices (NDRE, CI_{RE}, REDVI, MTCI, and DATT) within each developmental stage and then averaging across the three developmental stages with equal stage weighting. (C) Relationship between fertilizer-N yield

response and overall red-edge N detectability. **(D)** Relationship between grain-protein response and overall detectability of red-edge N.



3.4 Agronomic Relevance of Spectral N Detectability and N-Rich Normalization

UAV-derived spectral N detectability was strongly associated with the agronomic response to fertilizer N across environments (Figure 3.5C). Overall red-edge N detectability increased almost proportionally with relative grain-yield response, resulting in a perfect rank-order relationship across the seven environments (Spearman, ρ=1.00, P<0.001) and a very strong linear association (Pearson, r=0.995, P<0.001). Thus, environments with the greatest yield response to fertilizer N also showed the strongest UAV-observed spectral response, whereas sites with little or no yield response exhibited weak spectral separation.

This relationship was robust to both the choice of red-edge vegetation index and the expression of yield response. When individual red-edge indices were analyzed separately,

Spearman correlations with relative yield response ranged from 0.964 to 1.000. Relationships also remained strong when yield response was expressed as absolute grain-yield gain, with ρ values ranging from 0.929 to 0.964. Leave-one-location-out sensitivity analysis further showed that the relationship was not driven by any single environment: correlations for absolute yield gain ranged from 0.943 to 1, and the relative yield response was 1 regardless of which location was omitted.

Grain-protein response showed a distinctly different relationship with UAV-derived N detectability. Although grain protein concentration increased with fertilizer N at all seven environments, the magnitude of protein response was unrelated to overall red-edge N detectability (Spearman $\rho = 0.00$, $P = 1.000$; Pearson $r = -0.014$, $P = 0.976$; Figure 3.5D). Protein response was also unrelated to profile soil NO_3-N (Spearman $\rho = 0.18$, $P = 0.702$; Pearson $r = 0.07$, $P = 0.877$). Thus, unlike grain-yield responsiveness, which closely tracked UAV-based N detectability across environments, the magnitude of final grain-protein response did not correspond with either background soil nitrate supply or seasonally summarized red-edge N detectability.

N-rich normalization further strengthened the relationship between UAV spectral measurements and agronomic N responsiveness (Figure 3.6). For NDVI, Spearman correlations between absolute VI values and yield response were -0.36, -0.79, and -0.75 at tillering, stem elongation, and grain formation, respectively, whereas the corresponding correlations using the N-rich sufficiency index were -0.89, -0.96, and -0.93. A similar pattern was observed for NDRE: correlations strengthened from -0.71, -0.75, and -0.86 using absolute NDRE to -0.82, -0.86, and -0.96 after N-rich normalization at tillering, stem elongation, and grain formation, respectively.

The sufficiency-index relationships were also stronger and more consistent across developmental stages. For NDVI, the N-rich SI was significantly associated with yield response

at tillering ($P=0.007$), stem elongation ($P<0.001$), and grain formation ($P=0.003$). For NDRE, significant relationships were observed at tillering ($P=0.023$), stem elongation ($P=0.014$), and grain formation ($P<0.001$). In contrast, several corresponding relationships based on absolute VI values were weaker or only marginally significant. These results indicate that expressing canopy spectral measurements relative to a local N-rich reference improved the consistency of the relationship between UAV sensing and agronomic response to fertilizer N across environments.

Extending the sufficiency-index analysis across the full fertilizer-N gradient further demonstrated the agronomic relevance of local N-rich normalization (Figure 3.7). Within environments, spectral sufficiency indices were positively associated with the yield ratio to the 202-kg N ha^{-1} reference across N rates from 0 to 168 kg N ha^{-1}, and relationships strengthened with crop development. Median within-location Spearman correlations across vegetation indices ranged from 0.60 to 0.71 during tillering, 0.83 to 0.94 during stem elongation, and 0.89 to 0.94 during grain formation. All eight vegetation indices showed positive relationships in all seven environments during stem elongation and grain formation, indicating that locally normalized spectral measurements consistently tracked the crop's progression toward the N-rich yield level across the fertilizer-N gradient.

Figure 3.6. Comparison of the relationships between fertilizer-N yield response and absolute vegetation-index values versus N-rich sufficiency-index values for (A) NDVI and (B) NDRE across the tillering, stem elongation, and grain formation stages of winter wheat. Points represent

Spearman's rank correlation coefficients ***ρ*** across seven environments. More negative ***ρ*** values indicate a stronger inverse association with fertilizer-N yield response.

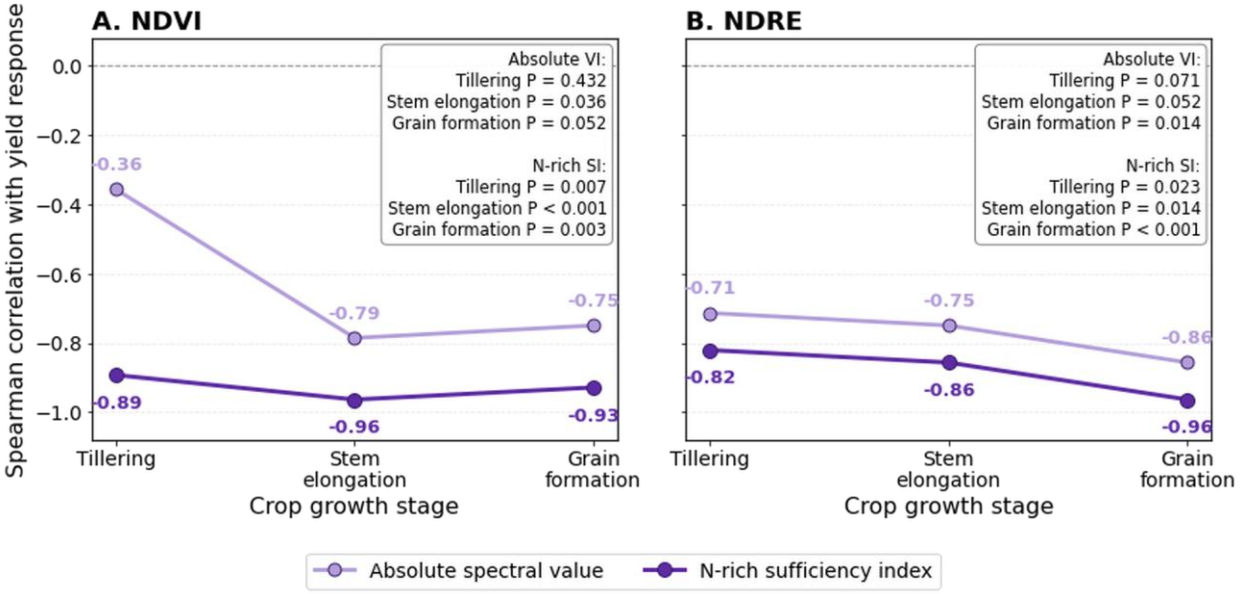
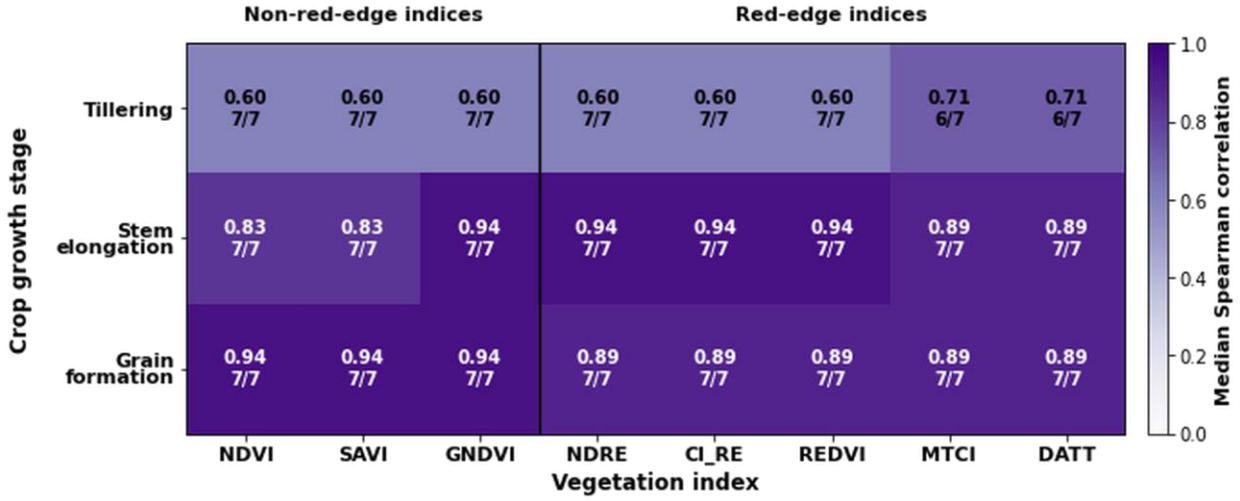


Figure 3.7. Association between vegetation-index sufficiency indices and grain-yield response across the full fertilizer-N gradient and winter wheat developmental stages. Cell values represent the median within-location Spearman rank correlation (ρ) between the vegetation-index sufficiency index and the yield ratio to the 202 kg N ha^{-1} reference across fertilizer-N rates from 0 to 168 kg N ha^{-1}. Values below each correlation coefficient indicate the number of environments with a positive association out of seven total environments.



4 Discussion

4.1 Stage- and Index-Specific Sensitivity to Fertilizer N

UAV-based detection of fertilizer-N response varied with both crop developmental stage and spectral formulation. N detectability was relatively low during tillering and increased as the crop developed, but the clearest differentiation among vegetation indices occurred during stem elongation. At this stage, NDRE, CI_{RE}, MTCI, and DATT were more detectable than NDVI after adjustment for multiple comparisons, whereas by grain formation, most indices showed relatively strong N responses. This suggests that the stem-elongation period provided the clearest advantage of red-edge information rather than simply representing the period of maximum spectral response.

The practical value of the stem-elongation window is that it still overlaps the period when an in-season N application can be acted on. Grain number, the main determinant of wheat yield, is set largely during stem elongation and the weeks before anthesis (Fischer, 1985; Slafer et al., 2014), and wheat can recover from early N deficiency when fertilizer is applied before this window closes. Across 12 Oklahoma site-years, delaying N application until Feekes 7 did not reduce grain yield even when N deficiency was visible earlier, whereas applications after Feekes 8 reduced yield relative to pre-plant N (Souza et al., 2022; see also Morris et al., 2006). Ravier et al. (2017) similarly reported that early N deficiencies can favor high yield, grain protein content, and N use efficiency. Red-edge detection of N response during stem elongation could therefore still inform a corrective N application, whereas detection during grain formation comes too late to recover yield and is more relevant to late-season N decisions aimed at grain protein (Giordano et al., 2023).

The greater sensitivity of red-edge indices is consistent with previous studies showing that red-edge wavelengths remain responsive to variation in chlorophyll and crop N status as wheat canopy biomass increases, while conventional red-NIR indices become less sensitive (Cao et al.,

2015; Prey and Schmidhalter, 2019; Argento et al., 2021). Recent work by Vuille-dit-Bille et al. (2025) further demonstrated the value of red-edge information across two winter wheat sites and three growing seasons using a 10-band UAV multispectral sensor. Their NDRE based on a 740-nm red-edge band predicted yield with $R^2 = 0.92$ at heading and $R^2 = 0.89$ during grain filling, while a 705-nm red-edge index was particularly informative for crop N before heading. These results agree with the findings from our study by showing that red-edge sensitivity is strongly stage-dependent and that additional red-edge spectral information may improve the characterization of wheat N response.

The relatively weak detectability at tillering observed here may reflect limited canopy development, greater soil-background influence, and insufficient time for fertilizer-N effects to become fully expressed. Conversely, by grain formation, treatment differences were large enough to be detected by most indices, reducing the relative advantage of red-edge formulations. Because UAV acquisitions occurred at different intervals after N application among environments, the seasonal increase in detectability likely reflects both crop development and the progressive expression of fertilizer response. For this reason, the stage comparisons describe how detectability changed with crop development rather than how quickly the fertilizer response appeared after application.

The limited cultivar-by-N interaction during stem elongation (Appendix Figure A.2) further supports this period as a robust sensing window. Although cultivar-related differences in canopy structure, phenology, and senescence likely contributed to greater genotype dependence during tillering and, particularly, grain formation, the spectral response to fertilizer N was remarkably consistent among cultivars during stem elongation. Thus, the stem-elongation advantage of red-edge indices was not primarily driven by a small subset of cultivars.

4.2 Soil N Supply and Agronomic Responsiveness Associated with UAV-Based N Detectability

A major finding of this study was the large variation among environments in both agronomic and spectral response to the same fertilizer-N gradient. Profile soil NO_3-N ranged from approximately 29 to 270 kg N ha^{-1}, while grain-yield response to fertilizer N ranged from essentially zero to more than 100%. Environments with higher profile nitrate generally exhibited weaker yield responses and reduced red-edge N detectability. This pattern indicates that fertilizer rate alone does not define crop N limitation. If substantial residual N is already available in the soil, even unfertilized treatment may maintain relatively high canopy growth and yield, resulting in limited spectral contrast between low- and high-N treatments. Wang et al. (2025) reported similar results in their recent study. They reported that variation in soil mineral N significantly affected canopy structure, grain yield, and N-use efficiency, and identified red-edge spectral features among the most informative remotely sensed indicators. Their results support the idea that local environmental conditions, such as residue N, soil moisture, and other local factors, can modify the canopy characteristics observed by UAV sensing and, therefore, need to be considered when interpreting spectral indicators of crop N status.

The greater correlation between spectral detectability and agronomic N responsiveness further strengthens this interpretation. Study sites with large fertilizer-induced yield responses also developed strong red-edge spectral separation, whereas environments with little yield benefit from additional N showed weak spectral differentiation. However, grain-protein response did not see the same trend as the yield-based relationship. Protein concentration increased with fertilizer N across all environments, yet the magnitude of this increase was unrelated to either background soil nitrate or overall red-edge N detectability. Similar results were reported in previous studies,

showing that canopy spectral measurements are generally more closely related to biomass and grain yield, and less so to grain protein content (Bogard et al., 2010; Walsh et al., 2022).

4.3 Local N-Rich Normalization and Implications for Precision N Management

Using vegetation indices relative to a local N-rich reference substantially strengthened their association with agronomic N responsiveness. For NDVI, the magnitude of the correlation with yield response increased from $|\rho| = 0.36–0.79$ using absolute values to $|\rho| = 0.89–0.96$ using the N-rich sufficiency index. NDRE showed a similar pattern, increasing from $|\rho| = 0.71–0.86$ to $|\rho| = 0.82–0.96$. The consistency of this response across developmental stages suggests that local normalization reduces environmental differences that complicate the interpretation of absolute vegetation-index values.

The N-rich reference provides a locally relevant benchmark because both the target crop and reference experience similar weather, soil, cultivar, and developmental conditions. This principle is consistent with previous N-rich strip and sufficiency-index approaches in which locally fertilized reference areas improve the interpretation of canopy sensing (Mullen et al., 2003; Raun et al., 2005; Clay et al., 2012). In the present study, high sufficiency-index values indicated that the unfertilized canopy closely resembled the N-rich reference and were generally associated with limited yield response, whereas lower values corresponded with greater agronomic responsiveness. These findings are particularly relevant for precision N management because soil N supply can vary substantially within producer fields, causing different areas of the same field to respond differently to fertilizer N (Reese et al., 2002; Schwalbert et al., 2019; Argento et al., 2021).

5 Limitations and Future Research

One limitation of this study was that the soil nitrate profile was characterized at each site level rather than for individual plots or within-field management zones. Unmeasured within-site

variation in residual nitrate may therefore have contributed to plot-to-plot variation in spectral response that was attributed to fertilizer rate. This would add noise to the site-level detectability estimates but is unlikely to have changed the ranking of environments, which was driven by large among-site differences in profile nitrate (29 to 270 kg N ha^{-1}). The relationships among soil N, yield response, and spectral detectability were also based on only seven site-level observations, so they should be viewed as indicative rather than definitive, as reflected by P values near 0.09 for soil nitrate. Future work should consider real field-scale evaluation of N-rich strips combined with spatial soil N sampling, which would be particularly useful for determining whether the relationships observed across experimental environments also occur in heterogeneous production fields. In addition, the interval between N application and each UAV flight differed among locations (Table 3.1), so detectability within a developmental stage partly reflects how long the crop had been responding to fertilizer; stage-level differences should be interpreted together with that timing. The sensor also provided a single red-edge band near 735 nm. Vuille-dit-Bille et al. (2025) found that an index based on a 705-nm red-edge band was particularly informative for crop N before heading, so the weak detectability observed during tillering may partly reflect the band available here rather than a true absence of N response. Additional evaluation using sensors with multiple red-edge or hyperspectral bands may further improve discrimination of crop N status, particularly during earlier developmental stages. Recent evidence that multiple red-edge bands improve the characterization of wheat N and yield across sites and seasons supports this direction (Vuille-dit-Bille et al., 2025). Ultimately, calibration against economically optimal N rates will be required before the approach can be translated from identification of N-responsive zones to direct variable-rate fertilizer prescriptions.

Moreover, the contrasting behavior of grain yield and grain protein also highlights the need for a separate framework for sensing and managing grain-quality responses to N. Future studies should evaluate at the real-world farmer-field scale, considering landscape-induced variation in soil nutrients, moisture, and local growing conditions and their interactions. Future studies should also consider whether late-season multispectral or narrow-band hyperspectral sensing can characterize post-anthesis N uptake, N remobilization, and subsequent grain protein accumulation.

6 Conclusion

UAV-based detection of fertilizer-N response in winter wheat depended on both spectral characteristics and the agronomic environment in which the crop is growing. Red-edge-based indices provided greater detectability of N stress. Differences in spectral detectability among environments were closely associated with differences in agronomic N responsiveness, while greater profile soil nitrate tended to correspond with weaker spectral and yield responses. Normalizing UAV-derived VIs to a local N-rich reference further strengthened their relationship with fertilizer-induced yield response. Together, these findings indicate that UAV sensing should be interpreted as a measure of crop response to N availability within its local soil and environmental context rather than as a direct indicator of fertilizer rate alone. In spatially variable fields, integrating soil N information, red-edge UAV imagery, and local N-rich references may provide a practical framework for identifying N-responsive management zones and improving site-specific N management.

References

Argento, F., Anken, T., Abt, F., Vogelsanger, E., Walter, A., & Liebisch, F. (2021). Site-specific nitrogen management in winter wheat supported by low-altitude remote sensing and soil data. Precision Agriculture, 22, 364–386. https://doi.org/10.1007/s11119-020-09733-3

Barnes, E. M., Clarke, T. R., Richards, S. E., Colaizzi, P. D., Haberland, J., Kostrzewski, M., Waller, P., Choi, C., Riley, E., Thompson, T., Lascano, R. J., Li, H., & Moran, M. S.

(2000). Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data. In *Proceedings of the Fifth International Conference on Precision Agriculture* (pp. 1–15). American Society of Agronomy, Madison, WI.

Bausch, W. C., & Delgado, J. A. (2005). Impact of residual soil nitrate on in-season nitrogen applications to irrigated corn based on remotely sensed assessments of crop nitrogen status. *Precision Agriculture, 6*(6), 509–519. https://doi.org/10.1007/s11119-005-5641-9

Bogard, M., Allard, V., Brancourt-Hulmel, M., Heumez, E., Machet, J.-M., Jeuffroy, M.-H., Gate, P., Martre, P., & Le Gouis, J. (2010). Deviation from the grain protein concentration–grain yield negative relationship is highly correlated to post-anthesis N uptake in winter wheat. Journal of Experimental Botany, 61(15), 4303–4312. https://doi.org/10.1093/jxb/erq238

Cao, Q., Miao, Y., Feng, G., Gao, X., Li, F., Liu, B., Yue, S., Cheng, S., Ustin, S. L., & Khosla, R. (2015). Active canopy sensing of winter wheat nitrogen status: An evaluation of two sensor systems. Computers and Electronics in Agriculture, 112, 54–67. https://doi.org/10.1016/j.compag.2014.08.012

Cao, Q., Miao, Y., Wang, H., Huang, S., Cheng, S., Khosla, R., & Jiang, R. (2013). Non-destructive estimation of rice plant nitrogen status with Crop Circle multispectral active canopy sensor. Field Crops Research, 154, 133–144. https://doi.org/10.1016/j.fcr.2013.08.005

Clay, D. E., Kharel, T. P., Reese, C., Beck, D., Carlson, C. G., Clay, S. A., & Reicks, G. (2012). Winter wheat crop reflectance and nitrogen sufficiency index values are influenced by nitrogen and water stress. Agronomy Journal, 104(6), 1612–1617. https://doi.org/10.2134/agronj2012.0216

Clay, D. E., Kim, K.-I., Chang, J., Clay, S. A., & Dalsted, K. (2006). Characterizing water and nitrogen stress in corn using remote sensing. Agronomy Journal, 98(3), 579–587. https://doi.org/10.2134/agronj2005.0204

Colaço, A. F., & Bramley, R. G. V. (2018). Do crop sensors promote improved nitrogen management in grain crops? Field Crops Research, 218, 126–140.

Colaço, A. F., Richetti, J., Bramley, R. G. V., & Lawes, R. A. (2021). How will the next-generation of sensor-based decision systems look in the context of intelligent agriculture? A case-study. Field Crops Research, 270, 108205.

Dash, J., & Curran, P. J. (2004). The MERIS terrestrial chlorophyll index. *International Journal of Remote Sensing, 25*(23), 5403–5413. https://doi.org/10.1080/0143116042000274015

Datt, B. (1999). A new reflectance index for remote sensing of chlorophyll content in higher plants: Tests using Eucalyptus leaves. Journal of Plant Physiology, 154(1), 30–36. https://doi.org/10.1016/S0176-1617(99)80314-9

Fischer, R. A. (1985). Number of kernels in wheat crops and the influence of solar radiation and temperature. Journal of Agricultural Science, 105(2), 447–461. https://doi.org/10.1017/S0021859600056495

Flowers, M., Weisz, R., & Heiniger, R. (2003). Quantitative approaches for using color infrared photography for assessing in-season nitrogen status in winter wheat. Agronomy Journal, 95, 1189–1200. https://doi.org/10.2134/agronj2003.1189

Giordano, N., Sadras, V. O., & Lollato, R. P. (2023). Late-season nitrogen application increases grain protein concentration and is neutral for yield in wheat: A global meta-analysis. Field Crops Research, 290, 108740. https://doi.org/10.1016/j.fcr.2022.108740

Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sensing of Environment, 58(3), 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A. A., Viña, A., Ciganda, V., Rundquist, D. C., & Arkebauer, T. J. (2005). Remote estimation of canopy chlorophyll content in crops. Geophysical Research Letters, 32(8), L08403. https://doi.org/10.1029/2005GL022688

Guillard, K., Miele, K. M., Sürer, I., & Morris, T. F. (2021). Guiding fall fertilization of cool-season turfgrass lawns with NDVI sufficiency index. Crop Science, 61, 784–803. https://doi.org/10.1002/csc2.20301

Hawkesford, M. J. (2014). Reducing the reliance on nitrogen fertilizer for wheat production. Journal of Cereal Science, 59(3), 276–283. https://doi.org/10.1016/j.jcs.2013.12.001

Hirel, B., Le Gouis, J., Ney, B., & Gallais, A. (2007). The challenge of improving nitrogen use efficiency in crop plants: Towards a more central role for genetic variability and quantitative genetics within integrated approaches. Journal of Experimental Botany, 58(9), 2369–2387. https://doi.org/10.1093/jxb/erm097

Holland, K. H., & Schepers, J. S. (2010). Derivation of a variable rate nitrogen application model for in-season fertilization of corn. Agronomy Journal, 102, 1415–1424. https://doi.org/10.2134/agronj2010.0015

Huete, A. R. (1988). A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment, 25(3), 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Joshi, D. R., Clay, S. A., Sharma, P., Rekabdarkolaee, H. M., Kharel, T., Rizzo, D. M., Thapa, R., & Clay, D. E. (2024a). Artificial intelligence and satellite-based remote sensing can be used to predict soybean (Glycine max) yield. Agronomy Journal, 116(3), 917–930. https://doi.org/10.1002/agj2.21473

Joshi, D. R., Clay, D. E., Clay, S. A., Moriles-Miller, J., Daigh, A. L. M., Reicks, G., & Westhoff, S. (2024b). Quantification and machine learning based N2O–N and CO2–C emissions predictions from a decomposing rye cover crop. Agronomy Journal, 116, 795–809. https://doi.org/10.1002/agj2.21185

Kharel, T. P., Tyler, H. L., Mubvumba, P., Huang, Y., Bhandari, A. B., Fletcher, R. S., Anapalli, S., Joshi, D. R., Mengistu, A., Birru, S., Adhikari, K., Dhakal, M., Maskey, L., Reddy, K. N., & Clay, D. E. (2025). Machine learning on multi-spectral imagery to estimate nutrient yield of mixed-species cover crops. Agricultural & Environmental Letters, 10, e70009. https://doi.org/10.1002/ael2.70009

Large, E. C. (1954). Growth stages in cereals: Illustration of the Feekes scale. Plant Pathology, 3(4), 128–129.

Li, F., Miao, Y., Zhang, F., Cui, Z., Li, R., Chen, X., Zhang, H., Schroder, J., Raun, W. R., & Jia, L. (2014). In-season prediction of grain yield and protein content of winter wheat using multispectral imagery. Remote Sensing of Environment, 142, 1–13.

Li, F., Mistele, B., Hu, Y., Chen, X., & Schmidhalter, U. (2014). Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression. European Journal of Agronomy, 52, 198–209. https://doi.org/10.1016/j.eja.2013.09.006

Li, F., Zhang, H., Jia, L., Bareth, G., Miao, Y., & Chen, X. (2010). Estimating winter wheat biomass and nitrogen status using an active crop sensor. Intelligent automation and soft computing, 16(6), 1221-1230.

Li, Y., Miao, Y., Zhang, J., Cammarano, D., Li, S., Liu, X., Tian, Y., Zhu, Y., Cao, W., & Cao, Q. (2022). Improving estimation of winter wheat nitrogen status using random forest by integrating multi-source data across different agro-ecological zones. Frontiers in Plant Science, 13, 890892. https://doi.org/10.3389/fpls.2022.890892

Lucero, M., Adee, E., Ruiz Diaz, D., Sullivan, T., Jha, G., Simon, L., & Joshi, D. R. (2026). Machine learning insights into long-term soil fertility legacy and seasonal weather effects in corn–soybean rotation. Soil Science Society of America Journal, 90(5), e70334.

Morris, K. B., Martin, K. L., Freeman, K. W., Teal, R. K., Girma, K., Arnall, D. B., Hodgen, P. J., Mosali, J., Raun, W. R., & Solie, J. B. (2006). Mid-season recovery from nitrogen stress in winter wheat. Journal of Plant Nutrition, 29(4), 727–745. https://doi.org/10.1080/01904160600567066

Mullen, R. W., Freeman, K. W., Raun, W. R., Johnson, G. V., Stone, M. L., & Solie, J. B. (2003). Identifying an in-season response index and the potential to increase wheat yield with nitrogen. Agronomy Journal, 95(2), 347–351. https://doi.org/10.2134/agronj2003.3470

Prey, L., & Schmidhalter, U. (2019). Sensitivity of vegetation indices for estimating vegetative N status in winter wheat. Sensors, 19(17), 3712. https://doi.org/10.3390/s19173712

Python Software Foundation. (2024). Python (Version 3.13) [Computer software]. https://www.python.org

Raun, W. R., & Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. Agronomy Journal, 91(3), 357–363. https://doi.org/10.2134/agronj1999.00021962009100030001x

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., Mullen, R. W., Freeman, K. W., Thomason, W. E., & Lukina, E. V. (2002). Improving nitrogen use efficiency in cereal grain production with optical sensing and variable rate application. Agronomy Journal, 94(4), 815–820. https://doi.org/10.2134/agronj2002.8150

Raun, W. R., Solie, J. B., Stone, M. L., Martin, K. L., Freeman, K. W., Mullen, R. W., Zhang, H., Schepers, J. S., & Johnson, G. V. (2005). Optical sensor-based algorithm for crop nitrogen fertilization. Communications in Soil Science and Plant Analysis, 36(19), 2759–2781. https://doi.org/10.1080/00103620500303988

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

Reese, C. L., Clay, D. E., Beck, D., Liu, Z., & Carlson, C. G. (2002). The influence of landscape position and nitrogen rate on hard red spring wheat (Triticum aestivum L.) quality and quantity. In P. C. Robert et al. (Eds.), Precision agriculture: Proceedings of the 6th International Conference (Minneapolis, MN, 14–17 July 2002). ASA, CSSA, and SSSA.

Rouse, J. W., Jr., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. In Third Earth Resources Technology Satellite-1 Symposium, Vol. 1, Section A (pp. 309–317). NASA Goddard Space Flight Center

Ruiz Diaz, D. A., et al. (2023). Soil test interpretations and fertilizer recommendations in Kansas (MF2586). Kansas State University Agricultural Experiment Station and Cooperative Extension Service.

Schwalbert, R. A., Amado, T. J. C., Reimche, G. B., & Gebert, F. (2019). Fine-tuning of wheat (Triticum aestivum, L.) variable nitrogen rate by combining crop sensing and management zones approaches in southern Brazil. Precision Agriculture, 20(1), 56–77. https://doi.org/10.1007/s11119-018-9581-6

Singh, R., Kaur, S., Joshi, D. R., Iboyi, J., Dar, E., Sharma, L. K., & Singh, H. (2025). Estimating cotton biomass and nitrogen content using high-resolution satellite and UAV data fusion with machine learning. Smart Agricultural Technology, 12, 101191.

Slafer, G. A., Savin, R., & Sadras, V. O. (2014). Coarse and fine regulation of wheat yield components in response to genotype and environment. Field Crops Research, 157, 71–83. https://doi.org/10.1016/j.fcr.2013.12.004

Souza, J. L. B., Antonangelo, J. A., Silva, A. de O., Reed, V., & Arnall, B. (2022). Recovery of grain yield and protein with fertilizer application post nitrogen stress in winter wheat (Triticum aestivum L.). Agronomy, 12(9), 2024. https://doi.org/10.3390/agronomy12092024

Vuille-dit-Bille, N., Levy Häner, L., Strebel, S., Burton, A., Schaad, N., Pellet, D., Wuyts, N., Treier, S., de F. Bongiovani, P., & Herrera, J. M. (2025). Evaluation of winter wheat varieties' responses to nitrogen supply supported by 10-band multispectral aerial imaging: Reproducibility over crop seasons and sites. *Smart Agricultural Technology, 12*, 101528. https://doi.org/10.1016/j.atech.2025.101528

Walsh, O. S., Marshall, J., Jackson, C., Nambi, E., Shafian, S., Jayawardena, D. M., Lamichhane, R., Owusu Ansah, E., & McClintick-Chess, J. R. (2022). Wheat yield and protein estimation with handheld- and UAV-based reflectance measurements. Agrosystems, Geosciences & Environment, 5, e20309. https://doi.org/10.1002/agg2.20309

Wang, J., Meyer, S. T., Xu, X., Weisser, W. W., & Yu, K. (2025). Drone multispectral imaging captures the effects of soil mineral nitrogen on canopy structure and nitrogen use efficiency in wheat. Computers and Electronics in Agriculture, 235, 110342. https://doi.org/10.1016/j.compag.2025.110342

Xu, G., Fan, X., & Miller, A. J. (2012). Plant nitrogen assimilation and use efficiency. Annual Review of Plant Biology, 63, 153–182. https://doi.org/10.1146/annurev-arplant-042811-105532

Chapter 4 - Conclusions

Summary of Findings

This thesis evaluated UAV multispectral sensing of hard red winter wheat across seven Kansas environments during the 2024–2025 growing season for two related purposes: predicting grain yield and protein at locations not used for model development (Chapter 2), and detecting crop response to fertilizer N across environments with contrasting background soil N supply (Chapter 3).

In Chapter 2, combining stage-specific UAV vegetation indices with agronomic information supported prediction of grain yield and protein at held-out locations under nested leave-one-location-out validation. The final partial least squares regression (PLSR) models achieved $R^2 = 0.55$ and CCC = 0.72 for grain yield and $R^2 = 0.47$ and CCC = 0.67 for grain protein, and PLSR outperformed random forest and XGBoost under the same framework. GNDVI at stem elongation was the dominant yield predictor and was selected in all seven outer folds, followed by N rate, whereas MTCI at grain formation, N rate, and plant height were the most important protein predictors. Cultivar information made a modest contribution to yield prediction but a large contribution to protein prediction, with protein R^2 decreasing from 0.47 to 0.14 when cultivar indicators were removed.

In Chapter 3, spectral detectability of fertilizer-N response increased with crop development, and the clearest advantage of red-edge-based indices occurred during stem elongation (Feekes 6–9), when NDRE, CI_{RE}, MTCI, and DATT were more detectable than NDVI and spectral response to N was consistent among cultivars. Across environments, profile soil NO_3-N ranged from 29 to 270 kg N ha^{-1}, and grain-yield response to fertilizer N ranged from −0.7% to 105%. Environments with greater background soil nitrate tended to show weaker yield response

and weaker red-edge detectability ($\rho = -0.68$, $P = 0.094$), and overall red-edge detectability ranked environments in the same order as relative yield response ($\rho = 1.00$, $P < 0.001$). Grain protein increased with fertilizer N at all locations, but the size of that increase was unrelated to background soil nitrate or red-edge detectability. Expressing vegetation indices relative to a local N-rich reference strengthened their association with yield response; for NDVI, the magnitude of the correlation increased from $|\rho| = 0.36–0.79$ using absolute values to $|\rho| = 0.89–0.96$ using the sufficiency index.

Synthesis Across Chapters

Together, the two studies point to several common conclusions. First, the value of UAV information depended on crop developmental stage. Imagery collected during stem elongation was the most informative in both studies: GNDVI at stem elongation was the most stable transferable yield predictor in Chapter 2, and stem elongation provided the clearest red-edge advantage and the most consistent cultivar response in Chapter 3. Red-edge information during grain formation, represented by MTCI, was more closely associated with grain protein.

Second, grain yield and grain protein were related to UAV spectral information in different ways. Canopy reflectance was more closely tied to grain yield and its response to N, whereas grain protein depended more strongly on management and genetic information. In Chapter 3, the protein response to fertilizer N was unrelated to spectral N detectability, and in Chapter 2, protein prediction relied on N rate, cultivar, and plant height in addition to MTCI at grain formation. These results indicate that multispectral canopy sensing alone is not sufficient for estimating grain protein and that protein-focused applications require additional information.

Third, UAV measurements need to be interpreted within their local agronomic context. In Chapter 3, background soil N supply influenced both the agronomic and the spectral response to

the same fertilizer-N gradient, and normalization to a local N-rich reference improved the agronomic interpretation of vegetation indices across environments. In Chapter 2, prediction accuracy at held-out locations was lower than values reported in studies that randomly split observations, but these predictions were obtained without any data from the target location. The simpler PLSR model also outperformed the more flexible machine-learning models, suggesting that, with a limited number of independent environments, greater model complexity did not improve transferability.

Implications for Nitrogen Management

For winter wheat N management in Kansas, these findings suggest that UAV multispectral imagery is most useful when it is combined with information on local soil N supply and interpreted relative to a local N-rich reference. Red-edge imagery collected during stem elongation provided the clearest indication of crop response to N, while profile soil nitrate helped explain why some environments responded little to additional fertilizer. In spatially variable fields, integrating soil N information, red-edge UAV imagery, and N-rich references may help identify N-responsive management zones. For pre-harvest prediction, combining UAV vegetation indices with known agronomic information such as N rate, cultivar, and plant height provides a more transferable basis for estimating grain yield and protein than spectral information alone.

Limitations and Future Research

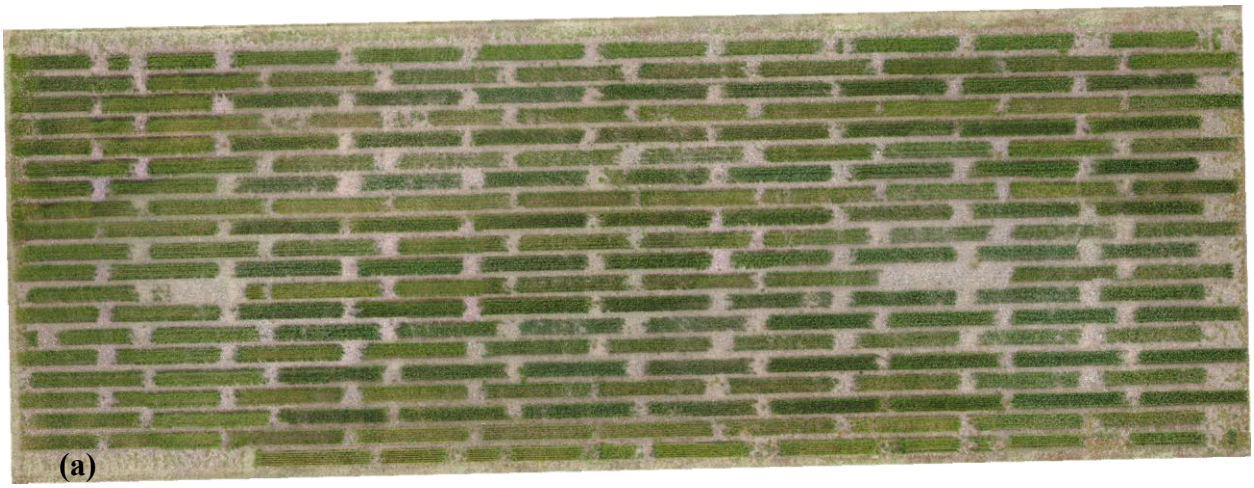
Both studies were conducted in replicated research plots at seven environments during a single growing season. The results therefore describe variation among locations within one season and do not account for year-to-year variation in weather, phenology, and crop response. Soil nitrate was also characterized at the site level rather than within plots or management zones, and the multispectral sensor provided a single red-edge band. Future research should test these approaches

across multiple years and in commercial production fields with field-scale soil and weather information; evaluate N-rich strips combined with spatial soil N sampling in heterogeneous fields; assess whether sensors with multiple red-edge or hyperspectral bands improve the detection of crop N status during early development and the characterization of grain protein; and examine whether locally calibrated or transfer-learning approaches improve prediction in environments that differ from those used for model development. Calibration of spectral indicators against economically optimal N rates will also be required before these approaches can be translated into variable-rate fertilizer prescriptions.

Overall, this thesis shows that UAV multispectral imagery can provide useful information for winter wheat N management and for pre-harvest prediction of grain yield and protein in Kansas. That information was most reliable when interpreted in its agronomic and environmental context, considering crop developmental stage, background soil N supply, a local N-rich reference, and known management and cultivar information, rather than used in isolation.

Appendix A - Supplementary Figures for Chapter 3

Appendix Figure A.1. Field layout and split-plot experimental design used for the winter wheat nitrogen study. The upper panel shows a representative UAV image of the experimental field, and the lower panel illustrates the treatment arrangement within one replicate block.



(b) Block I

90	60	180	150	120	30	0
LCC 101	SYM 109	LCC 117	4269 125	SYW 133	4401 141	SYW 149
4401 102	LCC 110	4401 118	SYW 126	SYM 134	BD 142	4699 150
4269 103	4699 111	4269 119	SYM 127	4401 135	SYM 143	4269 151
SYW 104	LCJ 112	BD 120	LCJ 128	4699 136	4699 144	SYM 152
4699 105	BD 113	LCJ 121	4699 129	4269 137	LCC 145	4401 153
BD 106	4269 114	4699 122	BD 130	LCC 138	4269 146	LCC 154
LCJ 107	SYW 115	SYM 123	LCC 131	LCJ 139	SYW 147	BD 155
SYM 108	4401 116	SYW 124	4401 132	BD 140	LCJ 148	LCJ 156

Whole plot: N rate, lb N ac^{-1} (kg N ha^{-1})

0 (0) | 30 (34) | 60 (67) | 90 (101) | 120 (135) | 150 (168) | 180 (202)

Subplot: cultivar

BD	Bob Dole	**SYM**	SY Monument	**4401**	WB4401
LCC	LCS Chrome	**SYW**	SY Wolverine	**4699**	WB4699
LCJ	LCS Julep	**4269**	WB4269		

Small numbers = plot ID. Subplot: 7 rows at 19.1 cm (7.5 in) spacing × 9.1 m (30 ft). Schematic; not to scale.

Appendix Figure A.2. Cultivar modification of UAV spectral response to fertilizer N across winter wheat developmental stages. (A) Percentage and number of environments exhibiting a significant N rate × cultivar interaction for each vegetation index. Values within cells indicate the percentage of the seven environments with a significant interaction; the corresponding number of environments is shown below. (B) Overall percentage of significant N rate × cultivar interaction tests across eight vegetation indices and seven environments within each developmental stage. Spectral response to fertilizer N was most consistent among cultivars during stem elongation.

