Received: 5 December 2020 | Accepted: 26 January 2021

DOI: 10.1002/agj2.20620

Agronomy Journal

ARTICLE

Soil Fertility & Crop Nutrition

Active optical sensor measurements and weather variables for predicting winter wheat yield

Lawrence Aula[1] | **Peter Omara**[1,2] | **Eva Nambi**[1] | **Fikayo B. Oyebiyi**[1] | **Jagmandeep Dhillon**[3] | **Elizabeth Eickhoff**[1] | **Jonathan Carpenter**[1] | **William R. Raun**[1]

[1] Dep. of Plant and Soil Sciences, Oklahoma State Univ., Stillwater, OK, 74078, USA

[2] Dep. of Agronomy, Faculty of Agriculture and Environment, Gulu Univ., Gulu, Uganda

[3] Dep. of Plant and Soil Sciences, Mississippi State Univ., Mississippi State, MS, 39762, USA

Correspondence
Lawrence Aula, Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK, 74078, USA.
Email: aula@okstate.edu

Associate Editor: Maryse Bourgault

Abstract

Accurate winter wheat (*Triticum aestivum* L.) grain yield prediction is vital for improving N management decisions. Currently, most N optimization algorithms use in-season estimated yield (INSEY) as a sole variable for predicting grain yield potential (YP). Although evidence suggests that this works, the yield prediction accuracy could be further improved by including other predictors in the model. The objective of this work was to evaluate INSEY, pre-plant N rate, total rainfall, and average air temperature from September to December as predictors of winter wheat YP. An 8-yr (2012–2019) data set for grain yield was obtained from Experiment 502, Lahoma, OK. The experiment was designed as a randomized complete block with four replications and N applied at 0, 45, 67, 90, and 112 kg ha^{-1}. Weather data was obtained from the Oklahoma Mesonet (http://mesonet.org). The data were analyzed using *R* statistical computing platform. The best model was selected using least absolute shrinkage and selection operator. Root mean square error (RMSE) was obtained using *k*-fold cross-validation. The model selection algorithm produced the full model as the best model for yield prediction with an R^2 of .79 and RMSE of 0.54 Mg ha^{-1}. The best one-variable model – as expected – used INSEY as the predictor and had the highest RMSE of 0.72 Mg ha^{-1} and an R^2 of .62. Mid-season YP prediction accuracy could be improved by including pre-plant N rate, mean air temperature, and total rainfall from September to December in a model already containing INSEY.

1 | INTRODUCTION

Winter wheat (*Triticum aestivum* L.) is one of the most important cereal crops produced around the world. Globally, it is grown on more than 19% (220 million ha) of the crop production area (Dhillon et al., 2019) and yielded 0.74 billion Mg in 2015 (Omara et al., 2019). Because of its importance, researchers are constantly exploring new opportunities to improve wheat production and productivity. A substantial amount of N is applied each year to sustain or improve yield levels. In 2015 alone, 61.2 million Mg of N was applied worldwide for cereal production (Omara et al., 2019). In some cases, producers apply N beyond requirements for the maximum yield of a given crop (Cui et al., 2008; Ju et al., 2004).

Abbreviations: GDD, growing degree days; INSEY, in-season estimated yield; LASSO, least absolute shrinkage and selection operator; NDVI, normalized difference vegetation index; RMSE, root mean square error; YP, yield potential.

This increases the likelihood of a significant loss of some of the applied N. This is why understanding the yield level is crucial for a variety of reasons including N management decision for a particular field. In the past – and in some cases today – yield goal was a common method for estimating the yield potential (YP) of the crop-growing environment before planting (Dahnke et al., 1988; Raun et al., 2017). This approach involves averaging grain yield for the past 3–5 yr plus 10–20% of the mean grain yield (Raun et al., 2017). Based on this yield estimate, an amount of N to apply is arrived at using the assumption that 1 Mg wheat grain yield requires 33 kg N (Zhang & Raun, 2006). However, in a comprehensive evaluation of yield goal, Raun et al. (2017) concluded that yield goal explains only a small proportion of the variability in winter wheat grain yield (<0.01–16%), making it less effective at estimating YP and consequently the quantity of N to apply. Nonetheless, it is still an approach used by producers around the world.

In order to overcome the inadequacy of relating yield goal to N rate, other researchers evaluated alternative approaches such as economic or agronomic optimum N rate. In this approach, the amount of N to apply is estimated from a linear plateau, quadratic, or quadratic plateau fit (Makowski et al., 1999; Ortuzar-Iragorri et al., 2010; Thomason et al., 2011). These methods assume that there is a point (N rate) beyond which application of more N results in no added yield benefit and may begin to decline. It is this rate that these methods assume will result in grain yield that generates the best economic return to the producer. Sometimes, soil samples are taken to quantify the amount of N present in the soil prior to planting (Bundy & Andraski, 2004). This allows for adjustment of the amount of N recommended by some of the methods highlighted above. While these approaches may improve estimates for crop N demand, they do not extensively address changes in soil N supply that may occur during the course of the crop growing season.

Several other methods exist for predicting YP for different crops (Ransom et al., 2020) but perhaps the one that gained the most widespread attention involves the use of optical sensors (Cao et al., 2015; Dhillon, Eickhoff, et al., 2020; Franzen et al., 2016; Li et al., 2009; Liu et al., 2017; Shanahan et al., 2008). In this approach, spectral reflectance values are collected mid-season and then normalized to give a normalized difference vegetation index (NDVI). By obtaining NDVI mid-season, producers are able to determine whether there is a need for additional N or not. According to Raun et al. (2002), NDVI for winter wheat obtained between Feekes 4 (FK_4) and Feekes 6 (FK_6) growth stages improves the precision of prediction of YP without added N (YP_0). When these NDVI values are divided by growing degree days (GDD) to obtain the in-season estimated yield (INSEY) and used to predict observed yield, the accuracy with which yield is predicted is greatly improved. As much as 54% of the grain yield variability can be explained using INSEY compared to the mass balance method such as yield goal that one study found to account for no more than 16% of the grain yield variations (Raun et al., 2017). Raun et al. (2005) also reported that crop response to N (response index or RI) can be obtained by dividing NDVI from N-rich strip by NDVI from plots with less or no N applied (farmer practice). Knowing the RI then facilitates the prediction of YP with added N (YP_N) (Raun et al., 2002). Using this approach, it is possible to more accurately recommend and apply N mid-season to obtain a high yield in a given wheat-growing environment (Morris et al., 2006; Mullen et al., 2003). In a recent review of active optical sensors, Aula et al. (2020) found that this approach for N management could save as much as 53 kg N ha^{-1} while simultaneously sustaining grain yield level.

Core Ideas

- The model that used INSEY alone had a high RMSE of 0.72 Mg ha^{-1}.
- Yield prediction improved for a model that used INSEY, N, and weather variables (0.54 Mg ha^{-1}).
- After INSEY, N rate was the second most important predictor for yield prediction.

Despite the improved yield prediction accuracy, the GreenSeeker (Trimble Navigation Limited for GreenSeeker) algorithm has opportunities for improvement (Aula et al., 2020). This is evidenced by the 54% of the variability in grain yield explained by the algorithm using INSEY as a sole predictor (Raun et al., 2005). This suggests that approximately 46% of the variability in grain yield remains unexplained by the model. When Colaço and Bramley (2019) analyzed data from multiple years in two locations, INSEY was found to explain even a much lower proportion of the variability in grain yield with an R^2 of .25. This is possibly because there are numerous environment variables that need to be included in future algorithms for an improved yield prediction and N estimation (Raun et al., 2019). This indicates that INSEY as a sole predictor may become less effective at accurately predicting YP due to increased entropy over time. Yet, most optical sensor algorithms that rely on INSEY do not include other predictors of YP (Colaço & Bramley, 2019; Li et al., 2009; Raun et al., 2001). They assume that much of the variability in yield is explained by INSEY. As highlighted above, INSEY may not explain all the variations in grain yield, making it necessary to consider other variables that could improve the predictive capability of future algorithms. Our study hypothesizes that pre-plant N rate, total rainfall, and average temperature from September to December could play a major role in improving yield prediction and estimate of the quantity of N to apply mid-season. Colaço and Bramley (2019) also reported the need to include other variables that may improve yield prediction accuracy. In a maize (*Zea mays* L.) study, Dhillon, Aula, et al. (2020) revealed that including weather variables such as fractional water index, soil, and air temperatures increases the possibility of accurately predicting YP mid-season. As such, the inclusion of more predictors in the algorithm for winter wheat presents an opportunity to improve its robustness. The variables that this study intends to evaluate may influence wheat grain yield in different ways and each may have a unique contribution in predicting grain yield. For example, winter wheat grain yield is likely lower as temperature increases above the optimum level (Batts et al., 1997; Gibson & Paulsen, 1999; Kristensen et al., 2011). This is particularly true if CO_2 concentration in the atmosphere is low (Batts et al., 1997). Generally, temperature has been increasing over the decades and more rapidly during colder periods (Stewart et al., 2018). The recent rise in winter temperatures may make yield prediction more challenging with one independent variable (Raun et al., 2019). During the September–December period, winter wheat also undergoes the vernalization process in order to promote flowering later (Streck et al., 2003) and any change in temperature is likely to influence the final grain yield, making it potentially necessary to improve the robustness of an algorithm for yield prediction. Winter wheat also requires enough soil moisture for proper establishment and development (He et al., 2016) and this possibly begins with promoting high seed germination and emergence to achieve a uniform plant stand. Stone and Schlegel (2006) showed that soil water content at emergence was able to explain as much as 70% of the variation in grain yield. This suggests that adequate rainfall to replenish soil moisture during germination and emergence could play a role in grain yield prediction. Meanwhile, other authors reported that winter precipitation and spring temperature have less influence on grain yield (Kristensen et al., 2011). Embedding the contribution from each of these predictors in the algorithm may increase the likelihood of accurately predicting mid-season YP, a key step for estimating the quantity of N to apply.

Therefore, this study aims to improve the accuracy of prediction of YP using INSEY and selected weather variables.

2 | MATERIALS AND METHODS

2.1 | Site description and experimental and treatment designs

Data sets from Experiment 502 (E502) located in Lahoma, OK (36˚23′15″ N, 98˚06′30″ W) were used to build and validate a model for yield prediction. This experiment was established in 1971 under continuous winter wheat. The soil at the experimental site is a Grant silt loam with a 1–3% slope (fine-silty, mixed, superactive, thermic, Udic Argiustoll). The site was under conventional tillage from 1971 to 2010 and was changed to no-till in 2010–2011 and that continues today.

The experiment was set up as a randomized complete block design with 14 treatments and four replications. For this study, six treatments where N was the only factor with variable levels (0, 22, 45, 67, 90, and 112 kg N ha^{-1}) were selected. Urea (46–0–0) was applied pre-plant as the source of N. The study also selected 8 yr of data for developing the model, that is, from 2012 to 2019. These years were selected because they represented a period where full records of NDVI data for FK_4 and FK_5 required for algorithm development were available. These stages correspond to between 90 and 110 GDD needed to accurately predict grain YP using active optical sensor measurements (Dhillon, Figueiredo, et al., 2020). Normalized difference vegetation index collected using the GreenSeeker sensor was calculated as shown in Equation 1 (Bushong et al., 2016):

$$\text{INSEY} = \frac{\text{NDVI at } (\text{FK}_4 + \text{FK}_5)}{\text{GDD at FK}_4} \tag{1}$$

where NIR, near-infrared surface reflectance measured at a specific wavelength of 780 nm while Red was surface reflectance measured at a wavelength of 660 nm in the visible region of the electromagnetic spectrum.

Each year, grain harvest was accomplished using a self-propelled combine and grain yield adjusted to 12.5% moisture content. From each experimental unit in each year, NDVI data was collected at FK_4 and FK_5. A total of two NDVI values were collected from each crop growth stage associated with a particular treatment within a period of 60 s. These values were averaged to produce a single NDVI value. The time for collection of NDVI was between 9:00 and 10:00 a.m., U.S. central time. Growing degree days at each of the growth stages were also obtained from the Oklahoma Mesonet (2020). These stages were selected because FK_4 accounts for plant growing conditions and its health while the second measurement at FK_5 corrects for radiometric errors (Raun et al., 2001). Rainfall and air temperature data from September to December were also obtained from the Oklahoma Mesonet (2020). The cropping season was from September to June. As a result, the grain yield recorded in 2012 was obtained using rainfall from September 2011 to June 2012. However, rains used in model development and validation were those that occurred early in the season (September–December) as they could permit prediction of the final grain yield well ahead of harvest and, as noted by Raun et al. (2002), use that information to estimate the quantity of N to apply mid-season. For instance, air temperature and rainfall used to build a model or make a prediction of the observed grain yield in 2012 was the average air temperature and sum of rainfall from September

TABLE 1 Treatment design for the regional experiment located at Hennessey, OK

Treatment	N applied		
	Pre-plant	Topdress	Total
	kg ha^{-1}		
1	0	0	0
2	168	0	168
3	28	0	28
4	28	28	56
5	28	56	84
6	28	84	112
7	28	112	140
8	28	140	168
9	56	56	112
10	112	0	112
11	56	0	56
12	84	0	84
13	140	0	140
14	225	0	225

to December 2011. The technical details of the Oklahoma Mesonet (2020) are well documented by Brock et al. (1995) and McPherson et al. (2007).

In-season estimated yield was obtained using Equation 2 below:

$$\text{INSEY} = \frac{\text{NDVI at } (\text{FK}_4 + \text{FK}_5)}{\text{GDD at FK}_4} \quad (2)$$

where GDD, growing degree day computed as $(\frac{T_{\max+}T_{\min}}{2} - T_{base} > 0)$ with $T_{\max}$, $T_{\min}$, and T_{base} defined as daily maximum, daily minimum, and base (4.4 ˚C) temperatures, respectively.

A more detailed description and evolution of yield prediction and N fertilization algorithms are contained in research work accomplished by Raun et al. (2001), Raun et al. (2002), and Raun et al. (2005).

Additional data was obtained from an experiment located at Hennessey, OK (36˚06′58.3″ N 97˚54′02.6″ W) – about 50 km from Lahoma experimental site – and used to further evaluate the performance of the model. This data used was pooled from two different years, that is, 2012 and 2013. The soil at Hennessey is a Bethany silt loam with a 0–1% slope (fine, mixed, superactive, thermic Pachic Paleustoll). This particular experiment had a randomized complete block design comprising of 14 treatments, each replicated four times. The treatment design consisted of various pre-plant N rates and a combination of pre-plant and topdress N rates ranging from 0 to 225 kg ha^{-1} (Table 1). Only the pre-plant N rate was used in the algorithm for yield prediction.

2.2 | Statistical analysis, model development, and cross-validation

Statistical analyses and algorithm development were performed using R statistical computing platform (R Core Team, 2020). The packages used included ggplot2 within the tidyverse package for data visualization (Wickham et al., 2019), glmnet for implementing the least absolute shrinkage and selection operator (LASSO) algorithm (Friedman et al., 2010), ggpmisc for generating and labelling p value and r^2 on the graphs (Aphalo, 2020) and readxl for importing MS Excel data into R (Wickham & Bryan, 2019). The predictors were evaluated and variables that best fit a model were selected using LASSO. In implementing LASSO, the 8-yr data set from the Lahoma experimental site was split into two equal parts, that is, training and validation sets.

The best one-variable, two-variable, and three-variable models were also fitted using LASSO after identifying them based on Bayesian Information Criterion (BIC) generated via best subset selection algorithm (James et al., 2013). This allowed for the evaluation of the performance of the most commonly used one-variable model that works with INSEY as a sole independent variable against models that used more than one predictor. The precision with which the model accurately predicted grain yield was obtained using root mean square error (RMSE). Root mean square error was obtained using k-fold cross-validation approach with k groups = 10 (James et al., 2013).

TABLE 2 Selection and evaluation of models for winter wheat grain yield prediction using data from Lahoma, OK

Model category	Equation[a,b]	RMSE[c]	R^2[d]
		Mg ha^{-1}	
1-variable model	Yield = −0.46 + 309.7 × INSEY	0.716	.620
2-variable model	Yield = −0.83 + 277.2 × INSEY + 0.013 × N rate	0.605	.729
3-variable model	Yield = −3.11 + 282.8 × INSEY + 0.013 × N rate + 0.18 × Temp	0.571	.758
4-variable model	Yield = −4.8 + 202.0 × INSEY + 0.014 × N rate+0.31 × Temp + 0.0041 × Rain	0.539	.785

[a]Model selection and parameter estimates were achieved using least absolute shrinkage and selection operator (LASSO). The predictor(s) that constituted each model category was determined using best subset selection algorithm followed by fitting a model using LASSO to obtain RMSE and R^2.

[b]The predictors included INSEY, in-season estimated yield; N rate, amount of nitrogen applied pre-plant (kg ha^{-1}); Rain, total rainfall from September to December (mm); Temp, mean air temperature from September to December (˚C).

[c]RMSE, root mean square error obtained using *k*-fold cross-validation.

[d]R^2 was obtained from the relationship between predicted and observed grain yield.

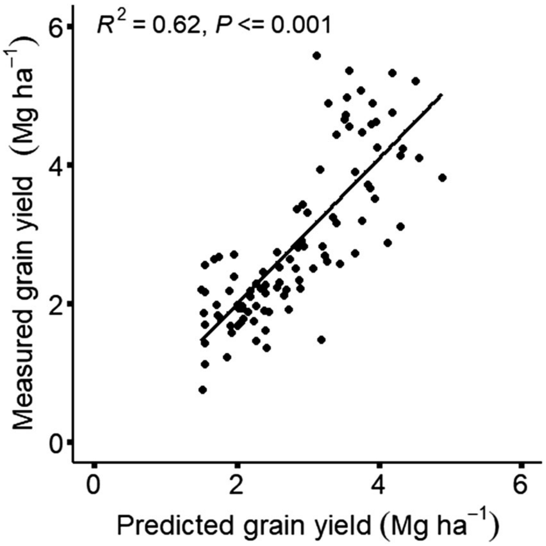

FIGURE 1 The relationship between predicted and measured grain yield for a model that used in-season estimated yield (INSEY) as the sole predictor (Yield = −0.46 + 309.7 × INSEY)

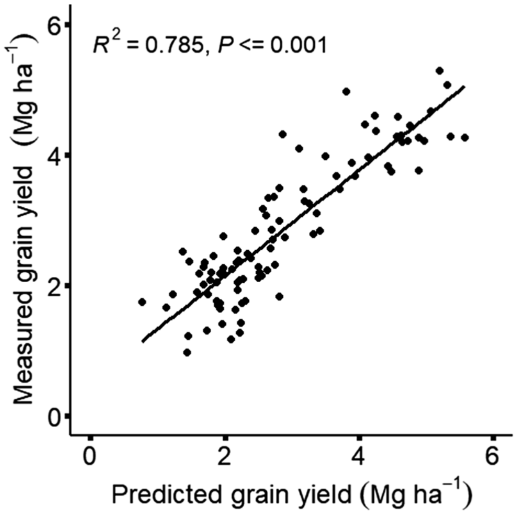

FIGURE 2 The relationship between predicted and observed grain yield at Lahoma, OK, obtained using an algorithm that selected all the four predictors (Yield = −4.8 + 202.0 × INSEY + 0.014 × N rate + 0.0041 × Rain + 0.31 × Temp). INSEY, in-season estimated yield; N rate, amount of nitrogen applied pre-plant; Rain, total rainfall from September to December; Temp, average air temperature from September to December

3 | RESULTS AND DISCUSSION

This study evaluated four predictors – INSEY, pre-plant N rate, total rainfall and average temperature from September to December – with the aim of selecting key variables for use in an algorithm to predict winter wheat grain yield. The results are shown in Table 2. Cross-validation showed that the model with the highest yield prediction accuracy used all the four predictors (full model). This model had a root mean square error (RMSE) of 0.54 Mg ha^{-1} (Table 2). This suggests that on average, the model predicted a grain yield that differed from the measured grain yield by ± 0.54 Mg ha^{-1}. Much of the information required to predict winter wheat grain yield was provided by INSEY. This was why it was the best one-variable model with an R^2 of .62 (Table 2, Figure 1). However, the best one-variable model also had the highest RMSE of 0.72 Mg ha^{-1}. The 0.62 coefficient of determination is within the range reported in other studies. For instance, Raun et al. (2005) and Li et al. (2009) indicated that a model using INSEY as a single independent variable explained between 43.0 and 72.9% of the variability in winter wheat grain yield.

When the model included all the predictors, RMSE was reduced from 0.72 Mg ha^{-1} when INSEY was the only predictor to 0.54 Mg ha^{-1} (Table 2). This was accompanied by an increase in R^2 from .62 to .79 (Figures 1 and 2). This is an indication that including pre-plant N rate, rainfall, and temperature to the model that already contained INSEY, improves winter wheat yield prediction accuracy.

This provided evidence that while INSEY explained about 62% of the variation in winter wheat grain yield, it is important to include other relevant predictors in order to improve the model's robustness to predict YP. In our study, including pre-plant N rate, rainfall, and temperature in the model allowed for more than 78% of the yield variability to be explained. A similar increase in precision of yield prediction was also reported by Zhang et al. (2017) where they found RMSE to reduce from as high as 0.60 Mg ha^{-1} when NDVI was the sole predictor to 0.39 Mg ha^{-1} when NDVI was used in the model alongside other predictors. In the past, other scholars also proposed including other predictors in the model that uses INSEY in order to increase the precision of yield prediction. For instance, Bushong et al. (2016) showed that using soil moisture and INSEY led to a lower RMSE of 0.92 Mg ha^{-1} compared to 0.95 Mg ha^{-1} generated with a model having INSEY as a sole predictor. Colaço and Bramley (2019) advocated for the inclusion of deep soil moisture data collected between 50 and 90 d after planting to realistically account for variability in wheat grain yield. They found INSEY from multiple years and locations to account for only 25% of the variation in grain yield. Walsh et al. (2013) multiplied soil moisture at 5-cm depth with INSEY and used the product as a predictor for grain yield and found it to have a higher R^2 of .61 when compared to .31 R^2 value associated with a model that used INSEY as the only independent variable. While these authors proposed including INSEY together with a different variable (soil moisture) to the ones used in our study, all the studies showed that yield prediction accuracy is higher when INSEY is used alongside other predictors.

Nonetheless, the model selection algorithm only considered weather variables when pre-plant N rate was already added to the model (Table 2). This meant that after INSEY, pre-plant N rate was the second most important variable included in the model. Including the pre-plant N rate to the model already containing INSEY, reduced RMSE from 0.72 to 0.61 Mg ha^{-1} (Table 2). In other words, the ability of the model to explain variation in wheat grain yield increased from 62.0 to 72.9% (Table 2). This illustrates the relevance of pre-plant N rate in improving our ability to accurately predict wheat grain YP.

Although NDVI used to generate INSEY reflected the plant N uptake (Cabrera-Bosquet et al., 2011; Tremblay et al., 2009), the quantity of N applied pre-plant may provide additional information that is relevant for predicting the final grain yield. Berntsen et al. (2006) found wheat grain yield to be associated with N applied. Raun et al. (2002) indicated that some doses of N applied pre-plant followed by the sensor-based recommended N rate is necessary for maximizing winter wheat grain YP. This pre-plant N could help to increase the number of tillers while at the same time minimizing competition that may occur among tillers without additional N application (Efretuei et al., 2016).

The algorithm for model selection further added mean temperature and total rainfall from September to December to the model that already had INSEY and pre-plant N rate (Table 2). This led to three-best and four-best variable models, respectively. Average temperature for September and December was first added to the two-variable model to form the best three-variable model followed by total rainfall to make the four-variable model (Table 2). Compared to the best two-variable model with INSEY and pre-plant N rate as the predictors, three- and four-variable models reduced RMSE from 0.61 to 0.57 and 0.54 Mg ha^{-1}, respectively (Table 2). The proportion of variance in yield that was explained also increased from 72.9 to 75.8 and 78.5%, respectively (Table 2). Rainfall and temperature from September to December are possibly important for the germination and early development of winter wheat crop plants. Cold hardiness tends to be completed in December (Yoshida et al., 1998), suggesting that temperature received during that period may be important for winter survival of wheat. Winter wheat needs vernalization that can only occur at low temperatures during winter to allow the reproductive phase particularly flowering to progress smoothly later in the season (Dixon et al., 2019; Streck et al., 2003). Yet, global temperature has been rising (Stewart et al., 2018) and apparently this needs to be captured in the model for an accurate yield prediction. Adequate soil moisture is also required for proper establishment and development of winter wheat (He et al., 2016). This likely initiates good seed germination and emergence to achieve a uniform plant stand. In demonstrating the importance of soil moisture, Stone and Schlegel (2006) found soil water content at emergence to have a relationship with grain yield that was as high as 70%. Under rainfed conditions, rainfall is likely to be the only source for replenishing soil moisture, potentially making it important for predicting grain yield. Thus, grain yield prediction accuracy was improved by including temperature and rainfall in the model. As such, weather variables are factors that should be evaluated in building models for yield prediction in different regions (Colaço & Bramley, 2019).

Raun et al. (2019) pointed out that because of randomness in the environment, our ability to accurately predict grain yield and make a recommendation for N will require using more environment-specific variables in the model to account for entropy or disorderliness in agricultural systems. In this study, including pre-plant N rate, rainfall, and temperature in the model appear to reduce this disorderliness in the system and improve the ability to predict the final grain yield.

The best model (lowest RMSE) was further validated with data from a neighboring experiment station in Hennessey, OK. The proportion of variation in measured wheat grain yield that was explained by the predictors that constituted the full model was found to be 69.8% (Figure 3). When the same data was used with the best one-variable model using INSEY as the predictor, only 47.6% of the variability in yield was explained (Figure 4). This is an illustration of the limitation of using INSEY as a single independent variable for predicting winter wheat grain yield. The full model explained

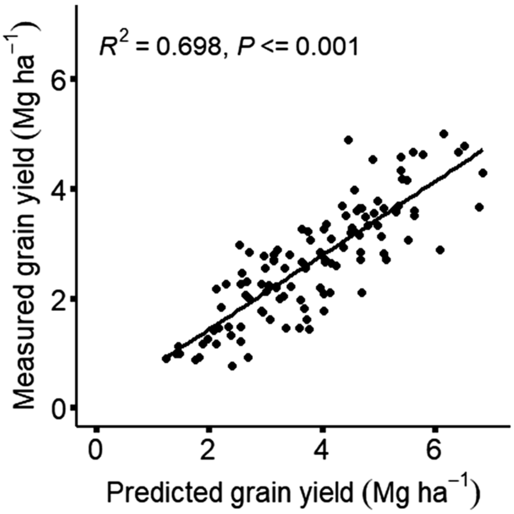

FIGURE 3 The relationship between predicted and observed grain yield generated using an independent data set from Hennessey experiment station, OK. This was a result of prediction using a full model (Yield = −4.8 + 202.0 × INSEY + 0.014 × N rate + 0.0041 × Rain + 0.31 × Temp). INSEY, in-season estimated yield; N rate, amount of nitrogen applied pre-plant; Rain, total rainfall from September to December; Temp, average air temperature from September to December

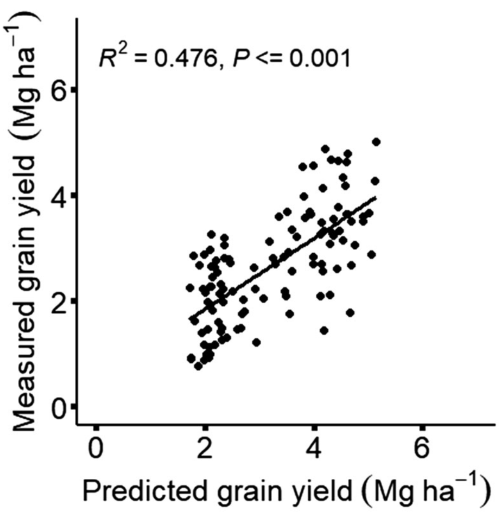

FIGURE 4 The relationship between predicted and observed grain yield generated using an independent data set from Hennessey experiment station, OK. This was a result of prediction using the best one-variable model (Yield = −0.46 + 309.7 × INSEY). INSEY, in-season estimated yield

much of the variation in grain yield at Hennessey, OK, and exceeded that explained in work done by Bushong et al. (2016) by approximately 20.8% when they included only INSEY and soil moisture in the model. The full model also produced results that were much higher than the yield goal approach that was found to explain no more than 16% of the variation in wheat grain yield (Raun et al., 2017). The improved yield prediction accuracy over the model using INSEY alone increases the likelihood that an accurate quantity of N may be estimated to match crop N needs. Nonetheless, these results are a further exemplification that including INSEY in a model may be superior to the yield goal approach for yield prediction even if it is the sole predictor. While LASSO generated a well validated model by splitting the data set into training and test sets, applying the full model to this additional data from Hennessey, OK, to predict the final grain yield provided further evidence that the model might serve the intended purpose in the real world.

Once the best model has been selected and used to predict grain yield (YP_0) mid-season, Raun et al. (2005) detailed the steps to follow in order to make a recommendation for N. In brief, response index (RI) is first computed to determine if crops will respond to fertilizer N using Equation 3:

$$RI = \frac{NDVI_{NRS}}{NDVI_{FP}} \tag{3}$$

where $NDVI_{NRS}$, normalized difference vegetation index value from a non-limiting nitrogen plot (nitrogen-rich strip); $NDVI_{FP}$, normalized difference vegetation index value from a plot with less or no nitrogen applied (farmer practice).

This is followed by estimating the grain yield that would be obtained if nitrogen is applied (YP_N) using Equation 4 below:

$$YP_N = YP_0 \times RI \tag{4}$$

Finally, nitrogen rate (NR) is computed as indicated in Equation 5:

$$NR = GN\left(\frac{YP_N - YP_0}{\eta}\right) \tag{5}$$

where GN, wheat grain nitrogen concentration (%), a value that can be obtained by analyzing the grain in the laboratory. The GN for wheat was also assumed to be 2.13% in a worldwide computation of NUE for cereal crops (Raun & Johnson, 1999); η, efficiency factor which is assumed to be as high as 70% when N is applied mid-season.

4 | CONCLUSION

The accuracy with which winter wheat grain yield was predicted improved by including pre-plant N rate, total rainfall, and average temperature from September to December in a model that already had INSEY as a predictor. In-season estimated yield resulted in the best one-variable model but had the highest RMSE of 0.72 Mg ha^{-1}. The full model selected by the model selection algorithm reduced this RMSE value to 0.54 Mg ha^{-1}. Nearly, 80% of the variation in wheat grain yield was explained by using a full model when compared to 62% achieved with INSEY as the only independent variable for yield prediction.

This is an indication that relying on INSEY as a sole predictor for grain YP may be bettered by a model that includes both INSEY and weather variables. This also means that the mid-season N management decision is likely to be based on a more accurate estimate of winter wheat grain yield. Since weather variables can easily be obtained from a vast network of local and national weather stations, including them in the model for yield prediction could assist producers in making more informed decisions as to when and how much N to apply.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Lawrence Aula https://orcid.org/0000-0001-8792-9063
Peter Omara https://orcid.org/0000-0002-3167-2286
Fikayo B. Oyebiyi https://orcid.org/0000-0002-1930-9345
Elizabeth Eickhoff https://orcid.org/0000-0002-9723-2646
William R. Raun https://orcid.org/0000-0002-1206-1105

REFERENCES

Aphalo, J. P. (2020). *ggpmisc: Miscellaneous extensions to ‘ggplot2’*. Retrieved from https://CRAN.R-project.org/package=ggpmisc

Aula, L., Omara, P., Nambi, E., Oyebiyi, F. B., & Raun, W. R. (2020). Review of active optical sensors for improving winter wheat nitrogen use efficiency. *Agronomy*, *10*(8), 1157. https://doi.org/10.3390/agronomy10081157

Batts, G., Morison, J., Ellis, R., Hadley, P., & Wheeler, T. (1997). Effects of CO_2 and temperature on growth and yield of crops of winter wheat over four seasons. *European Journal of Agronomy*, *7*(1-3), 43–52. https://doi.org/10.1016/S1161-0301(97)00022-1

Berntsen, J., Thomsen, A., Schelde, K., Hansen, O., Knudsen, L., Broge, N., Hougaard, H., & Hørfarter, R. (2006). Algorithms for sensor-based redistribution of nitrogen fertilizer in winter wheat. *Precision Agriculture*, *7*(2), 65–83. https://doi.org/10.1007/s11119-006-9000-2

Brock, F. V., Crawford, K. C., Elliott, R. L., Cuperus, G. W., Stadler, S. J., Johnson, H. L., & Eilts, M. D. (1995). The Oklahoma Mesonet: A technical overview. *Journal of Atmospheric Oceanic Technology*, *12*(1), 5–19. https://doi.org/10.1175/1520-0426(1995)012⟨0005:TOMATO⟩2.0.CO;2

Bundy, L. G., & Andraski, T. W. (2004). Diagnostic tests for site-specific nitrogen recommendations for winter wheat. *Agronomy Journal*, *96*(3), 608–614. https://doi.org/10.2134/agronj2004.0608

Bushong, J. T., Mullock, J. L., Miller, E. C., Raun, W. R., Klatt, A. R., & Arnall, D. B. (2016). Development of an in-season estimate of yield potential utilizing optical crop sensors and soil moisture data for winter wheat. *Precision Agriculture*, *17*(4), 451–469. https://doi.org/10.1007/s11119-016-9430-4

Cabrera-Bosquet, L., Molero, G., Stellacci, A., Bort, J., Nogues, S., & Araus, J. (2011). NDVI as a potential tool for predicting biomass, plant nitrogen content and growth in wheat genotypes subjected to different water and nitrogen conditions. *Cereal Research Communications*, *39*(1), 147–159. https://doi.org/10.1556/crc.39.2011.1.15

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

Colaço, A., & Bramley, R. (2019). Site-year characteristics have a critical impact on crop sensor calibrations for nitrogen recommendations. *Agronomy Journal*, *111*(4), 2047–2059. https://doi.org/10.2134/agronj2018.11.0726

Cui, Z., Zhang, F., Chen, X., Miao, Y., Li, J., Shi, L., Xu, J., Ye, Y., Liu, C., Yang, Z., Zhang, Q., Huang, S., & Bao, D. (2008). On-farm evaluation of an in-season nitrogen management strategy based on soil Nmin test. *Field Crops Research*, *105*(1-2), 48–55. https://doi.org/10.1016/j.fcr.2007.07.008

Dahnke, W., Swenson, L., Goos, R. J., & Leholm, A. (1988). *Choosing a crop yield goal*. Fargo, ND: North Dakota State University Extension Service.

Dhillon, J., Aula, L., Eickhoff, E., & Raun, W. (2020). Predicting in-season maize (*Zea mays* L.) yield potential using crop sensors and climatological data. *Scientific RepoRtS*, *10*(1), 1–8. https://doi.org/10.1038/s41598-020-68415-2

Dhillon, J., Eickhoff, E., Aula, L., Omara, P., Weymeyer, G., Nambi, E., Oyebiyi, F., Carpenter, T., & Raun, W. (2020). Nitrogen management impact on winter wheat grain yield and estimated plant nitrogen loss. *Agronomy Journal*, *112*(1), 564–577. https://doi.org/10.1002/agj2.20107

Dhillon, J. S., Eickhoff, E., Mullen, R., & Raun, W. R. (2019). World potassium use efficiency in cereal crops. *Agronomy Journal*, *111*(2), 889–896. https://doi.org/10.2134/agronj2018.07.0462

Dhillon, J. S., Figueiredo, B. M., Eickhoff, E. M., & Raun, W. R. (2020). Applied use of growing degree days to refine optimum times for nitrogen stress sensing in winter wheat. *Agronomy Journal*, *112*(1), 537–549. https://doi.org/10.1002/agj2.20007

Dixon, L. E., Hepworth, J., & Irwin, J. A. (2019). Vernalization. In *eLS* (pp. 1–11). Chichester: John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470015902.a0002048.pub4

Efretuei, A., Gooding, M., White, E., Spink, J., & Hackett, R. (2016). Effect of nitrogen fertilizer application timing on nitrogen use efficiency and grain yield of winter wheat in Ireland. *Irish Journal of Agricultural and Food Research*, *55*(1), 63–73. https://doi.org/10.1515/ijafr-2016-0006

Franzen, D., Kitchen, N., Holland, K., Schepers, J., & Raun, W. (2016). Algorithms for in-season nutrient management in cereals. *Agronomy Journal*, *108*(5), 1775–1781. https://doi.org/10.2134/agronj2016.01.0041

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*(1), 1–22.

Gibson, L., & Paulsen, G. (1999). Yield components of wheat grown under high temperature stress during reproductive growth. *Crop Science*, *39*(6), 1841–1846.

He, G., Wang, Z., Li, F., Dai, J., Li, Q., Xue, C., Cao, H., Wang, S., & Malhi, S. S. (2016). Soil water storage and winter wheat productivity affected by soil surface management and precipitation in dryland of the Loess Plateau, China. *Agricultural Water Management*, *171*, 1–9. https://doi.org/10.1016/j.agwat.2016.03.005

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning with applications in R*. New York: Springer. https://doi.org/10.1007/978-1-4614-7138-7

Ju, X., Liu, X., Zhang, F., & Roelcke, M. (2004). Nitrogen fertilization, soil nitrate accumulation, and policy recommendations in several agricultural regions of China. *AMBIO: A Journal of the Human Environment*, *33*(6), 300–305. https://doi.org/10.1579/0044-7447-33.6.300

Kristensen, K., Schelde, K., & Olesen, J. E. (2011). Winter wheat yield response to climate variability in Denmark. *The Journal of Agricultural Science*, *149*(1), 33. https://doi.org/10.1017/S0021859610000675

Li, F., Miao, Y., Zhang, F., Cui, Z., Li, R., Chen, X., Zhang, H., Schroder, J., Raun, W., & Jia, L. (2009). In-season optical sensing improves nitrogen-use efficiency for winter wheat. *Soil Science Society of America Journal*, *73*(5), 1566–1574. https://doi.org/10.2136/sssaj2008.0150

Liu, X., Ferguson, R. B., Zheng, H., Cao, Q., Tian, Y., Cao, W., & Zhu, Y. (2017). Using an active-optical sensor to develop an optimal NDVI dynamic model for high-yield rice production (Yangtze, China). *Sensors*, *17*(4), 672. https://doi.org/10.3390/s17040672

Makowski, D., Wallach, D., & Meynard, J. M. (1999). Models of yield, grain protein, and residual mineral nitrogen responses to applied nitrogen for winter wheat. *Agronomy Journal*, *91*(3), 377–385. https://doi.org/10.2134/agronj1999.00021962009100030005x

McPherson, R. A., Fiebrich, C. A., Crawford, K. C., Kilby, J. R., Grimsley, D. L., Martinez, J. E., Basara, J. B., Illston, B. G., Morris, D. A., & Kloesel, K. A. (2007). Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. *Journal of Atmospheric Oceanic Technology*, *24*(3), 301–321. https://doi.org/10.1175/JTECH1976.1

Morris, K. B., Martin, K., Freeman, K., Teal, R., Girma, K., Arnall, D., Hodgen, P., Mosali, J., Raun, W., & Solie, J. (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

Oklahoma Mesonet. (2020). *Daily data retrieval*. Retrieved from http://www.mesonet.org/index.php/weather/daily_data_retrieval 10.15763/dbs.mesonet (accessed 5 June 2020).

Omara, P., Aula, L., Oyebiyi, F., & Raun, W. R. (2019). World cereal nitrogen use efficiency trends: Review and current knowledge. *Agrosystems*. *Geosciences & Environment*, *2*(1), 1–8. https://doi.org/10.2134/age2018.10.0045

Ortuzar-Iragorri, M., Castellón, A., Alonso, A., Besga, G., Estavillo, J., & Aizpurua, A. (2010). Estimation of optimum nitrogen fertilizer rates in winter wheat in humid Mediterranean conditions, I: Selection of yield and protein response models. *Communications in Soil Science and Plant Analysis*, *41*(19), 2293–2300. https://doi.org/10.1080/00103624.2010.508094

R Core Team. (2020). *R: A language and environment for statistical computing* (Version R-4.0.3). Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Ransom, C. J., Kitchen, N. R., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernández, F. G., Franzen, D. W., Laboski, C. A., Nafziger, E. D., & Sawyer, J. E. J. A. J. (2020). Corn nitrogen rate recommendation tools' performance across eight US Midwest Corn Belt states. *Agronomy Journal*, *112*(1), 470–492.

Raun, W. R., Dhillon, J., Aula, L., Eickhoff, E., Weymeyer, G., Figueirdeo, B., Lynch, T., Omara, P., Nambi, E., & Oyebiyi, F. (2019). Unpredictable nature of environment on nitrogen supply and demand. *Agronomy Journal*, *111*(6), 2786–2791. https://doi.org/10.2134/agronj2019.04.0291

Raun, W., Figueiredo, B., Dhillon, J., Fornah, A., Bushong, J., Zhang, H., & Taylor, R. (2017). Can yield goals be predicted? *Agronomy Journal*, *109*(5), 2389–2395. https://doi.org/10.2134/agronj2017.05.0279

Raun, W. R., & Johnson, G. V. (1999). Improving nitrogen use efficiency for cereal production. *Agronomy Journal*, *91*(3), 357–363.

Raun, W. R., Solie, J. B., Johnson, G. V., Stone, M. L., 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. https://doi.org/10.2134/agronj2001.931131x

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., Solie, J., Stone, M., Martin, K., Freeman, K., Mullen, R., Zhang, H., Schepers, J., & Johnson, G. (2005). Optical sensor-based algorithm for crop nitrogen fertilization. *Communications in Soil Science and Plant Analysis*, *36*(19-20), 2759–2781. https://doi.org/10.1080/00103620500303988

Shanahan, J., Kitchen, N., Raun, W., & Schepers, J. S. (2008). Responsive in-season nitrogen management for cereals. *Computers and Electronics in Agriculture*, *61*(1), 51–62. https://doi.org/10.1016/j.compag.2007.06.006

Stewart, B., Thapa, S., Xue, Q., & Shrestha, R. (2018). Climate change effect on winter wheat (*Triticum aestivum* L.) yields in the US Great Plains. *Journal of Soil Water Conservation*, *73*(6), 601–609. https://doi.org/10.2489/jswc.73.6.601

Stone, L. R., & Schlegel, A. J. (2006). Yield–Water supply relationships of grain sorghum and winter wheat. *Agronomy Journal*, *98*(5), 1359–1366. https://doi.org/10.2134/agronj2006.0042

Streck, N. A., Weiss, A., & Baenziger, P. S. (2003). A generalized vernalization response function for winter wheat. *Agronomy Journal*, *95*(1), 155–159.

Thomason, W., Phillips, S., Davis, P., Warren, J., Alley, M., & Reiter, M. (2011). Variable nitrogen rate determination from plant spectral reflectance in soft red winter wheat. *Precision Agriculture*, *12*(5), 666–681. https://doi.org/10.1007/s11119-010-9210-5

Tremblay, N., Wang, Z., Ma, B.-L., Belec, C., & Vigneault, P. (2009). A comparison of crop data measured by two commercial sensors for variable-rate nitrogen application. *Precision Agriculture*, *10*(2), 145. https://doi.org/10.1007/s11119-008-9080-2

Walsh, O. S., Klatt, A., Solie, J., Godsey, C., & Raun, W. (2013). Use of soil moisture data for refined GreenSeeker sensor based nitrogen recommendations in winter wheat (*Triticum aestivum* L.). *Precision Agriculture*, *14*(3), 343–356. https://doi.org/10.1007/s11119-012-9299-9

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Grolemund, G., Alex, H., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019).

Welcome to the Tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2019). *readxl: Read Excel Files (Version 1.3.1)*. Retrieved from https://CRAN.R-project.org/package=readxl

Yoshida, M., Abe, J., Moriyama, M., & Kuwabara, T. (1998). Carbohydrate levels among winter wheat cultivars varying in freezing tolerance and snow mold resistance during autumn and winter. *Physiologia Plantarum*, *103*(1), 8–16. https://doi.org/10.1034/j.1399-3054.1998.1030102.x

Zhang, H., & Raun, B. (2006). *Oklahoma soil fertility handbook*. Department of Plant and Soil Sciences, Oklahoma Agricultural Experiment Station, Oklahoma Cooperative Extension Service, Division of Agricultural Sciences and Natural Resources, Oklahoma State University.

Zhang, N., Zhao, C., Quiring, S. M., & Li, J. (2017). Winter wheat yield prediction using normalized difference vegetative index and agro-climatic parameters in Oklahoma. *Agronomy Journal*, *109*(6), 2700–2713. https://doi.org/10.2134/agronj2017.03.0133

How to cite this article: Aula L, Omara P, Nambi E, et al. Active optical sensor measurements and weather variables for predicting winter wheat yield. *Agronomy Journal*. 2021;1–10.
https://doi.org/10.1002/agj2.20620