.

work is properly cited.

Received: 13 March 2020 DOI: 10.1002/csc2.20293

ORIGINAL RESEARCH ARTICLE

Crop Breeding & Genetics

Application of the factor analytic model to assess wheat falling number performance and stability in multienvironment trials

Stephanie M. Sjoberg¹ | Arron H. Carter¹ | Camille M. Steber^{1,2,*} Kimberly A. Garland Campbell^{1,2,*}

¹ Dep. of Crop and Soil Sciences, Washington State Univ., PO Box 646420, Pullman, WA 99164, USA

² Wheat Health, Genetics and Quality Research Unit, USDA-ARS, PO Box 646420, Pullman, WA 99164, USA

Correspondence

Kimberly A. Garland Campbell, Wheat Health, Genetics and Quality Research Unit, USDA-ARS, and Dep. of Crop and Soil Sciences, Washington State Univ., PO Box 646420, Pullman, WA 99164, USA. Email: kim.garland-campbell@usda.gov

*The last two authors contributed equally to this work.

Assigned to Associate Editor Jochum Wiersma.

Funding information

Washington Grain Commission, Grant/Award Numbers: 7336, 7599; National Institute of Food and Agriculture, Grant/Award Number: 2015-05798; United States Department of Agriculture Agriculture Research Service Project 2090-21000-033-00D

Abstract

A factor analytic model was used to characterize data generated with the Hagberg-Perten falling number (FN) method, a measure of wheat (Triticum aestivum L.) quality influenced by genotype-by-environment interactions. The FN method detects starch degradation due to the presence of the enzyme α -amylase in wheat grain such that a low FN indicates high α -amylase activity and high risk of poor end-product quality. Because farmers receive severe discounts for low FN, FN data have been collected over multiple years for the Washington State University multilocation variety trials to help farmers and breeders identify lower risk varieties. Analysis of these data to objectively rank varieties is challenging because the dataset is unbalanced and because FN is subject to complex genotype-by-environment interactions. Low FN can result from environmental differences at multiple stages in grain development because there are two major causes of α -amylase accumulation in grain, latematurity α-amylase (LMA) and preharvest sprouting (PHS). A five-factor analytic model extracted explicit measures of overall performance and of stability in variable environments from historical FN data from the multilocation trial, providing a basis for breeding and planting decisions. Whereas a linear model explained 70.3% of the variation, the five-factor analytic model accounted for 92.5% of variation in the data. Examination of factor loadings enabled us to separate environments and genotype response to either PHS or LMA, specifically. This is the first application of a factor analytic model to evaluate the end-use quality trait FN, providing a method to rank varieties for grower decisions and breeder selections.

INTRODUCTION 1

Wheat (Triticum aestivum L.) growers in the northwestern United States produce some of the highest quality soft white wheat in the world, the majority of which is exported to Asian markets for use in specialty products such as sponge cakes and noodles (FAO, 2017). One important determinant of quality

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original

Abbreviations: AIC, Akaike information criterion: AMMI, additive main effects and multiplicative interaction; BIC, Bayesian information criterion; FA, factor analytic; FN, falling number; G×E, genotype-by-environment; LMA, late-maturity α-amylase; PHS, preharvest sprouting; WSU, Washington State University.

[©] The Authors. Crop Science published by Wiley Periodicals LLC on behalf of Crop Science Society of America. This article has been contributed to by US Government employees and their work is in the public domain in the USA

is the level of starch degradation present in the flour. Certain weather conditions induce the enzyme α -amylase in the grain prior to harvest, potentially leading to starch digestion and decreased end-use quality. The wheat industry measures the level of α -amylase and starch degradation using the Hagberg– Perten falling number (FN) method (Ross & Bettge, 2009). Briefly, the FN test measures the gelling capacity of a mixture of flour and water by counting the number of seconds it takes for a stirrer to drop through the heated mixture. Growers can receive as much as a US\$0.01 per 27 kg (1 bu) discount for every second below a FN of 300 s (USDA, 2017).

Falling number is a complex trait affected by multiple environmental and genetic factors. There are two main causes of high α -amylase and, therefore, low FN: preharvest sprouting (PHS) and late-maturity α -amylase (LMA) (Mares & Mrva, 2014). The first, PHS, is the germination of the mature grain on the mother plant in response to rainy conditions before harvest. Starch is degraded during germination to provide fuel for embryo and seedling growth. The degree of PHS tolerance is determined by genetics but is affected by environmental factors during seed development and germination. The second potential cause of low FN is LMA, the production of α -amylase in response to a cool temperature shock or continuous cool temperatures during late grain filling (Derkx & Mares, 2020; Mrva & Mares, 1996). Late-maturity α-amylase is also influenced by both genetic and environmental factors. Moreover, low FN can also be caused by either low grain protein content or variation in starch composition (Ross, Flowers, Zemetra, & Kongraksawech, 2012; Zeng, Morris, Batey, & Wrigley, 1997). The multiple genetic and environmental influences on the FN phenotype present a challenge to plant breeders and agronomists to identify tolerant wheat cultivars.

Starting in 2013, FN tests have been performed on grain harvested from the Washington State University (WSU) Cereal Variety Testing Program to provide farmers with information about the relative FN susceptibility among varieties (Sjoberg, Carter, Steber, & Campbell, 2020). Although the data can be viewed by variety over multiple locations or by location within a year on a publicly accessible database (http://steberlab.org/project7599.php), there have been inconsistencies in ranking of varieties among environments, suggesting the presence of genotype-by-environment $(G \times E)$ interactions. Another factor contributing to the complicated data analysis was that the number of FN replicates that were performed for different varieties and environments varied due to resource limitations. Moreover, the dataset was unbalanced because old varieties were slowly replaced with newly released varieties and because not all varieties were grown in all locations due to variability in adaptation.

The WSU Cereal Variety Testing Program conducts trials over a wide range of rainfall zones, ranging from <300 to >500 mm of annual precipitation. Vari-

Core Ideas

- The FA model better accommodates a dataset of unbalanced nature with large G×E effects.
- This study showed that the FA model can be applied to FN, a wheat end-use quality trait.
- The FA model quantified genotype performance and environmental stability.

eties are grown in low- or high-precipitation environments depending on their projected adaptation. Trials in various precipitation zones exhibited heterogeneous variances when compared across the entire dataset. Because of the unbalanced structure of the varieties over years, and the multiple causes of low FN in this dataset, variety means over environments do not accurately reflect differences among varieties. Although reporting data based on precipitation zone groups can account for heterogenous variances for a trait like grain yield, these groups do not accurately reflect variety performance for FN because the different environments have problems with low FN in different years.

Approaches for analyzing this multienvironment trial dataset using weather patterns in order to determine whether LMA or PHS were causing the low FN in each experiment were explored in Sjoberg et al. (2020) and provided a means to identify stable varieties with good FN. In these analyses, Sjoberg et al. (2020) showed that the variance components associated with genotype were of similar magnitude to those associated with G×E. While useful, methods to dissect G×E interaction like linear models, the Finlay–Wilkinson regression, additive main effects and multiplicative interaction (AMMI), and genotype plus G×E models are limited in their ability to accommodate the unbalanced nature of the dataset (Sjoberg et al., 2020).

The current research evaluated an alternative to the methods described above, the factor analytic (FA) model, as a more robust way to tease apart the G×E interactions affecting FN in this unbalanced dataset. Although the AMMI model explored in Sjoberg et al. (2020) is one type of FA model, the model introduced by Piepho (1998) and Smith, Cullis, and Thompson (2001) considers genotype and G×E effects to be random and accommodates the heterogeneity of genetic variance across environments and between pairs of environments. The FA approach models the G×E interaction effects to account for covariances among them in terms of a small number of common factors.

Previous studies have used the FA model to handle the complex $G \times E$ associated with unbalanced multienvironment trial datasets in a variety of crops including, but not limited to, wheat, barley (*Hordeum vulgare* L.), sorghum [*Sorghum*]

bicolor (L.) Moench], lentil (*Lens culinaris* Medik.), and maize (*Zea mays* L.) (Burgueño, Crossa, Cornelius, & Yang, 2008; Dawson et al., 2013; Kelly, Smith, Eccleston, & Cullis, 2007; Mengesha et al., 2019). Smith and Cullis (2018) developed a simple method of summarizing the results of the FA model to compare varieties, and it has been successfully adopted by the Australian government organization responsible for conducting small grain and legume field trials.

Studies that have explored the genetics of variation in large datasets of wheat quality data have not been able to establish a good measure of stability or adaptability. Laidig et al. (2017) used a linear model approach to estimate the genetic and nongenetic sources of variation present in 32 yr of wheat quality data, including FN, finding that in their case FN had a genetic gain of only 5.8% over 32 yr. Hernández-Espinosa et al. (2018) evaluated end-use quality traits for stability in the International Maize and Wheat Improvement Center (CIM-MYT) wheat breeding program using a fixed-effect linear model and rank shift across environments to classify varieties as stable or unstable. Yabwalo et al. (2018) used the genotype plus G×E biplot method to measure stability and mean performance of grain morphology traits in wheat. Our study represents a unique application of the FA model to the grain quality trait FN.

Our goal was to evaluate the FA model as an approach to rank varieties for response to FN within an unbalanced dataset from multienvironment trials. The resulting model allowed varieties to be ranked both according to overall performance and stability in challenging environments based on RMSD (Smith & Cullis, 2018). The objective of this analysis was to provide objective methods to rank varieties based on FN to better inform farmers making planting decisions and breeders making crossing decisions.

2 | MATERIALS AND METHODS

2.1 | Field materials and the Hagberg–Perten falling number method

Statistical analyses were performed on a dataset consisting of 3,570 FN datapoints representing 129 soft white winter wheat genotypes, including released and unreleased cultivars and breeding lines from the University of Idaho, WSU, Oregon State University, Limagrain Cereal Seeds, WestBred Wheat, USDA-ARS, and AgriPro Wheat (reported in Sjoberg et al., 2020). The WSU Cereal Variety Testing Program provided samples from WSU Extension Cereal Variety Testing Program Soft Winter Wheat Trials conducted in 2013, 2014, and 2016 at 35 unique location and year combinations (defined as environments) (Guy, Jitkov, Lauver, & Horton, 2013; Guy, Jitkov, Lauver, Horton, & Higginbotham, 2014; Higginbotham, Jitkov, & Horton, 2016). These specific years and locations were selected because low FN occurred across multiple testing environments and entries. Cultural practices and the FN tests were performed as described in Sjoberg et al. (2020). An α -lattice incomplete block design with three replications was used in each field experiment, and only trials where FN tests were conducted on more than one field replication were used in this study. Because samples from all three field replicates were not always available, samples were generally taken from the first two replications. Therefore, a biological replicate was defined as samples from different field replicates, analyzed separately. The α -lattice blocking was maintained for the sampled field replications when the FN were assayed.

2.2 | Mixed linear model

The full dataset was analyzed initially using a standard linear model of the *l*th block within the *k*th replicate of the *i*th variety observed in the *j*th environment, written as

$$Y_{ijkl} = \mu + V_i + E_j + V E_{ij} + R(E)_{k(i)} + b(RE)_{l(ki)} + e_{ijkl}$$
(1)

where Y = the plot FN, $\mu =$ the overall mean (intercept), V_i = the genotype effect, E_j = the environment (location × year) effect, VE_{ij} = the G×E interaction, $R(E)_{k(i)}$ = the replication effect within each environment, b = the block effect within each replication, and e = the residual variance. Variance components were estimated using residual maximum likelihood (REML). All effects were considered random.

Packages used for these analyses included ASReml-R (V4) and ggplot2 in R version 3.5.1 (Butler, 2020; Wickham, 2016).

2.3 | Factor analytic model

Factor analytic models are referred to as FAk where k = the number of factors defined by the user. The G×E effect for variety *i* and environment *j* is written as

$$VE_{ij} = \lambda_{1_i} f_{1_i} + \lambda_{2_i} f_{2_i} + \dots + \lambda_{k_i} f_{k_i} + \delta_{ij}$$
(2)

where λ_{rj} were the effects (also known as loadings) for the latent environmental factors and f_{ri} were effects (also known as scores) for the variety responses to those factors. The error is represented by δ_{ij} . Loadings and scores were derived from the between-environment additive genetic variance matrix as described in Smith et al. (2001). Overall performance of varieties was measured using the following equation:

$$\bar{\lambda}_1 \, \tilde{f}_{1i}^* = \, \frac{1}{p} \sum_{j=1}^p \hat{\lambda}_{1j}^* \tilde{f}_{1i}^* \tag{3}$$

where $\hat{\lambda}_1$ was the mean of the rotated estimated loadings (as indicated with an asterisk) for individual environments for the first factor across *p* environments (Smith & Cullis, 2018).

Empirical best linear unbiased predictions (EBLUPs) of the random variety-by-environment effects can be obtained directly from estimations of Equation 3, as

$$\widetilde{VE}_{ij} = \tilde{\beta}_{ij} + \tilde{\delta}_{ij} \tag{4}$$

where $\tilde{\beta}_{ij}$ is the predicted regression component (also called the common variety by environment effect) and $\tilde{\delta}_{ij}$ is the error (Smith, Ganesalingam, Kuchel, & Cullis, 2015). For FA models, higher overall performance values are better and lower RMSD values reflect a higher level of stability.

The global stability of a variety, as defined by Smith and Cullis (2018), was measured by the RMSD, calculated as

$$RMSD = \sqrt{\frac{1}{p} \sum_{j=1}^{p} \tilde{\epsilon}_{ij}^{*2}}$$
(5)

where $\tilde{\varepsilon}_{ij}^{*2} = \tilde{\beta}_{ij} + \hat{\lambda}_{1j}^* \tilde{f}_{1i}^*$. The $\tilde{\varepsilon}_{ij}^{*2}$ were therefore deviations from the first factor predictions in a plot where the *x* axes were the first factor loadings and *y* axes were the common effects.

Responsiveness of varieties to each factor was calculated as in Smith and Cullis, (2018):

$$T_r = \left(\bar{\lambda}_{r+} - \bar{\lambda}_{r-}\right)\tilde{f}_{ri}^* \tag{6}$$

where T_r was a variety's response for factor r, and $\bar{\lambda}_{r+}$ and $\bar{\lambda}_{r-}$ were, respectively, the mean of the positive and negative estimated loadings for factor r.

A linear model with genotype, environment, G×E, replication, and block (Equation 1) was used as a baseline to compare with the FA model using from one to nine factors. All effects were considered random except the environment and replicates within the environment, which were fitted as fixed effects for both models. For the FA model, a separate variance for blocks within replicates within each environment was fitted as random; there was also a separate error variance for each environment. The maximum number of factors was limited to nine due to the large processing time required for additional factors. The Akaike information criterion (AIC), Bayesian information criterion (BIC), and residual log-likelihood tests were used to choose the most parsimonious model over the higher order models.

3 | RESULTS

3.1 | Phenotypic data analysis

This study characterized the G×E interactions affecting the FN of soft white winter wheat in the WSU Cereal Variety testing Program in 2013, 2014, and 2016. The FN dataset was highly unbalanced because not every variety was grown in every experiment. The overlap between the varieties grown is shown using a heat map, ranging from the lowest in green to the highest in white (Figure 1). The minimum number of overlapping varieties between experiments was nine, the maximum was 60, and the mean was 28.4. The groupings shown in Figure 1 reflect the nature of the WSU Cereal Variety Testing Program in that locations that fall in the same rainfall zone and year will have a similar set of adapted varieties planted, in contrast with those planted in another rainfall zone. This dataset had widespread incidence of low FN. Of the 35 experiments, 32 experienced at least a minor low FN event, defined by the incidence of at least five varieties with a FN < 300 (Table 1). A moderate, 13-fold increase was observed in error variance among experiments, demonstrating that the amount of unexplained error in experiments is inconsistent and heterogeneous (Table 2). Moreover, differences in CV were due to a combination of genetic and G×E variability, further strengthening the argument to apply the FA model as a means of understanding the G×E variability.

3.2 | Factor analytic model estimates

To find the best fit for our dataset, 10 different models were compared. The AIC, BIC, and residual log-likelihood tests supported the use of the FA5 model as the best fit over the higher order models to compare with the linear model with genotype, environment, $G \times E$, replication, and block as random effects. Although the higher order models explained slightly more variation, they may have been overparameterized, so the five-factor model was used in the rest of the study (Table 3).

The loadings by factor in the FA5 model were decomposed for further examination using basic summary statistics (Table 4). Environmental loadings for the first factor of the model were all positive, ranging from 11.5 to 58.9 with a mean across varieties of 29.1 and a median of 25.1. The remaining four factors had a range extending into negative values. The percentage variation accounted for by the FA5 model within each environment was calculated by factor and overall. The range, mean, median, and the number of environments that had over and under 80% of their variation explained by all factors in the model were reported (Table 5). The mean individual trial percentage variance accounted for by the first factor in the FA5 model was 55.8%, but the mean percentage

TABLE 1 Decomposition of 3,570 falling number (FN) datapoints representing 129 soft white winter wheat genotypes, based on incidence of low FN

Year	No. locations	No. varieties	No. locations with FN <300
2013	9	65	9
2014	8	73	5
2016	18	79	18

TABLE 2 Summary of Washington State University cereal variety testing program winter wheat trial falling number (FN) data

Statistic	Min.	Max.	Mean	Median
Mean FN, s	64.5	489	334.3	340
Error variance, s ²	0.35	4.64	1.42	1.03
CV, %	7.76	41.34	15.25	12.92

TABLE 3 Summary of fitted models, including the linear model method and factor analytic (FA) models with one factor to nine factors

Model	No. of parameters	Residual log- likelihood	AIC	BIC	vaf
					%
Linear model	4	-14,826.0	29,659.9	29,684.6	70.3
FA1	70	-14,110.3	28,500.6	29,364.5	59.1
FA2	104	-13,949.6	28,247.3	29,320.9	82.7
FA3	137	-13,895.3	28,204.6	29,481.9	87.1
FA4	169	-13,863.9	28,205.8	29,680.5	89.5
FA5	200	-13,839.3	28,218.7	29,884.7	92.3
FA6	230	-13,814.0	28,228.0	30,079.1	94.9
FA7	259	-13,788.5	28,235.1	30,265.1	97.5
FA8	287	-13,770.6	28,255.2	30,458.1	98.5
FA9	314	-13,762.2	28,292.4	30,661.9	98.0

Note. The percentage of variance accounted for is indicated as vaf. AIC and BIC refer to Akaike and Bayesian information criteria, respectively.

TABLE 4 Summary of environmental loadings for the FA5 model

Statistic	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Min.	11.5	-27.0	-23.2	-16.4	-15.0
Max.	58.9	46.6	21.1	15.6	13.6
Mean	29.1	-5.1	-1.3	1.3	-0.2
Median	25.1	-12.5	-3.0	1.0	0.1

TABLE 5 Summary of percentage of variation accounted for by each factor in the FA5 model

Statistic	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	All factors
Min.	21.2	1.4	0.1	0.00	0.00	66.16
Max.	91.5	51.1	52.4	24.2	18.4	100
Mean	55.8	22.1	7.4	4.1	3.4	92.9
Median	56.6	24.0	3.1	1.9	1.8	100
No. of experiments >80%						31
No. of experiments <80%						4



FIGURE 1 Connectivity of experiments based on genotypes planted. Green indicates the least connected experiments, and white indicates the most connected experiments

of variance accounted for by the five factors combined was 92.9%, and the number of experiments with <80% of variance accounted for was just four.

Varieties were ranked according to their overall performance and according to their stability. Using FA5, the overall performance of a variety was plotted against the RMSD, which represented a measure of global stability (Figure 2). A set of varieties with high FN and good stability or environmental resilience was identified. For this dataset specifically, those within the blue box had a high level of overall performance for FN and also had a high stability over the varying environments. For example, WB 456 had the highest overall performance (49.5) and a better than average level of stability (RMSD = 21.4). Variety Pritchett had a similar level of stability to WB 456 (20.7), but its overall performance was much lower (-30.0). In contrast, variety Coda had a high level of overall performance (39.7) but is one of the more unstable varieties (29.2). The four worst varieties based on this graph are Bruehl, SY Ovation, Xerpha, and ARS-Selbu. This analysis suggests that growers should be cautious when considering these varieties, since they may not meet export receival standards for FN.

3.3 | Variety comparisons using the factor analytic model

Latent regression plots compared five varieties: WB 456, Xerpha, WA 8226, Coda, and WB 1376CLP (Figure 3). WB 456 was chosen for this comparison because it had the highest overall performance value. Xerpha and WA 8226 were the most unstable based on the RMSD. Coda was of interest because it had the fourth highest overall performance value but had a high RMSD value indicative of lower stability. WB 1376CLP had the second highest overall performance value but had a low RMSD value indicative of higher stability. The y axis represents the common variety effect, and based on what we know of this model, it can be concluded that the larger the common variety effect, the more desirable that variety was in terms of FN. Note that the percentage variation explained by the FA5 model was unique by environment. The x axis was the (rotated) estimated first environment loadings. The slope of the line was the effect (also known as the score) for the variety response to the first factor. WB 456 was the superior variety because it had a high common variety effect across all estimated first loadings. In contrast, Coda did not perform well





FIGURE 2 Overall performance vs. stability measure (RMSD) for falling number for the Washington State University Cereal Variety Testing Program winter wheat trials. The box highlights the area with varieties that perform well and are relatively stable. Varieties are represented by blue circles, but only released cultivars are labeled by name



FIGURE 3 Latent regression plot for the first factor for five varieties. REML, residual maximum likelihood

consistently, as evidenced by the high residual error between the points and the line.

3.4 | Trends and variety responses to components of the factor analytic model

Interestingly, experiments that had a high response to the (positive) estimated loading for the second factor corresponded to those found to have low FN due to PHS in a previous study (Figure 4; Sjoberg et al., 2020). The original characterization of PHS events were based on the incidence of precipitation within the 2 wk prior to harvest. This trend suggests that the second factor differentiates PHS events from other low FN events, so from here on it will be referred to as the PHSfactor. By plotting the overall performance values against the variety's estimated loading for the PHS-factor, pertinent



FIGURE 4 Residual maximum likelihood (REML) estimates of the second factor vs. REML estimates of the third factor, by environment. The second factor is also referred to as the PHS-factor because environments with high second factor loadings experienced severe PHS events



FIGURE 5 The response, by variety, to the preharvest sprouting (PHS)-factor on the *x* axis, vs. the overall performance on the *y* axis. Each point corresponds to a variety, but only 45 of the 129 varieties are annotated

information about the varieties may be observed, especially if the factor is associated with a meaningful environmental factor (Figure 5). Varieties WA 8202 and SY Ovation had the greatest PHS-factor loading, meaning they performed well in environments with positive loadings to the second factor. Keeping in mind that environments with a positive loading for the PHS-factor suffered a PHS event, it can be concluded that WA 8202 and SY Ovation have some tolerance to sprouting. In contrast, varieties such as Bruneau, Xerpha, and Bruehl had some of the most negative loadings to the PHS-factor. These varieties have previously been identified as susceptible to sprouting (Martinez et al., 2018; Tuttle et al., 2015). The PHS-factor appears to identify varieties with higher loadings as tolerant to PHS but still not generally great performers. This could indicate that varieties with high PHS-factor loadings, such as WA 8202 and SY Ovation, are susceptible to another cause of low FN, such as LMA. As estimated loading for the PHS-factor decreased, the trend disappeared. Estimated loading of experiments to the third factor did not appear to be associated with any environmental trend.

4 | DISCUSSION

Developing varieties with reduced risk of experiencing low FN is an important breeding objective because farmers receive substantial discounts for low FN. To achieve this objective, breeders and farmers need a tool to assess the relative risk when growing different varieties. This is difficult to do because FN is controlled by complex G×E interactions. The FA model and the subsequent use of selection tools developed by Smith and Cullis (2018) address this problem by using the overall performance value as an indicator of genetic contribution and the RMSD as the estimator of stability under challenging environmental conditions. By parsing out these two elements, the FA model provides two objective measures to assess relative risk.

The best varieties for FN have a high overall performance value and a low RMSD value (Figure 2). Based on this criterion, the most desirable varieties for FN performance were WB 1376CLP and WB 456. Other good alternatives for FN resistance included Norwest Duet, Mary, and Skiles. However, it is essential to examine the fit of the model before using such data to make planting recommendations or use in breeding schemes. Note that there was roughly a linear relationship between FN overall performance and stability based on RMSD (Figure 2). With some exceptions, most of the varieties with higher overall FN performance had better stability, whereas poor-performing varieties tended to have less stability. Thus, breeders should be able to concentrate mainly on breeding for improved FN test performance.

The FA model improves upon other graphical tools to discern stability (Piepho, 1998; Smith et al., 2001). The Finlay–Wilkinson as well as Eberhart–Russell methods do not account for nonlinear G×E interactions (Sjoberg et al., 2020; Smith et al., 2015). The AMMI model maintains all effects as fixed, so it cannot account for variance heterogeneity. In addition to accounting for variance heterogeneity, the FA models can also handle large, unbalanced datasets. The FA model has been successfully used for a wide range of phenotypes. Kelly et al. (2007) applied the FA model to multienvironment 380

trial datasets of wheat, barley, sorghum, and lentil and found it to outperform unstructured models in several categories including robustness, prediction accuracy in the presence of G×E, and handling of heterogeneous genetic variances and crossovers in genotype ranking. In maize datasets, Burgueño et al. (2008) concluded that the FA model could be used for modeling association among environments or genotypes and that conclusions drawn from it were valid for a broader inference space. Dawson et al. (2013) made genomic predictions within groups of environments clustered using several methods, and of those reported, the FA model had superior accuracy across environments with differing levels of genetic correlation. Mengesha et al. (2019) used the FA model to identify stability and adaptability of maize varieties as it related to both grain yield and provitamin A content. This study showed that the FA model can be applied to FN, a wheat end-use quality trait. The variety trial FN data analyzed was highly unbalanced, subject to large variety-by-environment effects, and underpinned by two different biological processes, PHS and LMA. The FA model proved a highly powerful statistical tool for analyzing this data. Although the utility of the FA model is clear, it also is complex and requires a reasonably high level of statistical understanding to both fit the model and interpret the results. The FA model might become more broadly applied to breeding and variety characterization with development of a tutorial on the postprocessing of values extracted from it, to produce graphs such as Figure 2 and 3. These graphs would be useful to breeders who may not have in-depth knowledge of advanced statistical models.

Previous work showed that the FA model performs similarly to or better than other models used to investigate $G \times E$. Piepho (1998) proposed the use of a FA model with common specific variance, finding no significant advantage over the traditional variance component model. Nevertheless, Kelly et al. (2007) reported that the unstructured FA model was superior to variance component models because it handles the G×E and heterogeneous genetic variances present in the datasets. Meyer (2009) reviewed the efficacy of the FA models in comparison with the linear mixed model framework, finding them to be more capable of modeling G×E in all types of datasets, including those with missing observations or multiple random effects. An extension of this work enables plant breeders to use FA for selection based on multiple traits in multiple environments and for genomic selection models (Dias et al., 2018; Ward et al., 2019).

The goal of this study was to find a model that could better accommodate the unbalanced nature of the dataset and large G×E effects. The FA model provided an excellent tool to examine G×E interactions by quantifying genotype performance and environmental stability (Figure 2). Under a linear model structure, 70.3% of variation was accounted for by the model, whereas the FA5 model accounted for 92.5% of variation in the data (Table 3).

The FA5 model allowed us to observe the trends in experiments and glean information about biological responses. The environment loadings from the second factor allowed us to see some distinction between independent causes of low FN, specifically to separate severe PHS events from all others, and to identify varieties (WA 8202 and SY Ovation) that were tolerant to PHS but susceptible to LMA. This demonstrates the power of this model in detecting environmental effects on FN. This result also suggests that FA modeling of FN data from variety trials could be used to identify varieties as being likely affected by PHS and/or LMA. This could be useful in choosing varieties for specific research or breeding goals. Although the second factor was able to successfully model PHS, there was no clear connection between a factor and LMA. Latematurity α -amylase is more difficult to model because there is a short window of susceptibility to cold induction during soft dough stage of grain filling (Mares & Mrva, 2014). Thus, a variety may express LMA in one environment but not another. Also, the FN test measures grain deterioration due to either cause. Since PHS and LMA are induced by specific weather conditions that may occur separately or overlap during the same season, it will continue to be difficult to determine the minimum number of seasons and environments needed to define relative genetic susceptibility to low FN from each cause until we develop methods to discern the true cause of low FN in harvested grain samples. Clearly, one needs to look at multiple years in which one or the other phenomenon occurred. Because low FN is environmentally controlled, it is difficult to sample enough environments to obtain estimates of environmental stability. We have continued to add to this dataset and plan to use it to conduct power analyses with the goal of developing predictive models for varietal risk for low FNs.

Latent regression plots, as in Figure 3, are more informative than *t* tests because the FA model can give an estimate of variety performance in environments the variety was not grown in, and it also shows all the variety by environment effects for those chosen varieties. Ideally, latent regression plots should be used to further investigate varieties identified as interesting based on the graph of overall performance vs. RMSD (Figure 2).

Extracting valuable information from unbalanced, multienvironment trial data is a challenge for breeding programs and variety testing programs around the world. This study demonstrated the utility of FA models in examining FN, an end-use quality trait with high variability due to G×E interactions. The FA models also allowed us to separate the G×E interactions into similar groups, such as those affected by PHS, or LMA, even when we did not have additional information about the environments. The FA models thus have utility to improve the efficiency of selection for a given trait that can be influenced by multiple environmental effects and even by multiple plant physiological pathways.

ACKNOWLEDGMENTS

The authors thank those who conducted the trials of Washington State University's Cereal Variety Testing Program, including Vadim Jitkov, Stephen Guy, and Ryan Higginbotham. Thank you also to Tracy Harris, Rehana Parveen, and Xavian Thompson for generating FN data. The authors sincerely thank Alison Smith for providing excellent guidance for running the FA model in ASReml-R. We could not have completed this analysis without her generous assistance. This study was supported by the Washington Grain Commission Projects 7599 and 7336, by the USDA-ARS, by the USDA National Institute of Food and Agriculture Project 2015-05798, and by USDA-ARS In-House Project 2090-21000-033-00D.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Stephanie M. Sjoberg ¹⁰ https://orcid.org/0000-0003-0351-7518

Arron H. Carter https://orcid.org/0000-0002-8019-6554 *Camille M. Steber* https://orcid.org/0000-0001-6255-7670 *Kimberly A. Garland Campbell* https://orcid.org/0000-0003-4747-3270

REFERENCES

- Burgueño, J., Crossa, J., Cornelius, P. L., & Yang, R.-C. (2008). Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. *Crop Science*, 48, 1291–1305. https://doi.org/10.2135/cropsci2007.11.0632
- Butler, D. (2020). asreml: Fits the linear mixed model. R package version 4.1.0.126. VSN International. Retrieved from www.vsni.co.uk
- Dawson, J. C., Endelman, J. B., Heslot, N., Crossa, J., Poland, J., Dreisigacker, S., ... Jannink, J.-L. (2013). The use of unbalanced historical data for genomic selection in an international wheat breeding program. *Field Crops Research*, 154, 12–22. https://doi.org/10.1016/j. fcr.2013.07.020
- Derkx, A. P., & Mares, D. J. (2020). Late-maturity α-amylase expression in wheat is influenced by genotype, temperature and stage of grain development. *Planta*, 251. https://doi.org/10.1007/s00425-020-03341-1
- Dias, K., Gezan, S. A., Guimarães, C. T., Nazarian, A., da Costa E Silva, L., Parentoni, S. N., ... Pastina, M. M. (2018). Improving accuracies of genomic predictions for drought tolerance in maize by joint modeling of additive and dominance effects in multi-environment trials. *Heredity*, 121, 24–37. https://doi.org/10.1038/s41437-018-0053-6
- FAO. (2017). The state of food security and nutrition in the world 2017: Building resilience for peace and food security. Rome: FAO.
- Guy, S. O., Jitkov, V. A., Lauver, M. A., & Horton, A. (2013). Cereal and Grain Legume Variety Performance Trials 2013. Washington State University Extension Variety Testing Program.
- Guy, S. O., Jitkov, V. A., Lauver, M. A., Horton, A., & Higginbotham, R. (2014). *Cereal and Grain Legume Variety Performance Trials 2014*. Washington State University Extension Variety Testing Program.

- Hernández-Espinosa, N., Mondal, S., Autrique, E., Gonzalez-Santoyo, H., Crossa, J., Huerta-Espino, J., ... Guzmán, C. (2018). Milling, processing and end-use quality traits of CIMMYT spring bread wheat germplasm under drought and heat stress. *Field Crops Research*, 215, 104–112. https://doi.org/10.1016/j.fcr.2017.10.003
- Higginbotham, R., Jitkov, V., & Horton, A. (2016). Cereal Variety Performance Trials 2016. Washington State University Extension Variety Testing Program.
- Kelly, A. M., Smith, A. B., Eccleston, J. A., & Cullis, B. R. (2007). The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. *Crop Science*, 47, 1063– 1070. https://doi.org/10.2135/cropsci2006.08.0540
- Laidig, F., Piepho, H.-P., Rentel, D., Drobek, T., Meyer, U., & Huesken, A. (2017). Breeding progress, environmental variation and correlation of winter wheat yield and quality traits in German official variety trials and on-farm during 1983–2014. *Theoretical and Applied Genetics*, 130, 223–245. https://doi.org/10.1007/s00122-016-2810-3
- Mares, D. J., & Mrva, K. (2014). Wheat grain preharvest sprouting and late maturity alpha-amylase. *Planta*, 240, 1167–1178. https://doi.org/ 10.1007/s00425-014-2172-5
- Martinez, S. A., Godoy, J., Huang, M., Zhang, Z., Carter, A. H., Campbell, K. A. G., & Steber, C. M. (2018). Genome-wide association mapping for tolerance to preharvest sprouting and low falling numbers in wheat. *Frontiers in Plant Science*, 9. https://doi.org/10.3389/fpls. 2018.00141
- Mengesha, W., Menkir, A., Meseka, S., Bossey, B., Afolabi, A., Burgueno, J., & Crossa, J. (2019). Factor analysis to investigate genotype and genotype × environment interaction effects on pro-vitamin A content and yield in maize synthetics. *Euphytica*, 215. https://doi.org/10. 1007/s10681-019-2505-3
- Meyer, K. (2009). Factor-analytic models for genotype × environment type problems and structured covariance matrices. *Genetics, Selection, Evolution, 41*. https://doi.org/10.1186/1297-9686-41-21
- Mrva, K., & Mares, D. J. (1996). Inheritance of late maturity alphaamylase in wheat. *Euphytica*, 88, 61–67. https://doi.org/10.1007/ BF00029266
- Piepho, H.-P. (1998). Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. *The*oretical and Applied Genetics, 97, 195–201. https://doi.org/10.1007/ s001220050885
- Ross, A. S., & Bettge, A. D. (2009). Passing the test on wheat end-use quality. In B. F. Carver (Ed.), *Wheat science and trade* (pp. 455–493). Oxford, UK: Wiley-Blackwell.
- Ross, A. S., Flowers, M. D., Zemetra, R. S., & Kongraksawech, T. (2012). Effect of grain protein concentration on falling number of ungerminated soft white winter wheat. *Cereal Chemistry*, 89, 307–310. https://doi.org/10.1094/CCHEM-07-12-0085-RC
- Sjoberg, S. M., Carter, A. H., Steber, C. M., & Campbell, K. A. G. (2020). Unraveling complex traits in wheat: Approaches for analyzing genotype × environment interactions in a multienvironment study of falling numbers. *Crop Science*. https://doi.org/10.1002/csc2.20133
- Smith, A. B., & Cullis, B. R. (2018). Plant breeding selection tools built on FA mixed models for multi-environment trial data. *Euphytica*, 214. https://doi.org/10.1007/s10681-018-2220-5
- Smith, A. B., Ganesalingam, A., Kuchel, H., & Cullis, B. R. (2015). Factor analytic mixed models for the provision of grower information from national crop variety testing programs. *Theoretical and Applied Genetics*, 128, 55–72. https://doi.org/10.1007/s00122-014-2412-x
- Smith, A., Cullis, B., & Thompson, R. (2001). Analyzing variety by environment data using multiplicative mixed models and adjustments

for spatial field trend. *Biometrics*, 57, 1138–1147. https://doi.org/10. 1111/j.0006-341X.2001.01138.x

- Tuttle, K. M., Martinez, S. A., Schramm, E. C., Takebayashi, Y., Seo, M., & Steber, C. M. (2015). Grain dormancy loss is associated with changes in ABA and GA sensitivity and hormone accumulation in bread wheat, *Triticum aestivum* (L.). *Seed Science Research*, 25, 179– 193. https://doi.org/10.1017/S0960258515000057
- USDA. (2017). Commodity data specific to wheat, feed grains, and oilseeds. Washington, DC:USDA Farm Service Agency.
- Ward, B. P., Brown-Guedira, G., Tyagi, P., Kolb, F. L., Van Sanford, D. A., Sneller, C. H., & Griffey, C. A. (2019). Multienvironment and multitrait genomic selection models in unbalanced early-generation wheat yield trials. *Crop Science*, 59, 1–17. https://doi.org/10.2135/cropsci2018.03.0189
- Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. New York: Springer.
- Yabwalo, D. N., Berzonsky, W. A., Brabec, D., Pearson, T., Glover, K. D., & Kleinjan, J. L. (2018). Impact of grain morphology and the

genotype by environment interactions on test weight of spring and winter wheat (*Triticum aestivum* L.). *Euphytica*, 214. https://doi.org/10.1007/s10681-018-2202-7

Zeng, M., Morris, C. F., Batey, I. L., & Wrigley, C. W. (1997). Sources of variation for starch gelatinization, pasting, and gelation properties in wheat. *Cereal Chemistry*, 74, 63–71. https://doi.org/10.1094/ CCHEM.1997.74.1.63

How to cite this article: Sjoberg SM, Carter AH, Steber CM, Garland-Campbell KA. Application of the factor analytic model to assess wheat falling number performance and stability in multienvironment trials. *Crop Science*. 2021;61:372–382. https://doi.org/10.1002/csc2.20293