

Kansas State University Libraries

New Prairie Press

Conference on Applied Statistics in Agriculture

1998 - 10th Annual Conference Proceedings

STATISTICAL ANALYSIS OF FIELD WHEAT VARIETAL PERFORMANCE TRIALS

A. M. Feyerherm

R. G. Sears

J. J. Higgins

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

Recommended Citation

Feyerherm, A. M.; Sears, R. G.; and Higgins, J. J. (1998). "STATISTICAL ANALYSIS OF FIELD WHEAT VARIETAL PERFORMANCE TRIALS," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1274>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact cads@k-state.edu.

STATISTICAL ANALYSIS OF FIELD WHEAT VARIETAL PERFORMANCE TRIALS

A.M. Feyerherm, R.G. Sears, and J.J. Higgins¹

Abstract

The purpose of this research was to formulate statistical models and assumptions to apply to the problem of comparing wheat varieties for yielding ability among locations within seasons and over seasons. The methodology could just as well be applied to field testing of other crops for yield or other characteristics of interest (test weight, protein level, etc.)

The methodology approaches the problem of comparing varieties by comparing how well each “measures up” when matched against some common checks. For each variety, the basic data are differences in yield between the variety and the average yield of the checks at different testing locations within a season and over seasons. The differences are assumed to be “nature-randomized” sample values from a population of differences created by different environments within seasons and over seasons.

The methodology is illustrated by application to hard red spring wheat varieties in the U.S. Northern Plains. Results showing varieties in descending order by differential yielding ability, together with standard errors and probabilities when testing null hypotheses, provide a consolidated summary of elite varieties in testing programs.

KEY WORDS: varietal testing, differential yielding ability.

1. Introduction

Performance trials to compare wheat varieties have been a part of research programs in the U.S. for over a century, Georgeson, et al. (1890). During that time, simple means have been the acceptable statistic to measure level of performance among locations within seasons and among seasons, or among seasons at a particular location. On the other hand statistical methods to do inference from sample to population have been ill-defined for handling variation in results over locations and seasons.

Many of the location/season variation problems encountered, when applying statistical methods to varietal performances trials (VPTs), were recognized early on. Salmon and Laude

¹Professor Emeritus of Statistics, Professor of Agronomy, and Professor of Statistics, Kansas State University; Manhattan, KS.

(1932) looked at 20 years of winter wheat VPTs and concluded that seasonal variation was a dominating factor when interpreting yield differences among varieties and that this was well known to agronomists. Yates and Cochran (1938) pointed out the limited utility of a single trial, either for picking a best variety or elucidating underlying scientific laws. They also indicated the need for multiple locations and seasons when calculating proper standard errors for sets of VPTs, and added the warning that uncritical application leads to erroneous conclusions. Salmon (1951) discussed the dilemma of waiting through many seasons for statistical significance to occur versus the need for a farmer to make decisions in a shorter time frame. He also indicated that season variation in yields is more important than soil variation in the Great Plains when comparing varieties. Baker (1969) stressed the need for multiple locations/seasons in selection programs in Canada and found that a reduction of reps per location resulted in minor loss in efficiency when selecting varieties for yields.

Kempton and Talbott (1988) reviewed the role played by statistics in the breeding of new varieties and referenced research on stability analysis relative to the effects of variety \times environment ($G \times E$) interactions. They also referenced use of checks, neighboring plots, and incomplete block designs for improving efficiency in single trials. However, they acknowledged that a single trial is of limited value for predicting performance for varieties grown in other fields and seasons. Talbot and Robinson (1987) considered allocation of resources between locations and seasons as well as referencing work on estimating variance components in multiple VPTs using REML (restricted maximum likelihood) methods. Methods for estimating $G \times E$ effects were compared by Zobel et al. (1988). Brennam et al. (1981) used pattern analysis to group locations but, when applied, suggested groupings varied over seasons in Australia.

Both past and present annual reports of VPTs in the U.S. Great Plains show computation of C.V.'s (coefficients of variation) and LSD's (least significance differences) at each location for a current year even though the statistical population is confined to the immediate environment in that one year. Two and three-year means are reported for each location sometimes with and sometimes without LSD's. If LSD's are reported, they usually are smaller than those for a single trial indicating use of an error term without a variety \times season interaction component. The result is an underestimate of standard errors of means.

To enhance reporting of VPTs for multiple locations/seasons, we developed a methodology which, within a specified geographical/climatological region, produces an estimator $\overline{\overline{D}}$ which measures how well each variety performs relative to a standard. In addition to $\overline{\overline{D}}$ values for each variety, standard errors of $\overline{\overline{D}}$ and probabilities of rejecting the null hypothesis are part of the output. Development of our method involved basic elements of statistical theory and practice; namely defining units of observation, sampling schemes, definitions of populations, randomization concepts, linear models, variance components, estimators and their standard errors, and probability distributions.

2. Data Set

Rather than finding data to fit a theory, our investigation began with a data set of VPTs that needed a statistical theory to interpret its contents. The set had wheat yield, by variety, for the past 50-80 years from a 13-state region in the U.S. Great Plains at a variable number, roughly 3-15, locations within a state. The trials were laid out in replicated plots on university-affiliate land or in fields of cooperating farmers. Most of the varieties included in the trials had survived an elimination process and were “released” for commercial planting. The experimental design for each VPT had some form of blocking (randomized complete block or lattice). Varieties chosen for a trial were the prerogative of research and extension specialists or private companies and the set of varieties often changed not only from season to season but sometimes location to location within a state. Varietal yields were reported as averages over reps at a location.

3. Environmental Units and Populations

Our basic unit of observation was an environmental unit (EU). Physically, it encompassed the total environment above and within the soil. Hence it included not only the soil into which wheat was sown but also weather conditions, presence of pests, and management practices employed throughout a season of growth.

We considered an EU to be a sampling unit and applied statistical methods for one-way random effects linear models in two-stage sampling with primary environmental units (PEU) and secondary environmental units (SEU). A PEU was defined in space/time by the environment that existed in a geographical/climatological region (state, portions of a state, multiple states, or portions of multiple states) during a growing season. An SEU was defined by the environment that existed at a location within a specified region during a growing season. A population of seasons, conceptually infinite, generates a population of PEUs while a population of locations within a region, conceptually infinite, generates a population of SEUs within a PEU. A sample of five locations per season over four seasons generates 4 PSUs and 20 SEUs.

4. Role and Choice of Checks

Varieties that were used as checks played an important role in our methodology because they were used to remove some of the variation in grain yields due to variation in environments from PEU to PEU and SEU to SEU. For each SEU (a location-year), the yield of a check, or better yet, the average yield of multiple checks, was a standard for that particular SEU. Each variety present in that SEU was then compared with that standard. The check or checks were common to all locations and in all seasons covered by a given analysis for the

region of interest. They played the role of an environmental measure much like the role of a VPT mean yield over varieties in some analyses of stability and G x E interactions (Finlay and Wilkinson, 1963). Unlike a VPT mean they were not dependent on the set of varieties present.

The best choice of checks were those which were popular with producers according to USDA surveys of varieties planted. These varieties were usually included in almost all SEUs. Such was the case for the checks shown in Tables 1 and 2.

5. Models and Assumptions

Consider a sample of n locations within seasons and N seasons. For a variety, call it A , we express its differential yielding ability, relative to a standard, by the one-way random effects linear model:

$$D_{ij}(A) = \mu(A) + S_i(A) + L_{ij}(A); i = 1, 2, \dots, N, j = 1, 2, \dots, n,$$

where

$D_{ij}(A)$ = yield of variety (A) minus the standard at location j within season i where the standard equals the average yield of the checks in the ij^{th} SEU;

$\mu(A)$ = population mean difference in yield between variety (A) and the standard for the population of PSUs (seasons) associated with the region of interest;

$\mu(A) + S_i(A) = \mu_i(A)$ = population mean of $D_{ij}(A)$ values over SEUs (locations) for the i^{th} PEU (season),

where $S_i(A)$ was assumed to be independently and identically distributed $N(0, \sigma_S^2(A))$;

$L_{ij}(A)$ = departure of $D_{ij}(A)$ from $\mu_i(A)$ for the j^{th} SEU (location) in the i^{th} PSU (season),

where $L_{ij}(A)$ was assumed to be independently and identically distributed $N(0, \sigma_L^2(A))$.

Lastly, we assumed $S_i(A)$ and $L_{ij}(A)$ were independent random variables.

To support the independence assumptions we examine the factors that contribute to variation in the $L_{ij}(A)$ and $S_i(A)$ values. The G x E interaction effects are prime contributors to both terms mainly because of weather variation from EU to EU both within seasons (among

locations) and among seasons (Feyerherm, et al., 1992). Not only can EU to EU variation in soil moisture and soil and ambient temperature profiles for a growing season effect variety (A) and the checks differently, but also it can create further variation in D-values by interacting with genetic characteristics (resistance to viruses, diseases, and insects; winter hardiness, straw strength, rate of development, tillering habit, response to fertilizers). Some of this interactive variation may contribute to $\mu(A)$ (e.g. rust resistance may increase the yield of variety (A) over the checks by several bushels/acre on the average over seasons). The remaining portion is distributed between $L_{ij}(A)$ and $S_i(A)$.

Weather is not the only force contributing to variation in the error terms. Presence of viruses, diseases, and insects play a role. Together with weather their effects vary uncontrollably and unpredictably both among locations within a season and among seasons. Thus “nature” becomes a randomizing agent and makes the independence assumptions reasonable.

The dominance of weather in the variance of D values suggest the need for two components of variance to model D-values. Even cursory examination of weather data demonstrates that, within a climatological region, weather-related factors vary more in intensity from season to season than from location to location within a season. This does not mean that $\sigma_S^2(A)$ is necessarily larger than σ_L^2 . It may be zero for some varieties. It does mean that the variance of a randomly chosen D(A) is $\sigma_L^2(A) + \sigma_S^2(A)$.

6. Estimators, Standard Errors, and Inference

Since we use a one-way random effects model to describe $D_{ij}(A)$, computation of estimators and their standard errors is well documented (Snedecor and Cochran, 1967). We use the sample mean $\bar{D}_i(A) = \sum_{j=1}^n D_{ij}(A)/n$ to estimate $\mu_i(A)$ and

$$SE(\bar{D}_i(A)) = \left[\sum_{j=1}^n (D_{ij}(A) - \bar{D}_i(A))^2 / n(n-1) \right]^{1/2} \text{ to estimate the standard error of } \bar{D}_i(A). \text{ To}$$

estimate the parameter $\mu(A)$, we use $\bar{\bar{D}}(A) = \sum_{i=1}^N \bar{D}_i(A)/N$ and to estimate its standard error use

$$SE(\bar{\bar{D}}(A)) = \left[\sum_{i=1}^N (\bar{D}_i(A) - \bar{\bar{D}}(A))^2 / N(N-1) \right]^{1/2}. \text{ } SE^2(\bar{D}_i(A)) \text{ estimates } \sigma_L^2(A)/n \text{ and}$$

$$SE^2(\bar{\bar{D}}(A)) \text{ estimates } [\sigma_L^2(A)/n + \sigma_S^2(A)]/N.$$

Based on our model for $D_{ij}(A)$, Student's t-distribution is appropriate for affixing confidence intervals and testing hypotheses about $\mu_i(A)$ and $\mu(A)$. If we want to compare varieties (A) and (B), the difference $(\bar{\bar{D}}(A) - \bar{\bar{D}}(B))$ can be used as a measure of their differential yielding ability with associated precision measured by the conservative estimate of its standard error:

$$SE(\bar{\bar{D}}(A) - \bar{\bar{D}}(B)) = [SE^2(\bar{\bar{D}}(A)) + SE^2(\bar{\bar{D}}(B))]^{1/2}.$$

Student's-t with $[(N_A - 1) + (N_B - 1)]$ d.f., where N_A and N_B are the number of seasons A and B have been tested, would provide a good approximation for inferences about $[\mu(A) - \mu(B)]$.

In practice, it would not be unusual for the number of locations per season (n) to vary. The unweighted mean should be used to calculate $\bar{\bar{D}}(A)$ and the $SE(\bar{\bar{D}}(A))$ should be calculated using the same formula as in the equal sample size case. Probabilities using the t-distribution will be approximate but the error should be reasonable unless the n 's are very disparate. Clearly, seasons with $n = 1$ should be dropped from the sample of seasons.

7. Application

The methodology presented here has been applied to three of the major wheat regions in the central United States (hard red winter, soft red winter, hard red spring). In all regions, the component $\sigma_s^2(A)$ was significantly different from zero for some but not all varieties.

Obviously, the component $\sigma_t^2(A)$ was greater than zero.

As an example, consider the hard red spring region (Tables 1-2). Calculations were performed in SAS (SAS Institute Inc., 1985). First PROC MEANS was used to generate and output \bar{D}_i -values, for each season and variety, from the D_{ij} -values at locations within regions. Then PROC MEANS was applied to the \bar{D}_i -values ($i = 1, 2, \dots, N$) to generate $\bar{\bar{D}}$, the standard error of $\bar{\bar{D}}$, and the probability (PRT) for each variety.

Checks were chosen on the basis of their popularity with growers as shown in USDA surveys. For example, 'Amidon', '2375', and 'Grandin' were the most popular varieties in the western 2/3rds of the Dakotas (Table 1). This made them "naturals" for other varieties to compete against in head-to-head competition at each location. All three varieties appeared together at almost all locations over the nine years (seasons) 1989-1997 so few VPTs had to be dropped. The results in Tables 1-2 are typical of what we found in other wheat regions and suggest that use of $\alpha = .05$, as a decision point for rejecting a null hypothesis, is not practical considering the number of seasons it may take to overcome variance with replication. Increasing α to, say .20 to .40 (Carmer, 1976 and Carmer and Walker, 1988) would decrease the probability of committing the error of passing up varieties which may have a several bushel advantage over popular checks. By the same reasoning one might consider confidence intervals of the form $\bar{\bar{D}}(A) \pm 1 * SE(\bar{\bar{D}}(A))$ when estimating $\mu(A)$.

Our formulation and proposed solution to the problem of modeling VPTs for statistical analysis allows considerable latitude to testing hypotheses of interest. One can compare varieties within regions, or a variety with itself across regions, if common checks were used. Additionally, one can adjust regional boundaries making them larger or smaller even to the size of one location. However, for the one-location region with a small number of seasons the $SE(\bar{\bar{D}}(A))$ may be too large to obtain a reasonably small PRT value.

Acknowledgments

We are indebted to all the wheat breeders, wheat specialists, cooperating farmers, experiment station personnel, and anyone else who had a hand in generating yield data from wheat varietal tests over a 13-state area over the past 50 years. The data set has been invaluable in developing and applying the statistical concepts and models used in this research.

References

- Baker, R.J. 1969. G x E interactions in yield of wheat. *Can. J. Plant Sci.* 49, 743-751.
- Brennan, P.S., D.E. Byth, D.W. Drake, I.H. DeLacy, and D.G. Butler. 1981. Determination of the location and number of test environments for a wheat cultivar evaluation program. *Aust. J. Agric. Res.* 32, 189-201.
- Carmer, S.G. 1976. Optimal significance levels for application of the least significant difference in crop performance trials. *Crop Sci.* 16:95-99.
- _____ and W.M. Walker. 1988. Significance from a statistician's viewpoint. *J. Prod. Agric.* 1:27-33.
- Feyerherm, A.M., R.G. Sears, and G.M. Paulsen. 1992. Utilizing genotype x weather interactions in cultivar selection. *Agron J.* 84:391-394.
- Finlay, K.W. and G.N. Wilkinson. 1963. The analysis of adaption in a plant-breeding programme. *Aust. J. Agric. Res.* 14:742-754.
- Georgeson, C.C., H.M. Cottrell, and W. Skelton. 1890. Experiments with wheat. Kansas State Agric. College. Dept. of Agric., Manhattan, KS, Bulletin No. 11.
- Kempton, R.A. and M. Talbot. 1988. The development of new crop varieties. *Jour. Royal Stat. Soc., Series A.* 151:327-341.
- Salmon, S.C. 1951. Analysis of variance and long-time variety tests of wheat. *Agron. J.* 43:562-570.
- _____ and H.H. Laude. 1932. Twenty years of testing varieties and strains of winter wheat. *Kansas Agric. Expn. Sta., Kansas State College, Manhattan, KS.*
- SAS®. 1985. *SAS User's Guide: Basics. Version 5 Edit.* SAS Institute Inc., Cary, N.C.

Snedecor, G.W. and W.G. Cochran. 1967. *Statistical Methods*, Sixth Edit. The Iowa State University Press. Ames, IA.

Talbot, M. and D.L. Robinson. 1987. Planning a series of variety trials. *Biuletyn Oceny Odimian*. Tom XII, Zeszyt 1-2 (17-18).

Yates, F. and W.G. Cochran. 1938. The analysis of groups of experiments. *J. Agric. Sci., Camb.* 28:556-580.

Zobel, R.W., W.J. Wright, and H.J. Gauch Jr. 1988. Statistical analysis of a yield trial. *Agron. J.* 87:388-393.

Table 1. Ordering of hard red spring wheat varieties by differential yielding ability, relative to checks, in dryland varietal trials in Minnesota and the Dakotas (1989-1997). Units are bushels/acre.

Western 2/3rds of DAKs					Minnesota & East 1/3rd DAKs				
Variety	\bar{D}^\dagger	S.E. [§]	PRT ^{††}	N [‡]	Variety	\bar{D}	S.E.	PRT	N
Lars	4.6	1.9	.09	4	Russ	4.2	0.4	.01	3
Verde	3.8	1.7	.15	3	Oxen	3.9	0.1	.01	2
Nordic	3.2	0.8	.01	6	Hamer	3.7	0.7	.02	4
2398	2.8	0.7	.06	3	Lars	3.4	0.9	.03	4
Hamer	2.6	2.7	.41	4	Verde	3.1	0.8	.06	3
McNeal	2.4	1.3	.12	5	^c 2375	2.0	0.3	.00	9
Norm	1.9	1.1	.17	5	Norlander	2.0	1.0	.15	4
Gunner	1.7	1.5	.45	2	2370	1.6	0.6	.03	8
Prospect	1.5	0.5	.02	7	Nordic	1.0	1.0	.35	8
Oxen	1.5	0.6	.24	2	Trenton	0.8	0.9	.42	3
Russ	1.3	0.9	.27	3	Barrie	0.8	0.1	.03	2
^c 2375	1.1	0.6	.10	9	Keene	0.5	0.1	.08	2
Cora	0.8	0.2	.15	2	2398	0.2	0.7	.84	3
Keene	0.5	1.0	.72	2	Sharp	0.1	0.3	.59	8
Barrie	0.3	0.5	.61	2	Gunner	0.0	2.0	.99	2
^c Amidon	-0.1	0.6	.94	9	^c Butte 86	-0.7	0.3	.03	9
Trenton	-0.3	1.0	.78	3	Kulm	-1.0	0.9	.32	4
Sharp	-0.9	0.8	.32	8	Amidon	-1.3	1.2	.30	9
2371	-1.0	0.9	.32	7	Wheaton	-1.3	0.8	.12	9
Stoa	-1.0	0.2	.00	9	^c Grandin	-1.3	0.4	.01	9
^c Grandin	-1.0	0.3	.00	9	Norm	-1.6	1.2	.25	6
Guard	-1.2	0.6	.10	6	Marshall	-3.4	1.1	.01	9
Butte 86	-1.5	0.6	.04	9	2371	-3.5	0.5	.00	7
2370	-1.6	1.2	.22	8	Gus	-3.8	1.2	.01	9
Norlander	-1.7	1.1	.21	4	Backup	-10.4	2.1	.13	2
Gus	-1.8	0.2	.00	9	Chris	-14.8	1.4	.00	9
Kulm	-2.1	1.0	.13	4					
Ernest	-2.4	0.6	.05	3					
Len	-3.7	0.9	.00	9					
Glupro	-13.2	1.7	.02	3					
Bacup	-14.3	1.2	.05	2					

^c Check variety[†] \bar{D} = Mean of the seasonal means where a seasonal mean (\bar{D}) is the mean of differences, over locations within a season, between a variety and the average of check yields.[§]S.E. = Standard error of \bar{D} = standard deviation of seasonal means (\bar{D} s) divided by \sqrt{N} .[‡]N = Number of seasons.^{††}PRT = Probability that the absolute value of $[\bar{D}/S.E.]$ exceeds a Student's tabled t-value with N - 1 degrees of freedom if the population mean difference is zero.

Table 2. Ordering of hard red spring wheat varieties by differential yielding ability, relative to checks, in dryland varietal trials in Montana (1988-1997). Units are bushels/acre.

Variety	\bar{D}^{\dagger}	S.E. [§]	PRT ^{††}	N [‡]
McNeal	2.9	0.5	.01	4
^c Glenman	1.5	0.4	.00	10
^c Amidon	1.1	0.6	.08	10
WPB 936	1.0	1.1	.44	5
^c Newana	1.0	0.5	.10	10
Express	0.8	1.5	.63	5
WPB 926	0.5	1.1	.70	5
Fergus	0.0	0.8	.99	2
Hi-Line	-0.1	0.9	.90	6
Rambo	-0.4	0.5	.54	7
Trenton	-1.0	0.1	.02	2
Ernest	-1.1	0.7	.39	2
Len	-2.4	1.3	.10	8
^c Lew	-3.6	0.5	.00	10
Fortuna	-4.4	1.1	.00	10
Thatcher	-7.9	0.8	.00	10

C, †, §, ††, ‡ (See Table 1)