Kansas State University Libraries

New Prairie Press

Conference on Applied Statistics in Agriculture 2014 - 26th Annual Conference Proceedings

CHECK BASED STABILITY ANALYSIS METHOD AND ITS APPLICATION TO WINTER WHEAT VARIETY TRIALS

Jixiang Wu
South Dakota State University

Karl Glover
South Dakota State University

Nathan Mueller
UNL Extension in Dodge County

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

Wu, Jixiang; Glover, Karl; and Mueller, Nathan (2014). "CHECK BASED STABILITY ANALYSIS METHOD AND ITS APPLICATION TO WINTER WHEAT VARIETY TRIALS," *Conference on Applied Statistics in Agriculture*. https://doi.org/10.4148/2475-7772.1006

This Event 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.

Check Based Stability AnalysisMethod and Its Application to Winter Wheat Variety Trials

Jixiang Wu[1*], Karl Glover[1], and Nathan Mueller[2]

[1:]Plant Science Department, South Dakota State University, SD 57007 and 2: Cropping Systems & Agricultural Technologies Extension, UNL Extension in Dodge County, Fremont, NE 68025
*: Corresponding author (Jixiang.wu@sdstate.edu)

Abstract

Finley-Wilson (FW) regression based stability analysis is highly dependent on the testing varieties and environments being used. In this study, we proposed a check based regression method to determine yield stability. One advantage of this method is its capability to determine yield stability through widely acceptable varieties and thus to provide more meaningful information to evaluate the potential use of new varieties. In addition, with integration a resampling technique, bootstrapping method, yield stability can be compared among different varieties/genotypes from either the same or different testing environments. As a demonstration, we applied this method to analyze the 2009-2011 winter wheat CPT (crop performance test) data collected by South Dakota State University.

1. Introduction

Genotype-by-environment (G×E) interaction is a challenging issue in developing widely adapted crop varieties (Gray, 1982; M. S. Kang & Miller, 1984). A variety associated with high G×E interaction effectssuggests that it is sensitive to various environmental conditions and thus it is considered as less stable. Thus, fully investigating yield stability is critical to determining variety adaptation prior its release to the markets.

Crop trials or crop performance tests (CPT) at multiple locations and possibly for several years are required to generate experimental data for measuring stability of a genotype. Due to the complexity of field trials and various definitions of yield stability (Lin, Binns, & Lefkovitch, 1986), different statistical methods for measuring stability have been proposed. Based on the review paper (Lin et al., 1986), these methods can be clustered into two basic categories: variation-based (Francis & Kannenberg, 1978; Plaisted & Peterson, 1959; Shukla, 1972; Wricke, 1962) and regression-based (Eberhart & Russell, 1966; Finlay & Wilkinson, 1963; Perkins

&Jinks, 1968). In some later study both mean yield performance and yield stability are considered simultaneously (Fan et al., 2007; K. S. Kang & Mangari, 1995; M. S. Kang, 1993). Another commonly used approach is the additive main effects and multiplicative interaction (AMMI) method(Crossa, Gauch, & Zobel, 1990). The AMMI method is a principal component analysis based approach. Based on principal component analysis, as it treats multi-environmental data as multivariate data structures, this method considers these two important parameters as well.

In FW based yield stability analysis, two parameters are associated with yield stability: slope (b_1) and coefficient of determination (^{2}), where b_1 represents the mean response rate to environmental index, and ^{2} represents the consistence of response to environmental index. It seems that all genotypes are compared with environmental index, which are combinations of all genotypic means in the trials. Since each environmental index is a linear combination of all genotypic values under a specific environment, which is also equivalent to a specific environmental effect, the stability could be highly dependent on environmental indexes. If environmental conditions play a major role on yield, the stability parameters obtained by regression based methods could be overestimated.

A common issue associated with the above mentioned stability analyses methods is that yield stability assessment is highly dependent on the varieties in the trials. For example, environmental index (EI) in the simply linear regression method (Finlay & Wilkinson, 1963) is calculated by averaging varieties for each environment. Therefore, it is very likely that one variety is considered stable if it is similar to the most remaining varieties in the trials while being considered unstable if this variety is different from the most varieties in the same test. On the other hand, since varieties in the trials could be different from year to year, the comparison of yield stability among varieties in different years could be challenging. However, it is a very common practice that some check varieties are repeatedly used in different locations and years in variety trials and the results of these checks are not fully used for yield stability determination for other varieties. Thus, it has become our great interest if we can use these checks to evaluate yield stability for other varieties.

In this study, we proposed a new method that can be used to evaluate the stability of based on the performance of widely used checks. We also integrated a resampling technique, bootstrapping method, in this check-based yield stability method. As a

New Prairie Press
http://newprairiepress.org/agstatconference/2014/proceedings/6

demonstration, we applied this method to evaluate yield stability for the varieties in the winter wheat CPT conducted by South Dakota State University in 2009-2011. The purpose of this study was to provide a different way to reveal yield stability that could be used to determine the new genotypes to be released to farmers.

2. Materials and Methods

2.1. Materials

Three years of winter wheat CPT data (2009, 2010, and 2011) were used for this study. For a winter wheat CPT, cultivars normally are planted in fall and harvested in the coming summer. In order to reduce some confusion, we used the harvest years as the testing years in this study. For example, year 2009 in this study indicates that the testing winter wheat varieties were grown in 2008 but harvested in 2009. Only the variety mean values under each environment were used in this study. For detailed information, readers may refer to the website http://www.sdstate.edu/ps/extension/crop-mgmt/cpt/variety-trials-results.cfm. Summarized information regarding the data is provided in Results and Discussion of this study.

2.2. Statistical Methods

Several approaches can be used for analyzing these unbalanced trial data. One important strategy is to apply linear mixed model approaches to estimate variance components and to predict/estimate these effects. In this study, without losing our focus, we only emphasized on regression based yield stability analysis. In order to better understand the method we proposed in study, we would like to briefly repeat the description of the simple linear regression based stability analysis(Finlay & Wilkinson, 1963). We called it the FW-based method.

Given g genotypes tested in e environments, if we only use variety mean values in each environment for data analysis (Lin et al., 1986), then the linear model can be expressed in equation (1):

$$= + + + + \quad (1)$$

Where is the mean observation for genotype i in environment j, is the population mean, is a genotypic effect, is an environmental effect, is a genotype-by-environment interaction effect, which is related to the stability of a genotype, is a random error. In equation (1), and are confound if no field replications are provided. Based on equation (1), the environmental index (EI) for environment j can be calculated by using equation (2):

$$= \frac{1}{} \sum_{=1} \quad (2)$$

Once all EIsare obtained by using equation (2), a simple linear regression in equation (3) can be employed to estimate stability parameters for each variety:

$$= {}_0 + {}_1 \quad + \qquad j=1,\ldots e \qquad (3)$$

Equation (3) can be used to calculate two parameters, slope ${}_1$ and coefficient of determination ^{2}. The equation (3) can be repeatedly used to calculate these parameters for each variety. It is obviously that environmental indexes used for each genotype are the same if the data set is balanced. Stability among different varieties can be compared by confidence intervals obtained by bootstrapping methods (Efron, 1979; Manly, 2006;Wu, Glover, & Berzonsky, 2012).

Environmental effects in equation (1) are equivalent to the deviations of environmental indexes from population mean used in equations (2) and (3).Thus, the estimated stability parameters could be highly dependent on which environments and/or testing genotypesbeing used. Using the FW based method, it is possible that widely grown varieties such as check lines can be classified as unstable or less stable compared to other new testing genotypes which are genetically similar. Therefore, we proposed this check based stability analysis method in this study. Thecheck based stability analysis method proposed in this study is very similar to equation (3) and is detailed in equation (4):

$$= {}_0 + {}_1 \quad + \qquad j=1,\ldots e \qquad (4)$$

Where is called a check based environmental index, which is the mean value of one check under environment *j*.The remaining parameters in equation (4) were defined in equations (1) to (3). The major difference between this check based method and the FW – based method (Finlay & Wilkinson, 1963)is to usecheck means under different environments to replace the environmental index in the FW method. Therefore, the first major advantage of this check based method is that we can determine yield stability compared to one or more widely accepted checks.

Since the observations for check genotype(s) are random too, resampling approaches such as jackknife and bootstrapping techniques can be used to obtain standard errors and confidence intervals for all parameters of interest (Manly, 2006; Wu et al., 2012). By doing so, stability parameters can be compared among varieties using confidence intervals.All data analyses were conducted by the R programs developed by the senior author of this paper. These R functions have been packed in the R package qgtools and will be available for public use once this paper is accepted for publication.

3. Results and Discussion

3.1. SDSU Winter wheat CPT trials from 2009 to 2011

Before conducting yield stability analysis for these winter wheat genotypes evaluated in South Dakota environments, we would like to briefly provide some descriptions of these CPT trials in 2009, 2010, and 2011. There were 11, 13, and eight testing sites for 2009, 2010, and 2011, respectively. However, there were 15 common test siteswere for all three years, where intensive management studies (IMS) were also conducted in three locations: Brookings, Selby, and Winner (Table 1). The numbers of varieties planted in these 15 testing sites are provided in Table 1 as well. Only six testing sites: Brookings, Brookings.IMS (intensive management study in Brookings), Selby, Platte, Onida, and Pierre, were repeated for all three years, while the remaining nine testing sites were used one year or two years.

In total 36 winter wheat varieties were evaluated across these three years (Table 2); however, 22 varieties (Alice, Arapahoe, Art, Darrell, Expedition, Fuller, Harding, Hatcher, Hawken,Jagalene, Jerry, Lyman, Millennium, Overland, Radiant, SD05118-1, Settler CL, Smoky Hill, Striker, Wahoo, Wendy, and Wesley) were evaluated in all 32 testing environments. In addition, most of these varieties were under U.S. Plant Variety Protection (PVP), implying that seed sales of these varieties are restricted to classes of certified seed. The variety: Lyman, characterized with moderate resistance to Fursarium head blight, was reported to be adapted to all six areas in South Dakota as shown in Figure 1. If balanced data are used for stability analysis, fourteen (14) varieties in the trials would have to be dropped (Table 2). It suggests that all SD varieties except SD05118-1 developed from the winter wheat breeding program at South Dakota State University and many other varieties will be underutilized for yield stability assessment.

3.2. Yield stability for varieties in South Dakota environments

Each of these PVP varieties can be considered as a potential check and used to calculate relative yield stability for other varieties included in the three-year trials. The same procedure can be applied for calculating stability based on each of these PVP varieties or other varieties of interest. As a demonstration of this new method proposed in this study, weonly emphasized on Lyman as check to evaluate yield stability for all other varieties. For comparison, we also conducted stability analysis for these varieties using the FW based method.Both coefficients of determination (^{2}) and slopes ($_1$) for both methods were estimated and are reported here (Tables 3 and 4). In addition, the bootstrapping procedure was repeated for 1000 times for each variety to obtain 95% confidence of intervals, represented by 2.5% low-limit (LL) and 97.5% up-limit (UL). These results are summarized in Table 3 and Table 4 for ^{2}and $_1$, respectively.

A coefficient of determination(^{2}) in FW stability analysis measures average consistence of a variety over environmental indexes under the same environments, while an ^{2} in check-based stability measures average consistence of a variety over the performance of a check under the same environments.The coefficients of determination between two methods were highly correlated (0.885). As expected, the ^{2}forthe FW based method was generally greater than that for Lyman based because the performance of a variety is more correlated with EIs (Table 3). The results in Table 3 showed that the FW based ^{2} was either significantly greater than 0.90 or no significantly different from 0.90 for all varieties except variety Radiant, which had a low ^{2}(0.733), significantly less than 0.80 and those for most of other varieties (Table 3). Comparing with the performance of the variety Lyman across these testing environments, varieties Arapahoe, Art, Fuller, Infinity CL, Millennium, SD03164-2, SD05118-1, SD05W018, SD05W030, SD07056, SD07126, and SD07165had numerically high ^{2} (≥0.90), indicating that these varieties had similar stability with Lyman under the SD testing environments and thus they can be considered adapted to SD environments like Lyman (Table 3).

The estimate slope $_1$of a variety in FW analysis measures response sensitivity to the environmental indexes for this variety. If b_1 is close to 1, it suggests that this genotype has similar response sensitivity to environmental conditions compared to all other varieties. If b_1 is less than 1, it suggests that this genotype has less response sensitivity to these testing environmental conditions compared to all other varieties. If b_1 is greater than 1, it suggests that this genotype is more responsive to these environmental conditions on average. In the same principle, the estimate slope $_1$ in our check-based stability analysis measures response sensitivity to environmental conditions compared to the check used. Based on the FW method, varietiesExpedition, Fuller, SD05118-1, SD05W030, and Smoky Hill had more sensitivity to these environmental conditions compared other varieties in the trials while the varieties Hatcher, Millennium, and Wahoo showed less sensitivity to these environmental conditions (Table 4). Stability analysis based on the check Lyman showed that the varieties Arapahoe, Darrell, Harding,Hatcher, Hawken, Infinity CL, Wahoo, and Wesley had less sensitivity to these environmental conditions than Lyman (Table 4). The remaining varieties were not significantly different from Lyman regarding environmental response sensitivity.

With confidence intervals provided in Tables 3 and 4, parameters of interest like the coefficients of determination and slopes can be compared between any two varieties. For example, varieties Alice, Camelot, Darrell, Hatcher, Jagalene, Jerry, Radiant, SD06069, SD06158, Smoky Hill, Striker, Wahoo, Wendy, and Wesley had lower ^{2} than the variety Arapahoe, indicating these varieties were less stable under SD environments than the variety Arapahoe when Lyman was used as a check to evaluate yield stability in this study.

Many other pairwise comparisons the coefficients of determination and slopes could be conducted based on the confidence intervals provided in Tables 3 and 4.

On summary, the check based stability analysis method proposed in this study provides an option to investigate yield stability compared to a widely grown variety on the market. It could provide more useful information for farmers or breeders to choose which varieties to be used. This method can be used to evaluate yield stability in national variety tests, where several widely acceptable variety checks are used for comparison. In addition, with the resampling technique applied, yield stability parameters can be compared between varieties. The function has been integrated to our R package, qgtools, and will be available soon for public use.

References

Crossa, J., Gauch, H. G., & Zobel, R. W. (1990). Additive Main Effects and Multiplicative Interaction Analysis of Two International Maize Cultivar Trials. *Crop Science, 30*(3), 493-500.

Eberhart, S. A., & Russell, W. A. (1966). Stability parameters for comparing verirties. *Crop Science, 6*, 36-40.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Statist., 7*, 1-26.

Fan, X. M., Kang, M. S., Chen, H. M., Zhang, Y. D., Tan, J., & Xu, C. X. (2007). Yield stability of maize hybrids evaluated in multi-environment trials in Yunnan, China. *Agronomy Journal, 99*(1), 220-228.

Finlay, K. W., & Wilkinson, G. N. (1963). The analysis of adaptation in a plant-breeding programme. *Aust. J. Agric. Res., 14*, 742-752.

Francis, T. R., & Kannenberg, L. W. (1978). Yield stability studies in short season maize. I. A descriptive method for grouping genotypes. *Can. J. Plant Sci., 58*, 1029-1034.

Gray, E. (1982). Genotype × environment interactions and stability analysis for forage yield of orchardgrass clones. *Crop Science, 22*, 19-23.

Kang, K. S., & Mangari, R. (1995). Stable: A basic program for calculating stability and yield-stability statistics. . *Agronomy Journal, 87*, 276-277.

Kang, M. S. (1993). Simultaneous selection for yield and stability: Consequences for growers. *Agronomy Journal, 85*, 754-757.

Kang, M. S., & Miller, J. D. (1984). Genotype × environment interactions for cane and sugar yield and their implications in sugar breeding. *Crop Science, 24*, 435-440.

Lin, C. S., Binns, M. R., & Lefkovitch, P. L. (1986). Stability analysis: where do we stand? *Crop Science, 26*, 894-899.

Manly, B. F. J. (2006). *Randomization, bootstrap and Monte Carlo methods in biology.* (3 ed.). Boca Raton, FL: Chapman and Hall/CRC.

Perkins, J. M., & Jinks, J. L. (1968). Environmental and genotype-environmental components of variability. III. Multiple lines and crosses. *Heredity, 23*, 18.

Plaisted, R. L., & Peterson, L. C. (1959). A technique for evaluating the ability of selections to yield consistently in different locations or seasons. *Amer. Potato J., 36*, 381-385.

Shukla, C. K. (1972). Some statistical aspects of partitioning genotype-environmental components of variability. *Heredity, 29*, 237-245.

Wricke, G. (1962). Uber eine Methode zur Erfassung der okologischen Streubreite in Feldversuchen. *Z. Pflanzenzuchtg, 47*, 92-96.

Wu, J., Glover, K., & Berzonsky, W. (2012, April 29- May 01, 2012). *Statistical tests for stability analysis with resampling techniques.* Paper presented at the 25th Conference of Applied Statistics in Agriculture, Manhattan, KS.

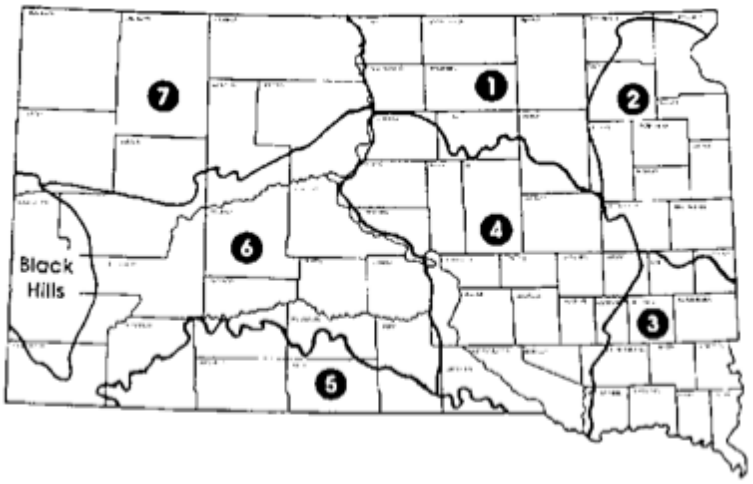

Figure 1. Crop adaption areas for South Dakota

Table 1.Number of cultivarsgrown in each site during three years.

Site	No. of genotypes†	Site	No of genotypes
Brookings	84	Wall	54
Brookings.IMS†	84	Winner	30
Hayes	54	Winner.IMS	30
Kennebec	54	Beresford	54
Onida	84	Bison	24
Pierre	84	Selby.IMS	54
Platte	84	Sturgis	24
Selby	84		

†: IMS=Intensive management study

Table 2.The number of environments where each of 38 winter wheat varieties being planted.

Variety	No. of testing environments†	Variety	No. of testing environments
Alice	32	Overland	32
AP503CL2	11	Radiant	32
Arapahoe	32	SD03164-2	11
Art	32	SD05118-1	32
Boomer	21	SD05W018	11
Camelot	21	SD05W030	8
Darrell	32	SD06069	19
Expedition	32	SD06158	19
Fuller	32	SD06163	11
Harding	32	SD07056	8
Hatcher	32	SD07126	8
Hawken	32	SD07165	8
Infinity CL	11	Settler CL	32
Jagalene	32	Smoky Hill	32
Jerry	32	Striker	32
Lyman	32	Wahoo	32
Millennium	32	Wendy	32
NuDakota	11	Wesley	32

†: 32indicates a variety grown in all three years, 21 indicates a variety grown in 2010 and 2011, 19indicates a variety grown in 2009 and 2011, 11 indicates a variety grown in 2009 only, and 8 indicates a variety grown in 2011 only.

Table 3.Estimated coefficients of determination with their corresponding 95% confidence intervals by two methods for each winter wheat variety.

	FW-based			Lyman-based		
	Estimate	LL	UL	Estimate	LL	UL
Alice	0.913	0.877	0.946	0.856	0.797	0.912
AP503CL2	0.931	0.881	0.977	0.876	0.723	0.962
Arapahoe	0.965	0.946	0.981	0.952	0.928	0.972
Art	0.920	0.880	0.956	0.907	0.862	0.943
Boomer	0.941	0.915	0.966	0.874	0.804	0.938
Camelot	0.936	0.904	0.964	0.828	0.735	0.913
Darrell	0.940	0.902	0.970	0.860	0.789	0.919
Expedition	0.950	0.923	0.974	0.874	0.781	0.942
Fuller	0.958	0.945	0.971	0.904	0.855	0.946
Harding	0.931	0.893	0.966	0.852	0.759	0.931
Hatcher	0.902	0.848	0.946	0.852	0.764	0.925
Hawken	0.919	0.885	0.951	0.867	0.786	0.930
Infinity CL	0.949	0.907	0.985	0.912	0.825	0.973
Jagalene	0.913	0.853	0.952	0.811	0.703	0.895
Jerry	0.886	0.830	0.935	0.829	0.726	0.910
Millennium	0.961	0.945	0.976	0.934	0.903	0.964
NuDakota	0.955	0.895	0.994	0.886	0.790	0.957
Overland	0.946	0.920	0.969	0.851	0.741	0.929
Radiant	0.734	0.625	0.837	0.557	0.377	0.723
SD03164-2	0.901	0.807	0.979	0.906	0.827	0.972
SD05118-1	0.955	0.933	0.972	0.910	0.869	0.950
SD05W018	0.910	0.838	0.981	0.902	0.836	0.964
SD05W030	0.996	0.992	0.999	0.952	0.887	0.994
SD06069	0.891	0.825	0.948	0.771	0.637	0.886
SD06158	0.929	0.883	0.968	0.781	0.646	0.894
SD06163	0.909	0.813	0.975	0.838	0.725	0.940
SD07056	0.953	0.899	0.989	0.947	0.882	0.993
SD07126	0.985	0.972	0.997	0.965	0.920	0.993
SD07165	0.973	0.947	0.993	0.955	0.900	0.994
Settler CL	0.948	0.922	0.970	0.855	0.742	0.935
Smoky Hill	0.928	0.884	0.968	0.841	0.759	0.909
Striker	0.911	0.878	0.941	0.785	0.670	0.882
Wahoo	0.896	0.829	0.951	0.778	0.664	0.872
Wendy	0.913	0.869	0.951	0.842	0.750	0.912
Wesley	0.898	0.850	0.942	0.842	0.774	0.905

Table 4.Estimated slopes with their corresponding 95% confidence intervals by two methods for each winter wheat variety.

	FW-based			Lyman-based		
	Estimate	LL	UL	Estimate	LL	UL
Alice	0.987	0.878	1.095	0.911	0.773	1.048
AP503CL2	1.020	0.834	1.190	0.867	0.653	1.076
Arapahoe	0.962	0.899	1.026	0.914	0.839	0.994
Art	1.059	0.951	1.173	1.010	0.899	1.125
Boomer	1.117	0.991	1.244	1.073	0.899	1.252
Camelot	0.976	0.860	1.089	0.906	0.726	1.080
Darrell	0.951	0.868	1.034	0.869	0.732	0.999
Expedition	1.095	1.005	1.184	1.008	0.878	1.152
Fuller	1.146	1.053	1.240	1.067	0.946	1.193
Harding	0.936	0.842	1.028	0.849	0.722	0.971
Hatcher	0.833	0.734	0.932	0.771	0.659	0.888
Hawken	0.907	0.812	1.005	0.842	0.735	0.952
Infinity CL	0.955	0.815	1.097	0.826	0.661	0.989
Jagalene	0.969	0.866	1.077	0.872	0.726	1.012
Jerry	0.930	0.817	1.046	0.860	0.721	0.994
Millennium	0.895	0.825	0.958	0.845	0.759	0.923
NuDakota	1.136	0.976	1.296	0.973	0.763	1.193
Overland	0.962	0.880	1.043	0.871	0.744	1.004
Radiant	0.916	0.718	1.110	0.766	0.547	1.006
SD03164-2	1.094	0.857	1.345	0.973	0.759	1.178
SD05118-1	1.185	1.096	1.281	1.105	0.973	1.228
SD05W018	1.006	0.793	1.207	0.881	0.691	1.076
SD05W030	1.110	1.062	1.172	1.067	0.873	1.255
SD06069	1.008	0.847	1.180	0.845	0.633	1.063
SD06158	1.058	0.915	1.196	0.871	0.658	1.095
SD06163	0.987	0.799	1.182	0.830	0.596	1.052
SD07056	0.968	0.801	1.123	0.951	0.783	1.120
SD07126	1.043	0.950	1.141	1.017	0.864	1.180
SD07165	0.973	0.841	1.099	0.942	0.785	1.112
Settler CL	1.048	0.958	1.141	0.946	0.805	1.087
Smoky Hill	1.167	1.053	1.285	1.059	0.898	1.222
Striker	1.043	0.930	1.158	0.928	0.758	1.097
Wahoo	0.876	0.765	0.976	0.778	0.636	0.915
Wendy	1.046	0.932	1.171	0.957	0.806	1.100
Wesley	0.904	0.803	1.002	0.833	0.710	0.961