

## Is the SEP Really as Good as it Looks?

**Summary** The standard error of prediction (SEP) is the standard deviation (SD) of differences between Near-infrared Spectroscopy (NIRS) predictions and the associated reference data. Because the results of NIRS testing are usually compared to those of reference testing on the same samples, the SEP may be strongly influenced by the precision of the reference test methods as indicated by the standard error of a single test (SET).

The standard deviation (SD) is an expression of the variability, or variance in data. It is the square root of the variance. The true meaning of the SD value itself is often overlooked. Only about 67% of the total population will fall within one SD above and one SD below the mean. Moreover, 95% of the population will fall between the limits of 1.96 times the SD above and below the mean and 98% of the population will have a value between 2.56 times the SD above and below the mean. The values of the mean  $\pm 1.96$  times the SD are referred to as the 95% confidence limits. The 95 % confidence limits are a more practical indication of the variance of a population than is the SD. Table 1 gives

examples of the influence of the magnitude of the SD on the values for protein and Falling Number (FN), two commonly tested parameters in wheat, that fall within the SD, the 95% and the 98% confidence intervals.

The range, estimated here as 3 times the SD, within which all samples will fall is surprisingly large, even at the low SDs (standard error per test) of 0.15% for protein and 20 seconds for the FN. Remember also that 5% of the population will fall outside of the 95% confidence limits. These are not “outliers”. They are part of the population and that is reflected in the SD.

**Table 1. Implication of the standard deviation**

Parameter	67% Mean $\pm$ SD		95% Mean $\pm$ 1.96*SD		98% Mean $\pm$ 2.56*SD		100% 3 x SD
	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Lower Limit	Upper Limit	Est. Range
<b>Protein</b>							
Mean 12.9; SD 0.15%	12.75	13.05	12.61	13.19	12.52	13.28	0.90 %
Mean 12.9; SD 0.25%	12.65	13.15	12.41	13.39	12.26	13.54	1.50 %
<b>Falling Number</b>							
Mean 350; SD 20 (sec)	330	370	311	389	299	401	120 sec
Mean 350; SD 25 (sec)	325	375	301	399	286	414	150 sec
Mean 350; SD 30 (sec)	320	380	291	409	273	427	180 sec

The SEP, also denoted as  $SD_{DIFF}$ , is the standard deviation of the differences between the NIR-predicted and the reference values for a set of samples. It is subject to the same implications as the SD. Table 2 shows the influence on the SEP of the standard error of the reference test where the standard error of the NIRS test is held constant. The standard error of a test (SET) is the standard deviation of the results of repeated tests on the same sample, including the sampling and all sample preparation steps. The examples are realistic for protein

testing by Kjeldahl or Dumas (both of which are equal in precision to the NIRS method), and for Falling Number testing. Notice that in this example, the means and maximum and minimum values were fairly constant in the three cases, although there the differences in the reference testing SET were marked. These figures illustrate the importance of knowing and maximizing the precision of reference testing before attempting the development of a NIRS calibration model.

**Table 2. Influence of reference error on SEP of NIRS predictions**

Statistic	Protein content			
	SET NIRS	Case 1 SET Ref.	Case 2 SET Ref.	Case 2 SET Ref.
SET*	0.116	0.116	0.151	0.210
Mean	13.66	13.66	13.66	13.66
SD	1.22	1.26	1.29	1.31
SEP ( $SD_{Diff}$ )	-	0.150	0.238	0.301
Max	16.0	16.1	16.1	16.2
Min	11.6	11.5	11.4	11.4
	Falling Number			
SET	16.0 sec.	20 sec.**	25 sec.	30 sec.
Mean	371	371	372	372
SD	20.2	27.3	33.2	39.6
SEP ( $SD_{Diff}$ )	-	30.1	37.6	45.1
Max	405	425	435	445
Min	340	315	310	305
SET = standard error of test method; ** SET of 16 seconds would be unrealistic for the Falling Number test.				

## Author

Dr. Phil Williams, PDK Grain  
PDK Projects, Inc., Winnipeg MB and Nanaimo B.C. Canada

**Reprinted with permission from Cereal Chemistry Division Newsletter, December 2005, Royal Australian Chemistry Institute**

Copyright © PDK Projects, Inc. 2006

PDK Projects, Inc. is incorporated in Manitoba and registered in B.C.

Business Office: 5072 Vista View Crescent, Nanaimo B.C. Canada  
Tel: 250-751-8890 Fax: 250-751-8225  
Laboratory: #2 – 12 McGillivray Place, Winnipeg MB R3T 1N4  
Email: philwilliams@pdkgrain.com  
[www.pdkprojects.com](http://www.pdkprojects.com)