Installing, Configuring, and Using Polytec's PSS-S-HOP Harvester Online Prediction Software

I. Introduction to Polytec's PSS-S-HOP Harvester Online Prediction Software 1
II. Understanding PSS-S-HOP 2
III. Implementing PSS-S-HOP 5
IV. Setting up the initial PC-to-Polytec NIR Spectrometer Communication 5
V. Installing PSS-S-HOP 6
VI. Configuring PSS-S-HOP Settings 12
VII. The Configuration "Data Input" Tab 13
VIII. The Configuration "Spectrometer" Tab 16
IX. The Configuration "Reference / Filter" Tab 18
X. The Configuration "Data Output" Tab 25
XI. The Configuration "Options" Tab 26
XII. The Configuration "Log- / Errorfile" Tab 28
XIII. The Configuration "About" Tab 29
XIV. Check Sample Testing 30
XV. Recommended Best Practices 33
XVI. Demonstrating PSS-S-HOP 38
XVII. Viewing and Understanding PSS-S-HOP Data and Results 43
XVIII. Routine Operation in the Harvester 46
XIX. PSS-S-HOP Operator Troubleshooting 47
XX. For more information 48

I. Introduction to Polytec's PSS-S-HOP Harvester Online Prediction Software

1) Overview

On-harvester NIR is used by plant breeders around the world to develop better varieties faster. Compared to lab NIR testing, on-harvester NIR measures traits faster, measures the entire plot instead of small subsamples, and measures every one of the thousands of plots in large trials. On-harvester NIR is used on harvesters to measure moisture, protein, oil, starch, sugar, and other traits in wheat, barley, oilseeds, silage, sugar beets and other crops.

Polytec's PSS-S-HOP Harvester Online Prediction software has been specially developed for small plot plant breeding trials. For every individual test plot, tens or hundreds of spectra of subsamples are measured in a few seconds. Sampling errors and instrument failures are automatically detected, and poor measurements of subsamples are automatically rejected. Before the final results are calculated, all the acceptable spectra are averaged so that results accurately measure the average for the entire plot.

Plots that are unusual compared to those in the calibration database are automatically and immediately identified as outliers so that operators and breeders can save the sample or take other steps. That way, users can improve their NIR calibrations and learn more about the varieties and effects of growing conditions.

Finally, all the results are stored in multiple, easily accessible formats. First, PSS-S-HOP immediately sends final calculated average results to the harvester software. Second, text logfiles are saved with detailed results for every plot. And third, PSS-S-HOP saves all the NIR spectra.

These instructions show new users every step of a complete installation and configuration. It uses an example calibration for protein and moisture in wheat, but it clearly shows users how the configuration can be modified for their specific purpose. Finally, it includes a complete demonstration of typical operation.

2) Who should read this document?

This document strives to cover "everything" about using PSS-S-HOP:

(a) Anticipating how NIR data will be most easily and accurately measured, evaluated, and saved
(b) Configuring the NIR, NIR-harvester interface, and PSS-S-HOP software, and
(c) Using PSS-S-HOP to operate the on-board NIR during harvest.

Therefore, this document has information for (a) NIR Technical Specialists (b) Harvester Equipment Specialists and (c) Harvester Operators, but different stakeholders will focus on different sections.

3) Current PSS-S-HOP version

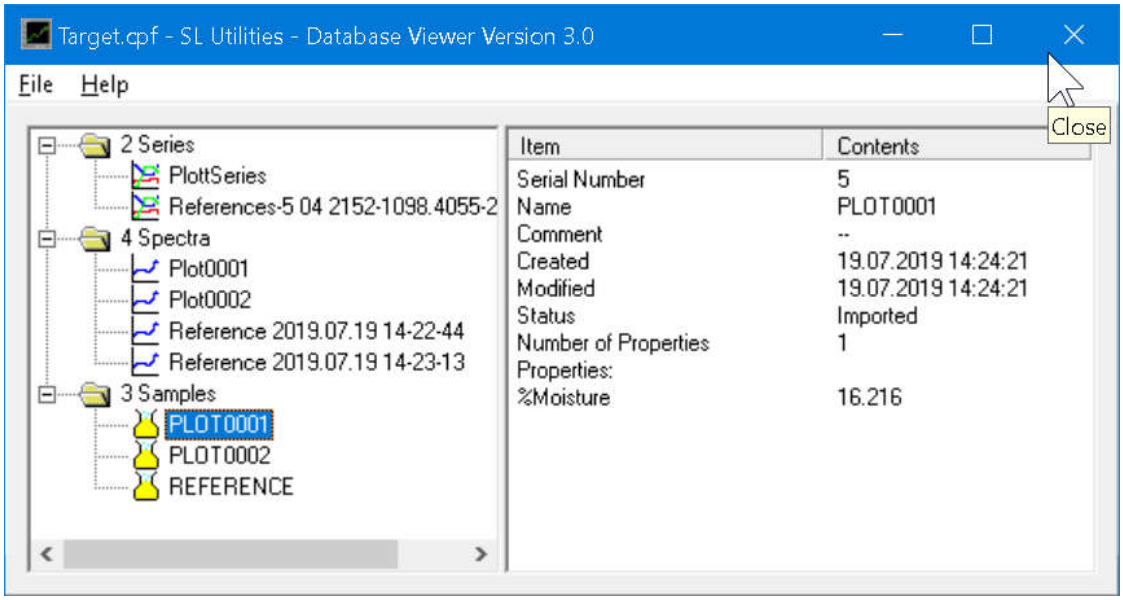
This document is based on using the version of PSS-S-HOP current in August 2019, ver 3.1.2.112.

II. Understanding PSS-S-HOP

To configure PSS-S-HOP, it is first important to understand the data that is read, measured and saved. PSS-S-HOP reads Plot IDs and a trigger signal to name, start and stop NIR measurements. PSS-S-HOP measures NIR spectra and calculates results for the test plots. Data and results are saved in 3 ways:

1) When PSS-S-HOP is controlled by harvester software like Mirus, Easy Harvest, or Harvest Manager, plot information and NIR results are exchanged between the harvester software and PSS-S-HOP for every plot
   - The Plot ID, NIR calibration name, NIR data filenames, and user defined "NIR Series" are all written in a "Logistic File" by the harvester software and read by PSS-S-HOP before every plot measurement
   - NIR results are written in an "Export file" by PSS-S-HOP and read by the harvester software after every plot measurement
2) All results including plot IDs, time stamps, calculated results and outlier statistics are tabulated in a *.CSV Results text file by PSS-S-HOP.
3) NIR spectra are saved in *.CPF databases as shown in the Figure below.
   - Each NIR *.CPF database is separated into "Spectra," "Samples," and "Series." Double-clicking will display Spectra and Series.
   - "Spectra" and "Samples" are named by the Plot ID.
   - "Samples" contain the measured result, for example, protein, oil, sugar, moisture, etc.
   - "Series" are user-named sets of spectra to organized data, for example "Eastern Field" and "Western Field" or "Site1" and "Site2" or "2019" and "2018".

Figure 1: The data stored by PSS-S-HOP in a *.CPF database includes Samples that include measured results, Spectra, and Series of Spectra



4) Three main *.CPF databases are saved

- The "Target" database stores the single averaged spectrum for each test plot after weak spectra and spectra that do not match the calibration are rejected. Spectra are rejected using "Subsample Filtering" and the "Similarity Check" as configured on the "Reference / Filter" tab
- The "Archive" database stores every individual subsample spectrum for each test plot, including spectra that are rejected from the average spectrum by "Subsample Filtering" and the "Similarity Check". Archive spectra are named with increasing increments. The "Archive" database shows the complete distribution of data in a plot and allow plots to be recalculated with new Filtering and Similarity thresholds.
- The "Reference" database is named with the spectrometer serial number and stores every individual Reference spectra. The types of Reference spectra and their purpose are as follows:
  - "Check Samples" are quick, routine tests by the Harvester Operator to confirm the NIR system works. The "Check Sample" can be as simple as an empty sensor or specific as a single stable material with a spectrum similar to test samples that can be used to track measurements over time. "Check Samples" are discussed in Section **XXX**.
  - "Dark Spectra" are references of perfectly black NIR and are routinely and automatically measured. They are flat noisy baseline spectra that remain below ~4000 counts should and be very consistent over a year
  - "References" are reference spectra of perfectly white NIR and are also routinely and automatically measured. They are shaped like a hill and have a peak of ~50,000 counts. The peak intensity should stay within ~10,000 counts over the 2,000 hour life of a lamp. Except for peak intensity changes, the shape of white references should appear consistent indefinitely.
  - The "Ur Reference" is the original white reference. New white references are automatically

compared to the "Ur Reference" to assure system performance. Except for peak intensity changes, a noticeable change in the shape suggests an instrument failure. The "Ur Reference" is discussed more in Section XXX.

- *Always Remember* - Only open *.CPF projects used by PSS-S-HOP when the software is closed. Do not view, change, or in any way open a *.CPF project used by PSS-S-HOP when also using PSS-S-HOP.

5) Operating the PSS-S-HOP Software Manually or Automatically

Plot IDs are read and measurements triggered either manually by harvester operators or automatically by the harvester software, for example, Wintersteiger's EasyHarvest or HarvestMaster's Mirus used by Zurn and others.

For manual measurements, harvester operators type Plot IDs or automatically name plots by date/time stamp. For automated measurements, the harvester software writes a text "Logistic" file read by PSS-S-HOP. For EasyHarvest and Mirus, the file extension must be *.US and the file format is:

ID_REC=Plot0001
ID_SERIES=EasternFields
CALNAME=Wheat Moisture 10-10
REF=JA
CPFNAME=2019Harvest
CSVNAME=2019Harvest

Where:

"ID_REC" is the Plot ID naming both the measured spectra and results
"ID_SERIES" is a user-named set of spectra used to organize data
"CALNAME" is the application stored in the PSS-S-HOP "Source Project" containing calibrations that is used to generate NIR results.
"REF" is "JA" when a new NIR reference should be measured and is blank when a reference should not be measured.
"CPFNAME" is the name of the *.CPF NIR database file where NIR spectra will be saved
"CSVNAME" is the name of the *.CSV result file where NIR results will be saved

The "Logistic" files for automated measurements using Harvest Manager, Bitzer, ProLeit, H2K and WUR plot management software are the *.PZL and *.WUR file formats and are specified in the PSS-S-HOP manual.

PSS-S-HOP writes the results of automated measurements made by the different plot management software in "PSSResults.txt," a two-line semicolon delimited text file listing the parameters in the table below.

Table 1: The NIR measurement parameters exported in the PSSResults.txt file

Column	Name	Description
1	Timestamp	Date and Time
2	ID_REC	The Plot ID naming both measured spectra and results. PlotIDs are limited to 30 characters. PlotIDs over 26 characters lead to complicated names for subsample spectra.
3	ID_SERIES	A user-named set of spectra to organize data, for example, the crop, location,

		or harvest year
4	CALNAME	The application stored in the PSS-S-HOP Source Project containing calibrations that is used to generate NIR results.
5	APPLICATIONNAME	The same as above.
6	Property_1_Name	The name of the measured property used the calibration file, e.g., Moisture, Protein, Oil, etc.
7	Property_1_Value	The value of the measured property in the calibrated units, e.g., % (w/w), etc.
8	Property_1_U	The uncertainty in the measured property value in the calibrated units, e.g., % (w/w), etc.
9	Property_1_H	The leverage outlier statistic describing whether the measured spectrum is accurately represented by the calibration data. The units are analogous to units of a standard deviation.
10	Property_1_S	The spectral reconstruction outlier statistic describing whether the measured spectrum contains features that are not in the calibration data. The units are analogous to units of a standard deviation.
11-15	Results for Property_2	
16-20	Results for Property_3	
...	...	
	NumSpec	The total number of measured spectra
	NumSpecOK	The total number of measured spectra that pass similarity tests and are not filtered as specified in the PSS-S-HOP configuration.
	Used%	The percentage of measured spectra used to calculate final NIR results.
	SpectrometerSerialNo	The Spectrometer Serial Number used for the measurements.

III. Implementing PSS-S-HOP

Using PSS-S-HOP in the field is designed to be simple and nearly automatic for non-experts. However, a clear understanding of the installation and configuration includes several additional topics:

- Setting up the initial PC-to-NIR communication
- Installing and configuring PSS-S-HOP, including automated testing for sampling problems and filtering weak spectra
- Making routine measurements, including routine instrument tests with "Check Samples"

IV. Setting up the initial PC-to-Polytec NIR Spectrometer Communication

1) Follow the sensor manual to set the supply voltage to 12.0 V one time after installation. This setting compensates for the exact length of the sensor electrical cable.

2) The computer controlling the spectrometer with the PSS-S-HOP Software must have an Ethernet adapter. If an internal adapter is not available, then a good alternative is any USB2.0-Ethernet adapter or better like the Linksys Model USB3GIG on Amazon for ~ $10.

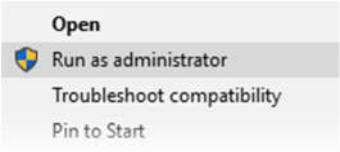
3) The easiest way to setup and test the PC-to-Polytec NIR Spectrometer communication is using the simple, basic, free PAS-DRV software. Follow the instructions "Installing and Running the Polytec PAS-DRV Demo Software." That procedure also explains how to view and understand typical NIR data using a free *.CPF

NIR Database Viewer. Installing the *.CPF NIR Database Viewer is important to view PSS-S-HOP data.

V. Installing PSS-S-HOP

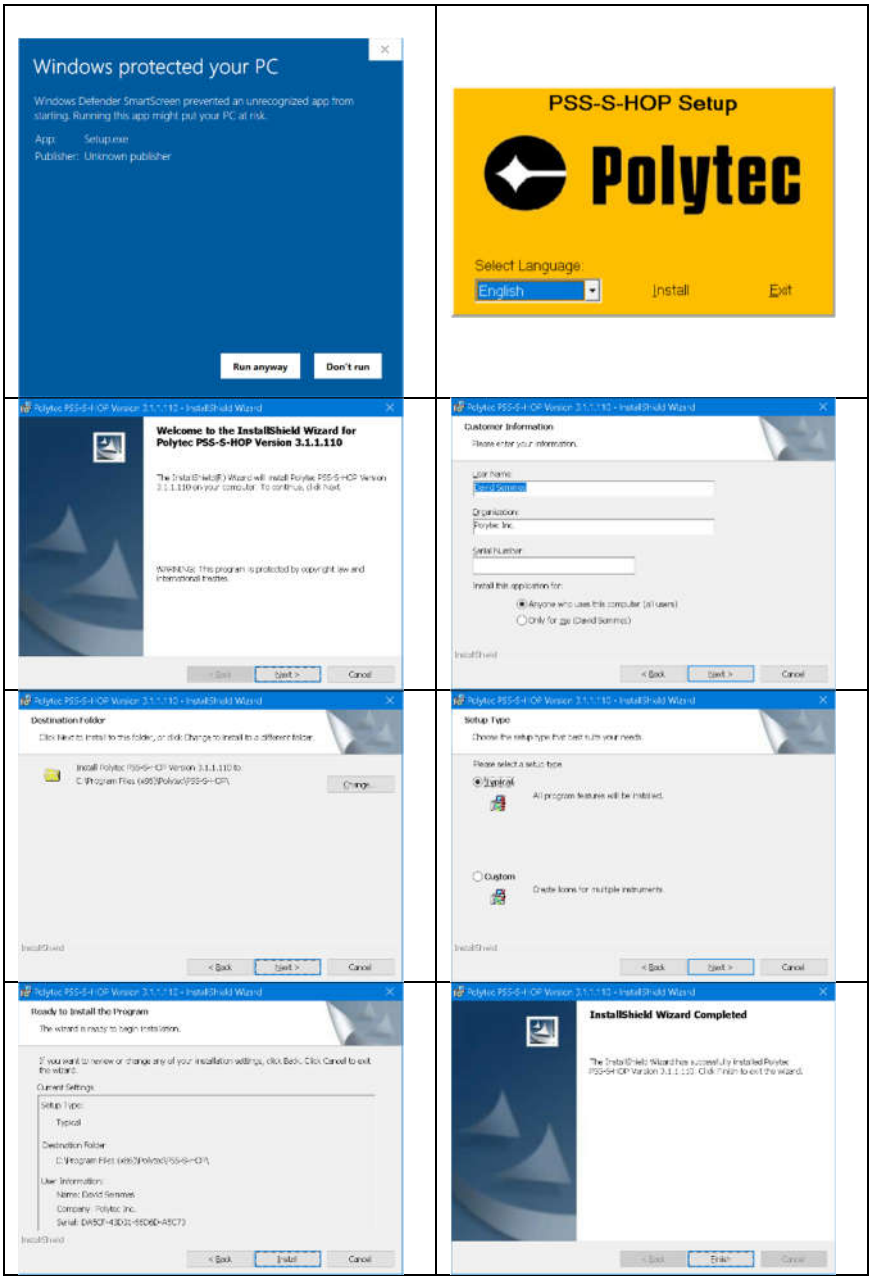
1) Insure that spectra can be measured and *.CPF files can be viewed before installing PSS-S-HOP.
2) Logon to the PC as an Administrator and save these files in C:\Polytec\NIR\Polytec Setup Files\:
   - "PSS-S-HOP 3.1.1.110 Setup.zip" from https://my.pcloud.com/publink/show?code=XZXugG7ZOI6IwHf17QuTQU54dgbf4XFXrzCX .
   - "PSS-S-HOP 2.1.2.100 Setup.zip" from https://my.pcloud.com/publink/show?code=XZpsol7ZpXrGTSruQBkAWPdo9O0DG4rpOElV .
   - "PSS-S-HOP-Manual .pdf" from https://my.pcloud.com/publink/show?code=XZUyuG7ZBY4Ce8gtuCV6hxRwC6y8M4Qtsn7k.
   - For PSS-212X spectrometers
     - The *.CPF NIR Database calibration file for PSS-212X spectrometers "Wheat Moisture 10-10 PSS-212X.cpf" from https://my.pcloud.com/publink/show?code=XZnlq37Zqq7OJF9UTXX0UOyDQ0TRrmrorDpX.
     - The PSS-S-HOP *.CFG configuration file "Wheat Moisture 10-10 PSS-212X.cfg" from https://my.pcloud.com/.
   - For PSS-172X spectrometers
     - The *.CPF NIR Database calibration file "Wheat Moisture 10-10 PSS-172X.cpf" from https://my.pcloud.com.
     - The PSS-S-HOP *.CFG configuration file for PSS-172X spectrometers "Wheat Moisture 10-10 PSS-172X.cfg" from https://my.pcloud.com/ .
   - The instructions "Using Polytec TeamViewer Quick Support" at https://my.pcloud.com/.
   - The latest version of these instructions "Installing, Configuring, and Operating Polytec's PSS-S-HOP Harvester Online Prediction Software" at https://my.pcloud.com/ .
3) Uninstall any old versions in Windows Control Panel's Programs and Features.
4) Unzip "PSS-S-HOP 3.1.1.110 Setup.zip". Use an elevated administrator prompt to execute “Setup.exe.” That is, select it, right click and select “Run as administrator”.

Figure 2: Right click and select “Run as administrator”



5) The PSS-S-HOP installation has the standard steps shown in the figures below. The PSS-S-HOP serial number from the CD case is required. The only unusual step is selecting a "Typical" or "Custom" Setup. "Custom" is for twin harvesters with two spectrometers.

Figure 3: PSS-S-HOP has a standard Windows installation, and requires the license serial number.



6) After rebooting, Control Panel Programs and Features will include "Polytec PSS-S-HOP Version 3.1.2.112" published by Polytec GmbH with 32.2 MB.

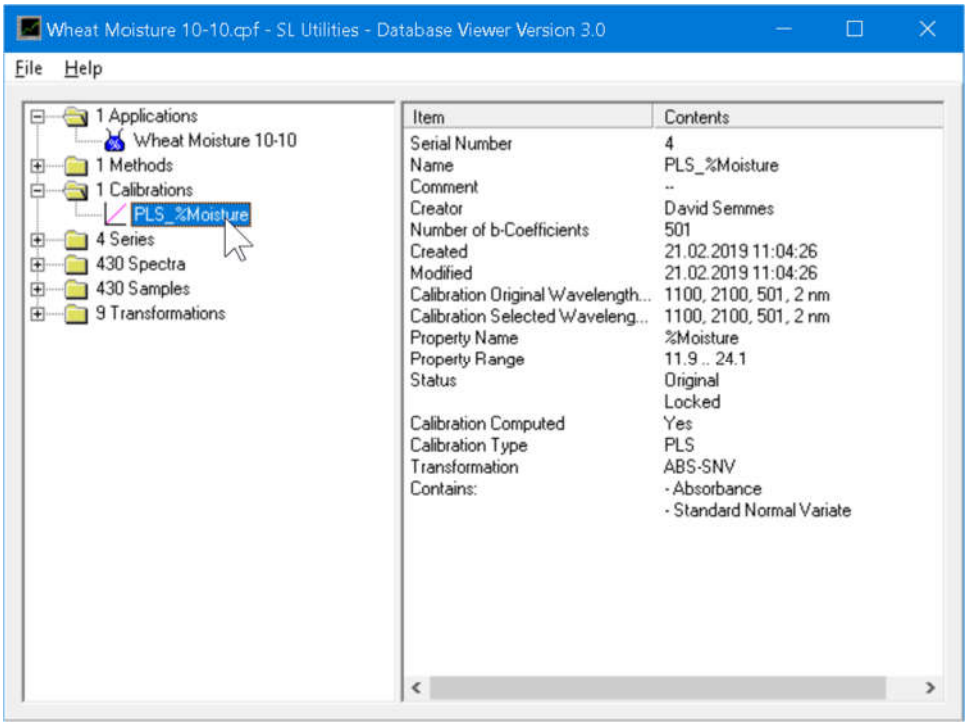
7) Create C:\Polytec\NIR\PSS-S-HOP Data\.

8) Copy the example *.CPF NIR Database file "Wheat Moisture 10-10.cpf" calibration file saved in Section IV.1 from C:\Polytec\NIR\Polytec Setup Files\ to C:\Polytec\NIR\PSS-S-HOP Data\. Choose the PSS-172X or PSS-212X version depending on your spectrometer.

   Open "Wheat Moisture 10-10.cpf" using the free SL Database Viewer and expand the folders for "1 Application" and "1 Calibrations" as in the figure below. You will see the application "Wheat Moisture

10-10" and the Calibration "PLS_%Moisture".

Figure 4: The file "Wheat Moisture 10-10 PSS-212X.cpf" displayed in the SL Database Viewer



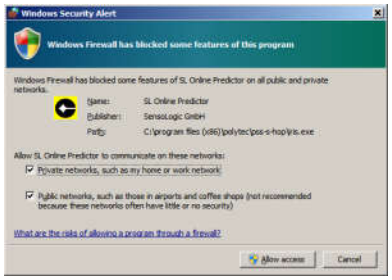
9) Copy "Wheat Moisture 10-10 PSS-212X.cfg" or "Wheat Moisture 10-10 PSS-172X.cfg" saved in Section IV.1 from C:\Polytec\NIR\Polytec Setup Files\ to C:\Polytec\NIR\PSS-S-HOP Data\. This is an example PSS-S-HOP *.CFG *configuration* file with default settings for the typical configuration described in these instructions. Again, choose the version appropriate for your spectrometer.

10) ***Before*** starting PSS-S-HOP the first time, be prepared to enable and allow the Ethernet network connection the first time PSS-S-HOP connects with the NIR spectrometer. ***Before running PSS-S-HOP, confirm*** successful use of the simple demo software as described in "Installing and Running the Polytec PAS DRV Demo Software". Also, ***note the Integration time*** to measure ~50,000 counts in the reference spectrum.

11) Use an elevated administrator prompt to start PSS-S-HOP the first time. That is, select it, right click and select “Run as administrator”.

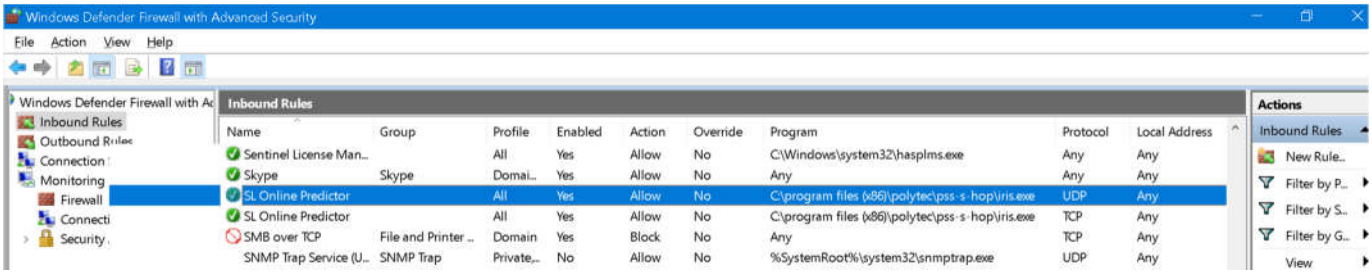
- After starting PSS-S-HOP, allow access to "SL Online Predictor" and "IRIS.exe" on all networks thru firewalls as in the figure below.

Figure 5: ALLOW network access to "SL Online Predictor" and "IRIS.exe" thru Firewalls.



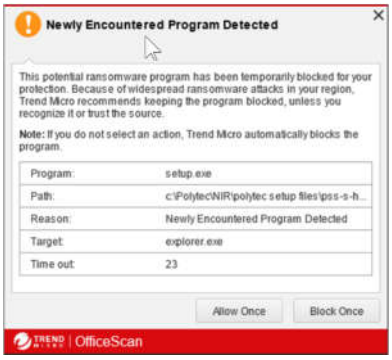
- If necessary, manually edit the Firewall Inbound and Outbound Rules to allow TCP and UDP protocols for "SL Online Predictor" and "IRIS.exe" in C:\Program Files (x86)\Polytec\PSS-S-HOP\ as in the Figure below.

Figure 6: The required rules in "Advanced settings" in "Windows Defender Firewall" in Windows10



- Allow connection thru other security software as in the figure below after starting PSS-S-HOP the first time.

Figure 7: Allow connection thru other security software if prompted

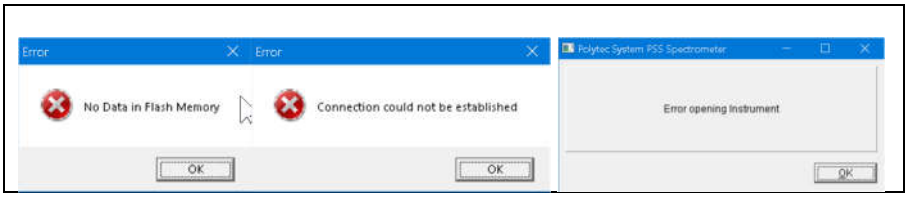


12) The message, "Please wait while connecting to Instrument…" is displayed when PSS-S-HOP tries to connect to the NIR, whether it can connect or not. It may be displayed very fast if it connects.

Figure 8: The message, "Please wait while connecting to Instrument…" …" is displayed when trying to connect

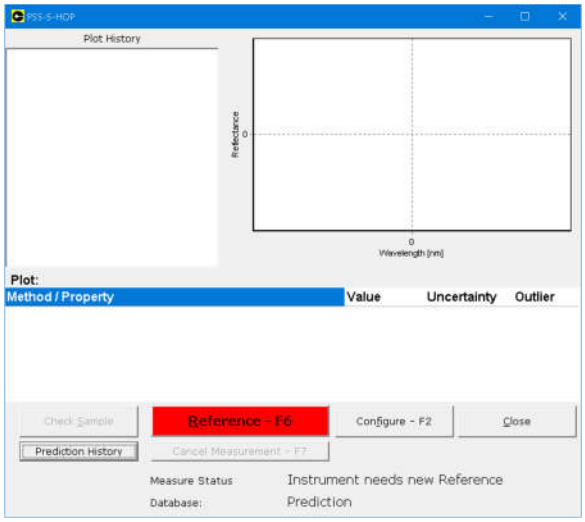
13) If PSS-S-HOP is not able to make an Ethernet connection, the messages, "No Data in Flash Memory," "Connection could not be established" and "Error opening Instrument" may displayed. The cause of these errors could be the EtherNet cable, the IP address, the Firewall, or security software. These problems should be resolved using the simpler, free demo in the PAS-DRV software. Follow the instructions "Installing and Running the Polytec PAS-DRV Demo Software" to establish a connection.

Figure 9: Software Errors caused by errors networking with the NIR



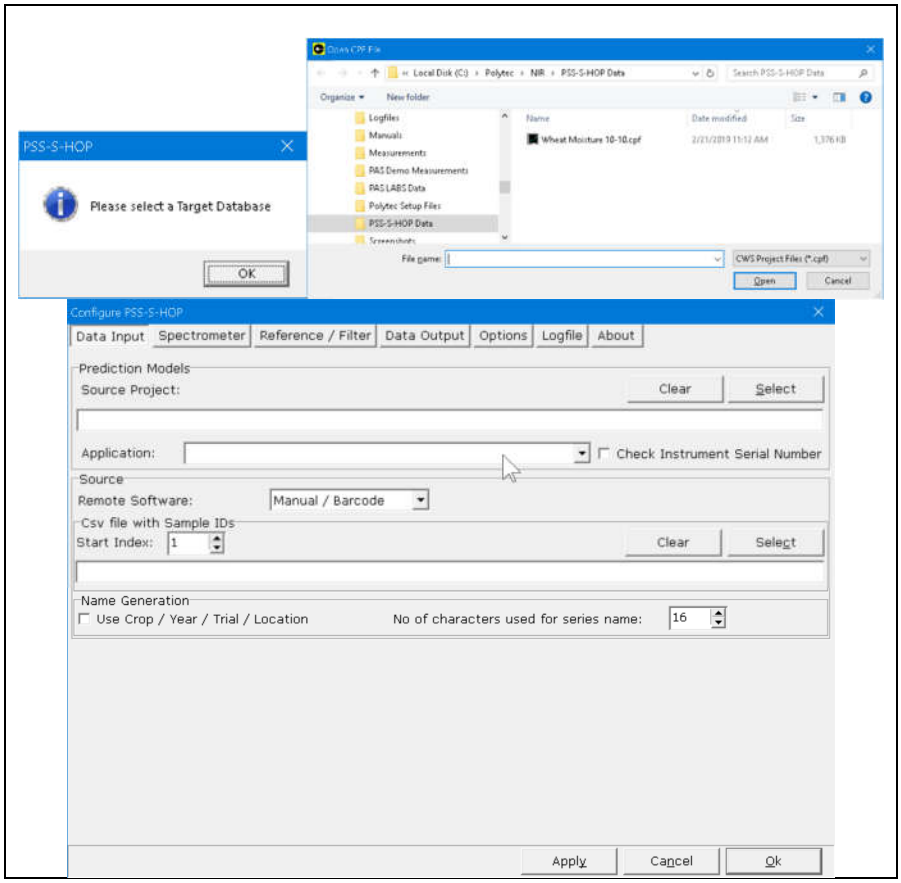
14) The PSS-S-HOP operating window shown in the figure below will be displayed without the error messages when the software connects the PC to the NIR.

Figure 10: The standard PSS-S-HOP operating window



15) Setup an initial configuration by clicking "Configure" and using the password "HOP". You will be prompted "Please select a Target Database" and a window allowing you to browse to files will open. Click "cancel" and the standard PSS-S-HOP configuration window shown in the figure below will be displayed.

Figure 11: The standard PSS-S-HOP configuration window



16) Upload an example configuration suitable for your system by clicking on the "Options" tab. In the "Configuration Files" section, click "Import Configuration". You will then be prompted to load a *.cfg configuration file in Windows Explorer. Select the version of "Wheat Moisture 10-10.cfg" for your spectrometer from C:\Polytec\NIR\PSS-S-HOP Data\. You will get the message, "Configuration will be overwritten. Continue?" When you allow the configuration to be overwritten, you will again see the message, "Please wait while connecting to Instrument…" while the new configuration is saved.

Figure 12: The "Options" tab and the "Import Configuration" button under "Configuration Files

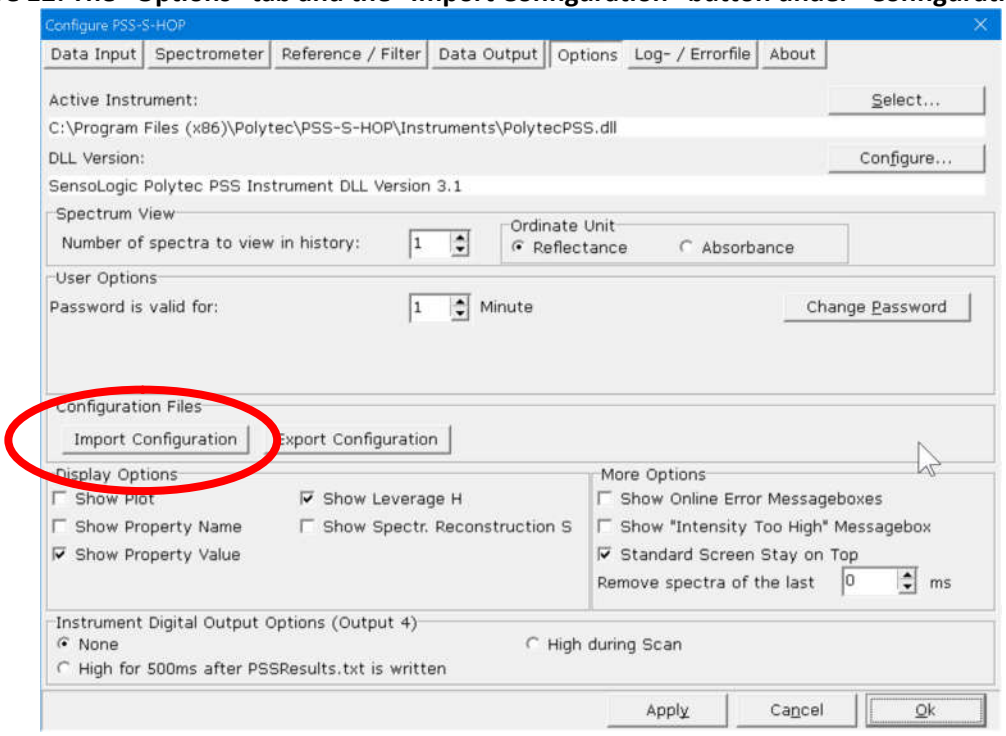
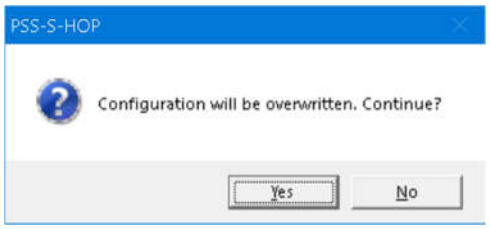


Figure 13: Users are warned before importing PSS-S-HOP configurations. Export configurations as backup.



17) Once communication between the PSS-S-HOP and the NIR has established, *do not edit the PSS-S-HOP settings without maintaining PC communication with the NIR*. For example, do not disconnect the NIR, edit the configuration, and then reconnect the NIR. If the synchronization between the spectrometer and PC is lost, it may be necessary to reset all PSS-S-HOP settings and recreate or reload a working configuration.

VI. Configuring PSS-S-HOP Settings

The purpose of the PSS-S-HOP configuration is to plan and specify settings for:

1) Measurements including for example the plots names, calibration, type of harvester software communication, number of scans, etc.
2) Automated testing of system performance to check the system performance and prevent errors. For example, the configuration includes threshold settings for the long term consistency of the reference spectra, and threshold settings for "bad" results that do not match the calibration, for example, for trash covering the sensor
3) Automated data processing, for example, threshold settings for filtering weak, noisy spectra before

averaging subsamples to generate final results

4) Organizing output data

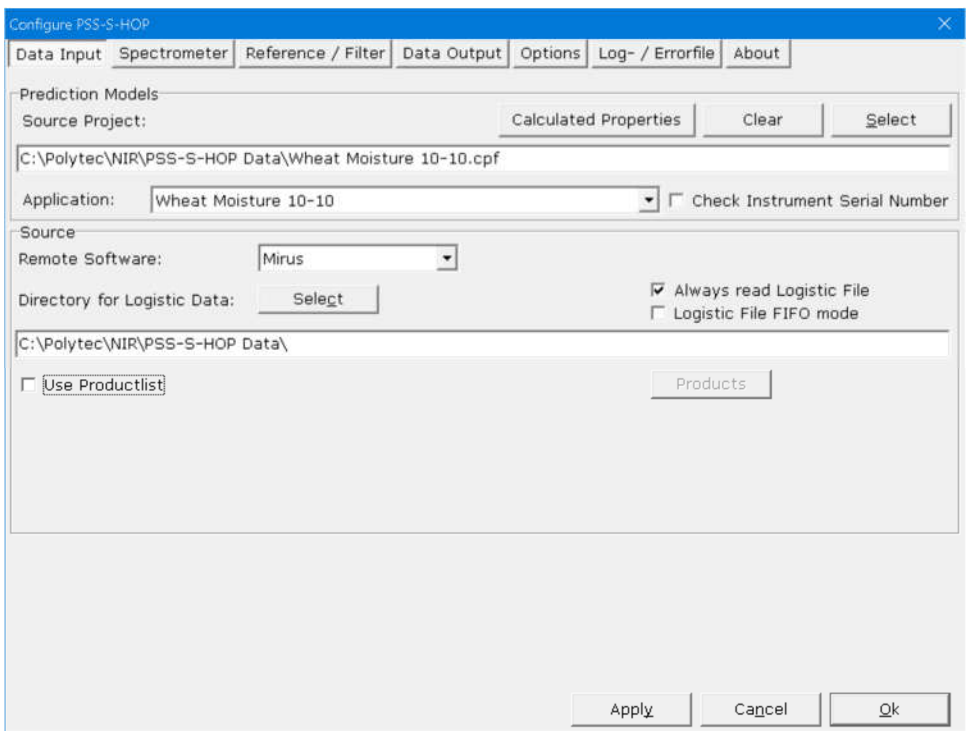
It is important to understand that the settings for automated testing of system performance and for automated data processing must be user judgments that reflect how carefully the user will control the system and how much knowledge and experience have users with their data. Initially, users are likely to avoid interruptions during harvest and therefore choose to test performance weakly and only reject clearly poor data. The more carefully users monitor performance and understand the reason for changes in field data, then the more carefully users will be able to control measurements and improve results.

The configuration settings are organized into the 7 tabs described below, "Data Input," "Spectrometer," "Reference / Filter," "Data Output," "Options," "Log- / Errorfile," and "About." The figures below show the settings in each tab of the example PSS-S-HOP configuration. Key features of each setting are briefly identified. More details are in the manual.

VII. The Configuration "Data Input" Tab

The "Data Input" tab is used to specify the calibration and the sample information.

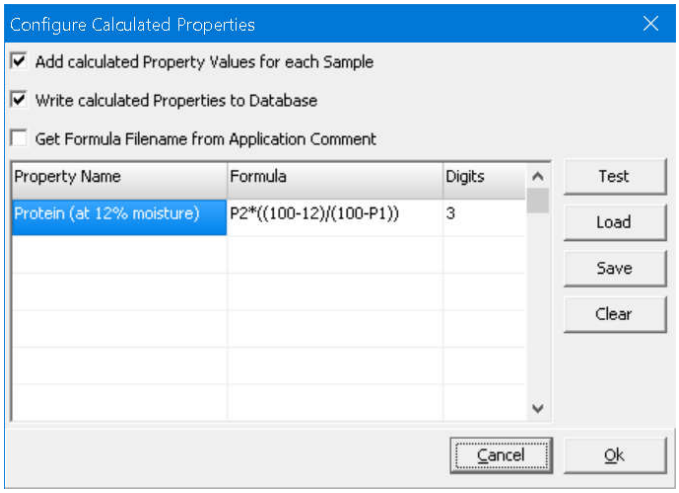
Figure 14: The "Data Input" tab of the PSS-S-HOP configuration window



1) **"Source Project"** and **"Application"** is the NIR calibration in the *.CPF NIR Database.

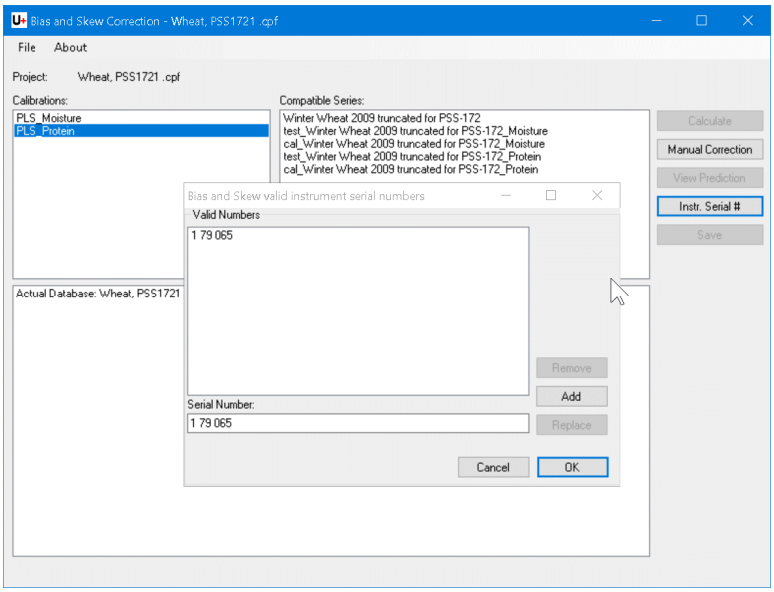
2) **"Calculated Properties"** allows users to include results calculated from measurement predictions, like protein with a specified constant moisture correction. For example, using a NIR calibration Application for "as-is" %protein and %moisture in wheat, "Calculated Properties" can be used to calculate the protein content on a standard 12% moisture basis with the formula "P2*((100-12)/(100-P1))" where P1 and P2 are the first and second NIR predictions, %Moisture and %Protein.

Figure 15: Use "Calculated Properties" to Generate and Save Results like Protein on a Standard Moisture Basis



3) **"Check Instrument Serial Number"** allows calibrations to be restricted to specific spectrometers. The permitted instrument serial numbers are specified in the calibration. To specify the serial number in the calibration, edit the CPF file with the SL Utilities Bias & Skew Adjustment tool. The serial number *must* be specified in the format "1 67 600", like in the file name of the Reference Project, not in formats like "1676 00."

Figure 16: Restrict calibrations to specific spectrometers using the SL Utilities Bias & Skew Adjustment tool



- An instrument without a permitted serial numbers will give an error like "The selected application 'Winter Wheat" is not valid for the current instrument."

4) **"Source" / "Remote Software"** allows user to measure spectra manually or automatically under control from harvester software using several different options.

5) **"Directory for Logistic Data"** identifies the directory where the Harvester Software writes the "logistic" file with the PlotID that is then read by PSS-S-HOP reads as discussed in Section II.

6) **"Always read Logistic File," "Logistic File FIFO mode," and "Auto Reference at Start up"** -- After the harvester software names a Plot ID in PSS-S-HOP, but before the plot is measured, HOP can be set to allow the pending name and reference to be easily updated with "Always read Logistic File" or HOP can retain the name even if an additional, new PlotID is named with "Logistic File FIFO mode."

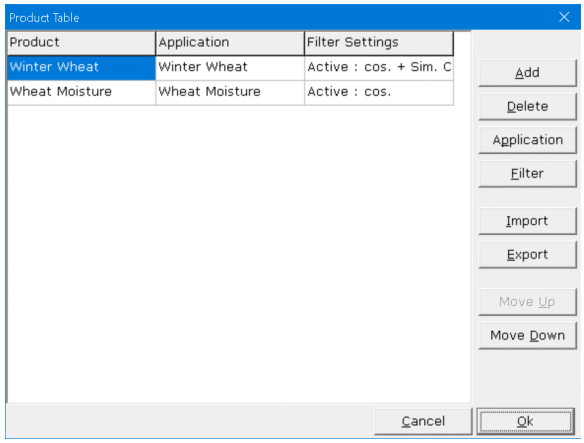
   Whichever options are selected, (1) *PlotIDs waiting to be triggered can be cancelled anytime* by clicking cancel and (2) *expired references can be manually remeasured anytime*.

7) **"Always read Logistic File"** -- Typically, plots are measured soon after naming, new names and references can be sent to HOP anytime, and the flexibility of renaming plots and measuring a new references is a benefit. In this case, select "Always read Logistic File". This option also always allows the harvester software to update an expired reference. In this case, PlotIDs that were previously read in PSS-S-HOP but not yet measured are ignored.

8) **"Logistic File FIFO mode"** -- Less typically, if the time between naming and starting a HOP measurement is long compared to the time between naming measurements, then users prefer to retain names and *prevent previous Plot IDs from being overwritten*. In this case, select "Logistic File FIFO mode." ("First-in, first-out"). However, FIFO mode may require an operator to manually remeasure expired references. Also, "Auto Reference at Start up" can be selected on the "Data Output" tab.

9) **"Use Productlist"** allows different products, calibrations, and spectral filtering to be controlled automatically by the harvester software. For example, two crops can be harvested in an intercropped field using Productlists as long as both calibrations are included in the Source *.CPF file. Similarly, fields with very different harvesting conditions can be harvested together using Productlists with different filtering conditions. When using Productlists, Logistic files must specify the Product ID like "Canola" or "Peas" instead of the Application Name, as normal.

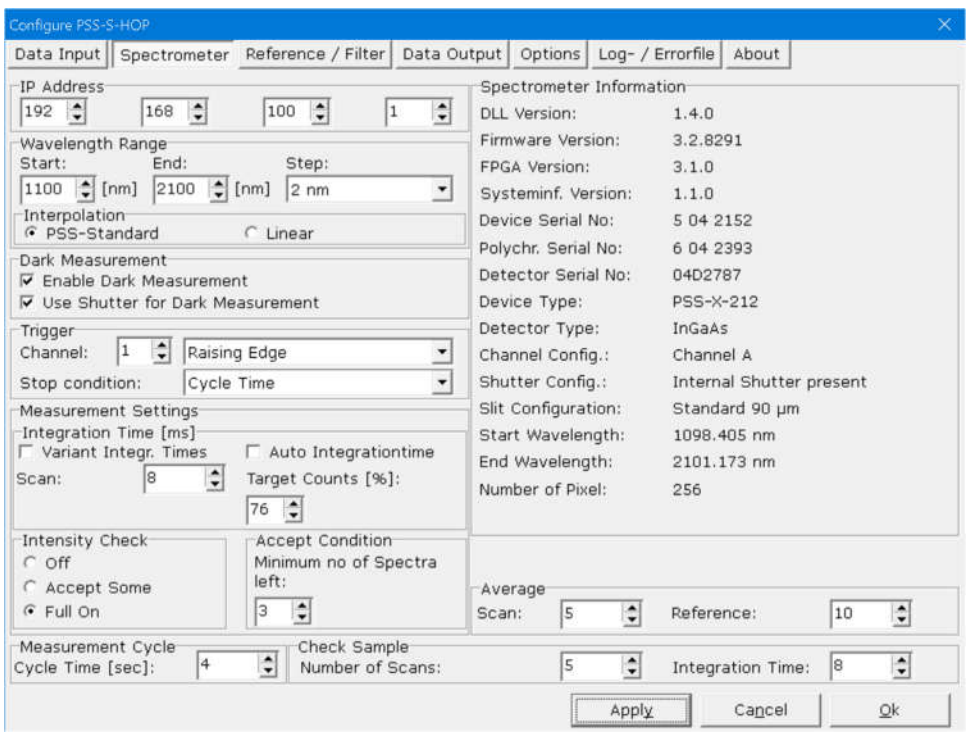
Figure 17: "Use Productlist" allows automatic control of different products or calibrations



VIII. The Configuration "Spectrometer" Tab

The "Spectrometer" tab is used to specify NIR settings like the number of measurement scans.

Figure 18: The "Spectrometer" tab of the PSS-S-HOP configuration window



1) **"IP Address"** is the static IP Address of the networked spectrometer. It can be changed, but is nearly always the default, 192.168.100.1. Confirm it and do not change it without clear understanding of the setting. When using Wintersteiger's EasyHarvest it is 192.168.50.30

2) **"Wavelength Range"** is nearly always 850-1650 nm for PSS-172 and 1100-2100 nm PSS-212 systems in 2nm steps. Confirm it and do not change it without clear understanding of the setting.

3) **"Interpolation=PSS-Standard"** is an important setting for instrument-to-instrument reproducibility. Confirm it and do not change it without clear understanding of the setting.

4) **"Enable Dark Measurement" / "Use Shutter for Dark Measurement"** is an important setting for typical operation. Confirm it and do not change it without clear understanding of the setting.

5) **"Trigger"** settings determine how a NIR Measurement starts and stops. Measurements start with a digital input and continues for either a Cycle Time or until the trigger is gone. The digital input usually comes from the Harvester software. It may also come from a button or sample detected by a photosensor.

   - The spectrometer manuals describe how two active-high standard TTL digital trigger inputs are used to start and stop measurements. More simply, connect PINs 1 and 2 on the 25 PIN TTL CONTROL Interface on the back of the spectrometer to trigger a sample measurement in PSS-S-HOP.

6) **"Integration Time" is a critical, system and application specific setting**. Important details about the integration time are listed below:

   - The integration time is specified in "Measurement Settings; Integration Time (ms); Scan"
   - The integration time is the measurement time for a single scan
   - The integration time should generally be set to the time that gives the best signal-to-noise without exceeding the detector range.
   - The detector range is determined by the 16 bit detector analogue-to-digital convertor and is 0-to-2^16 or 65,536.
   - For most samples, Polytec recommends integration times to give ~50,000 counts at the maximum for the internal white reference spectrum. 50,000 is chosen because it is large enough to optimize S/N, but small enough that it is unlikely to exceed the maximum detector range of 2^16 or 65,536.
   - The integration time that gives 50K counts for the internal white reference spectrum but is usually ~5-10 milliseconds for systems with a PSS-S-B01 contact sensor. That integration time is ~30-40 milliseconds for systems with a PSS-S-A03 distance sensor.
   - The integration time and maximum reference count rate is usually very consistent for a given system over time. For a single lamp used continuously for ~3 months, the maximum white reference intensity will typically change by <1000 counts.
   - For different lamps, the optimum integration time giving 50K white references may change by as much as ~30%. The best practice is for users to update the integration time when a new lamp is used.

7) **"Variant Integration Times"** allows users to use different integration times for samples and the internal reference. Samples that are unusually dark or shiny might be measured best with different integration times. Darker samples would require longer integration times to accurately measure the NIR peaks due to nutrients. Shiny samples might require shorter integration times to reduce high baselines from direct reflections.

   The best practice when working with a new type of sample is to make test measurements of realistic samples, on the harvester when possible. Work to measure high signal-to-noise NIR peaks without exceeding the maximum detector range by adjusting the number of scans, the Intensity Check/Accept condition, and the both integration times if appropriate. Polytec can guide customers for difficult or uncertain samples.

8) **"Auto Integration Time / Target Counts"** adjusts the integration time so that the white reference intensity meets a target, for example, 50K or 76% of the 65,536 maximum. This option allows users to automatically identify the integration time setting for new systems or systems with new lamps. It should not be used routinely because an automatically changing integration time might conceal a lamp or other hardware problem.

9) "**Intensity Check**" prevents errors caused by direct reflections in shiny samples, like moist glossy black seeds. Shiny samples sometimes lead to direct, intense reflections that exceed the maximum detector range of 65,536 counts. Intensities over the maximum detector range distort spectra and may cause inaccurate results. **"Intensity Check/Full On"** omits subsample spectra if one scan exceeded that range. **"Intensity Check/Accept Some"** allows you to retain spectra if the "**Minimum no of Spectra left** is above a limit. For example, you might retain a 5 scan subsample spectrum if only 1 or 2 individual scans exceeded the maximum detector range in an inhomogeneous shiny sample.

10) **"Measurement Cycle / Cycle Time"** is the alternative "Stop Condition" for the "Trigger" above. With the example settings of a 4 second cycle time, an 11 msec integration time and 5 scans per subsample, a maximum of 72 subsample could be measured. In fact, only 61 spectra were measured using one system since extra time is required for processing data.

11) **"Average / Scan and Reference"** specifies the number of individual scans to be averaged for a single subsample or reference spectrum. For many samples, averaging many scans to reduce noise is not necessary and only ~ 5 scans may be averaged for a single subsample spectrum. *Instead of averaging many scans to reduce noise, it may be better to average many subsamples to reduce inhomogeneity and filter unusual spectra*. Whether to average many scans or many subsamples depends on whether variation is largest from spectral noise or sample inhomogeneity. The best solution will depend on the samples and testing details. For example, the optimum balance of scans per subsample and number of measured subsamples will likely differ significantly between many kilos of shiny (reflective) chopped fruit pulp and a small plot of dry, matte black colored seeds.

    - The number of Reference scans is usually higher, ~100, since the Reference is homogeneous, only measured every hour or two and since 100 scans with a 10 msec integration time only takes 1 second.

12) **"Check Sample / Number of Scans and Integration Time"** specify measurement conditions for a "Check Sample." A "Check Sample" is a powerful tool for operators to easily check the system anytime.

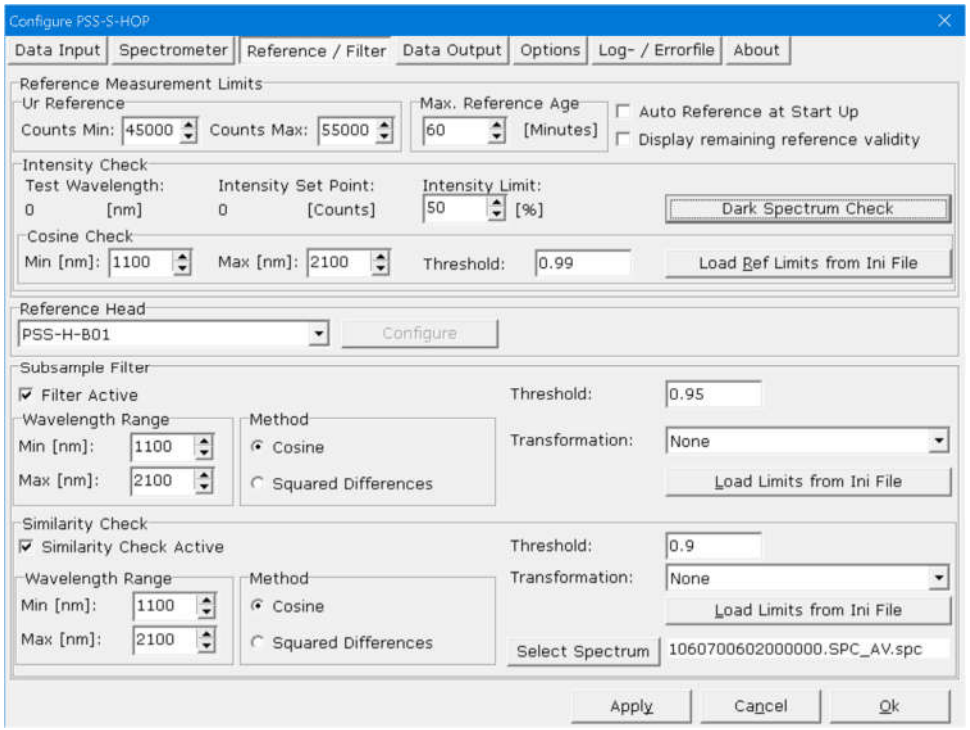
    A "Check Sample" is a test sample that is easy to measure a predictable characteristic result. Even the empty sensor could be a "Check Sample". A stable piece of plastic that reproducibly gives a characteristic spectrum similar to normal samples would be a better "Check Sample" since the consistency of the shapes of peaks could be monitored over time.

    The best practice is to measure the same Check Sample at the beginning and end of every day. See Section XXX, "PSS-S-HOP Operator Troubleshooting" for some details and examples.

IX. The Configuration "Reference / Filter" Tab

The "Reference / Filter" tab specifies settings to detect instrument and sampling errors

Figure 19: The "Reference / Filter" tab of the PSS-S-HOP configuration window



1) **"Ur Reference Counts Min / Counts Max"** are the limits allowed when the original "Ur" reference spectrum is measured. The intensity and shape of that original "Ur Reference" is then used to indefinitely track and judge system performance. *These limits only apply to that original "Ur" reference spectrum*.

   Depending on the integration time and other settings in the "Spectrometer" tab discussed in the previous section, the limits are often set so that the Ur reference will be near the ~50K count target. If the Ur Reference intensity is beyond these tight limits, you will get the error message, "Reference Scan failed. Intensity too low (or high)." You can modify the integration time until the Ur reference is successfully measured and stored. Also, you can use "Auto Integration Time" discussed previously or use the PAS Demo software to identify the integration time that give 50,000 counts at the maximum of the internal white reference. Once the "Ur Reference" is measured, the shape of measured references are monitored to automatically test system performance.

   To measure and use a new Ur Reference, change the Reference Project specified in Section XX. For example, rename the file "Ref_5 04 2152.cpf" where "5 04 2152" is the spectrometer serial number as " Ref_5 04 2152 YYYYMMDD.cpf," the Reference Project saved that date. A new Reference Project and Ur Reference will be automatically created when the next reference is measured.

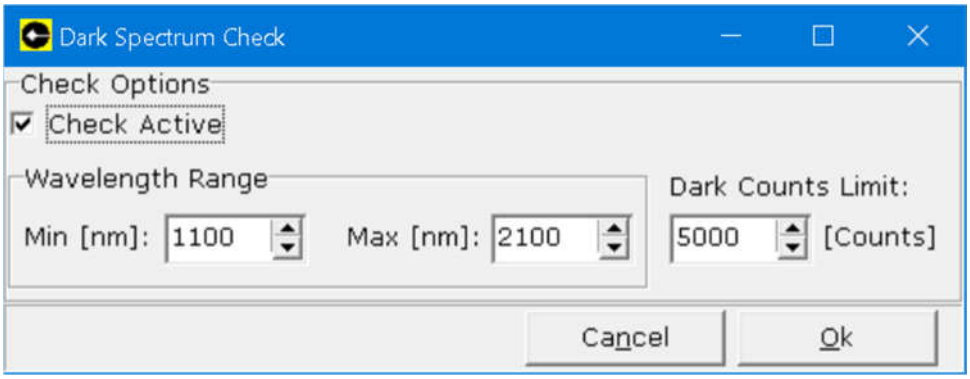
2) **"Max. Reference Age"** limits the time a reference is valid and is usually set to about 60 minutes. When using harvester software to control PSS-S-HOP, it is usually best to control the reference age with the harvester software and set this limit to a slightly longer time.

3) **"Auto Reference at Start Up"** -- when "Always Read Logistic File" is off and the Reference has expired, the harvester software cannot automatically run a Reference. In this case, users can manually Reference, or this option can be used to allow complete control of HOP by the harvester software.

4) **"Intensity Check"** tests whether white reference peak intensities fall below a fraction of the initial Ur reference. Usually, (1) results are insensitive to changes even over ~20%, (2) reference intensity decreases by <~10% during the lifetime of a lamp, and (3) lamps may differ by over ~20%. A default intensity limit that is unlikely to interrupt harvest unnecessarily is ~75%.

5) **"Dark Spectrum Check"** is an important test of a rare but possible failure. The dark spectrum depends on the integration time, but normally is below 5000 counts. A reasonable threshold is twice the maximum measured in historical data.

Figure 20: Normally the Dark Spectrum Check should be active. The limit depends on integration time, but 5000 is a good default value for typical setups.



6) **"Cosine Check," "Subsample filter"** and **"Similarity Check"** are all tests that use a match score threshold to test whether measured spectra are similar to a reference. In each case, the best way to determine a threshold is to use the SL Utilities software to calculate the distribution of match scores for spectra of known "good" samples and known "bad" samples of the risk you would like to detect. If the threshold is too tight, the tests may give "false alarms" during harvest. If the threshold is too loose, the test could miss inconsistencies that lead to decreased accuracy.

7) **"Cosine Check"** is a test of the shape of the white reference, and therefore a test of instrument consistency. Evaluate the threshold by comparing many reference spectra from a single system over a long period of time using multiple lamps. Potential failures that would be detected are dirt *in* the sensor, unusual humidity, and a dark shutter failure. For example, the first figure below shows data for multiple measurements over two harvests. The higher and lower spectra were at different integration times. Nearly 300 reference spectra were above 0.9997. Polytec recommends: (a) periodically check Reference Spectra match scores and setting the threshold to prevent any observed problems (b) purposely measure References on cold humid mornings, and setting the threshold accordingly (c) set the threshold low enough that you accept the risk of a failure interrupting testing. The second figure below shows the error "Reference Scan failed. Reference Identity Check failed." With that message, an operator could repeat the reference or lower the threshold. A typical threshold is 0.99 and failures are rare.

Figure 21: Harvest Data for two years gives cosine matches above 0.9997

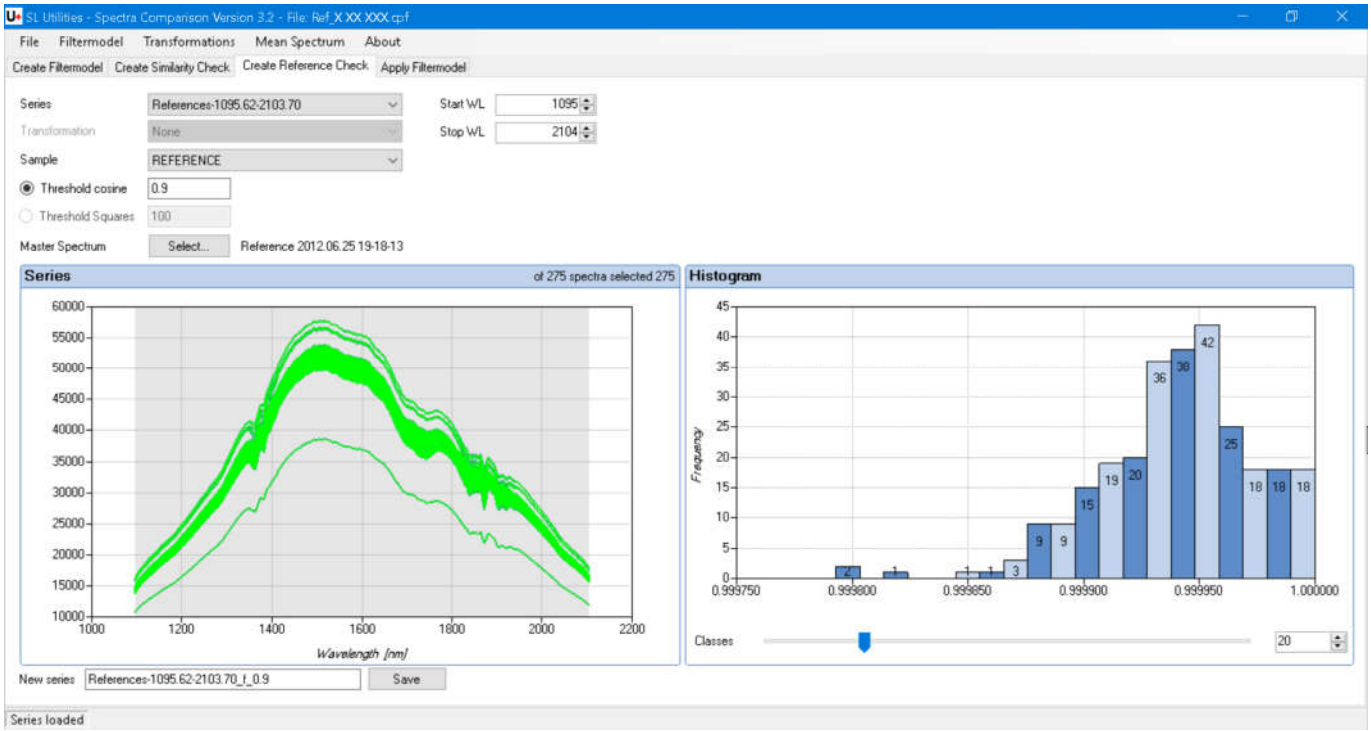
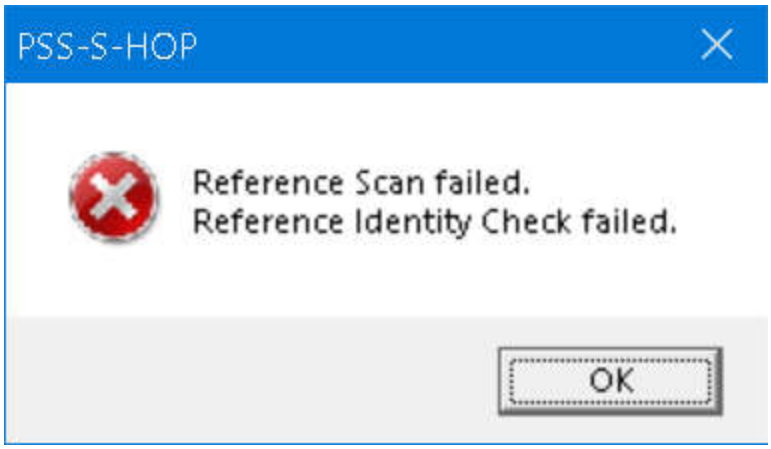


Figure 22: The Error Message for a Failed Reference Match Test



8) **"Subsample filter"** is an important feature that enables users to improve the accuracy of results by automatically eliminating weak spectra of a partly covered sensor. In some crops, it is also possible to eliminate spectra with chaff or soil.

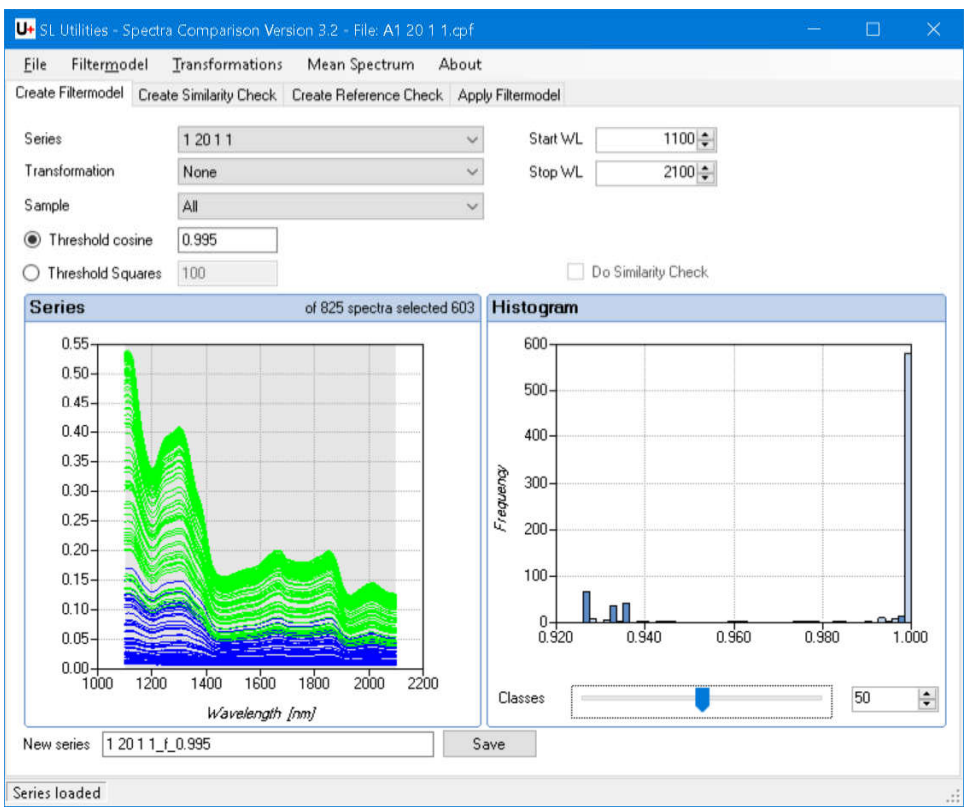
- A key setting for the "subsample filter" is the threshold used to distinguish samples that are acceptable or not.
- To choose a threshold, use the SL Utilities "Spectra Comparison & Filtering" and select the "Create Filtermodel" tab. This software is usually purchased from Polytec with the "Calibration Bundle" software.

- "Create Filtermodel" analyzes a series of inconsistent spectra. The simplest example would be a series measured while a sensor is covered with a single, small, homogeneous sample. A more interesting, relevant sample would be all the spectra of small subsamples from a field trial plot that varies in quality traits, chaff, and soil in addition to the coverage on the sensor.
- The figure below from "Create Filtermodel" shows how it works for the simple example of slowly pouring a small sample of wheat grain into a sample cup.

  The utility is used to choose a threshold for acceptable spectra for subsample predictions. The best threshold is selected by calculating the cosine or squared threshold match score of all the spectra with the spectrum that assumed to the best, the most intense spectrum. In the example, in this plot, the spectra in green have cosine match scores above 0.995 compared to the most intense spectrum in that set. In fact, 99 of the 165 "subsample" spectra are above 0.999. Most of the remaining spectra have cosine match score of ~0.926. According to this plot, any threshold between about 0.926 and 0.999 would distinguish "good" and "bad" spectra. Experienced users can identify the two types of spectra as spectra of wheat and of the empty sensor.

  Be certain to select "Sample" = "All" to view all the subsample spectra of all the samples at one time.

Figure 23: Using SL Utilities/"Spectra Comparison & Filtering"/"Create filtermodel" to choose a Subsample Filtering Threshold



- An additional option is to use a spectral transformation to improve the selectivity of the NIR to distinguish the spectra that are acceptable.
- It is important to interpret this data carefully. The match score does not directly reflect the accuracy

of the NIR results. Also, the plot in this examples compares the full wavelength spectra without transformations. The calibration is likely to use transformed spectra, possibly over a narrow spectral range.

- Also, this simple example shows a small, clean, homogeneous sample. A sample from a real field trial is likely to be ~100 times larger, to vary in quality traits like moisture and protein and match score, and to vary in chaff and soil. Just as a user might want to try a tighter threshold to eliminate spectra with chaff, it might also be necessary to loosen the threshold instead to include all the clean sample spectra with varying moisture content. It is inevitable that any threshold will only be a balance of the risk of the error of eliminating spectra that should be retained and the corresponding error of retaining spectra that should be eliminated.
- This plot is only a tool to attempt to "filter poor quality" spectra. Ultimately, the goal is for analysts to use tools like these to investigate why spectral measurements vary, how to eliminate spectral variation due to experimental error and how to analyze spectral variation reflecting the traits of interest. The more analysts succeed, then the more accurately they can learn about their varieties. Generally, Polytec recommends that NIR analysts use the Spectra Comparison & Filtering utility to analyze Archive spectra from multiple plots to try to identify spectral transformations and match score thresholds to filter irrelevant experimental variation from the variation of interest in plot measurements.

9) **"Similarity Check"** works nearly the same as Subsample Filtering, but has a different purpose. Both:

- Use a calculated Match Score to compare subsample spectra and a reference spectrum
- Use the Match Scores to test for "bad" spectra
- Identify "bad" spectra as those with a Match Score outside a user-set threshold
- Set the threshold based on Match Scores distributions of samples identified as "good" and "bad" by the user

The purpose of the two tests are different, however. Subsample Filtering uses the most intense subsample spectrum in each plot as the reference spectrum, and filters spectra of the empty sensor, small or weak samples, and hopefully samples contaminated with dirt or heavy chaff.

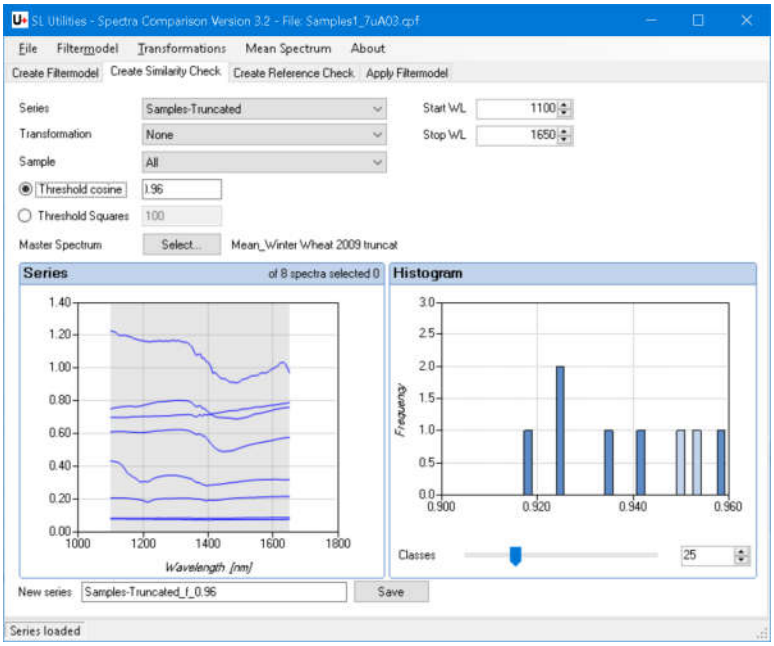
The Similarity Check instead uses the average spectrum from the calibration set as the reference spectrum. When PSS-S-HOP uses the most intense spectrum as the reference spectrum the Similarity Check is used to be certain that most intense Subsample spectrum matches a normal sample and not, for example, an empty sensor or trash stuck on the face of the sensor. More generally, the Similarity Check is used to set a threshold to reject any possible incorrect reference subsample.

Used together, the Similarity Check and Subsample Filtering improve results by making certain sample spectra (a) match the calibration and (b) are not weak, noisy, or contaminated.

- Set a threshold for Subsample Filtering to improve accuracy by eliminating spectra of the empty sensor and other inappropriate data. Use the SL Utility, "Spectra Comparison & Filtering" and the "Create Similarity Check" tab.
- Identify any alternative samples that might be measured in error and insure they are rejected by the threshold. Create a method with a threshold that will reject possible sampling errors like an empty sensor, paper or plastic debris, and leaves. The figure below shows match scores for the empty

sensor, rubber belts, leather belts, paper, plastic bags, rigid plastic, and leaves. All would be rejected by the Similarity Check.

Figure 24: Spectra of many materials have Match Scores below 0.96 compared to wheat grain

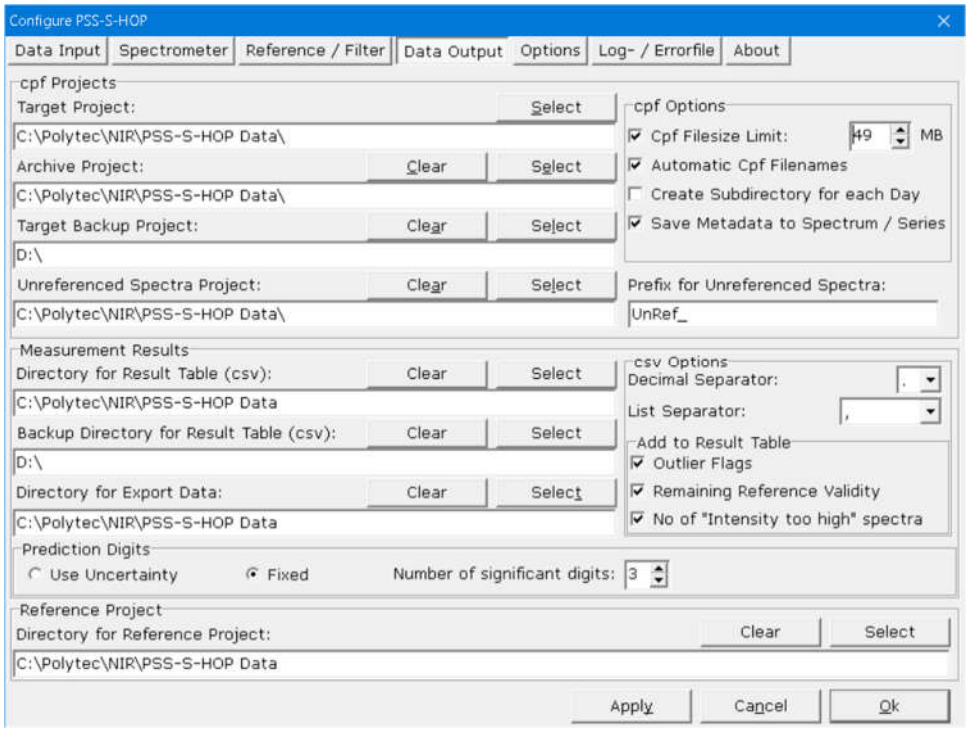


If the most intense subsample spectrum is rejected by the Similarity Check, the second most intense subsample spectrum will be tested, etc. A temporary sampling error caused, for example, by an empty sensor might then be corrected. However, if ***all*** the subsample spectra are rejected by the Similarity Check, for example by a sensor covered by tape or trash, then the error message, "Error -4" will be displayed like in the figure below.

Figure 25: The error messages displayed if *all* the subsample spectra are rejected by the Similarity Check

X. The Configuration "Data Output" Tab

Figure 26: The "Data Output" tab of the PSS-S-HOP configuration window specifies settings for saving data



1) The **"Target Project"** saves the final, averaged spectrum of all the subsample spectra for each plot that are not rejected by the "Subsample filter" and the "Similarity Check".

2) The **"Archive Project"** saves every individual subsample spectrum from every single plot, before testing similarity, filtering and averaging. Archive Projects can be reprocessed with different settings using the SL Utilities Software.

3) The **"Target Backup Project"** is a copy of the Target Project. Traditionally, one file might be saved on the PC hard drive and one on a USB drive that was removed and saved daily to a network. ***Using USB or Network drive is not recommended!*** Unreliable connections drive connections can lead to errors saving data and data loss. The default recommended practice is to copy all files daily for backup.

4) The **"Unreferenced Spectra Project"** stores unreferenced spectra of reflected light intensity. It shows when sample spectra are very weak and when they exceed the ~60,000 count detector limit. **When applications are developed, unreferenced project are used to adjust measurement conditions so that sample spectra are consistently below the maximum intensity.** After initial application development, the "Unreferenced Spectra Project" data is not usually saved.

5) The **"Result Table"** is a *.csv file listing complete results for every plot.

6) The **"Directory for Export Data"** is used for automatic testing controlled by the Harvester software. Typically it is the same as the **"Directory for Logistic Data"** specified on the "Data Input" tab.

7) The **"Reference Project"** saves all reference spectra. Data on reference changes over long times is helpful for understanding, monitoring, and controlling long-term instrument performance.

8) **"Prediction Digits"** allows users to choose whether the number of digits displayed in calculated results is

based on the precision of the calibration or is fixed.

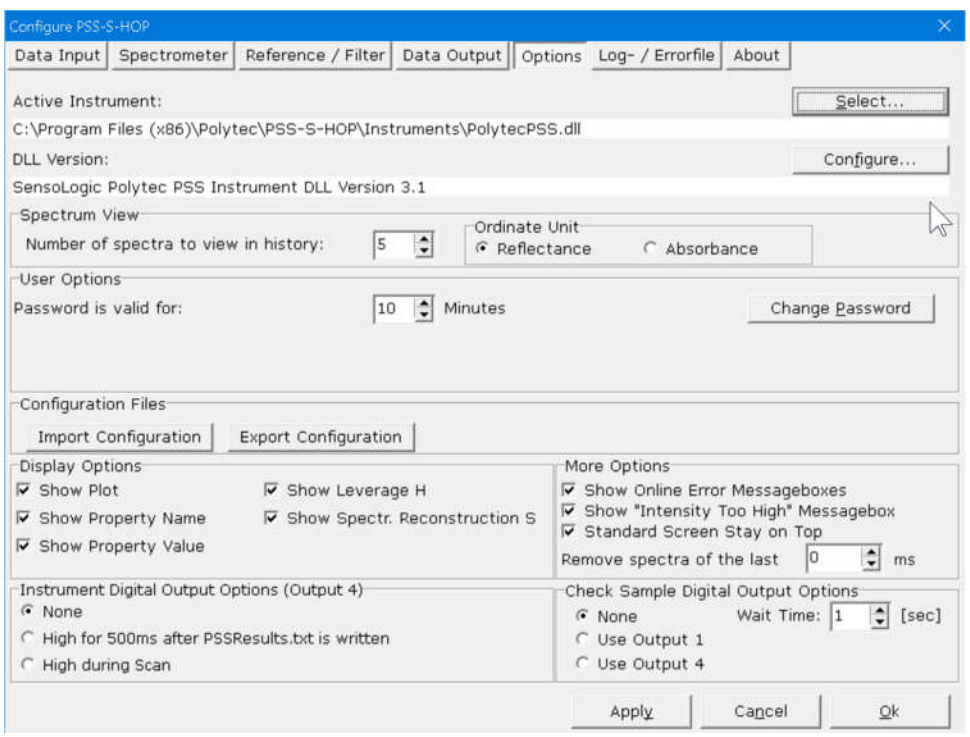
9) **"cpf Options / <u>Cpf Filesize Limit</u>"** should be small enough that graphics processing is not too slow in offline analysis. 49MB is a good default value.

10) **"cpf Options / <u>Automatic Cpf Filenames</u>"** -- Measured NIR spectra are stored in *.CPF database files. Those *.CPF files are named 3 ways:

- Manually named on the "Data Output" tab, for example as "Wheat Target.CPF," "Wheat Archive.CPF," "Wheat Backup.CPF," and "Wheat Unreferenced.CPF"
- Automatically named in manual operation or with no input. In that case, "Automatic Cpf Filenames" is checked on the "Data Input" tab and PSS-S-HOP is triggered without naming plots. *.CPF files are then named by date. The target and backup files are named with the syntax "YYYYMMDD.CPF." The archive and unreferenced projects are named the same with an "A" or "R" prefix, for example, "A20200307.cpf."
- Automatically named by Remote software like Mirus or EasyHarvest. Again, "Automatic Cpf Filenames" is checked on the "Data Input" tab. The Logistic File includes a line like "CPFNAME= Wheat Target". Again, the archive and unreferenced projects are named the same with an "A" or "R" prefix.

XI. The Configuration "Options" Tab

Figure 27: The "Options" tab of the PSS-S-HOP configuration window specifies display settings and allows configurations to be imported or exported



1) **"Number of Spectra"** -- Up to 10 spectra are displayed for comparison in an overlay plot. The last spectrum measured is always displayed in red and previous spectra are in green.

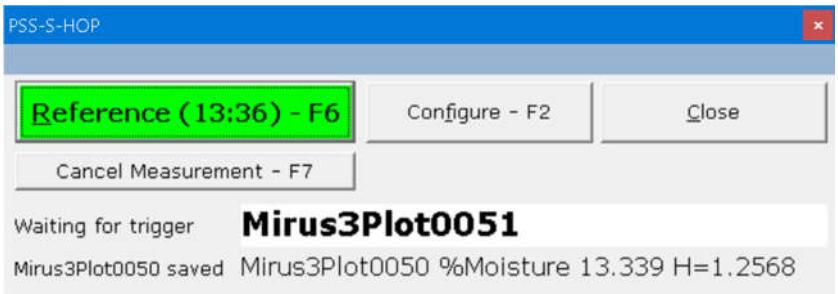
2) **"Reflectance / Absorbance"** - Reflectance and Absorbance about both common scales for viewing spectra that show different aspects of spectra. Reflectance normally varies between perfectly reflective materials giving ~1.0, and perfectly absorbing, dark materials giving ~0.0. On the other hand, Absorbance offsets show direct reflections from poor sample presentation, and values above ~1.5 often lead to less accurate results. *On either scale sample spectra should consistently be similar in offset, peak height, and shape.*

3) **"Password is valid for __ Minutes "** -- Set this to 0 minutes to disable the password restriction. The default password is "HOP".

4) **"Display Options/Show Plot/Show Property Name/Show Property Value/Show Leverage H/Show Spectr. Reconstruction S"** displays or removes these results in the smaller "Standard Screen" (F4).

5) **"More Options / Standard Screen Stay on Top"** keeps the smaller "Standard Screen" (F4) visible as in the figure below.

Figure 28: The smaller "Standard Screen" (F4) can be set to visible and display key results.



6) "More Options / Show Online Error Messageboxes / Show "Intensity Too High" Messagebox"

   - Online Error Messageboxes allows users to immediately interrupt operators in both manual and automatic operation with a message alerting them to a significant problems, like a low reference intensity from a broken bulb or fiber or a sample that does not match the expected calibration.

7) **"More Options / Remove spectra of the last X ms"** omits the last subsample spectra from the analysis. For example, if chaff typically rises to the top of a bin, and the subsample measurements are triggered off with a level sensor, then omitting the last spectra would omit chaff measurements.

8) **"Configuration Files / Import Configuration / Export Configuration"** allows all the settings in the PSS-S-HOP configuration to be saved in a small, single file both as a backup and for documentation. *ALWAYS INSURE YOUR CONFIGURATION IS SAFELY BACKED UP ANY TIME CHANGES ARE MADE.* Typically, name *.cfg configuration files with syntax like "YYYYMMDD David new thresholds.cfg". An ongoing text logfile listing changes would be the best practice.

9) **"Instrument Digital Output Options"** are sometimes used to provide a digital output signal indicating when measurements are made and when they are not.

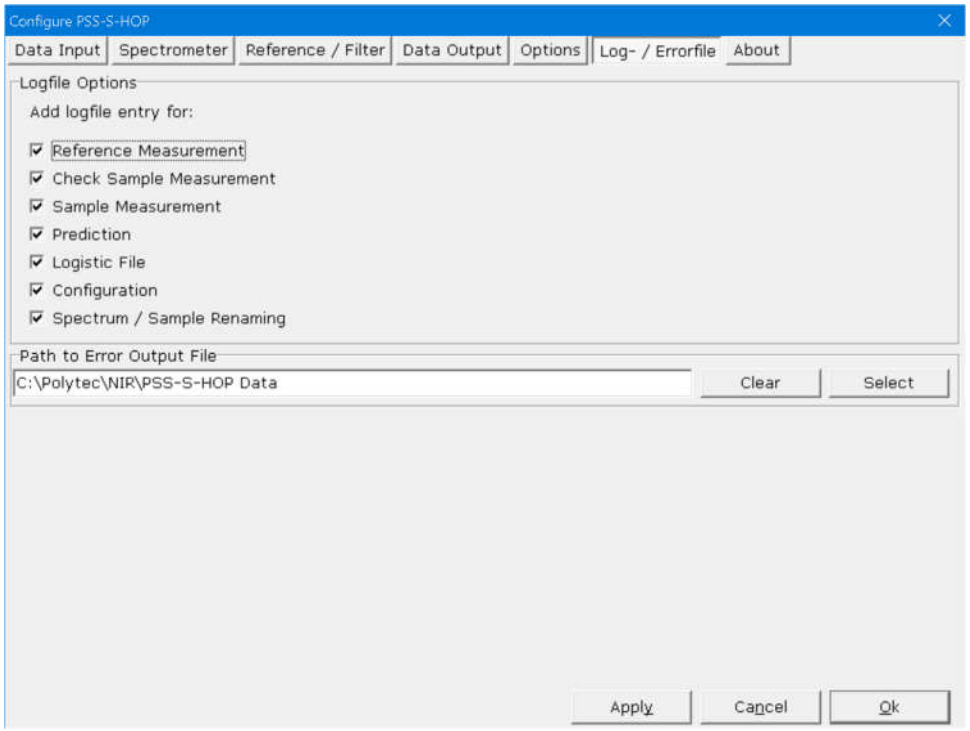
10) **"More Options / Show "Intensity Too High" Messagebox"** interrupts sample measurements to display an error message when the measured light intensity exceeds the digitization range. Generally, the measurement should be modified to avoid the high intensity, for example, by using a less reflective conveyor belt or background with a Distance Sensor. Typically the Messagebox should be shown and the measurement changed to eliminate the problem.

11) **"Check Sample Digital Output Options / Wait Time: X seconds"** is used to trigger automatic movement of Check Samples into the sample path. The wait time is the time between the digital output initiating the

Check Sample movement and trigger.

XII. The Configuration "Log- / Errorfile" Tab

Figure 29: The "Log- / Errorfile" tab of the PSS-S-HOP configuration window tab specifies logfile details.



1) A **Logfile** is created every day PSS-S-HOP is used.
   - The Logfile is named with the syntax "PSS-S-HOP-YYYY-MM-DD.log"
   - In Win10, the Logfile is saved in the folder "C:\Users\(username)\AppData\Local\VirtualStore\Program Files (x86)\Polytec\PSS-S-HOP\Logfiles\"
   - The logfile lists entries with a date and timestamp like, "Reference measurement started," "Reference measurement failed (Intensity too low)," "Reference measurement finished OK," "2019-07-19 13:21:09 Sample measurement started," "Logistics file read OK. Plot: Plot0001," "Prediction started," "Prediction finished OK," "Sample measurement finished," "Sample measurement started" etc. Searching logfiles allows events around the time of an unusual measurements to be investigated in the future.
2) An **Error Output file** is created after every PSS-S-HOP measurement:
   - The Error Output file is named "ResultError.txt" and is created after every measurement, overwriting existing earlier files.
   - It includes a Timestamp and codes for Measurement and Prediction Errors.

Table 2: An Example Error Output file, "ResultError.txt"

Measurement Error: 0
Prediction Error: -4
Timestamp: 01.10.2019 09:56:40

Table 3: Measurement Error Code Descriptions

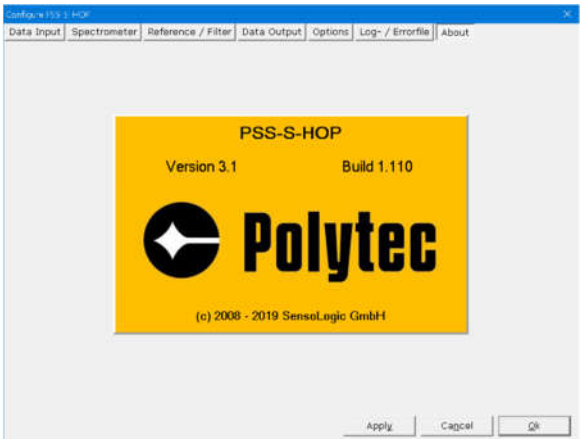
0	measurement OK
100	reference expired
103	intensity too low
104	intensity too high
105	Intensity Check: reference wavelength not found
106	Cosine (Identity) Check: Out of limits
107	Error saving reference
108	Dark Check: Wrong Wavelength Range
109	Dark Check: Counts too high
110	Reference Error: Automatic Integration could not be determined

Table 4: Prediction Error Code Descriptions

0	prediction OK
1	out of memory
2	unsupported calibration type
3	spectrum not compatible with calibration
4	outlier found
5	error while saving spectra

XIII.The Configuration "About" Tab

Figure 30: The "About" tab of the PSS-S-HOP configuration window identifies the PSS-S-HOP version

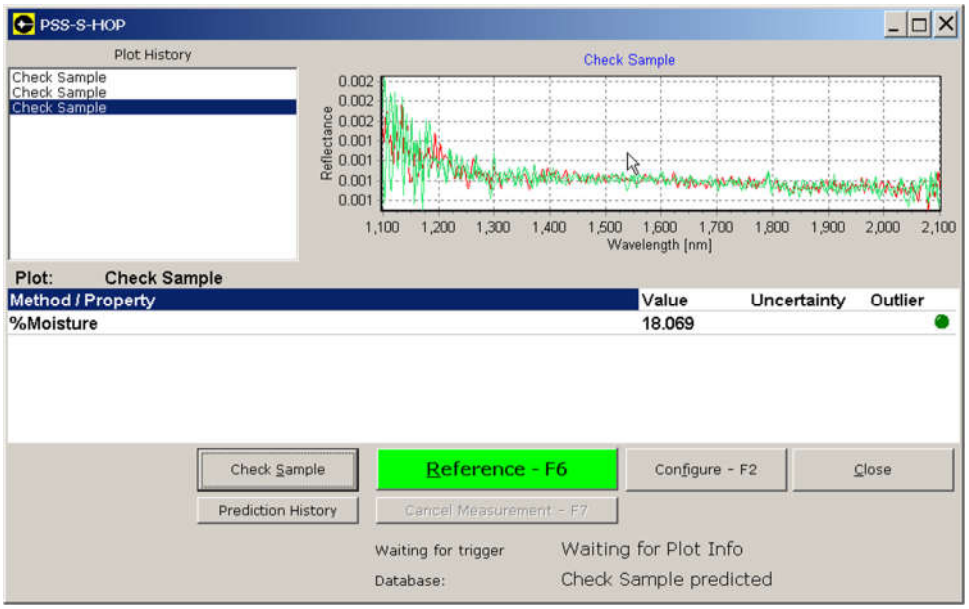


XIV. Check Sample Testing

"Check Sample" testing is an important method for confirming the long-term reliability of your NIR results. Measure a "Check Sample" at the start and end of every day and anytime you are not certain the NIR system operates as you expect. Confirm the NIR system works correctly by measuring a "Check Sample". Place your "Check Sample" directly on the sensor and click the "Check Sample" button in PSS-S-HOP. "Check Sample" spectra are saved in the Reference *.CPF project.

The purpose of the "Check Sample" is to confirm that consistent spectra are measured both as a routine check as a check any time you are not certain that the system works perfectly. The "Check Sample" you choose can be as simple as the empty sensor or as specific as a chemical or optical standard. For both, the operator would confirm the NIR works normally and the correct spectrum is consistently displayed. For example, using only an empty sensor as the "Check Sample," PSS-S-HOP would show the figure below.

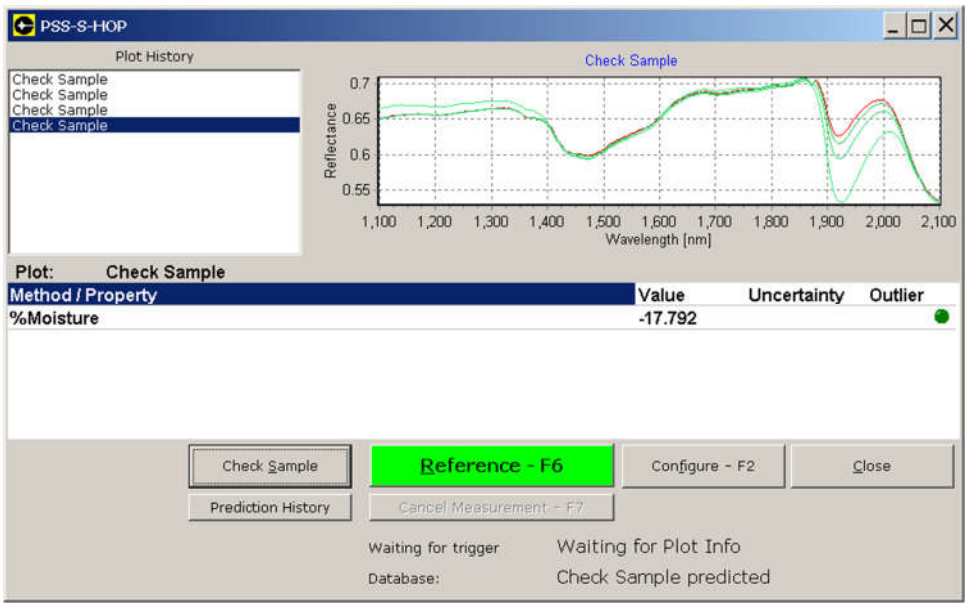
Figure 31: A "Check Sample" spectrum of a contact sensor with no sample is a flat noisy line at 0.001 Reflectance



In this case, the critical features the operator should confirm is that the spectrum with no sample is consistently a flat noisy line at ~0.001 Reflectance.

A better "Check Sample" would also show consistently display the characteristic NIR "fingerprint" of the "Check Sample". For example, even a piece of paper or random sample of harvested grain would consistently display spectra showing the characteristic NIR "fingerprint" of the material -- the approximate maximum, minimum, and shape of the NIR fingerprint would be as consistent as the sample. For example, the figure below shows repeated "Check Sample" tests of normal brown cardboard packaging.

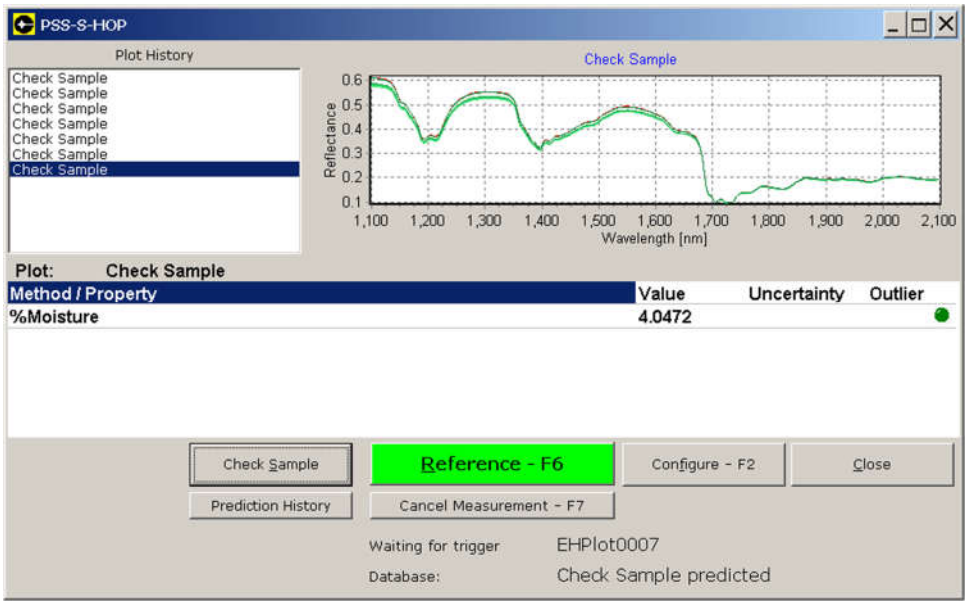
Figure 32: A "Check Sample" spectrum for most materials have a Reflectance between 0 and 1 and wide peaks pointing down. Changes in water content are seen at 1925 nm for materials that change with humidity and temperature.



While the shape around 1925 nm clearly changes, other features and the overall shape are generally consistent. The NIR feature at 1925 nm is very sensitive to water content, so changes like those in cardboard sensitively show changes in dryness or humidity and can be ignored when they are recognized and understood.

In general, the more stable and consistent the "Check Sample," the more consistent the NIR results will be and the more sensitively "Check Sample" results can test for changes, even over the course of years. A better "Check Sample" is the white plastic lid on a medicine bottle shown in the figure below. Spectra are consistent over time. Even the results from the NIR calibration -- for moisture in wheat! -- consistently give a result of ~4.1% water in the dry, water free plastic lid.

Figure 33: A "Check Sample" spectrum for a white plastic lid consistently shows a characteristic shape and apparent water content of ~4.05% with the calibration in use.



The critical features that operators should confirm in any check sample spectrum is that the spectrum is consistent and reasonable for the sample. Generally:

(a) The spectrum shape will be consistent from one day to the next

(b) NIR predictions from the Check Sample will be consistent, even if they do not make sense, when possible changes due to moisture, humidity and degradation are considered.

(c) The spectrum will show wide, negative peaks in Reflectance

(d) The spectrum will range over ~ tenths of Reflectance Y-values between 0 and 1.0 -- it will not range over only ~0.01 Reflectance unless it is an empty sensor or other unusual sample.

Check Sample spectra are stored in the Reference data project.

XV. Recommended Best Practices

Below are listed some recommended best practices to consider for using PSS-S-HOP in Field Trials. However, NIR and PSS-S-HOP are tools, not recipes, and application to your samples and purpose may require other decisions based on the specific application, goal, and results.

1) SELECT AN OPTIMUM INTEGRATION TIME FOR YOUR SAMPLE TYPE -- The optimum integration time will depend partly on the relative darkness or lightness *of your sample type in the NIR*. Select an integration time for your sample type that is long enough to measure a minimum Reflectance of ~0.2 R, but short enough to always be less than the maximum ~65,000 count intensity

   - Normally, use an integration time giving a peak intensity for the internal white reference of ~50,000 counts. If you are confident samples will not exceed the 65K maximum, a longer integration time to

decrease noise may be possible.

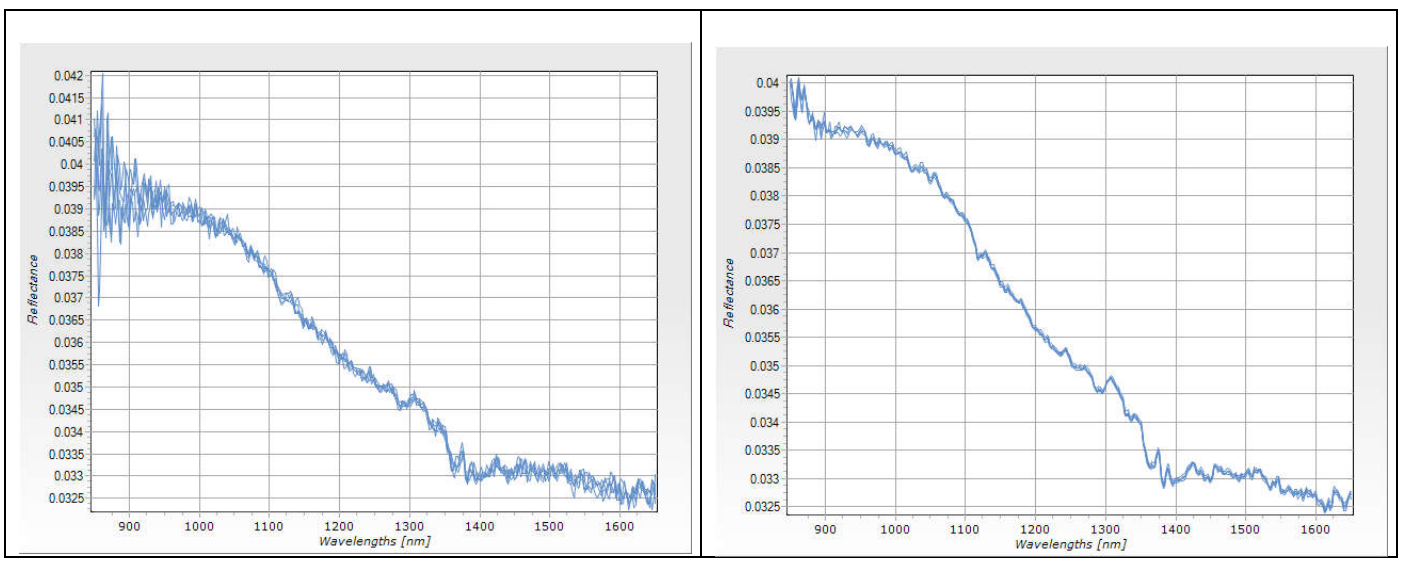
- If the peak Reflectance of a sample is then < ~0.2 R, consider increasing the integration time or selecting "Variant Integration Times" to use different integration times for samples and the internal reference.

2) RE-EVALUATE THE OPTIMUM INTEGRATION TIME AFTER REPLACING A SENSOR LAMP OR PERIODICALLY -- Lamps have a minimum 2000 hour lifetime. Over time, lamps intensities only decrease slightly, but intensities may differ as much as 30% between bulbs. Typically, the integration time might need to be updated with a new lamp, but not as a lamp ages.

3) SELECT AN OPTIMUM NUMBER OF SUBSAMPLE SCANS FOR YOUR SAMPLE TYPE -- The optimum number of scans in a subsample spectrum will depend on the relative importance of noise in a subsample spectrum and the benefit of averaging many subsamples in a plot. Many scans in each subsample have the advantage of decreasing noise, but the disadvantage that fewer subsample spectra are available, and each subsample spectrum will be more likely to include scans of dirty subsamples or the empty sensor. Fewer scans in each subsample may not effect noise significantly, but allow more and more uncontaminated subsample spectra to be averaged in the final result. Generally, spectra with a high reflectance, >~0.5 R will yield high quality, low noise spectra even with a single scan. Low reflectance, <~0.2 R spectra will yield lower quality, noisier spectra and require more scans.

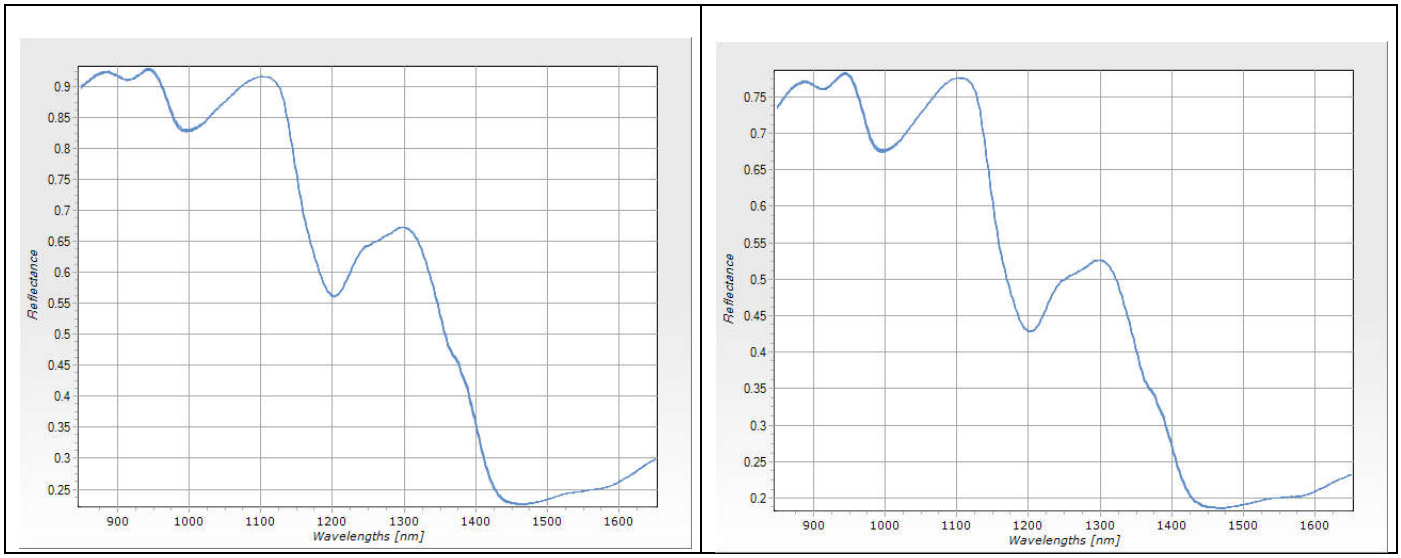
- The series of figures below show the relative importance of the number of scans in a subsample and the number of subsamples in a plot.
- The first pair of figures show the importance of the number show the importance of the number or scans to reduce noise in spectra of low reflectance black plastic pellets. (All spectra have 100 scan dark and reference spectra, as typical in PSS-S-HOP)

Figure 34: Comparing 5 repeat spectral measurements of 1 and 100 scan spectra of black plastic pellets shows multiple scans are necessary to measure a high signal-to-noise spectrum of this low reflectance sample



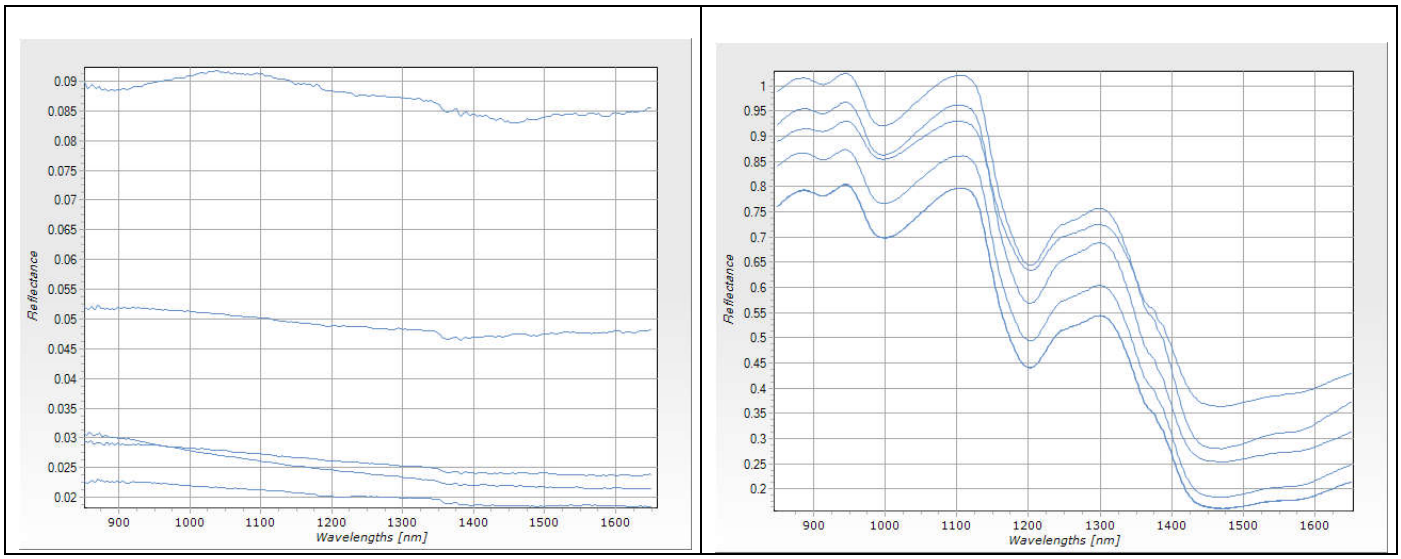
- However, a comparison of similar figures show the reduction in noise from multiple scans is *NOT* necessary to measure high signal-to-noise spectra for wheat grain, a more highly reflective sample.

Figure 35: Comparing 5 repeat spectral measurements of 1 and 100 scan spectra of wheat grain shows multiple scans are NOT necessary to measure a high signal-to-noise spectrum of this higher reflectance sample



- It is very important to recognize that the spectral noise reduced by scanning is not the only variability that can impact the accuracy of results. Sampling variability, the variability between repeat samplings is often large compared to spectral noise. The figures below of replicate samplings for both the black plastic and wheat show a large, different variability from the spectral noise shown above.

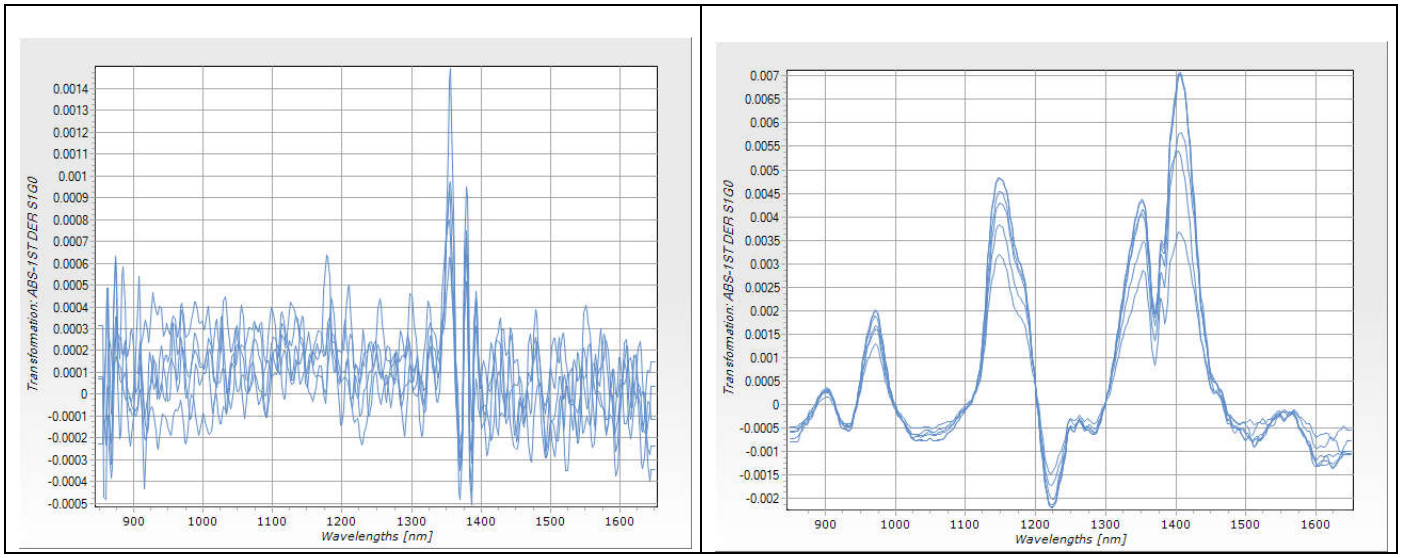
Figure 36: Replicate Samplings instead of Replicate Measurements show that for both black plastic and wheat, the sampling variability is large compared to noise



- The figures below show the same spectra of replicate samplings as above after spectral transformations that are commonly used to reduce spectral variability to create calibration models. After absorbance and first derivative transformations, the large variability between samplings is still

clearly visible.

Figure 37: Replicate Samplings show relatively large the sampling variability compared to noise even after typical spectral transformations

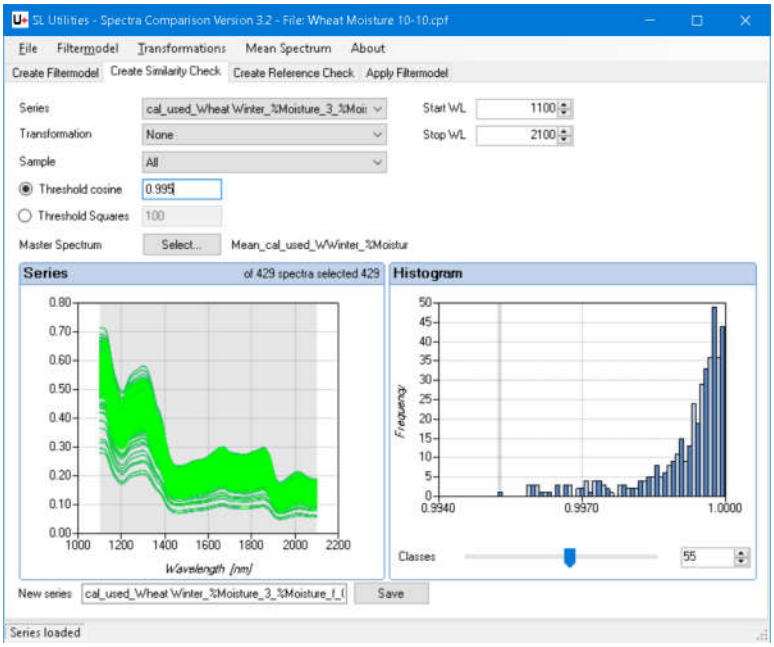


- The lesson of these examples is that users should generally set the integration time and number of scans for each subsample spectrum to limit the noise. Often, 5 scans or less with 50K counts in the white reference is sufficient. The rest of the measurement time can be used for multiple subsamples.

4) CREATE A SUBSAMPLE FILTERING METHOD -- Set a threshold for Subsample Filtering to improve accuracy by eliminating spectra of the empty sensor and other inappropriate data. Use the SL Utility, "Spectra Comparison & Filtering."

- First, assess the differences between subsample spectra of one, clean, representative sample flowing past the sensor using a total measurement cycle time that is ~50% longer than the time to measure tens of subsample spectra. Your method will need to discriminate the spectra of the nearly empty sensor. The figure in the earlier section on the configuration of the "Subsample filter" showed a threshold above ~0.94 was necessary for wheat. However, other grains will yield other results.
- Second, assess the range of match scores between spectra of many different samples at different content, moisture, etc. For example, view the distribution of match scores for many plots in the Target project or calibration. This distribution will give a realistic estimate of the tightest possible threshold -- the threshold should not eliminate samples that are acceptable, but with extreme content levels!

Figure 38: Spectra of the wheat with varying protein and moisture in the calibration set show that a threshold must be at least as low as ~0.995 to include normally varying spectra



- Some suggestions on how to better determine the threshold for Subsample Filtering are listed below:
  - Remember the threshold will likely be used on all harvested plots, so test it as carefully as possible first. If possible, evaluate it using the archive spectra of previously measured plots before using it.
  - Consider which risk is worse for your goals: setting a threshold that is too tight, filtering acceptable subsample spectra, and failing to calculate an accurate plot average? Or, setting a threshold that is too loose, including weak spectra and spectra of subsample potentially contaminated with chaff or dirt?
  - Carefully evaluate the threshold for actual harvest data under typical conditions. Try to identify the reason subsample spectra yield low match scores, and whether those subsample spectra should be included or not in an accurate plot average
  - Test the effect of known possible causes of inaccurate subsample spectra, like the presence of chaff and dirt.
  - Consider using spectral transformations similar to your calibrations.
  - Test whether a selected threshold reproducibly excludes unacceptable data?
  - Evaluate harvest spectra with Principal Component Analysis (PCA) to assess whether a single cosine threshold could distinguish acceptable and unacceptable subsample spectra

5) CREATE A SIMILARITY CHECK -- Set a threshold for the Similarity Check to confirm the correct sample is measured. Use the SL Utility, "Spectra Comparison & Filtering" and "Create Similarity Check."
   - Measure spectra of any samples that could possibly interfere with sample measurements in the field

in order to exclude them for certain. Measure the empty sensor, chaff, plastic bags, tape, and paper. Use the Similarity Check Utility to identify a threshold to distinguish them from the average of your calibration set.

6) IF USING A DISTANCE SENSOR WITH A CONVEYOR BELT, TEST THE REFLECTIVITY OF THE CONVEYOR BELT

- An occasional problem using distance sensors is conveyor belts that are partly mirror-like in the NIR. Such belts directly reflect bright NIR from the sample area instead of the usual, weaker, diffuse reflectance. The intense direct reflections can lead to poor results. Users should follow 3 steps using distance sensors with conveyor belts: (i) measure the reflectivity of the conveyor belt; (ii) use non-reflective belts; (iii) if necessary, use the "Configure/Spectrometer/Intensity Check" setting (not the "Intensity Check" setting on the "Reference / Filter" Tab).
- Make certain your belt has low reflectance by measuring it as Check Sample in PSS-S-HOP. The lower the reflectance, the lower the risk of a direct reflection causing an inaccurate measurement. It should at least be below ~ 0.5 R.
- Use PSS-S-HOP to measure whether sample intensity for the running belt ever gets close to 65,000 counts. Click "Configure / Data Output" and name the "Unreferenced Spectra Project" as "Unreferenced Belt." Run the belt and trigger PSS-S-HOP. Double click "Unreferenced Belt.cpf" to open it in the Database viewer. Double click the series for the sample you measured. The fewer the high reflectance measurements and lower the measurements, the better. Measurements should rarely be over 30,000 counts. There is no need to continue measuring Unreferenced spectra once a consistent, reliable method is determined.
- Set thresholds in both the Similarity Test and the Subsample Filter that exclude the spectrum of the empty belt. Test running the empty belt to be certain no result is listed. *Test running an empty belt with a small sample at one spot to be certain only acceptable sample measurements are used to calculate the result*.
- If some high reflectivity scans cannot be eliminated, then set them to be ignored by clicking "Configure / Spectrometer / Intensity Check" and "Full On or Accept Some."

7) WHEN REPLACING THE SENSOR LAMP,

- Re-evaluate the best integration time with the new lamp
- Check consistency over time of previous dark and white reference spectra
- Set the reference intensity limit and cosine threshold accordingly
- Set the "Dark Count Limit" in the "Dark Spectrum Check" accordingly

XVI. Demonstrating PSS-S-HOP

The goal of this section is to demonstrate using a configured system with an example calibration. In this section, users will:

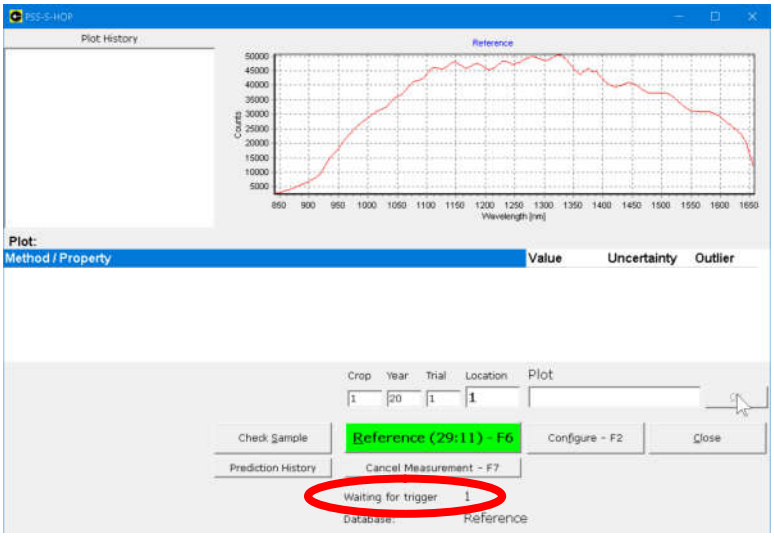
- Manually name, trigger, measure, and assess sample measurements
- "Automatically" name, trigger, measure, and report sample measurements using a *.US "Logistic" file, as if using Wintersteiger's EasyHarvest or HarvestMaster's Mirus
- Measure content in multiple wheat or other samples in ~ 1 second

- Identify outlier samples
- Automatically reject weak spectra and improve averaging to calculate a final plot average
- Warn the operator and prevent measurements when the sensor is covered by debris like debris jammed in a sample gate or debris accidentally on a conveyor belt
- Warn the operator and prevent measurements when the fiber optic is broken or disconnected or the instrument signal intensity is low

Follow these steps.

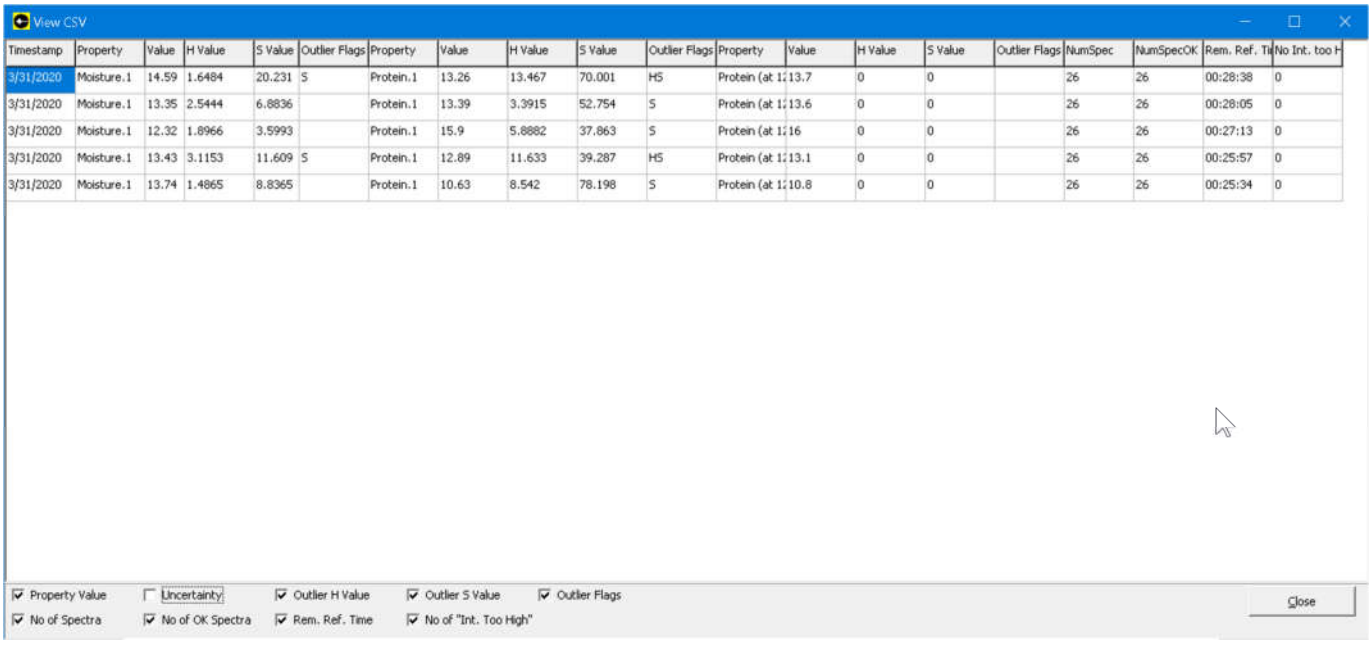
1) SET UP SPECTROMETER AND CONFIGURE PSS-S-HOP -- Follow instructions in previous sections to configure your system for manual testing, including (a) allowing the spectrometer connection thru the firewall etc, (b) setting the integration time to give 50K counts in the white reference, and (c) activating the filtering and similarity tests.
2) BACKUP YOUR INITIAL CONFIGURATION -- Export your configuration as a backup.
3) REFERENCE -- Measure a 50K reference
4) MEASURE "CHECK SAMPLE" -- Measure some "Check Sample" to confirm instrument performance. At a minimum, measure an empty sample and confirm consistent, insignificant reflectance. In the best case scenario, measure a single, stable, common sample that does not dry out or absorb moisture and is related to test samples.
5) MANUALLY MEASURE SAMPLES -- As in the figure below, manually name a sample with a syntax like "Crop=1 Year=YY Trial=1 Location=1 Plot=1," click "OK" and trigger the spectrum.
    - As the software changes from waiting for the Plot ID, to waiting for the trigger, to measuring scans, and then to waiting for the next Plot ID, the status message at the bottom of the PSS-S-HOP screen will change from "Waiting for Plot Info," to "Waiting for trigger" with the Plot ID displayed, to the number of scans measured, and finally back to "Waiting for Plot Info."

Figure 39: The PSS-S-HOP status shows whether the NIR is "Waiting for Plot Info," "Waiting for trigger" or is measuring scans.



- Such a measurement would create the results files named with the syntax, "1 YY 1 1.csv," the Target File "1 YY 1 1.cpf," Archive files "A1 YY 1 1.cpf" as well as the Reference filed named by serial number, for example, "Ref_X XX XXX .cpf"
- Similarly, without naming any sample, trigger the NIR to measure a plot. Results files will be named with the syntax, "YYYYMMDD.csv," Target Files as "YYYYMMDD.cpf," and Archive files as "AYYYYMMDD.cpf." Reference spectra continue to be saved in the file named by serial number, "Ref_X XX XXX .cpf"
- During a series of measurements, users can click "Prediction History" and view detailed results from the *.CSV results file as in the figure below.

Figure 40: Click "Prediction History" and view detailed results during a series of measurements



Timestamp	Property	Value	H Value	S Value	Outlier Flags	Property	Value	H Value	S Value	Outlier Flags	Property	Value	H Value	S Value	Outlier Flags	NumSpec	NumSpecOK	Rem. Ref. Ti	No Int. too H
3/31/2020	Moisture.1	14.59	1.6484	20.231	S	Protein.1	13.26	13.467	70.001	HS	Protein (at 1:	13.7	0	0		26	26	00:28:38	0
3/31/2020	Moisture.1	13.35	2.5444	6.8836		Protein.1	13.39	3.3915	52.754	S	Protein (at 1:	13.6	0	0		26	26	00:28:05	0
3/31/2020	Moisture.1	12.32	1.8966	3.5993		Protein.1	15.9	5.8882	37.863	S	Protein (at 1:	16	0	0		26	26	00:27:13	0
3/31/2020	Moisture.1	13.43	3.1153	11.609	S	Protein.1	12.89	11.633	39.287	HS	Protein (at 1:	13.1	0	0		26	26	00:25:57	0
3/31/2020	Moisture.1	13.74	1.4865	8.8365		Protein.1	10.63	8.542	78.198	S	Protein (at 1:	10.8	0	0		26	26	00:25:34	0

6) AUTOMATICALLY MEASURE SAMPLES AS IF USING HARVESTER SOFTWARE -- In PSS-S-HOP / Configure / "Data Input" Tab / Source / Remote Software, switch from "Manual / Barcode" to "EasyHarvest" or "Mirus"

- In Notepad or a similar app, create a text "Logistic file" with lines with the syntax:

  ID_REC=Plot0001
  ID_SERIES=EasternFields
  CALNAME=Wheat Moisture 10-10
  REF=JA
  CPFNAME=2020Harvest
  CSVNAME=2020Harvest

- Save the file with the name "Polytec.us" in C:\Polytec\NIR\PSS-S-HOP Data\ . This will demonstrate how all the remote harvester software work with PSS-S-HOP -- the harvester software writes a "Logistic" file with the PlotID and PSS-S-HOP reads it.
- When PSS-S-HOP reads the file, the "Measure Status" at the bottom of the screen will change from "Waiting for Plot Info" to "Waiting for Trigger" and the pending PlotID. Sometimes Windows will automatically change the filename extension to *.txt and PSS-S-HOP will NOT read it. In this case, name the file with quotation marks, e.g., "Polytec.us".
- After reading the Logistic file, trigger the system to measure the sample.
- After measuring the sample, PSS-S-HOP creates a file named "PSSResults.txt". This file lists the NIR results and is read by the harvester software into the breeding database.

7) GENERATE PROTEIN CONTENT ON STANDARD MOISTURE BASIS -- Follow instructions in the section on the configuration "Data Input" tab for "Calculated Properties" to directly generate and save results like

protein content on standard moisture basis that are calculated from the NIR results.

8) AUTOMATICALLY CHANGE CALIBRATION APPLICATIONS AS IF HARVESTING IN AN INTERCROPPED FIELD -- Follow instructions in the section on the configuration "Data Input" tab to "Use Productlist" to quickly, automatically, switch products or calibrations.

9) RESTRICT CALIBRATIONS TO SPECIFIC SPECTROMETERS -- Follow instructions in the section on the configuration "Data Input" tab to restrict calibrations to specific spectrometers. Since the serial number and restriction can be easily changed, this option is most useful preventing mix-ups rather than securing confidential calibration data. For security, use the SL Calibration Lock.

10) USE "SUBSAMPLE FILTERING" TO REJECT WEAK SPECTRA OF EMPTY OR SMALL SAMPLES -- Pour a sample into the sampling area of a sensor so that you can measure a long series of spectra beginning with a bad, empty sensor, and ending with a good, filled sensor. This could mimic, for example, an unusually small plot yield quickly flowing past a sensor. Use a long cycle time so that you can easily measure the data. Repeat the series of measurements multiple times.

    - View the distribution of match scores for your sample poured onto the sensor by following instructions from the "Reference / Filter" Section to use SL Utilities/"Spectra Comparison & Filtering"/"Create filtermodel" tab.

    - Make multiple measurements of poured samples with different thresholds, and view the spectral results, outlier statistics, number of filtered spectra and predicted content. Change the PSS-S-HOP configuration filtering threshold so that (1) it is extremely high, ~0.999 and filters all but the most intense samples; (2) it is midway between the two peaks, filtering all the spectra that are obviously weak; (3) and at ~0.90, below the lower peak so that no spectra are filtered. The displayed spectra of the filtered and averaged subsample spectra will appear most intense and the outlier statistics will be smallest. Other final sample spectra will be progressively less intense, and more outlying. Choose another threshold between the highest two. View the results, outlier statistics, number of filtered spectra and predicted content and decide the threshold you choose.

11) USE THE "SIMILARITY CHECK" TO CONFIRM THE CORRECT SAMPLE IS MEASURED.

    - Measure spectra of any samples that could possibly interfere with sample measurements in the field in order to exclude them for certain. Measure the empty sensor, chaff, plastic bags, tape, and paper. Use the Similarity Check Utility to identify a threshold to distinguish them from the average of your calibration set.

12) USE THE "INTENSITY CHECK" ON THE "SPECTROMETER TAB" TO REJECT "SATURATED" SAMPLE MEASUREMENTS.

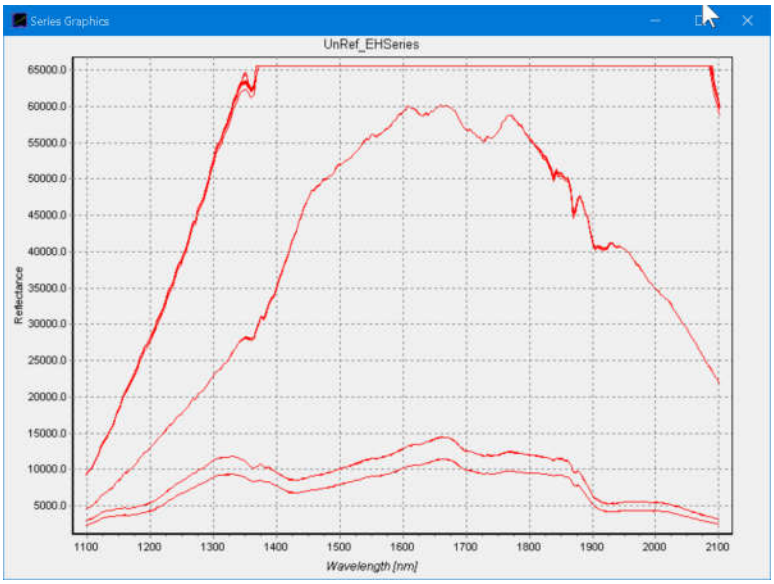
The "Intensity Check" feature allows users to reject occasional saturated sample scans. First, use a program like PAS Labs or PasSpectrometerLibraryDemo and shiny metal foil to cause and view saturated sample scans. Second, using PSS-S-HOP, use the "Spectrometer / Intensity Check" setting and the *.CPF project of Unreferenced Samples spectra to eliminate saturated sample scans. Then, users can set "Intensity Check" to reject and correct for occasional saturated sample scans in shiny samples, like moist glossy black seeds.

- Using the "Online" measurement in PAS Labs or PasSpectrometerLibraryDemo to view spectra while moving shiny metal foil in front of a sensor. Notice which orientations of the foil lead to the intense, high count measurements

- In PSS-S-HOP, set "Spectrometer / Intensity Check / Full On" to reject "saturated" measurements. Save the Unreferenced Samples *.CPF project. Move the foil to create reflections causing "saturated"

measurements. PSS-S-HOP will give the error message, "Measurement failed: Intensity too high!" The Unreferenced Samples *.CPF project will show off-scale intensity spectra like below.

Figure 41: Spectra in a *.CPF project of Unreferenced Samples from PSS-S-HOP showing direct, high-intensity reflections from a shiny metal sample



13) BACKUP YOUR FINAL CONFIGURATION -- Export your configuration as a backup.

XVII. Viewing and Understanding PSS-S-HOP Data and Results

This section describes how to view and understand PSS-S-HOP the data and results measured in the previous section.

- All of the NIR spectra are saved in SensoLogic *.CPF Database projects where:
- All data can be viewed with the free SL Database Viewer
- Spectra can be viewed individually or as a series by double-clicking on the spectrum or series
- The predicted content is saved in the as a sample.
- All the data can be moved, reorganized, and edited in SL Data Manager, part of the SL Calibration Suite

1) In "Remote Mode," results for every plot are written to "PSSResults.txt" a temporary file to be read by the plot management software. An example of "PSSResults.txt" is displayed below.

Figure 42: An example of "PSSResults.txt," the temporary "Remote Mode" output file written for each plot

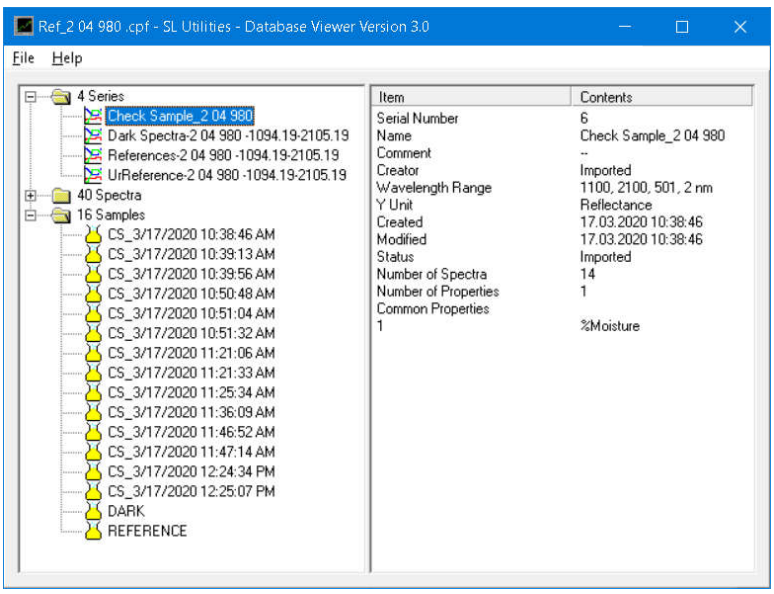
2) Results for every plot are listed in a *.CSV file as shown below. The name is specified in "Configure / Data Output."

Figure 43: An example of the *.CSV file listing all results for all plots.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Date	Time	Crop	Year	Trial	Location	Plot	PSS_S/N	Property_1_Name	Property_1_Value	Property_1_U	Property_1_H	Property_1_S	Property_1_Flags	NumSpec Total	NumSpec Used	Used%	Reference Remaining Time	Number of "Intensity Too High" spectra
2	3/17/2020	10:46:57 AM	1	20	1	1	poured wheat1	2 04 980	%Moisture	11.47	0.658	1.7954	2.9571		165	111	67.273	00:29:24	0
3	3/17/2020	10:47:27 AM	1	20	1	1	poured wheat2	2 04 980	%Moisture	11.63	0.66	2.5098	2.5323		165	111	67.273	00:28:54	0
4	3/17/2020	10:48:08 AM	1	20	1	1	poured wheat3	2 04 980	%Moisture	11.75	0.658	1.8869	4.9904		165	165	100	00:28:13	0
5	3/17/2020	10:48:41 AM	1	20	1	1	poured wheat4	2 04 980	%Moisture	11.66	0.66	2.6408	3.6717		165	106	64.242	00:27:41	0
6	3/17/2020	10:49:12 AM	1	20	1	1	poured wheat5	2 04 980	%Moisture	11.84	0.66	2.4472	3.6244		165	116	70.303	00:27:09	0

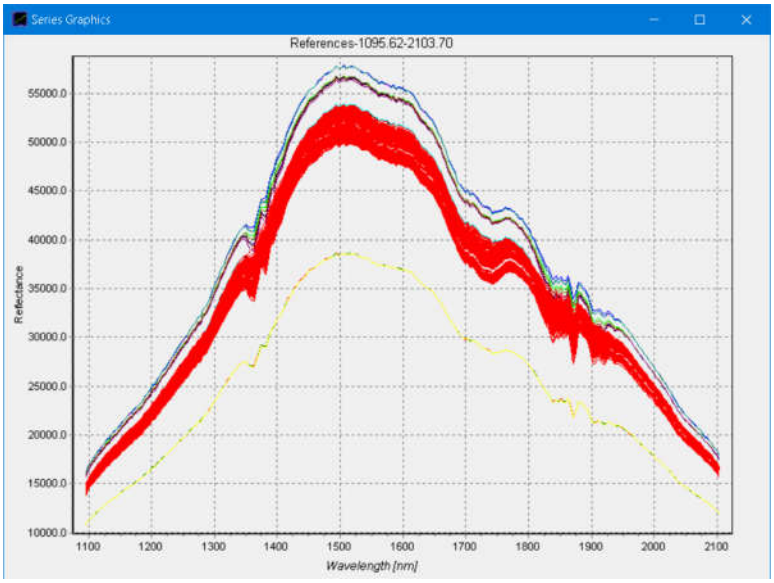
3) The Reference *.CPF Database project stores all Dark, Reference, and Check Sample spectra over long periods of time. This data is allows users to quantitatively analyze long term system performance.

Figure 44: The Reference *.CPF Database project stores all Dark, Reference, and Check Sample spectra



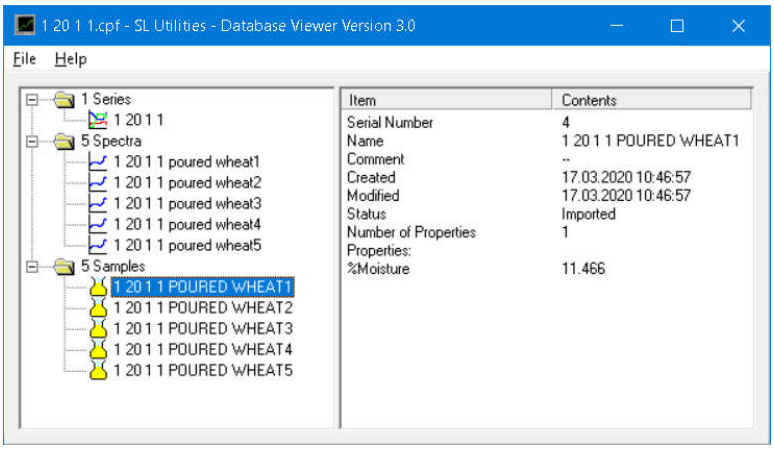
- In the SL Database Viewer, users can double-click on a series to see how hundreds of spectra change over time. For example the plot below shows 275 Reference spectra over the course of 2 years with maxima of ~ 51 +/- 2K counts that change little in intensity or shape. (The few spectra with higher and lower intensities were tests with different settings.) Similarly, the Dark spectra and Check Sample spectra can displayed.

Figure 45: Series of spectra measured over a long period of time are easily viewed in the SL Database Viewer



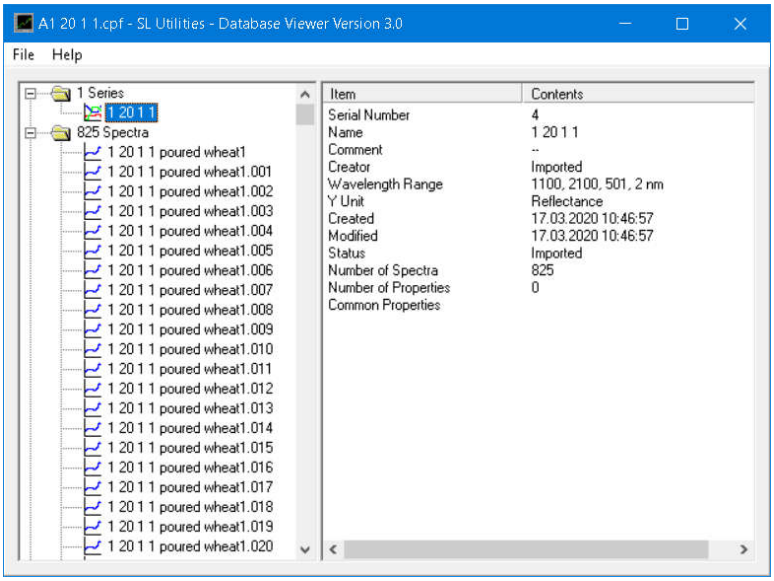
4) The Target *.CPF Database project stores the key spectral results -- averaged spectra for every plot after the similarity test and filtering. The predicted content is saved in the as a sample. As in all SL *.CPF Database projects, spectra can be viewed individually or as a series.

Figure 46: The Target Database stores averaged spectra for every plot after the similarity test and filtering.



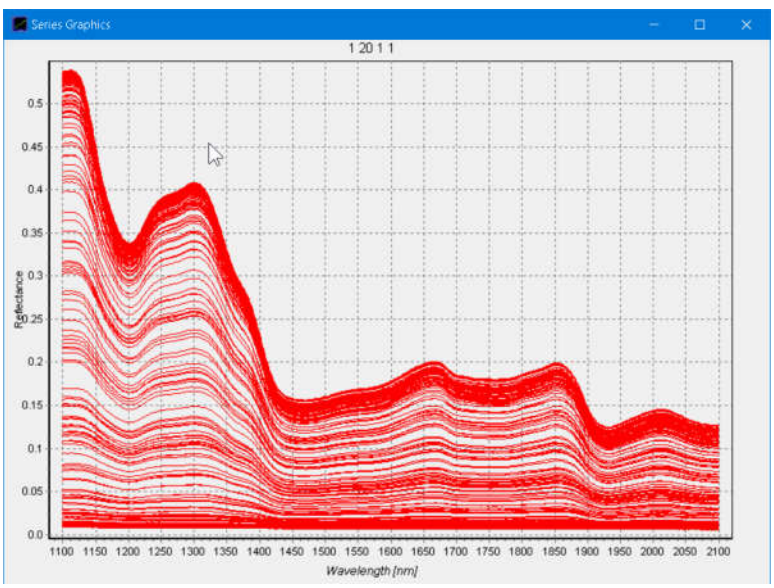
5) The Archive *.CPF Database Project stores all the subsample spectra for every plot. Since all the subsample spectra are saved, the data can be re-analyzed with different filtering methods. Sequential subsample spectra are saved with the indices ".001," ".002," ".003" etc. Archive projects are usually used to create filtering methods for harvest data.

Figure 47: The Archive Database stores all the subsample spectra for every plot.



- As with the other databases, spectra and series of spectra be easily viewed.

Figure 48: The series of archive spectra show how plot subsamples change over time.



XVIII. Routine Operation in the Harvester

After configuration, the process is much simpler for routine operation during harvest. Typical recommended best practices are:

- Before harvest, test new measurements and review history of changes in dark, reference, and "Check Sample" spectra over time. Save old data, test current signal levels, insure you have a backup spare bulb and fiber, update files, run check samples including sample for lab testing. Review standard

procedures. Review possible error messages and their follow-up.

- Insure normal, successful startup, including normal reference measurements and normal "Check Sample" measurements
- Verify PSS-S-HOP reads Plot IDs from the harvester software and measures NIR of plot in ~ 4 seconds
- Verify PSS-S-HOP displays results and sends results to the harvester software
- Identify and correct known errors including NIR communication problems, weak references, and sample-calibration mismatch and extreme outliers
- Occasionally verify that plots are correctly measured by NIR
- Occasionally verify that new NIR references are routinely measured automatically
- Measure a new "Check Sample" at the end of the day
- Periodically, back up all NIR data, especially a current "Configuration File"
- Periodically review historical data -- review changes in dark, reference, and "Check Sample" spectra over time. Compare changes with threshold settings, risks, and any measured problems.

XIX.PSS-S-HOP Operator Troubleshooting

The PSS-S-HOP software is designed to automatically prevent instrument problems, software or instrument communication problems, and sampling problems. These potential problems and quick solutions for harvester operators are listed in "PSS-S-HOP Operator Troubleshooting[DS1]." That guides summarizes solutions to possible error messages and simple but important steps that operators can take to insure reliable operation of the NIR system.

The exported configuration file is a combination of the Instrument Settings in PolytecPSS.ini and other configuration settings in the system file PredOnline.ini.

1) If the password is changed and then lost, find it in the “Win_Protect” line in the file “PredOnline.ini”.
2) Rarely, PSS-S-HOP cannot communicate with the spectrometer if configuration changes were made after spectrometer communication was disrupted. For example, during a power outage during set-up. Also rarely, PC configuration changes lead to new, unusual errors related to *.DLL mismatched or similar issues. Possible solutions in these cases are to reinstall PSS-S-HOP or to delete and *.ini files. PSS-S-HOP settings are saved in PredOnline.ini and PolytecPSS.ini. Depending on the PC, Windows version, PSS-S-HOP version, these *.ini files are saved as hidden files or in hidden folders like C:\Users\PC_UserName\IRIS\ , C:\Users\PC_UserName\AppData\Local\VirtualStore\Program Files (x86)\Polytec\PSS-S-HOP\, and C:\Program Files (x86)\Polytec\PSS-S-HOP\. PredOnline.ini has HOP settings and rarely becomes corrupted when communication is lost. PolytecPSS.ini has spectrometer settings.

XX. For more information

Contact:

David Semmes
Polytec Inc., Process Analytical
1 Cabot Road, Suite 102
Hudson, MA 01749
Tel (508) 417-1040
Cell (631) 601-5245
d.semmes@polytec.com