Introduction to Instrument Standardization and Calibration Transfer

Barry M. Wise

Eigenvector Research, Inc.
830 Wapato Lake Road
Manson, WA 98831

bmw@eigenvector.com
www.eigenvector.com
Motivation

- Calibration models for quantitation or classification often take advantage of relatively small changes in spectra.
- Instrument to instrument differences can be substantial, i.e. samples look different.
- Instruments may drift over time.
- Renders models invalid.
- Inconvenient to recalibrate instruments or may want to utilize a historical database.
Two Main Approaches

- Find a transformation that maps the response of the field instrument onto the standard instrument
  - Direct and piece-wise direct standardization
  - Neural network and other variants
- Process the data from both instruments in a way that makes the differences disappear
  - baselining and derivatizing
  - multiplicative scatter correction, FIR filtering
  - orthogonal signal correction
  - prediction augmented classical least squares
  - generalized least squares
  - explicit deresolution
**Piece-wise Direct Standardization (PDS)**

- Develop models which use windows on field instrument to predict single channels on standard.
Develop Transfer Matrix $F_b$

Difference between instruments modelled as:

$$S_1 = S_2 F_b + 1 b_s^T$$
Data Arrangement for PDS

Standard

Field

Window width
Data Arrangement for Double Window PDS

Standard Field

Window 1 = 5
Window 2 = 3
Second window can be spectrum full width = single model PDS
Direct Standardization

- Similar to PDS except $F_b$ matrix is full:
  \[ F_b = S_2 + S_1 \]
- Many more parameters in DS compared to PDS
Variations on PDS

- **Single model PDS**
  - widen second window in DWPDS until it is the width of the entire spectrum
  - model is the same for each channel in master instrument
  - transfer function not a function of wavelength

- **Single model PDS with index**
  - include the channel number as the parameter in the model
  - use non-linear model such as ANN
  - transfer function is a function of wavelength
Orthogonal Signal Correction

- OSC attempts to remove extraneous variation unrelated to the property of interest from the predictor variables.
- Principal components are calculated for the predictor variables then orthogonalized against the variable(s) to be predicted.
- Weighting vectors are determined with PLS which reproduce the orthogonal directions on new data.
- To use in standardization, apply to data measured on both instruments.
Example From NIR, Pseudo Gasoline Mixtures

Instrument number 1 spectras shown in red
Instrument number 2 spectras shown in blue
Difference Between Instruments

Difference between NIR Spectra from Instruments 1 and 2

Absorbance Difference vs. Wavelength (nm)
Instrument 1 Calibration

Actual versus Fit Concentrations Based on Instrument 1

Each analyte shown as different color
After Standardization

Difference between NIR Spectra from Instruments 1 and 2

Difference before correction shown in red
Difference after direct correction shown in green
Difference after piecewise correction shown in blue
Difference after OSC shown in magenta
Instrument 1 Calibration on Unstandardized Instrument 2

Actual Concentrations vs. Predictions Based on Instrument 2

Each analyte shown as different color
Instrument 1 Calibration on Standardized Instrument 2

Actual Concentrations vs. Predictions Based on Standardized Instrument 2

Each analyte shown as different color
- Piecewise direct standardized samples
  * Direct standardized samples
  + OSC standardized samples

Instrument 1 Fit Error: 0.33
Unstandardized: 7.49
Piecewise Standardized: 0.70
DWPDS Standardized: 0.77
Direct Standardized: 2.50
OSC Standardized: 1.45
Prediction Augmented Classical Least Squares

- If CLS is used for predictive model, new spectra can be added to prediction step to account for differences between instrument.
- Augmented spectra can include known new components or estimates of changes such as a baseline offset or mean difference.
- Eigenvectors of difference matrices can also be included.
**CLS: Predictions on Instrument 2 with Instrument 1 Spectra**

**Calibration Using Instrument 1 Spectra**

**Predictions Using Instrument 2 Spectra**
Estimated Pure Component Spectra and Additional Factors

Estimated Pure Component Spectra

Mean Difference and First 3 Eigenvectors
PA-CLS Predictions

Pure Spectra Augmented with Mean Difference

Fit Error   0.48
Uncorrected 18.05
PA-CLS w/mean 2.46
PA-CLS 1 EV 1.78
PA-CLS 3 EVs 1.15
NIR of Corn Samples

NIR Spectra

Absorbance

Wavelength (nm)
Calibration

Actual versus Fit Concentrations Based on Instrument 1

Each analyte shown as different color

Analytes are oil, moisture, starch and protein
Difference Before and After Standardization

Difference between NIR Spectra from Instruments 1 and 2

Absorbance Difference

Wavelength (nm)

Before
Direct
OSC
PDS
Effect of OSC on Spectra

Spectra Before and After Orthogonal Signal Correction

Original Spectra in Blue
OSC Transformed Spectra in Red

Absorbance vs. Wavelength (nm)
Results of Corn Standardization

Actual Concentrations vs. Predictions Based on Standardized Instrument

Each analyte shown as different color
- Piecewise direct standardized samples
  - Direct standardized samples
  - OSC standardized samples

Fit Error = 0.10
Unstandardized = 1.32
PDS = 0.33
DWPDS = 0.33
Direct = 0.45
OSC = 0.20
## Summary

<table>
<thead>
<tr>
<th>Method</th>
<th>Transforms?</th>
<th>Standards</th>
<th>Parameters</th>
<th>Uses Y</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direct</td>
<td>Yes</td>
<td>Real</td>
<td>Lots</td>
<td>No</td>
<td>Many samples</td>
</tr>
<tr>
<td>PDS</td>
<td>Yes</td>
<td>Anything</td>
<td>Few</td>
<td>No</td>
<td>Few samples</td>
</tr>
<tr>
<td>NN-PDS</td>
<td>Yes</td>
<td>Anything</td>
<td>Moderate</td>
<td>No</td>
<td>Non-linear</td>
</tr>
<tr>
<td>Derivative</td>
<td>No</td>
<td>None</td>
<td>None</td>
<td>No</td>
<td>Easy</td>
</tr>
<tr>
<td>MSC</td>
<td>Yes</td>
<td>Real, Few</td>
<td>Few</td>
<td>No</td>
<td>Easy</td>
</tr>
<tr>
<td>OSC</td>
<td>No</td>
<td>Real</td>
<td>Few</td>
<td>Yes</td>
<td>Requires Y</td>
</tr>
<tr>
<td>PA-CLS</td>
<td>No</td>
<td>Anything?</td>
<td>Few</td>
<td>No</td>
<td>Interpretable</td>
</tr>
<tr>
<td>GLS</td>
<td>No</td>
<td>Real</td>
<td>Moderate</td>
<td>No</td>
<td>New</td>
</tr>
<tr>
<td>Deresolution</td>
<td>No</td>
<td>None</td>
<td>Few</td>
<td>No</td>
<td>FTIR</td>
</tr>
</tbody>
</table>
Conclusions

◆ PDS still the method to “shoot for”
◆ DS more sensitive to number of transfer samples
◆ OSC produces especially good results in some data, also useful as a preprocessing technique
◆ FIR not adequate in situations we’ve seen
**PLS_Toolbox 2.0**

*for use with MATLAB*

- Version 2.0 for MATLAB 5 now available
- Wide selection of multivariate analysis tools
- Used in our Chemometrics Short Courses

**Barry M. Wise**
Neal B. Gallagher