Abstract

When synchronous fluorescence (SyF) spectroscopy is used for quantitative and qualitative analysis, selection of a useful wavelength interval between the excitation and emission wavelengths (Δλ) is needed. Presented is a fusion approach to combine Δλ intervals thereby negating the selection process. This study uses the fusion of SyF spectra for the quantitative analysis of the corn oil adulterant content. The SyF spectra were acquired by varying the excitation wavelength in the region 300-800 nm using Δλ wavelength intervals from 10 to 100 nm in steps of 10 nm producing 10 sets of SyF spectra. For quantitative analysis, two calibration approaches are evaluated with these 10 SyF spectral datasets. Multivariate calibration by partial least squares (PLS) and a univariate calibration (LV). The PLS approach creates numerous models with latent variables (LV). Each source is from a unique farm in Morocco. All Wavelengths Multivariate.

Model Measures Cont‘d

- Variance:
  - Euclidean 2-norm \( ||b||_2 \)
- U-Curves:
  - Bias-variance trade-off
  - \( C_1, c = \frac{||R_{C1} - R_{C2}||}{||R_{C1}||} \)
  - \( C_2, c = \frac{||R_{C1} - R_{C2}||}{||R_{C1}||} \)

MV Model Selection

- The PLS approach creates numerous models with latent variables (LV).
- A set of LV must be selected to form the model from the calibration set and then predict the validation set.
- To automate the LV selection a U-curve approach was used.
- Selected LV = \( \min(\text{mean}(C_1 + C_2)) \)
- \( C_1 \) balances prediction error for the calibration set, RMSE in conjunction with the model regression vector \( b \).
- \( C_2 \) balances the calibration model fit with the variance indicator \( \hat{y} \) and \( \hat{y} \) guard against selecting an over fitted or under-fitted model.

Argon Oil Data Set

- 5 sources from separate Moroccan farms.
- Each source was adulterated with corn oil purchased from a local store.
- Source 1: 27 samples
- Source 2: 30 samples
- Source 3: 29 samples
- Source 4: 33 samples
- Source 5: 34 samples
- SyF spectra obtained by simultaneously measuring excitation and emission wavelengths in the range 300 to 800 nm varying the Δλ interval from 10 to 100 nm in increments of 10 nm.

Modeling Approach

Partial Least Squares (PLS)

\( y = Xb \rightarrow b = (X^T X)^{-1} X^T y \)

\( \hat{y} = Xb \)

- Spectral responses are summed across the measured spectral ranges.
- \( X \): calibration samples
- \( y \): actual analyte concentration
- \( \hat{y} \): analyte concentration prediction
- \( b \): estimated model regression vector

Model Measures

- Bias:
  - \( R^2 \)
  - Root Mean Square Error (RMSE)
  - RMSEC, RMSEV

Total Synchronous Fluorescence Spectra

- Argon Excitation Region: 300 – 330 nm
- Corn Excitation Region: 300 – 410 nm
- Potential for differentiation within 300 – 410 nm range

Conclusions

- Low level fusion provided lower prediction errors.
- The bands provided more consistent prediction errors for individual Δλ.
- Full wavelength results were Δλ independent when the sources are fused.
- AUC performed better than PLS at the bands, in general
- Low level fusion was showcased to be an effective means to bypass the Δλ selection process.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. CHE-1506417 (co-funded by CDS&E) and is gratefully acknowledged by the authors.