

Chemometrics

Technical Note



Spectroscopy Best Practices

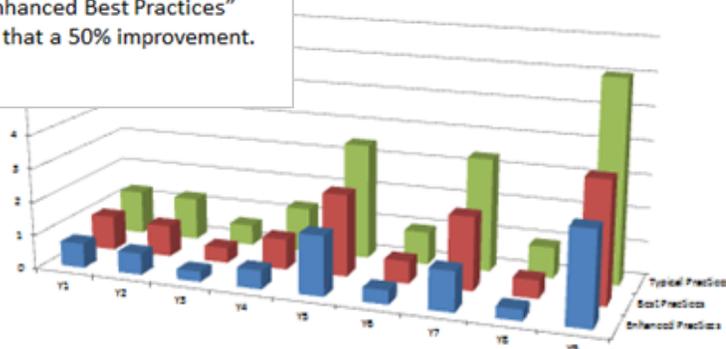
Abstract

Chemometric models used to process data from optical spectroscopy require oversight and occasional maintenance. If the initial calibration is performed carefully, the models may minimize this review effort, but it will not eliminate it. We recommend performing a detailed evaluation of the chemometric model in use by considering the following diagnostics, during the model construction, the validation and in prediction.

The result of this work will be an optimized model that can be integrated into any legacy system (hardware and software); there is no additional hardware or software that is needed. The example below illustrates the possibilities that can be achieved by application of best practices.

Modeled Property	Typical Practices	Best Practices	Enhanced Practices
Y1	1.31	1.03	0.74
Y2	1.27	0.93	0.64
Y3	0.61	0.45	0.31
Y4	1.32	0.92	0.56
Y5	3.46	2.45	1.78
Y6	0.98	0.67	0.42
Y7	3.37	2.18	1.19
Y8	0.92	0.54	0.35
Y9	6.02	3.61	2.77

A consistent application of "Best Practices" yields a 30% reduction in the error of analysis, given the same starting data. The "Enhanced Best Practices" garners more than a 50% improvement.



Current Best Practices

- Percent variance – a model diagnostic to help determine the relevant number of factors to include in the model;
- Comparison of predicted versus known – a model, a sample, and a prediction diagnostic to look for systematic patterns that deviate from the ideal line;
- Sample consistency – a sample diagnostic based on Robust PCA or Cooks ratio to identify samples that are in conflict with the model;
- Scores analysis – a sample diagnostic to determine the completeness of the training set and to identify constraints that should be placed on the model;
- Concentration residual analysis – a variable, a model and a sample diagnostic to look for systematic deviations and correlations with other parameters, identify non-instrument measurements that should be added to the model, and variables that are unusual across all samples (e.g., excessive noise);
- Studentized residual and leverage – a sample and a prediction diagnostic used to identify unusual samples or samples that have undue impact on the model;
- Loadings analysis – both a variable and a model diagnostic to support the evaluation of the degree of complexity (number of factors) to include in the model and to assess important and unusual variables;
- Measurement residual evaluation – a sample and a prediction diagnostic to identify outliers and look for structured patterns that may represent unmodeled chemical information;
- Root mean square error of prediction – a model diagnostic to establish the point where the error values are comparable to known (reference) values; and
- Fisher ratio – a sample and a prediction diagnostic that identifies unusual validation samples and ties closely to the reliability of future predictions.

Enhancements to Best Practices

- Topological regression – a following of complex, non-linear reaction models with an adaptive algorithm that determines the most appropriate model on a sample by sample basis using locally-weighted regression;
- Hierarchical modeling – a technique for classification and regression in any combination to follow reactions where the input material or the conditions may vary;
- Spectral TuneUp™ – a reduction of the impact of instrument imprecision that is not related to the chemistry of the blend by minimizing variance resulting from both temperature shifts and imprecision in spectral registration;
- Calibration transfer – an accommodation of instrumental differences when a chemometric model is built on one analyzer and then deployed for routine use on a daughter instrument; and
- Feed-forward chemometrics – a means of fusing information available in earlier parts of the process to guide selection of model attributes.

Accessing the Tools

Infometrix® has more than 30 years of experience in developing and optimizing multivariate models for use in quality control and production units. As we evaluate and validate new tools and concepts, we incorporate the improvements into our software. These tools can be implemented in the software of any partner company as well by incorporating our DLL.