Automatic NIR Model Development

Christoph Jansen

Metrohm Schweiz AG, CH Zofingen

Christoph.jansen@metrohm.ch

Near-infrared (NIR) spectroscopy offers clear advantages for routine analytical applications. It is fast, sustainable, easy to perform, requires no sample preparation, and generates no typical analytical waste. In addition, multiple parameters can be quantified simultaneously within seconds.

The development of identification models for raw materials is generally less effort and will not be addressed in this contribution.

In contrast, quantitative NIR methods still face several practical bottlenecks. For many applications, the main challenge is the effort required to develop models that allow the determination of all desired parameters. As a secondary analytical method, NIR analysers must be trained with spectra covering relevant concentration ranges and require reference values from primary analytical techniques. Collecting sufficient samples and reliable reference data generates cost and is often time-consuming.

For frequently used standard applications, instrument manufacturers often provide pre-calibrated systems. This represents the simplest way for end users to benefit from NIR spectroscopy and is widely established in markets such as food, fuels, polymers, and related industries.

The methods for creation of calibration models can be considered an early form of artificial intelligence. Compared to modern AI approaches, chemometrics typically achieves accurate results with relatively small data sets. Chemometrics is a statistical method that has been established since the 1960s.

One of the main challenges in model development is obtaining enough samples across the desired concentration range, as well as generating accurate reference values. While these steps remain demanding, the mathematical processing involved in building chemometric models can already be largely automated. This represents an important step toward automated method development.

If the final product is manufactured by mixing defined components, it is possible to artificially generate samples covering the required concentration range. This approach often results in surprisingly precise calibration models for liquid mixtures.

Using titration equipment, small amounts of concentrated components can be automatically added to the original matrix or diluted in a controlled manner. This enables the simulation of concentration distributions within the manufacturing tolerance range. If the final product is produced from the same components, matrix effects are minimal and do not limit model performance.

With this concept, high-quality NIR calibration models with low standard errors can be generated automatically, often within a single overnight run.