COMPARISON OF LINEAR AND NON-LINEAR NIR CALIBRATION METHODS USING LARGE FORAGE DATABASES

Paolo Berzaghi[*1,2], Peter C. Flinn[3], Pierre Dardenne[4], Martin Lagerholm[5], John S. Shenk[6], Mark O. Westerhaus[6] and Ian A. Cowe[5].

1 University of Padova, Agripolis, 35020 Legnaro Italy.
2 University of Wisconsin, 1925 Linden Dr., 53706 Madison, WI, USA.
3 Agriculture Victoria, Pastoral and Veterinary Institute, Private Bag 105, Hamilton, Victoria 3300, Australia.
4 Centre de Recherches Agronomiques de Gembloux CRAGx, 24, Chaussee de Namur, 5030 Gembloux, Belgium.
5 Foss Tecator AB, Box 70, SE-263 21 Höganäs, Sweden. 6 Infrasoft International, 109 Sellers Lane, 16870 Port Matilda, PA, USA.

The aim of the study was to evaluate the performance of 3 calibration methods, modified partial least squares (MPLS), local PLS (LOCAL) and artificial neural network (ANN) on the prediction of chemical composition of forages, using a large NIR database. The study used forage samples (n=25,977) from Australia, Europe (Belgium, Germany, Italy and Sweden) and North America (Canada and U.S.A) with information relative to moisture, crude protein and neutral detergent fibre content. The spectra of the samples were collected with 10 different Foss NIRSystems instruments, which were either standardized or not standardized to one master instrument. The spectra were trimmed to a wavelength range between 1100 and 2498 nm. Two data sets, one standardized (**IVAL**) and the other not standardized (**SVAL**) were used as independent validation sets, but 10% of both sets were omitted and kept for later expansion of the calibration database. The remaining samples were combined into one database (n=21,696), which was split into 75% calibration (**CALBASE**) and 25% validation (**VALBASE**). The chemical components in the 3 validation data sets were predicted with each model derived from **CALBASE** using the calibration database before and after it was expanded with 10% of the samples from **IVAL** and **SVAL** data sets. Calibration performance was evaluated using standard error of prediction corrected for bias (SEP(C)), bias, slope and R2. None of the models appeared to be consistently better across all validation sets. **VALBASE** was predicted well by all models, with smaller SEP(C) and bias values than for **IVAL** and **SVAL.** This was not surprising as **VALBASE** was selected from the calibration database and it had a sample population similar to **CALBASE,** whereas **IVAL** and **SVAL** were completely independent validation sets. In most cases, Local and ANN models, but not modified PLS, showed considerable improvement in the prediction of **IVAL** and **SVAL** after the calibration database had been expanded with the 10% samples of **IVAL** and **SVAL** reserved for calibration expansion. The effects of sample processing, instrument standardization and differences in reference procedure were partially confounded in the validation sets, so it was not possible to determine which factors were most important. Further work on the development of large databases must address the problems of standardization of instruments, harmonization and standardization of laboratory procedures and even more importantly, the definition of the database population.