

Eliminating External Interferences in Handheld NIR Spectroscopy for Grape Quality Assessment

Pascal Gauweiler¹, Xiaorong Zheng², Anna Kicherer² and Robin Gruna¹

One indicator for determining the perfect time for harvesting grapes to achieve the best quality wine is the sugar and acid content. Traditional methods are costly and time-consuming, whereas handheld NIR spectroscopy combined with chemometric methods are fast and non-destructive. To use this, a calibration process is required to extract the quality parameter from the spectrum. This involves first collecting spectral data and then determining the corresponding reference data, which is often done in a controlled laboratory environment. In practical application, external parameters from an uncontrolled environment such as the field, disturb the spectral signal, reducing the performance of the model and even making it unusable in some cases.

Types of Interferences

Internal interferences are caused by the diversity of fruits not fully included in the calibration set like differences in grape variety, cultivation region, surface texture, skin thickness [Zh24]. To compensate these influences, a sufficient data set with great diversity is needed or more complex or adaptive models can find a better correlation with less data.

External interferences include every influence besides the test sample itself whereas environmental influences play a significant role. Temperature introduces non-linearities due to spectral peak shifts, which lead to an altering prediction bias. Ambient light is another factor as it passes the sample, changing the measured spectra [Zh24]. Laboratory conditions have far less variation than outdoor areas, where sunlight and its spectral composition is changed by clouds.

Eliminating External Interferences

To eliminate external interferences, several methods have been proven. External Parameter Orthogonalization (EPO) allowing the removal of variations in the independent variable (the spectra) caused by external factors [Ro03]. These variations can be identified by comparing data from a controlled (lab) and uncontrolled (field) environment. The result is a correction matrix that is used to transform new spectra. A limitation is that the correction matrix is not adaptive and works only within the domains where the influenced spectra were recorded.

Test scenario for Grapes: Sample Acquisition and Modeling

By focusing on a single variety ('Weissburgunder'), the biological diversity is minimized. The samples were collected at 4 days each at the same time of day. The berries were spectrally captured in-field, picked and spectrally captured again in the laboratory using NIRONE S1.4 Sensor (1100-1350 nm, 511 channels) by Spectral Engines (Helsinki, Finland). The dataset comprises 127 sample pairs. For pre-processing, SNV and Savitzky-Golay filtering was used and Partial Least Squares Regression (PLSR) model was selected for its effectiveness and minimal parameters. To obtain reference values for glucose in individual samples, the samples were prepared and analyzed with HPLC-DIR.

Experiment 1

To show the effects of prediction results when the data are interfered by external parameters, a model was generated using all laboratory data and tested against the field data.

Experiment 2

The training set contained 64 lab and 64 field spectra, and test set contained 63 lab and 63 field spectra. The EPO matrix was fitted using the training data, and a model was developed trained on the transformed train data and tested on transformed test data.

Results

After applying corrections, Principal Component Analysis (PCA) shows a two-orders-of-magnitude reduction in variance, indicating that most of the variance due to external disturbances has been successfully removed.

The RMSEP for field data without correction is 142.18 g/L and R^2 of -5.15, while it improves to a RMSEP of 17.49 g/L and an R^2 of 0.74 with EPO correction (Fig. 1). This demonstrates that the model trained without any corrected data has significant bias and offset. An accuracy of 17.49 g/L remains slightly insufficient in practical applications, but it reflects the potential of the method. It highlights the opportunities for further improvements with more complex models and a greater data set.

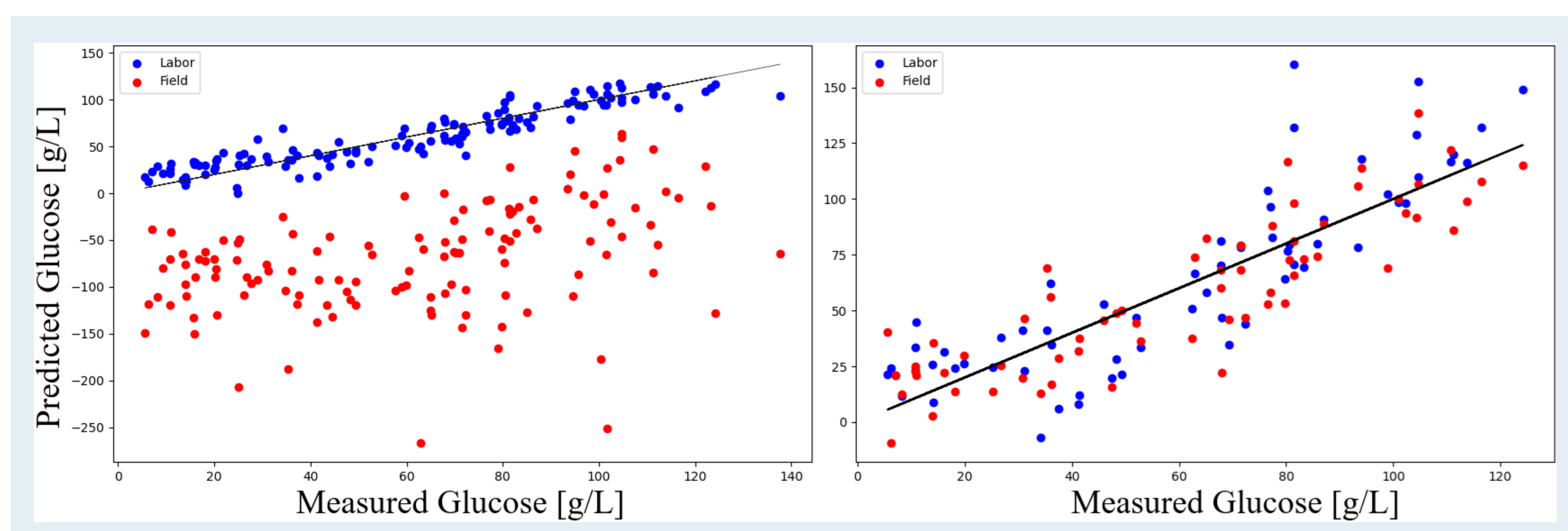


Figure 1: Scatterplots before and after EPO correction of test data of laboratory data (blue) and of field data (red).

Metrics without correction: RMSEP=142.18g/L R^2 =-5.15; with EPO-correction: RMSEP=17.49g/L R^2 =0.74

Contact

Pascal Gauweiler
Sichtprüfsysteme (SPR)
Tel. +49 721 6091-403
pascal.gauweiler@iosb.fraunhofer.de
Fraunhofer IOSB
Fraunhoferstr. 1
76131 Karlsruhe
www.iosb.fraunhofer.de

¹ Fraunhofer Institute for Optronics, System Technologies and Image Exploitation, Visual Inspection Systems SPR, Fraunhoferstr. 1, 76131 Karlsruhe, Germany

² Julius-Kühn-Institut: Institute for Grapevine Breeding, Geilweilerhof, 76833 Siebeldingen, Germany

[Zh24] Zhao, Yiyi: Visible/near-infrared Spectroscopy and Hyperspectral Imaging Facilitate the Rapid Determination of Soluble Solids Content in Fruits. Food Engineering Reviews 3/24, pp. 470-496, 2024

[Ro03] Roger, Jean-Michel; Chauchard, Fabien; Bellon-Maurel, Véronique: EPO-PLS external parameter orthogonalisation of PLS application to temperature-independent measurement of sugar content of intact fruits. Chemometrics and Intelligent Laboratory Systems, 2/03, pp. 191-204, 2003