IDC-improved direct calibration. Application to wine ethanol quantification in musts and wines

Jean Claude Boulet, Jean-Michel Roger

To cite this version:


HAL Id: hal-02750574
https://hal.inrae.fr/hal-02750574
Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
IDC-Improved Direct Calibration. Application to ethanol quantification in musts and wines

Jean-Claude BOULET*, Jean-Michel ROGER**
*INRA, UMR-SPO, F-Montpellier
**CEMAGREF, UMR-ITAP, F-Montpellier

Theory

\[
x'_{P} = y * k'_{P} + T_{x}\,*\,Q + K + T_{v}\,*\,A^*\,P
\]

Sample spectrum
Useful part
Chemical influence factors
Other influence factors

IDC principle:
to remove the influences from the useful part

Application

Quantification of Ethanol (% vol.) on clarified grape musts during alcoholic fermentation, using NIR spectrometry (500 to 1900 nm / 2nm)

- collect \(X_{G}\) (samples with constant \(y\))
- identify \(P\) with a SVD onto \(X_{G}\)
- project onto \(k\) orthogonally to \(K\) and \(P\)

\[
R = \begin{bmatrix} K \\ P \end{bmatrix}, \quad \Sigma = I - R'(RR')^{-1}R
\]

\[
b = \Sigma \, k \, (k' \, \Sigma \, k)^{-1}
\]

PRACTICAL ASPECTS

- IDC doesn't need any calibration dataset (contrary to PLSR) → possible application to hyperspectral images
- IDC is very flexible: chemical compounds whose pure spectra are unknown can be characterised by \(X_{G}\)
- Different ways to build \(X_{G}\) (see [1], [2], [3])
- IDC predictions can be as accurate as PLSR predictions
  - when \(X_{G}\) is calculated according to [2] or [3], constant but not zero influence factors are not taken into account, leading to IDC models with wrong slopes and/or offsets

THEORETICAL ASPECTS

The term \(\Sigma k\) corresponds to Lorber's definition of the NAS-Net Analyte Signal ([4]): « the net analyte signal may be computed as the part of its spectrum orthogonal to other coexisting constituents », extended all other -e.g. physical- influence factors.

Despite similar predictions, IDC and PLSR-b-coefficients vectors are different, and don’t converge (not shown).