Removing spatial effects of spectral dataset acquired into an experimental design by using multivariate analysis of variance

Maxime Ryckewaert, Anis Taleb Bendiab, Fabienne Henriot, Ryad Bendoula

Removing spatial effects of spectral dataset acquired into an experimental design by using multivariate analysis of variance.

Ryckewaert Maxime*[12], Lamour Julien[23], Taleb-Bendiab Anis[2], Fabienne Henriot[1], Bendoula Ryad[2]

1. *Limagrain Europe, Clermont-Ferrand, France;*

2. *UMR ITAP, IRSTEA, Montpellier, France*

3. *La Compagnie Fruitière, Marseille, France*

Keywords: Chemometrics, experimental design, maize, spectrometry, reflectance spectra

Abstract:

During an experiment, the presence of spatial autocorrelation can undermine plant response interpretation *(Ball et al., 1993; Legendre et al., 2004)*. Even by using blocking design and randomization, measurement can have a spatial dependence *(Oehlert, 2000)*. This may reduce the utility of applying the classical analysis of variance (ANOVA). Generalized linear mixed models (GLMMs) can handle the spatial dependence as random effects *(Bolker et al., 2009)*. For multiple responses, GLMMs can be used as in longitudinal analysis*(Liang and Zeger, 1986; Zeger and Liang, 1986)*. But GLMMs are generally inappropriate for multivariate data especially with the spectral data.

In agronomy, spectral data is increasingly being used *(Araus and Cairns, 2014)*. High-spectral resolution provides more information and can be valuable for disease detection *(Mahlein et al., 2012)*, for crops monitoring *(Ecarnot et al., 2013; Mahajan et al., 2016; Vigneau et al., 2011)* and early detection of water-stress *(Behmann et al., 2014; Römer et al., 2012)*. To take advantage of an experimental design with multivariate data such as spectral data, chemometrics methods have been developed *(Brereton et al., 2017)*. Among these methods, ANOVA-Simulatenous component analysis (ASCA) *(Smilde et al., 2005)* gives statistical significance of the factors and can reveal their interdependencies.

In this presentation, we describe the possibility to use such approaches in agronomy to study spatial effect of maize spectra included in an experimental design. An example is given by applying ASCA method on a dataset of 480 spectra where 10 genotypes are confronted by 2 irrigation treatment. Spectra are acquired with high-spectral resolution (256 bands) in the range of 310 nm – 1100 nm. The obtained results confirm that this class of methods is relevant to identify variability due to spatial effect. Describing the spectral region due to this effect extends the possibility to remove unwanted effects due to spatial dependence.

References

Araus, J.L., Cairns, J.E., 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19, 52–61. https://doi.org/10.1016/j.tplants.2013.09.008

Ball, S.T., Mulla, D.J., Konzak, C.F., 1993. Spatial heterogeneity affects variety trial interpretation. Crop Sci. 33, 931–935.

Behmann, J., Steinrücken, J., Plümer, L., 2014. Detection of early plant stress responses in hyperspectral images. ISPRS J. Photogramm. Remote Sens. 93, 98–111.

Blackburn, G., Ferwerda, J., 2008. Retrieval of chlorophyll concentration from leaf reflectance spectra using wavelet analysis. Remote Sens. Environ. 112, 1614–1632. https://doi.org/10.1016/j.rse.2007.08.005

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.-S.S., 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. 24, 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Brereton, R.G., Jansen, J., Lopes, J., Marini, F., Pomerantsev, A., Rodionova, O., Roger, J.M., Walczak, B., Tauler, R., 2017. Chemometrics in analytical chemistry—part I: history, experimental design and data analysis tools. Anal. Bioanal. Chem. 409, 5891–5899.

Hansen, P.M., Schjoerring, J.K., 2003. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Remote Sens. Environ. 86, 542–553.

Legendre, P., Dale, M.R., Fortin, M.-J., Casgrain, P., Gurevitch, J., 2004. Effects of spatial structures on the results of field experiments. Ecology 85, 3202–3214.

Liang, K.-Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.

Oehlert, G.W., 2000. A first course in design and analysis of experiments. W.H. Freeman, New York.

Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Panigada, C., Behmann, J., Léon, J., Thurau, C., Bauckhage, C., others, 2012. Early drought stress detection in cereals: simplex volume maximisation for hyperspectral image analysis. Funct. Plant Biol. 39, 878–890.

Smilde, A.K., Jansen, J.J., Hoefsloot, H.C.J., Lamers, R.-J.A.N., van der Greef, J., Timmerman, M.E., 2005. ANOVA-simultaneous component analysis (ASCA): a new tool for analyzing designed metabolomics data. Bioinformatics 21, 3043–3048. https://doi.org/10.1093/bioinformatics/bti476

Zeger, S.L., Liang, K.-Y., 1986. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 121–130.