Defining multivariate raw material specifications via SMB-PLS

J. Borràs-Ferrís[1] C. Duchesne[2] A. Ferrer[3]

[1] Multivariate Statistical Engineering Group, Department of Applied Statistics and Operational Research and Quality, Universitat Politècnica de València, València, Spain, joaborfe@eio.upv.es

[2] Chemical Engineering Department, Laval University, Quebec, Canada, Carl.Duchesne@gch.ulaval.ca

3 Multivariate Statistical Engineering Group, Department of Applied Statistics and Operational Research and Quality, Universitat Politècnica de València, València, Spain, aferrer@eio.upv.es

Keywords: Raw Material Multivariate Specifications, Sequential Multi-block Partial Least Squares Regression, Industry 4.0.

1 Introduction

Raw materials properties are usually considered as Critical Input Parameters (CIPs) because their variability has an impact on Critical Quality Attributes (CQAs) of the final product. Despite their importance, specifications are usually defined in an arbitrary way based mostly on subjective past experience, instead of using a quantitative objective description of their impact on CQAs. Furthermore, in many cases, univariate specifications on each property are designated, with the implicit assumption that these properties are independent from one another. However, multivariate specifications provide much insight into what constitutes acceptable raw material batches when their properties are correlated (as usually happens) [1]. To cope with this correlation several authors suggest using multivariate approaches, such as Partial Least Squares (PLS) regression, to improve the definition of raw materials specifications [1], [2]. Two approaches emerge from the literature when using PLS [3]. The first is based on a direct mapping in the latent space, and the second rather defines specifications by the PLS model inversion. The latter is used in this work as it also allows integrating the prediction uncertainty in the model inversion as we propose. Thus, specifications are defined analytically by considering a probabilistic approach being precisely linked with CQAs.

Nevertheless, not only raw material properties (Z) influence the quality of the final product, but also the process conditions (X) and, thus, raw material specification regions depend on these conditions [2]. Sequential Multi-block PLS (SMB-PLS) algorithm has very interesting advantages to address such a problem, since it allows considering the sequential nature of the blocks (Z and X) when modeling the relationship with the quality (Y) [4]. Thus, specification regions are defined considering the control actions of the process conditions taken to compensate for raw material variations. In this work, we propose a novel methodology based on SMB-PLS that not only considers the control actions from the past, but also improves them in the future. All of this will allow broader ranges in raw material specifications.

2 Material and methods

The SMB-PLS is a latent variable-based method that finds the main driving forces (i.e., latent variables) on the input space that are most related to the output space, being both spaces projected into a common latent space. Thus, it is used not only to model the inner relationships between the matrix of inputs (Z and X) and the matrix of output variables (Y), but also to provide a model for both. This fact gives them a very nice property: uniqueness and causality in the reduced latent space no matter if the data come either from a Design of Experiments (DOE) or daily production process (historical/happenstance data) typical in Industry 4.0 [5].

In addition, SMB-PLS imposes a sequential pathway between the regressor blocks according to the process flowsheet, and then uses orthogonalization to separate correlated information between the blocks from orthogonal variations. Hence, variations in process conditions that are correlated with raw material properties (e.g., control actions) are extracted in a first block of latent variables, and orthogonal variations in process conditions (i.e., uncorrelated with raw material properties) are then extracted in a second block (e.g., operator actions, known disturbances, etc.).

3 Results and discussion

The methodology developed was applied to a real case study from an industrial extrusion process [2]. In this case, 13 raw material properties, 2 process variables and a CQA were measured for 55 batches. A SMB-PLS model is fitted with two latent variables for the first latent block (raw material properties with correlated process conditions) and one latent variable for the second block (uncorrelated process conditions). The model presents a goodness of fit of 90.60% when considering both blocks.

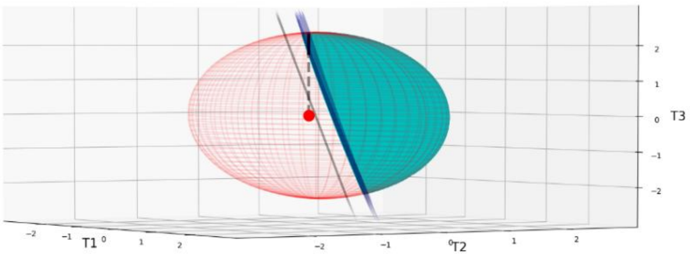

Figure 1 – Score plot showing the raw material specification region.

Figure 1 shows raw material specification in the latent space providing assurance of quality with a certain confidence level for the CQA (blue region). Besides, the red dot represents a new raw material batch when the control actions of process conditions remain in place as the past (i.e., the third component is null). In such a case, this batch would be outside the specification region. However, if the third component is modified orthogonally, such batch can become part of the specification region (black line). This is a batch that, a priori, would give place to an unacceptable value, but that by improving the control actions it would turn into an acceptable value.

4 Conclusion

In this work, we propose a novel approach for defining multivariate raw material specifications considering the possibility to modify process conditions to compensate for raw material properties variations. It would allow evaluating the capability of the raw material batches of producing product with CQAs within specification limits, before producing a single unit of the product.

5 References

[1] J. A. De Smet, "Development of Multivariate Specification Limits Using Partial Least Squares Regression," McMaster University, Hamilton, Ontario, Canada, 1993.

[2] C. Duchesne and J. F. MacGregor, "Establishing Multivariate Specification Regions for Incoming Materials," *J. Qual. Technol.*, 36, 1, 78–94, 2004.

[3] A. Paris, C. Duchesne, and É. Poulin, "Establishing Multivariate Specification Regions for Incoming Raw Materials Using Projection to Latent Structure Models: Comparison Between Direct Mapping and Model Inversion," *Front. Anal. Sci.*, 1, 1–15, 2021.

[4] K. Azari, J. Lauzon-Gauthier, J. Tessier, and C. Duchesne, "Establishing multivariate specification regions for raw materials using SMB-PLS," *IFAC-PapersOnLine*, 48, 8, 1132–1137, 2015.

[5] J. F. MacGregor, M. J. Bruwer, I. Miletic, M. Cardin, and Z. Liu, "Latent Variable Models and Big Data in the Process Industries," *IFAC-PapersOnLine*, 48, 8, 520–524, 2015.