Abstracts of the Ninth Workshop on Simulation (2018) 1-2

Run-length performance estimation of the functional regression control chart

Fabio Centofanti [1], Antonio Lepore [1] Alessandra Menafoglio [2], Biagio Palumbo [1] and Simone Vantini [2]

Because of advances in data acquisition technologies, it is increasingly common, in diverse fields of science, to gather data that are functional in nature. *Profile monitoring* [3] is an expanding area of statistical process control (SPC) whose methods allow monitoring and controlling process quality characteristics which can be modelled as *functional data* [6]. In many practical situations, measures of the quality characteristic are available jointly with those of other functional covariates. In this context, we propose a new control chart that takes into account additional information contained in those functional covariates to better monitoring the quality characteristic. This chart is referred to as *functional regression control chart* (FRCC) and is an extension of the *regression*, or *cause-selecting*, *control chart* which is used in the multivariate context [2].

Let $Y(t)$ represent the functional quality characteristic (hereinafter referred to as *response*) and $\boldsymbol{X}(t) = (X_1(t), \ldots, X_p(t))^\top$ be the functional covariates (hereinafter referred to as *predictors*) with $t \in \mathcal{S}$, a compact set in $\mathbb{R}$. We model the predictors as influencing the response according to the following multivariate non-concurrent (i.e., the response at time t is explained by the functional covariate values in $\mathcal{S}$ and not only at the same time t) functional-on-functional (i.e., functional response on functional covariates) linear regression model

$$Y^Z(t) = \int_{\mathcal{S}} (\boldsymbol{\beta}(s,t))^\top \boldsymbol{X}^Z(s)\, ds + \varepsilon(t), \tag{1}$$

with $Y^Z(t)$ and $\boldsymbol{X}^Z(t) = \left(X_1^Z(t), \ldots, X_p^Z(t)\right)^\top$ the point-wise standardized response and predictor variables, respectively, and $\boldsymbol{\beta}(s,t) = (\beta_1(s,t), \ldots, \beta_p(s,t))^\top$ the vector of the regression functional parameters.

The FRCC monitors the functional residuals of the model (1) adotpting the approach used in [4], where the coefficients from the Karhunen-Loève decomposition of the residual are monitored by means of the Hotelling's T^2 and the squared prediction error (SPE) control charts.

A simulation study is performed to quantify the FRCC performance, in terms of average run length (ARL), in identifying mean shifts in the functional response.

[1]Department of Industrial Engineering, Università di Napoli Federico II, Naples, P.le V. Tecchio 80, 80125, Italy, E-mail: fabio.centofanti@unina.it, antonio.lepore@unina.it, biagio.palumbo@unina.it.

[2]MOX - Dept. of Mathematics, Politecnico di Milano, Milan, Piazza Leonardo da Vinci 32, 20133, Italy, E-mail: alessandra.menafoglio@polimi.it, simone.vantini@polimi.it.

Simulated data are generated using pre-specified covariance structure and means reference model for the response and predictors, similar to simulation studies in [4]. Two different scenarios are considered: Scenario 1, where shifts affect only the conditional mean of the response given the predictors by means of changes in the response mean; Scenario 2, where both shifts in response and predictor means occurred. FRCC performance are compared with those of other two charts widely used to monitor profiles in the industrial context [4, 1]. Results show that the proposed control chart outperforms the competitor ones in both Scenario 1 and Scenario 2. However, in Scenario 2 shifts in the predictor means affect the FRCC performance, and conclusions about response mean shift must be drawn with caution.

Acknowledgements: The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [5]. In particular, the authors are deeply grateful to the engineer Fiorenzo Ambrosino, for his experienced and professional support in promptly solving technical issues. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see `http://www.cresco.enea.it/english` for information.

References

[1] Bianca Maria Colosimo and Massimo Pacella. On the use of principal component analysis to identify systematic patterns in roundness profiles. *Quality and reliability engineering international*, 23(6):707–725, 2007.

[2] BJ Mandel. The regression control chart. *Journal of Quality Technology*, 1(1):1–9, 1969.

[3] Rassoul Noorossana, Abbas Saghaei, and Amirhossein Amiri. *Statistical analysis of profile monitoring*, volume 865. John Wiley & Sons, 2012.

[4] Alessia Pini, Simone Vantini, Bianca Maria Colosimo, and Marco Grasso. Domain-selective functional analysis of variance for supervised statistical profile monitoring of signal data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 2017.

[5] Giovanni Ponti et al. The role of medium size facilities in the hpc ecosystem: the case of the new cresco4 cluster integrated in the eneagrid infrastructure. In *High Performance Computing & Simulation (HPCS), 2014 International Conference on*, pages 1030–1033. IEEE, 2014.

[6] J. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2005.