Preprints, 10th IFAC International Symposium on
Advanced Control of Chemical Processes
Shenyang, Liaoning, China, July 25-27, 2018

Multiresolution Analytics for Large Scale Industrial Processes

Marco S. Reis *. Tiago J. Rato *

* *CIEPQPF - Department of Chemical Engineering, University of Coimbra*
Pólo II, Rua Sílvio Lima, 3030-790 Coimbra, Portugal (e-mail: marco@eq.uc.pt)

Abstract: Data collected from Industry 4.0 scenarios present a variety of data structures, reflecting the evolution of industrial processes, measurement systems and IT infrastructures ("variety" is actually one of the 4 V's of Big Data, meaning that its existence is widely recognized). Data analytics platforms must adapt to this context and keep the pace of its evolution, in order to continue providing effective solutions to practitioners for dealing with the large data resources now available. In this context, one prevalent feature of industrial data has been largely overlooked: their multiresolution nature. The multiresolution nature of data is directly connected to their granularity in the time domain, an aspect that induces inner dependencies that current frameworks cannot address in a consistent and rigorous way. Furthermore, multiresolution has been often mistaken as a simple multirate scenario, where in fact the meaning of the observations is completely different. In this paper, we highlight such differences and discuss current multiresolution frameworks for effectively handling industrial data sets.

Keywords: Multiresolution data; Process monitoring; Soft sensors; Kalman filter; Batch processes; Continuous processes.

1. INTRODUCTION

With the evolution in sensing technology, delocalized acquisition systems, communication infrastructures (including the Internet of Things) and storage/retrieving facilities for huge amounts of data (using cloud technology), the nature of data presented to plant engineers and data scientists has been changing significantly. Modern data sets are high-dimensional, contain structured data of different types (sensor data, spectra, images, etc.) as well as unstructured data (fault and alarm tags, operators notes, etc.), noise, outliers, missing segments, dynamic features and multiresolution characteristics that need to be properly accounted for, in order to extract meaningful and useful information for the purposes of process monitoring, diagnosis, control and optimization.

The high-dimensionality is perhaps the most well-known characteristic and has been handled through the use of projection-based (or latent variable) methods such as Principal Component Analysis (PCA) (Jackson, 1959; Jackson and Mudholkar, 1979; Jolliffe, 2002) and Partial Least Squares (PLS) (Geladi and Kowalski, 1986; Jackson, 1991; Martens and Naes, 1989). These methodologies can also account for process dynamics by extending the original data matrices with time-shifted replicates (Kaspar and Harmon Ray, 1993; Ku et al., 1995; Rato and Reis, 2013a, 2013b). The presence of noise is also properly handled by these approaches, including heteroscedastic noise (Reis and Saraiva, 2006b; Wentzell et al., 1997) and outliers (Chiang et al., 2003). On the other hand, the multiresolution aspect has received far less attention. Furthermore, it has been wrongly treated as a multirate problem. This confusion can be explained by the resemblance of the data structures produced in multiresolution and multirate settings, but also by a certain lack of awareness for the importance to deal with multiresolution systems. In fact, in both cases variables are recorded at different rates. Apart from that, the inner data structure and meaning of the observations is completely different. In the multirate case, the recorded values are just the instantaneous measurements acquired from the process, which are collected at different sampling rates. In the multiresolution case, the recorded values contain information about the process with different levels of granularity (different resolutions). Granularity can arise from the implementation of aggregation rules that merge multiple (high resolution) samples into a single (low resolution) observation. Another source of granularity can be found in situations where variables represent measurements made on composite samples collected during a certain time period (e.g., production lot, working shift, etc.), and therefore the recorded values regard specific windows of time. These windows of time are here defined as the variable's "time supports". For a further discussion on the differences between multirate and multiresolution data sets we refer the reader to (Rato and Reis, 2017c).

Even though the multirate approaches underperform when applied to multiresolution data sets, it is worth reviewing the proposals made in this field, to underline their differences with the multiresolution frameworks. The main problem encountered in multirate data is the non-uniform sampling rate. To address this, a simple solution is to downsample the more frequently observed variables and then build a model with the remaining low sampling rate data (Dongguang et al., 2003; Li et al., 2001; Lin et al., 2009). However, by discarding the inter sample observations a considerable amount of information is lost. This is particularly critical when the process has dynamic characteristics or when the data is, in fact, multiresolution. Note that, in the latter case, a low resolution variable is aggregated over a given time support (e.g., through an averaging operation) and therefore it is likely to be related with the high resolution signal of the discarded samples of other variables. To preserve the relationship with past observations, multirate approaches based on finite impulse response (FIR) models have been proposed to weight past observations before including them in the model (Wu and Luo, 2010; Xie et al., 2013). In a similar manner, Shang et al. (2015) introduced a regularization approach to dynamic partial least squares (DPLS) in order to smooth out the coefficients related to past observations. However, none of these methodologies is able to handle the multiresolution structure of data in a consistent and rigorous way.

Regarding the signal processing methodologies for optimal estimation, namely those based on the Kalman filter (KF) (Kalman, 1960), it can be once again verified that multirate approaches only account for dynamic dependencies. Examples include the proposal of Roshany-Yamchi et al. (2013), where the missing observations are given zero weights during the KF state estimation; or the approach of Wu and Luo (2010), which employs a bank of KFs depending on the available data at each time instant.

The currently miscalled "multiresolution" approaches, fail to accommodate this data structure as well. In fact, this nomenclature does not reflect the analysis being done and should be called instead, "multiscale". Multiscale methods essentially apply a wavelet transformation in order to decompose single-resolution signals into several time-frequency scales (Basseville et al., 1992; Chou et al., 1994; Stephanopoulos et al., 2008a; Stephanopoulos et al., 2008b; Willsky, 2002). This leads to a set of detail coefficients, which represent the specific contribution of each scale, and approximation coefficients, with the coarsest approximation of the original signal (low frequency bands). As coefficients at different scales are analyzed, these methodologies should be called "multiscale". A multiresolution approach analyzes signals represented at different resolutions or granularity levels (i.e., only approximation coefficients at different scales are considered, and not the detail coefficients). This is the main difference between multiscale approaches and multiresolution approaches. For instance, multiscale statistical process control (MSSPC), is a multiscale, single-resolution approach for process monitoring (Bakshi, 1998; Reis et al., 2008); on the other hand (MR-MSSPC) is a multiscale, multiresolution approach (Reis and Saraiva, 2006c).

Following the above discussion, it is now both opportune and important to incorporate multiresolution analytics in the routine analysis of industrial data. This paper reviews the current multiresolution frameworks available for handling industrial data sets. These are divided into two main categories depending on whether the data already possesses a multiresolution structure or not. In the first case, the methodologies have to accommodate for the multiresolution structure in order to take advantage of the information embedded in it. On the other hand, for the cases in which the data set is single-resolution, we have found out that there may be a potential advantage in creating a multiresolution structure in order to optimize the analysis goal. Therefore, the motivation for the second case is to explore if there is any advantage on changing the original resolution of the variables in order to increase the performance of the analytics, for instance by producing better predictive models. The methodologies in each of these categories, as well as their application scopes, are summarized in Fig. 1 and will be further discussed in the following sections.

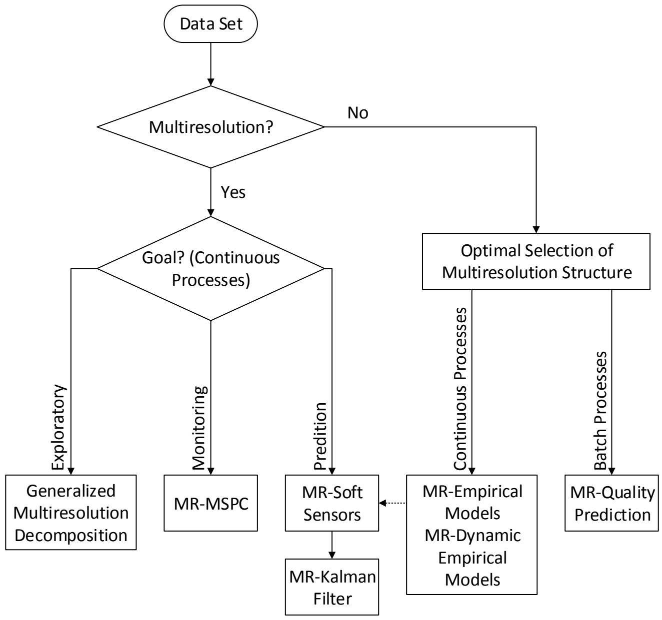

Fig. 1. Organogram of the available Multiresolution frameworks and their application scope.

The rest of this paper is organized as follows. In the next section the frameworks for analyzing data with a native multiresolution structure are discussed. Afterwards, in Section 3, we present the methodologies that actively introduce multiresolution structure into the data. The main conclusions of this work are summarized in Section 4.

2. ANALYSIS OF MULTIRESOLUTION DATA

2.1 Exploratory

Every data analysis task should start with an exploratory analysis of the collected data, with special focus on visual tools, such as graphs and diagrams. However, when data present multirate or multiresolution structures, a preprocessing stage is required, in order to make the analysis consistent in terms of the granularity of what is being portrayed. In this context, frameworks were developed that are able to project variables under analysis to the same resolution level (defined by the user). When the data is multirate, a Generalized Multiresolution Decomposition

(GMRD) framework was proposed that is able to extend Mallat's wavelet-based multiresolution decomposition (Mallat, 1989) to multirate contexts (Reis and Saraiva, 2006a). Therefore, the approximation coefficients obtained at different resolutions, provide the variable projections with different levels of granularity. Due to the multirate nature of data, the number of high-frequency observations used in the computation of a given coarser approximation coefficient may vary. This implies that these coefficients, for a given resolution, do not present all the same "quality" or uncertainty. Therefore, an uncertainty propagation computation is performed in parallel, in order to provide not only the values for the variables at the desired resolution (granularity), but also their associated uncertainty. With the values and associated uncertainties available, a detailed exploratory analysis can then be performed, using the existent rich graphical toolkit and good visualization practices (Tufte, 2001). Fig 2. illustrates the computational scheme adopted in the implementation of GMRD.

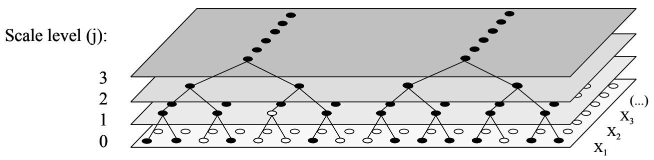

Fig. 2. The GMRD projection framework.

For multiresolution data, the approach to follow is similar in the sense of adopting a dyadic projection scheme (as the one depicted in Fig.1), but variables with coarser granularities can only be projected to resolutions even coarser than their own native resolution. The high resolution variables can be projected to all resolutions (Reis and Saraiva, 2006c). Therefore, depending on the resolution selected by the user to conduct the exploratory analysis, different sets of variables may be available, all of them consistent in terms of the variables' granularity.

2.2 Monitoring

The first (and only) methodology for handling multiresolution data for high-dimensional process monitoring was proposed by Reis and Saraiva in 2006 (MR-MSSPC) (Reis and Saraiva, 2006c). This approach extends multiscale statistical process control (Bakshi, 1998), where the process variables are simultaneously monitored at different time-scales, to a multiresolution scenario. MR-MSSPC begins with the specification of the native resolutions of each variable. Quite often there is a finest resolution, corresponding usually to the variables that are also collected at higher sampling rates. This is used to establish the finest grid of time (scale index $j = 0$). If variable X_i corresponds to averages over time supports of length 2^{J_i} times that of the finest resolution, than its scale index or resolution level is set to J_i . A variable at a resolution J_i can only be decomposed to scales coarser (i.e., higher) than J_i and therefore it does not contribute to the monitoring implemented at finer scales ($j \leq J_i$). Therefore, in MR-MSSPC not all variables are being monitored at all scales, but only at those that are coarser than their native resolutions. In this way, MR-MSPC is able to simultaneously handle the following "variety" aspects of industrial data: high-dimensionality, cross-correlation, multiscale dynamics and multiresolution structure.

2.3 Prediction

The key quality features of industrial processes are typically obtained offline with a considerable delay and by resort to expensive equipment. To avoid this experimental burden, soft sensors have been developed in order to predict the expensive quality variables based on the more frequently collected variables. However, the multiresolution structure of data raises several fundamental problems while extracting the relevant relationships between predictor and response variables. To accommodate for the presence of observations at different resolutions in soft sensor development, a new weighted PLS scheme was proposed (Rato and Reis, 2017c).

To introduce the modelling stages of multiresolution soft sensors (MR-SS), let us assume that the predictors are readily obtained at high resolution ($\mathbf{x}_t^{(0)}$), while the response ($y_t^{(r)}$) is a low resolution variable with time support 2^r (i.e., the response is only observed at every 2^r-th observation and the recorded value is the result of aggregating high resolution data over a window of 2^r time instants). The superscript "(r)" represents the variable's resolution.

Due to the multiresolution nature of the data, the observed low resolution response at time t ($y_t^{(r)}$) is inherently linked to an unknown high resolution version of itself ($\tilde{y}_t^{(0)}$) within its time support. This relationship is represented by,

$$y_t^{(r)} = w\tilde{y}_t^{(0)} + w\tilde{y}_{t-1}^{(0)} + \ldots + w\tilde{y}_{t-2^r+1}^{(0)}, \tag{1}$$

where w is a weighting constant (usually $w = 1/2^r$, representing an average operation). While the high resolution signal of the response is unknown, it can still be considered that the predictors (which are also at the highest resolution) can be used to estimate it through a linear relationship, such that (left side of (2)),

$$\begin{cases} y_t^{(r)} = \sum_{i=0}^{2^r-1} w\tilde{y}_{t-i}^{(0)} \\ \tilde{y}_t^{(0)} = \mathbf{b}\mathbf{x}_t^{(0)} \end{cases} = \begin{cases} y_t^{(r)} = \mathbf{b}\mathbf{u}_t^{(0)} \\ \mathbf{u}_t^{(0)} = \sum_{i=0}^{2^r-1} w\mathbf{x}_{j,t-i}^{(0)} \end{cases}, \tag{2}$$

where $\mathbf{b}$ is a vector of regression coefficients and $\mathbf{u}_t^{(0)}$ represents the sum of weighted predictors over the time support of the response. However, since the response's high resolution signal is unknown, the high resolution model cannot be directly fitted. For such, it is necessary to recast the model in its equivalent low resolution format, as presented in the right side of (2). The low resolution model can then be estimated by fitting a PLS model between $y^{(r)}$ and $\mathbf{u}^{(0)}$. This formulation can be further generalized for handling predictors at different resolutions, to model dynamic relationships as well as to address the use of unknown averaging weights (w) (Rato and Reis, 2017c).

The main advantages of this methodology is its ability to explicitly model the multiresolution structure of the data, through a more parsimonious model, which in turn translates into better estimates of the response. Furthermore, even though the model is fitted on the low resolution response, the equivalency between the high and low resolution models implies that the parameters of the high resolution model are also readily available. Therefore, it is possible to estimate (or reconstruct) the initially unknown high resolution signal of the response. This is a quite interesting aspect of the proposed framework, opening new perspectives to increase the resolution and frequency of the response estimates, through multiresolution analytics.

Another application of the weighted PLS is on the development of optimal multiresolution estimation frameworks that extend the single-resolution KF to the multiresolution scenario. In this regard, a multiresolution Kalman filter (MR-KF) can be used to optimally fuse the information conveyed by (i) the high resolution model (left side of (2)), (ii) the low resolution model (right side of (2)), (iii) the low resolution observations and (iv) the high resolution observations. Note that the high resolution observations are not strictly necessary for applying MR-KF, and in fact they are often unavailable. However, even in this scenario the MR-KF can still be used and benefit from the fusion of the other three sources of information.

The proposed implementation for the MR-KF has roots on the works of (Basseville et al., 1992; Chou et al., 1994; Stephanopoulos et al., 2008a; Stephanopoulos et al., 2008b; Willsky, 2002) and consist of applying two KF that exchange information with each other. The first stage of the MR-KF is a high-to-low sweep that employs a high resolution KF to merge the information at the high resolution level (i.e., high resolution observations and the estimates of the high resolution model). The output is a filtered high resolution estimate. This estimate is then send to a low resolution KF that merges it with the remaining low resolution information (i.e., low resolution observations and the estimates from the low resolution model) to generate an optimal low resolution estimate. Afterwards, a low-to-high smoothing stage is implemented in order to smooth the previous high resolution estimates based on the optimal low resolution estimate, by means of the Rauch-Tung-Striebel algorithm (Rauch et al., 1965).

The MR-KF was tested and compared against the standard single-resolution KF in simulated case studies (where the "real", noiseless response values are known) and showed to be substantially better than its counterpart, leading to estimation improvements for both the high and low resolution signals (Rato and Reis, 2017b).

3. OPTIMAL SELECTION OF THE MULTIRESOLUTION STRUCTURE

3.1 Continuous Processes

Even when the data is available at a single-resolution it is not guaranteed that their native resolution is the best resolution for achieving the analysis goals. This is particularly relevant in predictive analytics tasks. For instance, the variables may be collected at a high resolution in order to capture the local variability from the standpoint of process control or monitoring, while for process modelling it might be preferable to adopt low resolution representations in order to obtain more parsimonious and stable models. Therefore, during model building it is not only desirable to select the set of predictors most suitable for estimating the response (as in the case of classical stepwise selection methodologies), but also to consider at which resolution they should be included in the model. In this way, a framework was proposed for building multiresolution empirical model for continuous processes (MR-EMC) (Rato and Reis, 2017a) that contemplate the definition of the multiresolution structure as an additional degree of freedom for model building.

In brief terms, MR-EMC simultaneously selects the best variable to be included in the model and searches for the best resolution for each variable. The search space for the variables' resolution is constrained between the native resolution of the variables and a given maximum resolution defined by the user. One way to implement MR-EMC is by extending the predictors space with variable duplicates at all resolutions. This extended matrix is then the base for simultaneous variable and resolution selection, following a procedure similar to the standard stepwise forward algorithm (Draper and Smith, 1998). However, it should be noted that while the search space includes all duplicates at different resolutions, only one of them is included in the model (i.e., the same variable cannot included in the model at different resolutions).

To illustrate the advantage of MR-EMC, a simulated Continuous Stirred Tank Reactor (CSTR) was used to generate data and then build PLS, DPLS and MR-EMC models for estimating the output concentration. This simulator, returns readings for seven variables: (i) feed stream concentration of compound A, (ii) feed stream temperature, (iii) heating fluid inlet temperature, (iv) fluid level in the reactor, (v) outlet concentration of compound A, (vi) outlet stream temperature and (vii) heating fluid outlet temperature. The full specification of the simulated process can be found in (Rato and Reis, 2015). In this study, all variables are recorded at high resolution and subject to measurement noise. As mentioned before, for modelling purposes, the output concentration was selected as the response variable and the other six variables form the predictors set. To assess the consistency of the results, 100 replicates were made and for each replicate 1000 observations were generated for training the models and another 1000 observations for testing their performance. The performance was evaluated through the Mean Squared Error (MSE) based on the noise free values for the concentration. The median MSE over the 100 replicates is represented in Fig. 3. From this figure, it is clear that PLS presents the worst performance as it only makes use of the native resolution of the variables and has no information for modelling the process dynamics. When time-shifted replicated are added to the DPLS model, it is observed a significant improvement over PLS. However, the MSE of DPLS is still larger than that obtained with MR-EMC. As MR-EMC is not accounting for process dynamics, these results demonstrate that choosing the appropriate resolution for each variable is more important that including information about the process dynamics. In subsequent studies, it was also observed that a dynamic version of MR-EMC produces slightly better estimates, being again verified that resolution selection in the main driver for an accurate estimation of the response.

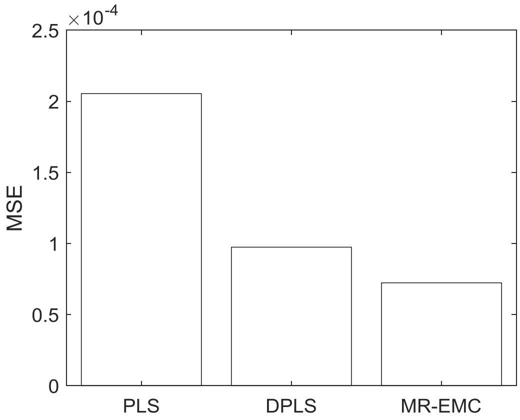

Fig. 3. Median MSE over 100 replicates of the CSTR case study. For reference, note that the variance of the noise is 4×10^{-6}.

3.2 Batch Processes

A multiresolution framework for predicting batch-end quality by exploiting the structured correlation in both the time and variables dimensions was also proposed, called *multi-resolution quality prediction* (MRQP) (Geert et al., 2017). This methodology lead to models that are often much more parsimonious than those derived from Batch Wise Unfolding (Nomikos and MacGregor, 1994; Nomikos and MacGregor, 1995), being theoretically guaranteed that they are *at least* as good as their single-resolution counterparts. From an interpretation standpoint, multi-resolution models are also more robust with respect to the selection of too many predictors, facilitating the identification of key process variables and providing information on the process time scales that influence final product quality, which can be further exploited for diagnosis, control, and optimization. This approach was tested with several systems, including simulated and real world processes. For the real world case, regarding an industrial batch polymerization process, the improvement achieved in prediction (*PRESS*) was of 54%.

4. CONCLUSIONS

Even though a large variety of industrial processes generate data with a multiresolution structure, the current modelling and analysis methodologies are not able to accommodate for this aspect, nor to take advantage of its presence. Furthermore, multiresolution structures are often erroneously taken as multirate data, leading to the adoption of inadequate analysis procedures. To address this situation, a series of multiresolution frameworks for data analysis have been proposed and are currently available. These approaches are applied to two distinct scenarios. In the first case, multiresolution is already present it the data and thus the focus is to explore the best way to incorporate it during model building. In the second scenario, data is originally single-resolution and the goal is to introduce multiresolution structure in order to improve the predictive performance. In both scenarios, the multiresolution approaches were compared against their single-resolution counterparts and consistently positive results were obtained in favor of the multiresolution frameworks.

ACKNOWLEDGEMENTS

The authors acknowledge financial support through project 016658 (references PTDC/QEQ-EPS/1323/2014, POCI-01-0145-FEDER-016658) co-financed by the Portuguese FCT and European Union's FEDER through the program "COMPETE 2020".

REFERENCES

Bakshi, B.R. (1998). Multiscale PCA with Application to Multivariate Statistical Process Control. *AIChE Journal,* 44 (7), 1596-1610.

Basseville, M., Benveniste, A., Chou, K.C., Golden, S.A., Nikoukhah, R., and Willsky, A.S. (1992). Modeling and estimation of multiresolution stochastic processes. *IEEE Transactions on Information Theory,* 38 (2), 766-784.

Chiang, L.H., Pell, R.J., and Seasholtz, M.B. (2003). Exploring process data with the use of robust outlier detection algorithms. *Journal of Process Control,* 13 (5), 437-449.

Chou, K., Willsky, A.S., and Nikoukhah, R. (1994). Multiscale systems, Kalman filters, and Riccati equations. *IEEE Transactions on Automatic Control,* 39 (3), 479-492.

Dongguang, L., Shah, S.L., Tongwen, C., and Qi, K.Z. (2003). Application of dual-rate modeling to CCR octane quality inferential control. *IEEE Transactions on Control Systems Technology,* 11 (1), 43-51.

Draper, N.R., and Smith, H. (1998). *Applied Regression Analysis* (3rd ed.). Wiley-Interscience, Hoboken, NJ.

Geert, G., Van Impe, J.F.M., and Reis, M.S. (2017). Finding the optimal time resolution for batch-end quality prediction: MRQP – a framework for Multi-Resolution Quality Prediction. *Chemometrics and Intelligent Laboratory Systems,* (Submitted, under revision).

Geladi, P., and Kowalski, B.R. (1986). Partial Least-Squares Regression: a Tutorial. *Analytica Chimica Acta,* 185, 1-17.

Jackson, J.E. (1959). Quality Control Methods for Several Related Variables. *Technometrics,* 1 (4), 359-377.

Jackson, J.E. (1991). *A User's Guide to Principal Components*. John Wiley & Sons, Inc., New York.

Jackson, J.E., and Mudholkar, G.S. (1979). Control Procedures for Residuals Associated With Principal Component Analysis. *Technometrics,* 21 (3), 341-349.

Jolliffe, I.T. (2002). *Principal Component Analysis* (2nd ed.). Springer, New York.

Kalman, R.E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME – Journal of Basic Engineering* (82 (Series D)), 35-45.

Kaspar, M.H., and Harmon Ray, W. (1993). Dynamic PLS modelling for process control. *Chemical Engineering Science,* 48 (20), 3447-3461.

Ku, W., Storer, R.H., and Georgakis, C. (1995). Disturbance detection and isolation by dynamic principal component analysis. *Chemometrics and Intelligent Laboratory Systems,* 30 (1), 179-196.

Li, D., Shah, S.L., and Chen, T. (2001). Identification of fast-rate models from multirate data. *International Journal of Control,* 74 (7), 680-689.

Lin, B., Recke, B., Schmidt, T.M., Knudsen, J.K.H., and Jørgensen, S.B. (2009). Data-Driven Soft Sensor Design with Multiple-Rate Sampled Data: A Comparative Study. *Industrial & Engineering Chemistry Research,* 48 (11), 5379-5387.

Mallat, S. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence,* 11 (7), 674-693.

Martens, H., and Naes, T. (1989). *Multivariate Calibration*. Wiley, Chichester.

Nomikos, P., and MacGregor, J.F. (1994). Monitoring Batch Processes Using Multiway Principal Component Analysis. *AIChE Journal,* 40 (8), 1361-1375.

Nomikos, P., and MacGregor, J.F. (1995). Multivariate SPC Charts for Monitoring Batch Processes. *Technometrics,* 37 (1), 41-59.

Rato, T.J., and Reis, M.S. (2013a). Advantage of Using Decorrelated Residuals in Dynamic Principal Component Analysis for Monitoring Large-Scale Systems. *Industrial & Engineering Chemistry Research,* 52 (38), 13685-13698.

Rato, T.J., and Reis, M.S. (2013b). Fault detection in the Tennessee Eastman process using dynamic principal components analysis with decorrelated residuals (DPCA-DR). *Chemometrics and Intelligent Laboratory Systems,* 125, 101-108.

Rato, T.J., and Reis, M.S. (2015). Multiscale and megavariate monitoring of the process networked structure: M2NET. *Journal of Chemometrics,* 29 (5), 309-322.

Rato, T.J., and Reis, M.S. (2017a). Multiresolution Frameworks for Empirical Model Building in Continuous Processes: MR-EMC and MR-DEMC. Under Submission.

Rato, T.J., and Reis, M.S. (2017b). Multiresolution Kalman Filter (MR-KF) for Optimal Fusion of Multiple Industrial Signals with Different Resolutions. Under Submission.

Rato, T.J., and Reis, M.S. (2017c). Multiresolution Soft Sensors (MR-SS): A New Class of Model Structures for Handling Multiresolution Data. *Industrial & Engineering Chemistry Research,* 56 (13), 3640–3654.

Rauch, H.E., Striebel, C.T., and Tung, F. (1965). Maximum likelihood estimates of linear dynamic systems. *AIAA Journal,* 3 (8), 1445-1450.

Reis, M.S., Bakshi, B.R., and Saraiva, P.M. (2008). Multiscale statistical process control using wavelet packets. *AIChE Journal,* 54 (9), 2366-2378.

Reis, M.S., and Saraiva, P.M. (2006a). Generalized Multiresolution Decomposition Frameworks for the Analysis of Industrial Data with Uncertainty and Missing Values. *Industrial & Engineering Chemistry Research,* 45, 6330-6338.

Reis, M.S., and Saraiva, P.M. (2006b). Heteroscedastic Latent Variable Modelling with Applications to Multivariate Statistical Process Control. *Chemometrics and Intelligent Laboratory Systems,* 80, 57-66.

Reis, M.S., and Saraiva, P.M. (2006c). Multiscale Statistical Process Control with Multiresolution Data. *AIChE Journal,* 52 (6), 2107-2119.

Roshany-Yamchi, S., Cychowski, M., Negenborn, R.R., Schutter, B.D., Delaney, K., and Connell, J. (2013). Kalman Filter-Based Distributed Predictive Control of Large-Scale Multi-Rate Systems: Application to Power Networks. *IEEE Transactions on Control Systems Technology,* 21 (1), 27-39.

Shang, C., Huang, X., Suykens, J.A.K., and Huang, D. (2015). Enhancing dynamic soft sensors based on DPLS: A temporal smoothness regularization approach. *Journal of Process Control,* 28, 17-26.

Stephanopoulos, G., Karsligil, O., and Dyer, M. (2008a). Multiscale theory for linear dynamic processes: Part 2. Multiscale model predictive control (MS-MPC). *Computers & Chemical Engineering,* 32 (4–5), 885-912.

Stephanopoulos, G., Karsligil, O., and Dyer, M.S. (2008b). Multiscale theory for linear dynamic processes: Part 1. Foundations. *Computers & Chemical Engineering,* 32 (4–5), 857-884.

Tufte, E.R. (2001). *The Visual Display of Quantitative Information* (2nd. ed. ed.). Graphics Press, Cheshire, Connecticut.

Wentzell, P.D., Andrews, D.T., Hamilton, D.C., Faber, K., and Kowalski, B.R. (1997). Maximum Likelihood Principal Component Analysis. *Journal of Chemometrics,* 11, 339-366.

Willsky, A.S. (2002). Multiresolution Markov models for signal and image processing. *Proceedings of the IEEE,* 90 (8), 1396-1458.

Wu, Y., and Luo, X. (2010). A novel calibration approach of soft sensor based on multirate data fusion technology. *Journal of Process Control,* 20 (10), 1252-1260.

Xie, L., Yang, H., and Huang, B. (2013). FIR model identification of multirate processes with random delays using EM algorithm. *AIChE Journal,* 59 (11), 4124-4132.