ON DATA PROCESSING AND RECONCILIATION: TRENDS AND THE IMPACT OF TECHNOLOGY

J.A. Romagnoli[a,1], P.A. Rolandi [b], Y.Y. Joe [c], Z.Q. Ding [c], K.V. Ling [d]

[a] *Department of Chemical Engineering, Louisiana State University, Baton Rouge, LA 70803, USA*
[b] *Process Systems Enterprise Limited, UK*
[c] *Singapore Institute of Manufacturing Technology, Singapore*
[d] *Nanyang Technological University, Singapore*

Abstract: The developments in technologies are expanding the boundaries and broadening the domain of what is technically and economically feasible to achieve in the application of data reconciliation activities in manufacturing plants. They also, naturally, incorporate additional issues and open the opportunities for new research activities. For example, recent developments on model-centric technologies to support plant operations based on advanced process modelling technologies opened the opportunities for performing large-scale parameter estimation – data reconciliation applications in complex dynamic industrial environments. On the other hand, new sensor technologies are becoming available based on recent advances in microprocessor-based instrumentation and digital communications. They provide opportunities for the realization of novel sensor network architectures towards a truly distributed environment for data processing and reconciliation. In this presentation we will discuss current research activities combining efforts in these areas towards the future operation of manufacturing plants.

Keywords: model-centric technologies, support of process operations, data processing, data reconciliation, sensor network, intelligent sensor.

1. INTRODUCTION

Rational use of the large volume of data generated by manufacturing plants requires the application of suitable techniques to improve their accuracy and to extract useful information about the operational status of the process. A number of technologies (data reconciliation, trend analysis, fault diagnosis, etc) have been the subject of active research during last 20 years with important advances. Among these data processing strategies, Data Reconciliation (DR) is one of the most typical approaches to obtain a consistent set of data.

Recent technological developments are expected to have a strong impact, broadening the domain of applications of data processing and reconciliation activities in manufacturing plants. In this presentation we will restrict the scope and will focus on the impact of two key technologies: the recent progress made towards model-centric approaches for support of manufacturing activities and the developments on sensor/sensor networks technologies expanding the capabilities of existing sensors.

During the last decade, general-purpose modelling tools have reached a level of maturity that allows the definition and solution of model-based problems of unprecedented complexity. Nowadays, state-of-the-art modelling, simulation and optimisation environments (MSOEs) have expanded their languages to account not only for the definition and solution of dynamic simulation activities, but also the declaration of dynamic optimisation and parameter estimation/data reconciliation activities with comparable generality and flexibility. However, while commercial and academic modelling technologies have largely engaged in developing frameworks and methodologies for tackling the model development process, complementary

[1] Corresponding author, e-mail: jose@lsu.edu

frameworks and mechanisms to help conceptualising and implementing "models" of process-engineering problems to support plant operation remain virtually unexplored. Progress in this direction, unquestionably, will expand the scope of what is technically feasible to achieve in the application of data processing and reconciliation activities in manufacturing plants. This will provide opportunities for performing large-scale applications within complex dynamic industrial environments and, additionally, integrating these capabilities with other activities for support of process operations into a single and consistent model-centric framework. In this work we will present a series of initiatives towards this vision.

On the other hand, recently sensors have received greater attention than in the past. This is due to: greater demands placed on all aspects of plant operation and improvements in technology. In terms of plant operation, competition has resulted in higher product quality and plant efficiency. Safety standards are constantly rising and measurements are the primary means of identifying potentially hazardous circumstances. In terms of improvements in technology, new sensor technologies are becoming available (extending the properties that can be measured, the environment in which they can be sampled). Microprocessors-based instrumentation and digital communications are having profound effect on the capability and/or functionalities of the sensor. These developments provide the opportunity for the realization of federated sensor network architectures towards a truly distributed environment for plant operation. In this presentation we will discuss a new conceptual model for the next generation of sensor devices, which incorporates activities such as DR at the sensor level thus improving diagnosis/classification and reducing the computational load at the controller levels. This type of architecture encompasses the extra capabilities required for the next generation of sensors and sensor networks and accommodates the additional demands required for modern manufacturing.

2. MODEL-CENTRIC TECHNOLOGIES AND DATA RECONCILIATION/ PARAMETER ESTIMATION

2.1 Background

Throughout the last decades, the computer-aided process engineering (CAPE) community made considerable progress in two strategic areas: the technical development and commercialisation of general-purpose modelling, simulation and optimisation environments; and the standardisation of open interface specifications for component-based process simulation. High-level equation-oriented declarative modelling languages have gained increased acceptance as the most appropriate framework to tackle the modelling process when full control over the scope and detail of the process model is required (Foss et al., 1998) because they provide the modeller with a series of sophisticated tools and mechanisms that contribute enormously to increase the efficiency of the modelling process.

An important advantage of equation-oriented modelling languages is the intrinsic independence between mathematical models and solution methods. By segregating the mathematical definition of any given model from structural, symbolic or numerical solution algorithms, a single model description can be used to accommodate for a potentially large number of complementary activities. Another major advance was the creation of high-level declarative languages to describe a wide range of advanced model-based problems such as dynamic optimisation and parameter estimation with a degree of generality and flexibility comparable to existing dynamic simulation languages. These days, commercial modelling languages have evolved into multi-purpose process-engineering modelling tools which we shall denote as "modelling, simulation and optimisation environments" (MSOEs).

As the CAPE community continues developing and validating individual process models, the incentive behind developing and implementing model-based technologies grows. In the mid 1990s, developers and end-users were confronted with the reality that the accessibility and usability of model descriptions embedded within modelling environments was very limited. To address this problem, the CAPE community initiated the CAPE-OPEN (CO) and Global CAPE-OPEN (GCO) projects. CO focussed on providing standard mechanisms to support a two-fold long term vision according to which: process modelling components (PMCs) built or wrapped upon the standard could be incorporated into process modelling environments (PMEs) straightforwardly; and model descriptions declared within PMEs supporting the standard would be accessible to external modelling tools.. This way, developers would be able to assemble software components from heterogeneous sources to solve complex model-based problems. The GCO consortium continued revising and updating existing standards and creating new ones for technologies beyond modelling and simulation. Within the scope of this work, the CO standards will be used as an enabling paradigm to support the creation of the advanced framework proposed later in this paper and as a point of reference to inspire some of its most innovative features.

2.2 *Model-Centric Framework for Support of Manufacturing Activities*

Following the previous discussion, it is clear that the creation of a model-centric framework that supports the definition of rigorous model-based activities and promotes the transfer of knowledge between complementary model-based software applications will extend the viability of model-centric technologies. In a series of papers, Rolandi and Romagnoli (2006a) presented a framework of such characteristics that enables the definition and implementation of model-based process-engineering problems typical of industrial environments.

The conceptual core of the framework was conceived according to the following vision. The framework was tailored to use mathematical models of process systems derived on the basis of mechanistic descriptions of natural phenomena. Although in principle the framework is applicable to a widespread of process systems, plant-wide models of industrial manufacturing plants were the main motivation of this work. Of course, as a result of rigorous mechanistic modelling of plant-wide industrial systems, the framework was tailored to deal with complex large-scale process models. In this work, the idea of modelling for multiple purposes was pursued, so that several model-based components were able to use a single fundamental model of the process to solve a widespread range of problems, implementing the notion of a model-centric framework. This integration crystallised the vision of a consistent solution of process-engineering problems, seeding synergistic interactions across model-based activities due to a consistent model formulation among the software components. Last but not least, the framework addressed a series of problems of relevance to industrial manufacturing operations, such as: model-based process simulation and optimisation, parameter estimation, data reconciliation and advanced process control. It is worth to emphasising, though, that the estimation/reconciliation component of interest to this work is just one of the modules of the entire framework discussed in Rolandi and Romagnoli (2006a).

Figure 1 provides a conceptual representation of how the different model-based components of the proposed model-centric framework are expected to support the operation of an industrial process system. As expected, the data pre-processing environment precedes all modules that make use of raw plant data, since it is imperative to obtain a consistent set of data by reconstruction of the process trajectories for the robust execution of any subsequent tasks. The estimation environment incorporates dynamic parameter estimation and dynamic data reconciliation activities, which make use of consistent data sets for the estimation of process operating parameters and evaluation of process measurement biases. The information gained from these activities is presented to the decision-makers, who then have a chance to make informed decisions on issues such as process instrumentation and equipment maintenance and inventory analysis. Consistent data sets are also provided to the simulation environment, which extracts meaningful information from past operating conditions. These process analysis activities are complemented by process improvement tasks such as process optimisation, transition planning studies and constrained real-time model-based optimisation and control, in a sequence of execution that that reshapes raw plant data into useful process knowledge and, hence, levers the chances for informed operative and supervisory decisions (see Rolandi and Romagnoli, 2006a).

In this work, we suggest extending the software architecture proposed by the CO standards (CAPE-OPEN Consortium, 2000) by introducing a new software object: the Problem Definition Environment (PDE). As sketched in Figure 2, the PDE manages the definition of advanced model-based problems by interacting with both the Process Modelling Executive (PMEs) and the user, while the PME performs the corresponding model-based activity by coordinating the calls to several Process Modelling Components (PMCs). These PMCs contain the mathematical description of the process model, and they also provide other services such as physical property calculations and numerical solution algorithms (Braunschweig et al., 2000). While the standardisation of open interfaces of the PME and PMCs has been the focus of the CO/GCO projects, the communication between the PDE and other elements of the architecture is regulated by a series of mechanisms intrinsic to the framework described in this work. These mechanisms entail the manipulation of the so-called "Data Model Templates" (DMTs) and "Data Model Definitions" (DMDs) (Rolandi and Romagnoli (2006a).

In the software architecture shown schematically in Figure 2, the MSOE (a PME and several PMCs) can be seen as a software tool for managing the development of mathematical models and, ultimately, coordinating the execution of the model-based activity, i.e. the MSOE is essentially a model builder and activity executive. On the other hand, the PDE is conceived as a software tool for supporting the definition of model-based problems, i.e. how to use plant data and the process model in the context of realistic process-engineering problems, which requires additional skills and expertise; in other words, the PDE is basically a problem builder.

2.3 *A Framework for Joint Parameter Estimation and Data Reconciliation*

As discussed above, a novel paradigm for the definition of rigorous model-based problems is now possible through the introduction of the PDE. The PDE manipulates the so-called Data Model Templates (DMTs) and Data Model Definitions (DMDs). In this section, we will briefly discuss the structure and purpose of the two data models relevant for the definition of hybrid data-driven/model-based parameter estimation data reconciliation problems. These data structures are the so-called Process Data Object data model (PDO) and the Dynamic Estimation Problem data model (DEP).

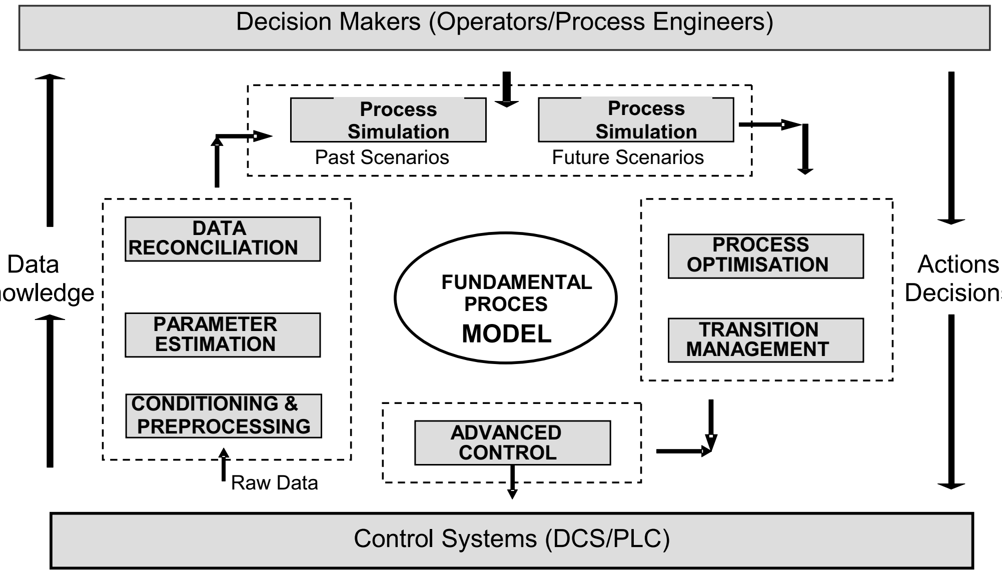


Fig. 1. The conceptual definition of the integrated framework for model-centric support of process operations.

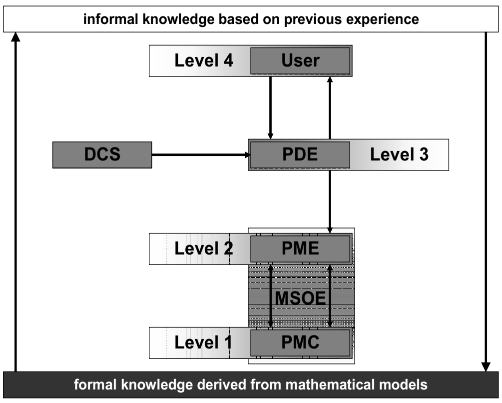


Fig. 2. The architecture of the framework.

- *DEP model:* Contains data determining the structure of a general dynamic estimation problem; for example, a valid instance of this data model links explicitly model variables to control/measured (input/output) process variables and corresponding process instrumentation, and determines the form of the objective function (estimator).

- *PDO model:* Contains data representing raw experimental process data in a form fit for hybrid empirical/mechanistic modelling; for instance, a valid instance of this data model associates the reconstructed dynamic trajectories of the input/output process variables (retrieved from process instrumentation) to model variables.

The distinction between these two structures is substantiated by the fact that it is not convenient to associate an experimental data set with a given estimation case-study, and vice versa. Overall, the DMT/DMD mechanism creates an innovative means to embed process knowledge and expertise on the definition of model-based problems, as well as increased opportunities for documentation and re-use of case-studies. The manipulation of these DMTs/DMDs is the scope of the PDE software object. In addition, the PDO and DEP data models encapsulate corporate expertise on the use of high-level declarative modelling languages, the process model and the process system, and they make possible the definition and execution of rigorous estimation/reconciliation problems of interest to operations personnel.

As described by Rolandi and Romagnoli (2005), an environment for the definition of generic dynamic estimation problems of the characteristics described above shares common characteristics with the simulation and optimisation environments of the integrated framework. Effectively, similarities with the simulation environment are due to the fact that estimation- reconciliation problems are based on experimental plant data and, therefore, there is a need for data pre-processing, conditioning and reconstruction of process trajectories. Similarities with the optimisation environment are given by the fact that dynamic estimation experiments are a particular type of dynamic optimisation problems and, hence, additional structural information (i.e. information apart from that contained in the mathematical model of the process) is needed to fully determine the nature of the estimation/reconciliation experiment/case-study. This is a point-of-synergy which can be exploited during the design, implementation and use of an integrated model-centric system for support of process operations.

2.4 *A Case Study: Application to the Pulping Section of a Pulp and Paper Mill*

The challenge associated to the joint parameter estimation and data reconciliation case-study proposed in this section lies on the complexity of both the industrial process system and the actual model-based problem. Unfortunately, the points-of-synergy between the solution of this process-engineering problem and other analysis and improvement tasks supported by the integrated model-centric framework is out of the scope of this manuscript.The goal of this case-study is to reconcile the process model and the plant data focussing on the closure of the general mass balance of the continuous pulping system. The estimation horizon is 1440min (24hr) and the window for reconstruction of process trajectories (Rolandi and Romagnoli, 2006b) is 30min for both input and output variables. A subset of 26 input process variables is used to imitate the input behaviour of the continuous process system. Among them 21 are controlled variables (set-points of PID control loops) and 5 are uncontrolled measured variables (disturbances).

The wood chip impregnation factor is a measure of the flowrate of steam condensate bounded to the interstitial void space between wood chips after the atmospheric pre-steaming step at the chip bin and before entering to the chip meter. Conventionally, the magnitude of this parameter would be obtained from the P&ID; however, changes in wood handling operations and operating conditions of the chip bin will change its nominal value. Since the magnitude of this parametric variable affects the closure of the mass balances, it will be chosen as a decision variable of the joint parameter estimation and data reconciliation problem. Additionally, we will estimate the magnitude of the pre-multiplier of the fundamental kinetic model of the Kraft pulping reactions occurring within the continuous cooking digester. Finally, we will also estimate the magnitude of the bias of three flow measurement devices: overall white liquor addition; wash filtrate addition to the digester's bottom; and black liquor extraction from the upper screens of the digester (see Table 1). The measurements of eight sensors are used for the purpose of estimation (Rolandi and Romagnoli, 2006b). The potential for model-based joint parameter estimation and data reconciliation of a large-scale complex industrial process system is demonstrated in this case-study: the problem results in the estimation of five parametric process variables (three of them are measurement biases) from an experimental data pool of eight measured variables and twenty-six control variables.

Table 1. Parametric variables of the continuous pulping area

DCS Tag	Variable Description
EE212.KinPreMult	Kinetic pre-multiplier
EE103.ChipImpFctr	Wood chip impregnation factor
FT212A.MB	Overall white liquor addition flow
FT212H.MB	Wash filtrate addition flow to digester bottoms
FT212C.MB	Upper extraction screen extraction flow

Table 2 shows the optimal estimates, confidence intervals and lower and upper bounds for the parametric variables. From this information we can calculate that the coefficient of variation for a 95% confidence on the individual estimates of the parametric variables EE212.KinPreMult and EE103.ChipImpFctr are 1.4% and 6.0% respectively; for all practical purposes, this is an indication of a satisfactory accuracy of estimation. The coefficient of variations of FT212A.MB and FT212H.MB based on a 95% confidence (Table 2) are reasonably small (3.4% and 4.0%, respectively), which is an indication of satisfactory accuracy of the estimates. On the other hand, the coefficient of variation corresponding to FT212C.MB is fairly large (46.7%), indicating a large uncertainty in the determination of this measuring device bias. In spite of this, the process variable can still be successfully estimated given the data pool used in this case-study. Figure 3 shows the fulfilment of the general mass balance of the continuous pulping system before and after reconciliation.

From a practical viewpoint, it was our aim to estimate those biases which have a strong impact on inventory analysis, or whose quantification is vital for other operational purposes (e.g. inferential soft-sensing). In the case of an industrial continuous pulping system, the most significant sources of revenue and expenses are likely to be the production of pulp, the cost of chip consumption and the cost of evaporation of weak black liquor (Rolandi and Romagnoli, 2006b). Fortunately, the cost of evaporation of weak black liquor can be partially reconciled from the estimate of the bias of the upper-screen extraction flow measurement. Interestingly, the 6.4% error of this process measurement (see Table 2) is associated to a material stream which accounts for nearly 32% of the overall weak black-liquor extraction flow from the continuous cooking digester at this nominal production level (~3.1m3/min). Additionally, the treatment of the black liquor in the evaporation area comprises approximately 56% of the variable costs of operation of the continuous pulping area (~ 88$/min). Hence, a 6.4% measurement error on such a critical process stream is equivalent to a production accounting miscalculation of approximately 0.50 million US$ per year, or an inventory analysis error of roughly 32 thousands cubic meters per year. This analysis demonstrates the economic incentive for advanced dynamic data reconciliation.

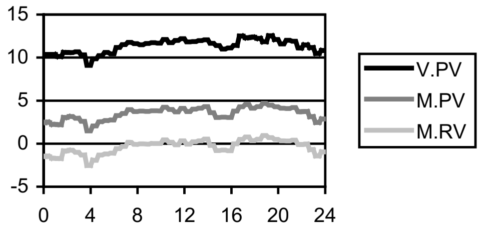


Fig. 3. Fulfilment of general mass balance (relative error [%] vs time [hr]); volumetric flow (measured process variables) and mass flow (calculated and reconciled process variables).

Table 2. Optimal estimates and confidence intervals

	EST	CI(95%)	LB	UB	CV [%]	ERR [%]
EE212.KinPreMult [adim]	3.34E-01	4.73E-03	3.29E-01	3.39E-01	1.42	n/a
EE212.ChipImpFctr [m3/kg]	1.83E-01	1.09E-02	1.72E-01	1.94E-01	5.96	n/a
FT212A.MB [m3/min]	7.32E-02	2.52E-03	7.06E-02	7.57E-02	3.44	7.1
FT212H.MB [m3/min]	1.69E-01	6.76E-03	1.62E-01	1.75E-01	4.01	8.2
FT212C.MB [m3/min]	6.51E-02	3.04E-02	3.47E-02	9.55E-02	46.7	6.4

Recently, Romagnoli and co-workers (Wang and Romagnoli, 2003) have presented a complete framework for Robust Data Reconciliation based on the generalized-t (GT) distribution, which can accommodate a family of distribution and thus providing extra flexibility as well as efficiency and optimality. The drawback of this approach is that actual statistical characterization of each sensor needs to be estimated, thus increasing the computation load if done at a centralized level.

In our proposed sensor network (Joe et al, 2005), some of these tasks can be delegated to the individual sensor (self-learning) in a decentralized manner, reducing the overall load at the higher level (control system) and at the same time providing extra functionalities for the sensor allowing the implementation of robust self-checking strategies at the local level. Two key things to be formulated in realising the sensor network are the architecture and the intelligence (or task allocation) of the sensor network. The following two subsections will discuss these, with a focus on GT-based data reconciliation. The last subsection presents application of the proposed sensor network to an integrated pilot-scale plant.

2.5 *Architecture of the Sensor Network*

The architecture of the sensor network defines the interconnection structure and functional relationships among the intelligent nodes in the network. The federated processing architecture is adopted in our proposed sensor network. Federated means that certain responsibilities are allocated to nodes at higher tier, but many functions are performed autonomously by nodes at lower tier (Sastry and Iyengar, 2005). Figure 4 shows the proposed network architecture. The network hierarchy consists of two tiers, each corresponding to an information processing level. Clusters are formed, i.e. each upper level node manages a few lower level ones. Each cluster corresponds to a process unit, i.e. the lower level nodes are none other than sensors measuring the variables of the process unit. We termed the lower level nodes as cluster members and the upper level ones as cluster heads. Nodes at the upper level can be considered as virtual sensors, in that they do not sense any physical phenomena, but mainly function as information processing units, performing tasks that are multivariate in nature such as data reconciliation. As such, cluster heads must collect information from their members in order to carry out their tasks. The cluster head and its members may require different computational capabilities.

The communication scheme is depicted by the lines connecting the nodes, i.e. cluster members communicate with their respective cluster heads, while cluster heads also communicate with one another. This results in different requirements of communication and networking capability and interface of the cluster heads as compared to the cluster members.

2.6 *Intelligence of the Sensor Network: Distributed Data Rectification*

The federated processing in a sensor network provides a potentially more efficient alternative implementation of the GT-based DR strategy than the centralized scheme. The GT-based DR strategy comprises two procedures that can be distributed in the federated sensor network: statistical characterization of sensor data (using GT distribution) and reconciliation of data using the obtained statistical characteristics. Since each sensor node (cluster member) is intelligent, the statistical characterization of data can be performed at the sensor level, resulting in self-*learning* of each sensor. Besides relieving the higher level from the computational burden and compressing the data to be communicated, self-learning also provides a signal model which is useful in reducing uncertainty in the measurement data and is the basic information that can be used for further processing that is not limited to DR only. The steps involved in sensor self-learning include the collection of a set of data points from which the sensor characteristics are extracted, and the estimation of the parameters of the statistical distribution of the data itself. This estimation can be mathematically expressed as:

$$\{\mu, p, q, \sigma\} = \arg\max \sum_{i=1}^{n} \log f_{GT}(u_i, p, q, \sigma) \qquad (1a)$$

where:

$$u_i = y_i - \mu \qquad (1b)$$

$$f_{GT}(u, p, q, \sigma) = \frac{p}{2\sigma q^{1/p} B(1/p, q)\left[1 + |u|^p / \sigma q^{1/p}\right]^{q+1/p}} \qquad (1c)$$

y is the i-th data point, n is the number of data points in the current data set, μ is the estimate of the process variable and {p, q, σ} are the parameters of the GT distribution function f_{GT}, i.e. the statistical characteristics of the sensor.

The second step of the distributed GT-based DR strategy, i.e. the reconciliation step, is performed at the level of cluster head (process unit) due to its multivariate nature. The cluster head is therefore responsible for consolidating data from each member sensor in the cluster, and subsequently performing the computation to reconcile the data. Mathematically, this computation can be expressed as:

$$\underset{x}{\text{Max}} \ -\log f_{\text{GT}}(\text{u} \mid \text{p}, \text{q}, \sigma) \ \text{ s.t. } \ g(x) = 0 \qquad (2)$$

where u = y-x, y is the measurements, T is the estimates of the p process variables and g(x) denotes the set of conservation equations. Note that the values of {p,q, σ} used in this estimation are the self characteristics communicated by each individual sensor.

2.7 *A Case Study: Application to an Integrated Pilot-Scale Plant*

Experiment Environment. An experimental platform comprising plant simulator, sensor (cluster member) simulator, and cluster head simulator is constructed.

- *Plant Simulator:* The virtual version of a process unit within an integrated pilot-scale plant is developed. This unit, a continuous stirred tank reactor (CSTR) with a cooling coil, is simulated using Matlab/Simulink. Measured variables include: Fin (feed flow rate), Tin (feed temperature), F (effluent flow rate), T (effluent temperature), Fc (cooling water flow rate), Tcin (inlet cooling water temperature), Tc (outlet cooling water temperature), Trx (reaction vessel temperature).

- *Cluster Member (Sensor) Simulator:* The cluster member consists of two parts: physical sensing and data processing. Accordingly, the simulator consists of noise generator and saturation function to mimic sensing, and a self-learning module to realize the data processing segment. For each of the eight measured variables of the CSTR, a cluster member is assigned. A cluster head manages and monitors these eight cluster members. To realize the mapping between cluster heads and cluster members, each cluster member is labelled by unique identification. The cluster members will transmit this identification and its estimated self/ signal characteristics {μ, p, q, σ} to the respective cluster heads and cluster head will use this information to perform data reconciliation.

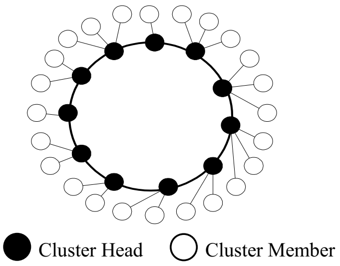


Figure 4: Federated sensor network architecture

- *Cluster Head Simulator:* The cluster head consists of communication controller, task controller and task modules. The communication controller is responsible for both external communication, i.e. with the cluster members and with the user interface, and internal communication, i.e. with the task controller. The task controller oversees all the tasks assigned to the cluster head. As such, the distributed data reconciliation as described in Section 3 is one of the underlying intelligence of the task controller. In our original proposal (Joe et al, 2005), besides data reconciliation, the intelligence also includes fault diagnosis and sensor reconstruction. The implemented framework is shown in Figure 5.

Experiment Results

- *Self-learning:* The results of self-learning for a few different noises are depicted in Figure 6, demonstrating considerably accurate characterization of the sensor data.

- *Data Reconciliation:* The formulated distributed GT-based data reconciliation is performed by cluster head using the characteristics obtained by self-learning in each cluster member. Figure 7 compares the estimation accuracy (ratio of absolute error of reconciled to that of measured data) of the proposed distributed method with the conventional centralized one according to Wang and Romagnoli (2003). Comparable performance is observed, hence demonstrating the viability of the distributed scheme.

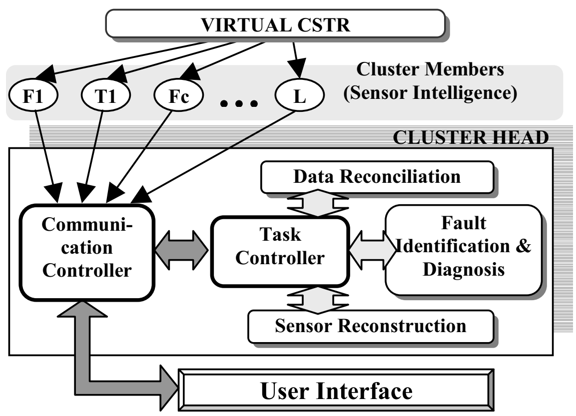
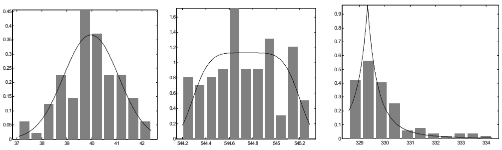


Figure 5: Overview of modules in the experimental setup



(a) F1, Gaussian (b) T1, Uniform (c) Fc, Laplacian

Figure 6: Statistical characterization of sensor data

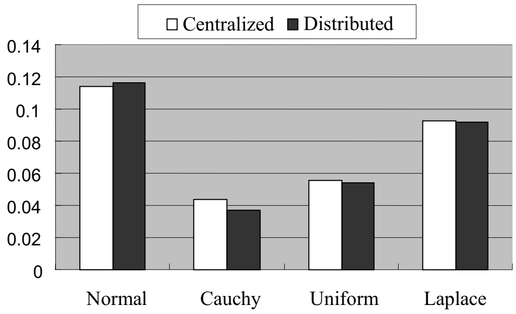


Figure 7: Estimation errors of the distributed and centralized data reconciliation approach

3. CONCLUSION & FUTURE WORK

The impact of emerging technologies on what is technically and economically feasible to achieve in the application of data reconciliation activities in manufacturing plants has been discussed and demonstrated.

Recent developments on model-centric technologies to support plant operations based on state-of-the-art process-engineering software tools were shown to provide the opportunities for performing large-scale parameter estimation – data reconciliation applications in complex dynamic industrial processing systems. The results of the industrial case-study were assessed from the perspective of production accounting and inventory analysis, and found a great incentive for the Process Industries to benefit from these advanced methodologies for plant data management. An environment for the definition of generic dynamic estimation problems of the characteristics described above shares common features with complementary simulation and optimisation environments. This is a point-of-synergy which can be exploited during the design, implementation and use of an integrated model-centric system for support of process operations.

New sensor technologies based on recent advances in microprocessor-based instrumentation and digital communications provided opportunities for the realization of novel sensor network architectures towards a truly distributed environment for data processing and reconciliation. The proposed federated architecture of intelligent sensor networks provides a flexible underlying framework for distributed data processing and rectification. We have presented a reformulation of the conventionally centralized robust partially adaptive data reconciliation scheme into a distributed scheme based on the federated sensor network. Experiment results showed highly comparable performance in terms of estimation efficiency, hence confirming the feasibility of the distributed scheme. Furthermore, in decentralizing the data reconciliation task, intelligence in the form of self-learning is incorporated into sensors. The presented distributed scheme serves as the first step towards a holistic distributed treatment of sensor data using the federated sensor network. This includes, but is not limited to, multi-resolution sensor data modeling, robust filtering and missing sensor data reconstruction.

REFERENCES

Braunschweig, B.L., Pantelides, C.C., Britt, H.I. and Sama, S. (2000). Process modeling: The promise of open software architectures. *Chemical Engineering Progress*, **Vol.96**, pp.65-76.

Global CAPE-OPEN Consortium (2000). Cape-open conceptual design document (cdd2). Available from: http://www.global-cape-open.org/02_CO_Conceptual_Design_Document_CDD2.pdf. Accessed on: 2004 Dec. 1.

Foss, B.A., Lohmann, B. and Marquardt, W. (1998). A field study of the industrial modeling process. *Journal of Process Control*, **Vol.8**, pp.325-338.

Joe, Y.Y, Ding, Z.Q, Ling, K.V. and Romagnoli, J.A. (2005). An intelligent sensor network for distributed data rectification and process monitoring. In *Proceedings of the 3rd International Conference on Industrial Informatics (INDIN 2005)*, IEEE, Perth, Australia.

Rolandi, P.A. and Romagnoli, J.A. (2005), An integrated environment for support of process operations, *AIChE Annual Meeting*, Cincinnati.

Rolandi, P.A. and Romagnoli, J.A. (2006a). Integrated model-centric framework for support of manufacturing operations. Part i: The framework. *Computers and Chemical Engineering*, Submitted for publication.

Rolandi, P.A. and Romagnoli, J.A. (2006b). Integrated model-centric framework for support of manufacturing operations. Part iii: Joint parameter estimation and data reconciliation. *Computers and Chemical Engineering*, Submitted for publication.

Sastry, S. and Iyengar, S.S. (2005). A taxonomy of distributed sensor networks. In *Distributed Sensor Networks* (Iyengar, S.S., Tandon, A., Brooks, R (Ed.)), pp.29-43. CRC Press, Boca Raton, Florida.

Wang, D. and Romagnoli J.A. (2003). A framework for robust data reconciliation based on a generalized distribution function. *Ind. Eng.Chem. Res.,* **Vol.42***,* pp.3075-3084.