Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Stochastic Flow Model using Kalman Filters for parameter estimation ⋆

Guillaume Graton * **Martin Guay** * **Jorge Arinez** **

* *Department of Chemical Engineering, Queen's University, Kingston, ON K7L 3N6, Canada (e-mail: guillaume.graton@chee.queensu.ca, martin.guay@chee.queensu.ca).*
** *General Motors R&D Center, Warren, Michigan, 48090, USA, (e-mail: jorge.arinez@gm.com)*

Abstract: Attaining high quality from manufacturing systems requires utilizing appropriate system-level quality performance modeling and analysis tools. This paper describes the application of the stochastic-flow-modeling (SFM) approach to represent the quality output behavior of a manufacturing system. To do this, a basic one-product type SFM is extended to that of a multiple-product manufacturing system. This work also provides a novel addition to the SFM approach through the use of a Kalman filter to estimate quality parameters. After a presentation of the reference manufacturing system, results are given for different examples and the effectiveness of the SFM model is examined in terms of accuracy and convergence.

1. INTRODUCTION

Quality is a critical component of overall manufacturing system performance that production managers are continually striving to improve. Achieving higher production quality levels requires an effort at all levels of the manufacturing enterprise. At the management level, several approaches, such as the Shewhart cycle (Plan-Do-Check-Act) (Shewhart (1939)), Quality Circles (Besterfield et al. (2003)), Six Sigma (Pande and Holpp (2002)) have been proven to be highly successful at motivating quality improvements throughout manufacturing organizations. These types of quality management techniques with their corresponding awards effect control at high management level. In addition, there are other approaches such as Statistical Quality Control (Woodhall and Montgomery (1999)) that focus on local, process level improvements.

These types of approaches as well as more specialized optimal control of individual process parameters also have achieved great improvements in quality levels. However, an area that has received less attention because of the increased modeling and computational complexity until know has been the intermediate level. This intermediate level may be thought of as the system or production line where the level of granularity corresponds to multiple machines controlled by the supervisor of the system. Examples of work that has been done at this manufacturing system level includes that of Cassandras and Lafortune (1999) or Brandin and Wonham (1994). Some of these approaches are based on the control optimization and optimization of queuing systems in manufacturing lines (Perkins and Skirant (1999), Brandin and Wonham (1994), Shu and Perkins (1998), Wardi et al. (2001)). In this paper, the focus is not on the management and optimization of buffers, but rather on the estimation of quality and the prediction of quality related parameters for a given manufacturing line (Graton et al. (2007)). System level quality is broadly defined as the ratio of products that exit the system to the total number that enter indicating the percentage of those products rejected by quality monitoring stations.

Quality model of a given manufacturing line with specific knowledge of the performance behavior can provide solutions of complex quality failures, decision-making insight and anticipate expected quality performances. Because of random features of reject occurrences on the manufacturing line, a stochastic modeling is preferred to a deterministic modeling. Stochastic flow models (SFM) take into account the random nature of process operations and provide a natural framework for quality estimation in stochastic processes Yu and Cassandras (2004), Wardi et al. (2001), Cassandras et al. (2002).

With stochastic flow models, distribution parameters corresponding to reject rates are estimated from data which are in turn used in the estimation of quality levels. Parameter estimation can be done using different algorithms, for example in Graton et al. (2007), a maximum likelihood approach and an adaptive approach are employed. These two approaches are essentially based on the mean estimation of reject rates. Over time, estimations become increasingly smoother and the instantaneous nature of reject rate variation becomes less evident. In this paper, a new approach is proposed that uses Kalman filters to estimate parameters associated with the instantaneous update of reject rates. The Kalman filter Kalman and Bucy (1961) has the advantage that it can incorporate a wide variety of data while providing an optimal estimation of the system's output parameter(s) of interest (assuming system linearity and measurement and system noise to be white and Gaussian).

In Jonsdottir et al. (2006), the parameter estimation for stochastic models is based on the likelihood function estimated by a Kalman filter. In Gaussian stochastic

⋆ This work is supported by General Motors.

process, the Kalman filter and the maximum likelihood have the same meaning. In this work, distributions are exponential and a corresponding approach to estimate such distribution parameters is presented. In Song and Speyer (1986), the approach used was based on a modified gain Kalman filter and where the parameter identification technique is similar to the one presented in this paper.

In this paper, the first section is devoted to the design of a closed-loop estimator for the stochastic flow model. The case of a single station is presented whereupon the approach is extended to that of an entire manufacturing line. The entire manufacturing line is modeled with an enhanced model incorporating quality feedback loops. Finally, results are provide regarding the accuracy of the approach based on simulation and convergence behavior.

2. CLOSED-LOOP ESTIMATOR

2.1 Single Station

Quality model This section proposes an approach to estimate the instantaneous value of the quality parameter of a single station within a production line. The knowledge of this quality parameter is important since it is required to implement a stochastic flow model. Prior to defining a parameter for instantaneous quality and its estimation approach, we first define the global quality of a station. Thus, the quality at a single station may be calculated using the following equation:

$$Q(t) = \frac{\int_0^t v(s)ds - \int_0^t r(s)ds}{\int_0^t v(s)ds} = \frac{V(t) - R(t)}{V(t)} \qquad (1)$$

where Q represents the quality at time t using the ouput function $v(s)$ and the reject function $r(s)$ over the interval $[0, t]$, the engine output of the station v is defined by $v(t) = 1$ if an engine exits the station, otherwise $v(t) = 0$ and the reject function r by $r(t) = 1$ if a reject appears, otherwise $r(t) = 0$.

Using the production data V and reject data R associated with the station equation (1) provides the quality measurement Q. The quality measurement can be seen as the average of the instantaneous quality q over the time interval $[0, t]$. The instantaneous quality parameter p is linked to the instantaneous quality as follows:

$$p(t) = \frac{1}{1 - q(t)} \qquad (2)$$

When $Q(t) = \int_0^t 1 - \frac{1}{p(s)} ds$, the direct expression of the quality parameter function $p(t)$ is defined as:

$$p(t) = \frac{1}{1 - t\dot{Q}(t) - Q(t)} \qquad (3)$$

To avoid possible chattering problems (see figure 1) of equation (3), an observer is implemented to estimate the instantaneous quality parameter.

Kalman filter The Kalman filter is defined around the quality measure $Q(t)$ from Equation (1) where the state

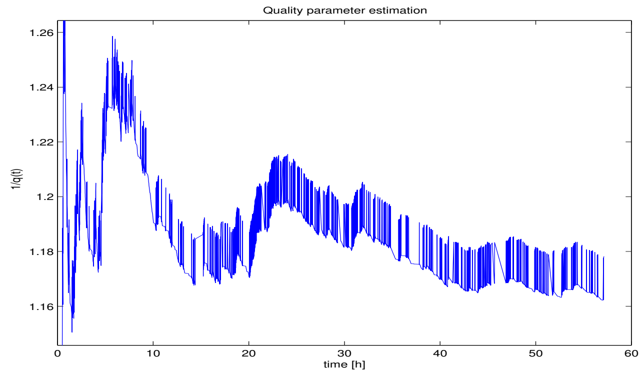

Fig. 1. Chattering problems from the direct estimation of equation (2)

space representation uses the differentiation of $Q(t)$ as follows:

$$\dot{x}(t) = A(t)x(t) + w_x(t) \qquad (4)$$

where $x(t) = [Q(t), \dot{Q}(t), \ddot{Q}(t), 1 - \frac{1}{p(t)}]$ represents the state, $A(t)$ the time-varying matrix and $w_x(t)$ the state noise.

The measurement equation is given by:

$$y(t) = Cx(t) + w_y(t) \qquad (5)$$

where $y(t)$ represents the measurement, C the measurement matrix and $w_y(t)$ the measurement noise.

Noises $w_x(t)$ and $w_y(t)$ are assumed to be non-correlated, white, gaussian with, respectively, a covariance matrix W_x and W_y.

Matrices A and C are defined by:

$$A(t) = \begin{pmatrix} -1/t & 0 & 0 & 1/t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 2 & t & 0 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}$$

The calculation of the eigenvalues of A gives the state-space representation as defined by equation (4) and is a non-stable system. Three poles are zero, where the last pole has a negative real part and depends on $\frac{1}{t}$. To satisfy convergence of the instantaneous quality estimations, an observer need to be implemented. The basic principle of an observer is to close the loop using the measurement as input information and to optimize the correlation between the measurements and the output estimations while minimizing the noise effects. In this work, an observer is given by:

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + K(t)(y(t) - C\hat{x}(t)) \qquad (6)$$

$$\hat{y}(t) = C\hat{x}(t) \qquad (7)$$

where $\hat{x}$ corresponds to the state estimation, $\hat{y}$ the output estimation, K the observer gain matrix.

The dynamic equation of the error $\tilde{x}$ between the state x and its estimation $\hat{x}$ can be calculated:

$$\begin{aligned}\dot{\tilde{x}}(t) &= \dot{x}(t) - \dot{\hat{x}}(t) \\ &= (A(t) - K(t)C)\tilde{x}(t) + w_x(t) - K(t)w_y(t) \qquad (8)\end{aligned}$$

From equation (8), the choice of matrix K yields the stability of the observer, if the real part of the $(A(t) - K(t)C)$-eigenvalues are negative. Moreover the choice of K can be realized by minimizing the covariance of the estimation error:

$$\begin{aligned}\dot{P}(t) &= Cov[\dot{\tilde{x}}(t)] \\ &= (A(t) - K(t)C)P(t) + P(t)(A(t) - K(t)C)^T + W_x \\ &\quad + K(t)W_y(t)K(t)^T \end{aligned} \quad (9)$$

To find the matrix K which minimizes the covariance error P, the influence of K on P is calculated from:

$$\frac{\partial trace(\dot{P}(t))}{\partial K(t)} = -P(t)C^T - P(t)C^T + 2K(t)W_y \quad (10)$$

If $\frac{\partial trace(\dot{P}(t))}{\partial K(t)}$ is equal to 0, that gives the necessary condition for the minimum of P, such that $K(t)$ is:

$$K(t) = P(t)C^T W_y^{-1} \quad (11)$$

If the expression of $K(t)$ given by equation (11) is replaced in equation (10), then:

$$\dot{P}(t) = A(t)P(t) + P(t)A(t)^T - P(t)C^T W_y^{-1} CP(t) + W_x \quad (12)$$

is a Riccati equation with the asymptotic convergence property of the eigenvalues and the greatest estimation error tends to 0.

Stochastic flow model using Kalman filter estimations Using the estimation $\hat{q}(t)$ of the state estimation $\hat{x}(t)$, the estimation of the instantaneous quality parameter gives the parameter of the distribution law $\mathcal{E}(\hat{p}(t))$. Quality is estimated by the Stochastic Flow Model (SFM) which uses information about the probability of the occurrences of reject codes for each engine model type and each production shift to compute the number of rejects on the line for any given period of time. Figure 2 shows a single-station discrete SFM.

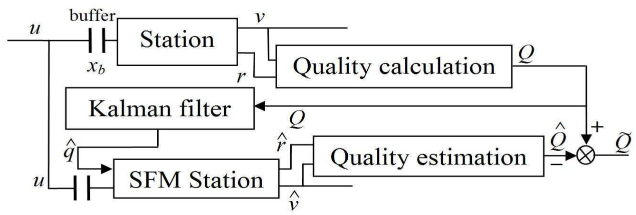

Fig. 2. Stochastic Flow Model

where $u(t)$ (parts/sec) is the rate of parts entering the station while $v(t)$ (parts/sec) is the output rate equal to the production rate $\rho(t)$ at which the station is able to produce parts. The buffer size is given by $x_b(t)$ (parts). Note that the station switches between two states, 0 (OFF) or 1 (ON), such that:

$$s(t) = \begin{cases} 1 \text{ machine working} \\ 0 \text{ otherwise} \end{cases} \quad (13)$$

where $s(t)$ denotes the operational state of the system.

The working and non-working states are triggered by random events that cause the station to change its current state. More elaborate logical statements may also be considered in practice.

The rate of parts entering the station, $u(t)$, is determined by production planning and scheduling. The rate at which the station produces parts, $v(t)$, is a more complex function of processing conditions. A simple rule-based operational constraint for the output rate is formulated as follows:

$$v(t) = \begin{cases} 0 & \text{if } s(t) = 0 \text{ or } x_b(t) = 0 \\ \rho(t) & \text{otherwise} \end{cases} \quad (14)$$

where the buffer content $x_b(t)$ is determined by the following ordinary differential equation:

$$\dot{x}_b(t) = u(t) - v(t) \quad (15)$$

The model uses information about the probability of the occurrences of reject codes for each engine type and each shift to compute the number of rejects on the line for any given period of time. At each station, the total number of rejects is calculated using the following ordinary differential equation:

$$\dot{x}_r(t) = \Delta(t)\delta(t) \quad (16)$$

where x_r is the number of rejects (with initial conditions $x_r(0) = 0$), $\delta(t)$ is a unit impulse.

Before defining the function $\Delta(t)$, additional definitions are required. First, a parameter called λ is introduced to represent the average rate of engines produced between the occurrence of two rejects.

Following the definition of the reject occurrence distribution, a threshold th must be introduced to isolate the reject probabilities from the non-reject probabilities. The exponential distribution $\mathcal{E}(\hat{p}(t))$ has a probability density function $f_{\mathcal{E}(\hat{p}(t))}(t)$ defined by:

$$f_{\mathcal{E}(\hat{p}(t))}(t) = \hat{p}(t)e^{-\hat{p}(t)t} \quad (17)$$

From the definition of the density function, the integral of $f_{\mathcal{E}(\hat{p}(t))}(t)$ from the limits 0 to ∞ is equal to 1 (100 %) but there is also $\frac{100}{\hat{p}(t)}$ % that corresponds to the reject rate and thus $\left(100 - \frac{100}{\hat{p}(t)}\right)$ % for the non-reject rate. The boundary between these two modes (reject and non-reject) is defined by the threshold th. In other words, this threshold, deduced from the exponential distribution $\mathcal{E}(\hat{p}(t))$, is defined such that: $\int_0^{th} f_{\mathcal{E}(\hat{p}(t))}(t)dt = \frac{1}{\hat{p}(t)}$ where the solution of this equation yields an expression for th:

$$th = \frac{1}{\hat{p}(t)} \ln\left(\frac{\hat{p}(t) - 1}{\hat{p}(t)}\right) \quad (18)$$

From the exponential distribution $\mathcal{E}(\hat{p}(t))$, a random sample, called $t_{\hat{p}(t)}$, is obtained for each engine exiting the work station. This sample is compared to the reject threshold th defined above. If the sample $t_{\hat{p}(t)}$ is under the threshold when the engine exits the station, then a reject flag is given to the engine. To complete this section, the function $\Delta(t)$ which generates rejects on the station is defined by:

$$\Delta(t) = \begin{cases} 1 \text{ if } v(t) \neq 0 \text{ and } t_{\hat{p}(t)} \leq th \\ 0 \text{ otherwise, i.e.: no reject} \end{cases} \quad (19)$$

This type of model just described can be used to provide high-level abstractions of discrete event systems (DES) models. In a DES (or a timed-DES), the ability to monitor single parts is limited since each action taken on each part must be tracked. As a result, representative DES models tend to be very complex and therefore this paper presents an alternative modeling approach based on SFMs.

2.2 Multiple Station Manufacturing System

Quality models In the previous section, quality is defined locally for each specific station. Local quality is defined by Q_i where i refers to the station number and is given by equation (1). Also as a global quality was introduced to provide insight regarding the first time quality of the entire manufacturing line. The modeling of the global quality is given for example in Graton et al. (2007) and is based on equation (1) where $v(t)$ is replaced by the output flow at the last station of the entire manufacturing line and $r(t)$ is defined as a boolean function which equals 1 when the engine has a reject in at least one station of the manufacturing line.

Stochastic flow model using Kalman filter estimations This section describes the use of the stochastic flow model to estimate the global quality of the entire manufacturing line. As explained before, each station uses a Kalman filter to obtain an estimation of the instantaneous quality parameter. This parameter is an estimation of the rate of engines with rejects which flow through the station and appears in the stochastic flow model as the parameter of exponential distribution $\mathcal{E}(\hat{p}_i(t))$, where i is the station number. Figure 3 illustrates the use of the filter in the parameter estimation of the manufacturing system.

Though the model for an entire plant is far more complex than that proposed in Figure 3, the model may be enhanced by incorporating quality repair loops to better represent quality behavior that could not otherwise be characterized solely with the use of inline stations. Figure 4 presents a slightly more complex configuration of a segment of a manufacturing line with five stations that are still simple enough to model yet more realistically represent the flow of parts that meet quality requirements and those that do not.

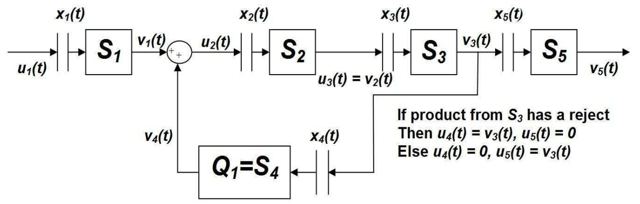

Fig. 4. Simple five-station SFM quality system

In Figure 3, each station is given an index i with $i \in \{1, 2, ..., n\}$ and the relation $v_i = u_{i+1}$ is used to link stations. However, this five-station model presents a further complication with the introduction of a parallel station which measures quality (station S_4 in Figure 4).

Thus, in Figure 4, stations S_1 and S_5 are defined as *processing stations*, stations S_2 and S_3 are *control stations* and finally, station S_4 is a *quality station* Q_1. Products can move from one station to the next (for example, from S_1 to S_2). But, in some instances, as at the exit of station S_3, two paths are possible. From S_3, engines can be directed to S_4 or station S_5. If the product has not triggered a quality rejects before station S_3, this product goes to station S_5. If a reject has occurred, the product must be repaired, and is therefore sent to station S_4. In station S_4, the problem identified at station S_3 on the engine is repaired. The product then returns to station S_2 to be processed and diagnosed a second time.

As the manufacturing system grows in size, it becomes more complex and modifications to the model are required. To complete the model, new definitions are added to generalize equations (13) to (18). Now, a set L_i is defined as the set of all stations linked with the input of station i, for example in Figure 4, $L_2 = \{S_1, S_4\}$. The input u_i of the station i can be defined as:

$$u_i(t) = \sum_{l \in L_i} b_{l,i}(t) v_l(t) \tag{20}$$

where $b_{l,i}$ is a boolean function (with the property $\sum_i b_{l,i} = 1$) that determines in which line the product is produced from station l to station i. In Figure 4, after station S_3 two choices are possible, if for example the product must be repaired then $b_{3,4}(t) = 1$ and also $b_{3,5}(t) = 0$.

The same functions, as defined previously in equations(13) to (18) can be defined for each station $i \in I$,

$$\dot{x}_{b_i}(t) = u_i(t) - v_i(t) \tag{21}$$

where the input u_i is defined by (20) and the output $v_i(t)$ follows:

$$v_i(t) = \begin{cases} 0 & \text{if } s_i(t) = 0 \text{ or } x_{b_i}(t) = 0 \\ \rho_{i,j,k}(t) & \text{otherwise} \end{cases} \tag{22}$$

and the operational state $s_i(t)$ is defined by:

$$s_i(t) = \begin{cases} 1 \text{ if a product is made in station } i \\ 0 \text{ otherwise} \end{cases} \tag{23}$$

These equations together give the dynamics of engines through the entire manufacturing line. Now, reject functions need to be introduced to complete the stochastic flow model. First, a reject function associated with an engine e is defined as:

$$R_{e,i} = \begin{cases} 1 \text{ if } s_i(t) = 0 \text{ and } t_{\hat{p}_i(t)} \leq th_i \\ 0 \text{ otherwise} \end{cases} \tag{24}$$

where the variable $t_{\hat{p}_i(t)}$ is a random variable with an exponential distribution $\mathcal{E}(\hat{p}_i(t))$ having a parameter $\hat{p}_i(t)$ defined for station i. The variable th_i is a threshold extracted from $\mathcal{E}(\hat{p}_i(t))$ as explained in (**??**). Each station has its own reject function $r_i(t)$ correlated to:

$$r_i(t) = R_{e,i} \delta(t) \tag{25}$$

where $\int_0^t r_i(s) ds$ is the number of rejects at station i over time $[0, t]$ (with initial conditions $r_i(0) = 0$).

When engine e exits the last station of the manufacturing line, the global reject function is updated:

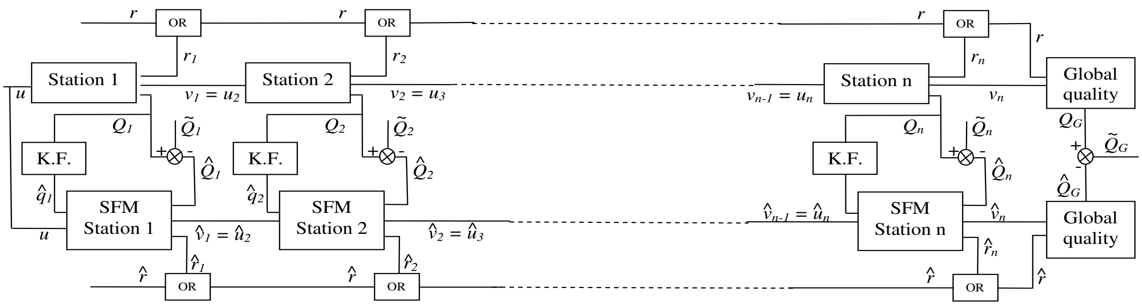

Fig. 3. Kalman filter with the SFM

$$r(t) = \begin{cases} 1 \text{ if } \bigwedge R_{e,i} \neq 0 \ i \text{ is not a quality station} \\ 0 \text{ otherwise} \end{cases} \quad (26)$$

where $\bigwedge_1^n R_{e,i}$ represents the OR operator on reject function over stations except quality stations.

3. RESULTS: FTQ SIMULATION

As explained previously, the calculation of the first time quality (global quality) is initiated when the first engine exits the last station of the manufacturing line. After the last station and as the engine leaves the manufacturing line, equations (22), (26) and (1) are computed. When an engine exits the line, one unit is added to $\int_0^t v_n(\tau)d\tau$. If the engine p has no reject then r stays constant otherwise one unit is added.

3.1 Quality estimation for single configuration

Prior to examining the results of the FTQ simulation for the entire manufacturing line, this section is dedicated to reviewing the accuracy of the instantaneous quality estimations performed on a single configuration. If this section were exhaustive, all stations would be tested one at a time. Thus, only important results are presented and the approach is limited to several stations which can attest to the effectiveness of the SFM model. As such, six stations are presented in each figure.

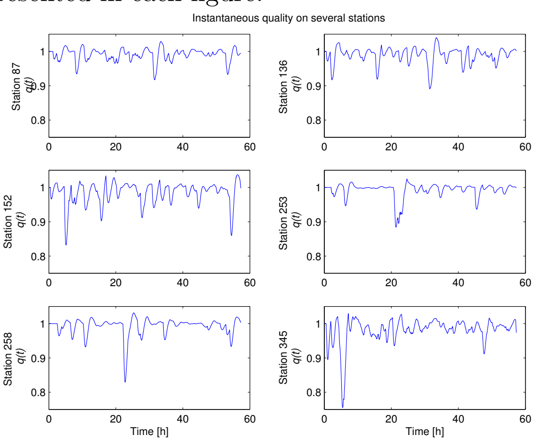

Fig. 5. Instantaneous quality at six stations.

In Figure 5, the estimations of the instantaneous quality is plotted. The instantaneous quality is, as its name indicates, a quality index and can not be greater than one. Notice that the estimations (figure 5) are sometimes greater than one, the estimations are not filtered and no thresholds are applied. But, for the SFM simulation, an upper threshold to the value one will be applied to avoid any physical problems. Figure 5 shows important quality variations over time. Moreover, the estimations seem not to be influenced by the differentiator properties of the state-space representation (equation 7).

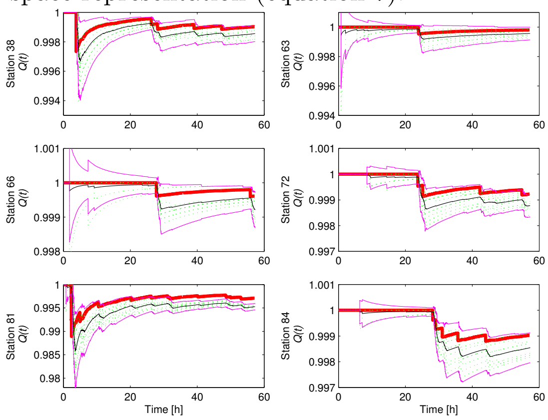

Fig. 6. Measured quality, SFM quality and expected value of the quality over 6 stations.

Figure 6 shows different simulations and is decomposed into six sub-figures corresponding to six specific stations. On each sub-figure, the measured quality represented is plotted as a bold line, and the ten simulations of the quality estimated by the SFM are represented by dotted lines. The expected values (normal line) along with the confidence interval of 5% of rejected values (upper and lower lines) are also shown. Notice that the expected value from ten simulations using the SFM are below the measured quality, however the relative error is less than 0.1%.

The figures above show that the instantaneous quality parameter estimations using a Kalman filter provide accurate results and the quality of each station estimated by SFM is close to that of the measured quality.

3.2 Global quality for an entire manufacturing line

The previous section provided results on the quality estimations at individual stations. In particular, Figure 5 showed that the estimated quality is less than the measured quality, which raises certain questions, such as: Does this error have an effect on the global quality? And if yes, what is the effect? Furthermore, are these errors additive?

Global quality is computed using equation 1 replacing function v by v_n for the last station where r is the boolean function obtained if the engine had one reject (or more) over the entire manufacturing line. In Figure 7, the measured quality is plotted as a bold line.

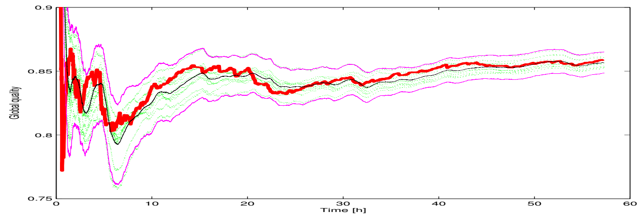

Fig. 7. Global quality comparisons over 60 hours

Using the measured quality of each station in the entire manufacturing line, the Kalman filter provides an estimation of the instantaneous quality parameter which then is used in the stochastic flow model. In this section, the global quality based on performing ten simulations is shown in Figure 7 (represented by the dotted lines). From these simulations, an average value is calculated and represented by a normal line in Figure 7. Also, with the standard deviation from the simulation, a confidence interval of 95% and 105% is given around the expected value (illustrated by the upper and lower normal lines).

Figure 7 also illustrates the accuracy of the estimation of global quality using the Kalman filter approach. The difference between the measured quality and the average quality from ten simulations tends to zero. Moreover, the confidence interval in Figure 7 appears to decrease over time. In addition, only 5% of estimations have an error greater than 1%, with the upper and lower bounds of the confidence interval separated by 2%. Figure 8 and 9 reflect the convergence characteristics of the quality estimations.

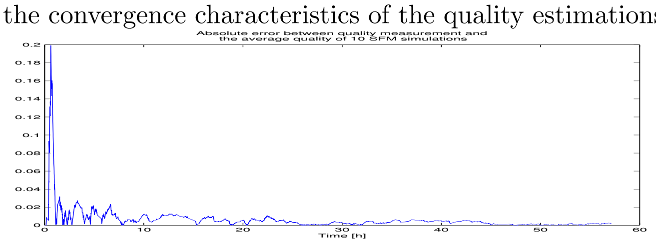

Fig. 8. Absolute error between the measured global quality and the average quality of 10 SFM simulations

In particular, Figure 8 illustrates the convergence to zero of the mean value of the global quality estimations. After less than 7 hours of simulation, the error is less than 2%, and after 23 hours it is less than 1%. At the end of the simulation, the error is around 0.2%.

Lastly, Figure 9 shows the standard deviation of the error between the global quality measurement and the global quality estimations. The asymptotic convergence of the standard deviation is clearly visible as convergence is provided by the Kalman filter. Finally, Figures 8 and 9 also provide evidence that using a Kalman filter is an effective approach to estimate the instantaneous quality parameters for global quality calculated using a stochastic flow model.

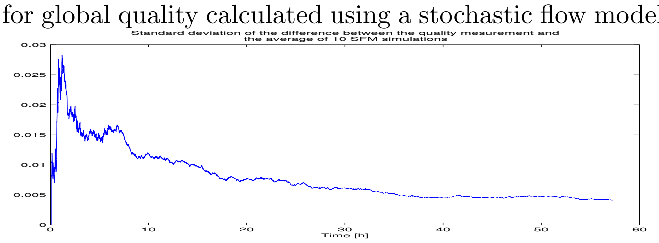

Fig. 9. Standard deviation of the difference between the quality measurement and the average quality of 10 SFM simulations

4. CONCLUSION

The results presented in this paper proof the good accuracy of the stochastic flow model estimation using the Kalman filter approach for the parameter estimation. The SFM provides a very good tool for quality modeling in assembly lines. This new approach provides a mechanism that accounts for non-stationary behavior of the model parameters. With the work, quality can be studied and analyzed station by station, or as a global quality (Fig. 3), this last point is useful for a good tracking of parameter values.

REFERENCES

D.H. Besterfield, C. Besterfield-Michna, G.H. Besterfield, and M. Besterfield-Sacre. *Total Quality Management.* Prentice Hall, Columbus, Ohio, 2003.

B.A. Brandin and W.M. Wonham. Supervisory control of timed discrete event systems. *IEEE Transactions on Automatic Control*, 39:329–342, 1994.

C.G. Cassandras and S. Lafortune. *Introduction to Discrete Systems.* Kluwer Academic Publishers, Massachusetts, USA, 1999.

C.G. Cassandras, Y. Wardi, B. Melamed, G. Sun, and C.G. Panayiotou. Perturbation analysis for online control and optimization of stochastic fluid models. *IEEE Transactions on Automatic Control*, 47:1234–1248, 2002.

G. Graton, M. Guay, P. Egbunonu, and J. Arinez. Stochastic flow model-based representation of system-level manufacturing quality. *ASME International Conference on Manufacturing Science and Engineering*, 2007.

H. Jonsdottir, H. Madsen, and O.P. Palsson. Parameter estimation in stochastic rainfall-runoff models. *Journal of Hydrology*, 2006.

R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Transactions of the ASME - Journal of Basic Engineering*, 1961.

P. Pande and L. Holpp. *What is six sigma?* McGraw-Hill, New York, 2002.

J.R. Perkins and R. Skirant. The role of queue length information in congestion control and resource pricing. *Proceedings 38th IEEE Conference on Decision and Control*, pages 2703–2708, 1999.

W.A. Shewhart. *Statistical Method from the Viewpoint of Quality Control.* Dover Publications, Incorporated, Mineola, NY, U.S.A., 1939.

C. Shu and J.R. Perkins. Optimal php control of multiple part-types on a failure-prone machine with quadratic buffer costs. pages 2149–2154, 1998.

T.L. Song and J.L. Speyer. The modified gain extended kalman filter and parameter identification in linear systems. *Automatica*, 1986.

Y. Wardi, C.G. Cassandras, and D.L. Pepyne. A backward algorithm for computing optimal controls for single-stage hybrid manufacturing systems. *Int. J. Prod. Res.*, 39:369–393, 2001.

W.H. Woodhall and D.C Montgomery. Research issues and ideas in statistical process control. *Journal of Quality Technology*, 31(4):376–386, 1999.

H. Yu and C.G. Cassandras. Perturbation analysis for production control and optimization of manufacturing systems. *Automatica*, 40:945–956, 2004.