Uncertainty quantification in system-level prognostics: application to Tennessee Eastman process

Ferhat Tamssaouet, Thi Phuong Khanh Nguyen, Kamal Medjaher, Marcos E. Orchard

Uncertainty Quantification in System-level Prognostics: Application to Tennessee Eastman Process

Ferhat Tamssaouet[1], TP Khanh Nguyen[1], Kamal Medjaher[1] and Marcos E. Orchard[2]

***Abstract*— This paper addresses the problem of uncertainty quantification in system-level prognostics. To this purpose, a three-step methodology, based on the inoperability input-output model, is presented. The first step concerns the estimation of the system inoperability, using a new adapted particle filtering method, while considering the interactions between its components. The second step focuses on the long-term prediction of the system inoperability in order to determine its evolution. Finally, in the third step, a method for calculating the remaining useful life of the system, based on the system configuration, is formulated. The proposed methodology is applied on data obtained from the Tennessee Eastman Process simulations to predict the shutdown due to violation of process constraints.**

I. Introduction

Prognostics and Health Management (PHM) is essential to ensure safe, reliable and correct operation of complex technical systems. Among the key elements of PHM, prognostics allows predicting the remaining useful life (RUL) of components, subsystems or systems before they become inoperable. Based on these predictions, effective actions can be taken to minimize losses, optimize maintenance, and extend components life.

According to practical requirements, prognostics received a great attention in literature. However, it has often been approached from a component view without considering interactions with other system components and the environment [1], [2]. Hence, for complex engineering systems, it is necessary to study the concept of failure prognostics at system-level by considering the mutual interactions between its elements.

Moreover, notwithstanding the increasing accuracy and precision of prognostics algorithms, their objects of study, i.e. degradation and failure mechanisms, remain stochastic phenomena and, therefore, the uncertainty cannot be eliminated totally [3]. Indeed, various sources contribute to make the estimation and prediction of one system state uncertain. The number of the uncertainty sources will rapidly increase when considering prognostics at system-level.

To provide a solution to the problems stated above, a new method that allows quantifying the uncertainty in the system remaining useful life (SRUL) prediction is proposed in this paper. It introduces, as a first contribution, a novel system-level prognostics framework based on the inoperability input-output model (IIM). This model allows tackling the issue related to the interactions between the system components. As a second contribution, a methodology to quantify uncertainty in the SRUL predictions based on the particle filtering is proposed. This methodology will be applied to the well-known Tennessee Eastman Process (TEP) to demonstrate its effectiveness.

The remainder of the paper is organized as follows. Section II presents the new system degradation model which is based on the inoperability input-output model. Section III describes the proposed methodology for uncertainty quantification and the SRUL determination. Section IV deals with uncertainty modeling and quantification for the Tennessee Eastman Process. A comparison of the obtained results with process real data, were made in order to show the effectiveness of the proposed method. Finally, Section V concludes the paper and gives some future works.

II. Inoperability Input-output Model

One of the main challenges for the SRUL prediction is to develop a model that allows taking into account the mutual components interactions and effects of the mission profile on the degradation evolution. For this purpose, a new model for the degradation of multi-component systems is proposed in this section, that is based on the inoperability input-output model (IIM). The proposed model is then used in Section III to quantify the uncertainty when predicting the SRUL.

The proposed model is inspired by the IIM which is an extension of the input-output model developed by Leontief Wassily in 1936 [4]. The IIM model and its variants are usually used to investigate the global effects of negative events on highly interdependent infrastructures or multi-sector economies [5], [6]. This is achieved by using the concept of inoperability, which is defined as the inability of a system to perform its intended functions. The aptitude of the IIM to consider mutual interactions between numerous elements offers a promising perspectives when applying it in PHM domain.

The IIM adapted to prognostics is proposed firt in [7] and is represented by the following formula:

$$q(t) = K(t).[A.q(t-1) + c(t)] \tag{1}$$

where:

- $q(t)$ is a vector representing the overall inoperability of the system components at time t. Each component of this vector is a value between 0 and 1, where $q_i(t) = 0$ corresponds to a healthy component (with an ideal performance) and $q_i(t) = 1$ to a faulty component (no longer able to perform its tasks).

[1]F. Tamssaouet, TPK Nguyen and K. Medjaher are with Laboratoire Génie de Production, LGP, Université de Toulouse, INPT-ENIT, Tarbes, France `ferhat.tamssaouet@enit.fr`

[2]M. Orchard is with the department of Electrical Engineering, University of Chile, Santiago, 8370451, Chile.

- A is a matrix representing the interdependencies between the system components. Each element a_{ij} of the matrix corresponds to the influence of the inoperability of component j on the inoperability of component i. The bigger a_{ij} is, the greater is the influence of j on i.
- $c(t)$ is a vector representing the internal inoperabilities of the system components at time t, i.e. the degradation of the component due to wear, corrosion or any other failure mechanism. The parameter $c_i(t)$ can be obtained by normalizing the health indicator of component i to its failure threshold.
  More details about the normalization of the component's health indicator can be found in [7].
- $A.q(t)$ represents the inoperability of a component due to its interdependencies. This quantity informs about the degradations caused by the interactions between components.
- $K(t)$ is a diagonal matrix representing the factors influencing the inoperabilities of components at time t with respect to the system inputs (mission profiles and environment conditions). Each element k_i is specific to only one component i.

As one can notice in (1), the degradation of component i, characterized by an inoperability $q_i(t)$, depends on its inherent natural degradation mechanisms expressed by $c_i(t)$ and the degradation induced by the interactions with other components through the term $A.q(t)$. By integrating these two types of degradation, IIM can estimate the health state of systems more accurately.

The advantages of using the IIM to model systems are multiple, among which: 1) the normalization of health indicators to obtain inoperability allows modeling systems with heterogeneous components (different health indicators, range values, degradation patterns and failure thresholds); 2) IIM describes a direct relationship between the mission profile effects and the degradation evolution, which eases the adaptation of the mission profile to extend the system life; 3) multiplying the inoperability by 100 gives a percentage of the component degradation relative to its failure threshold, which facilitates communication with the decision-makers.

III. Uncertainty Quantification in System-level Prognostics

The methodology proposed in this paper, and illustrated by Fig. 1, combines the estimated and the predicted system inoperabilities to compute the SRUL. The computational process requires the component-level degradation models, the interactions between the components, the thresholds related to each failure mode and the distributions associated to the uncertainties. The details of the three main steps of the proposed methodology is explained in the next subsections.

A. Inoperability uncertainty estimation

The objective of the first step of the methodology is to estimate the inoperability posterior density of the M system components at each time instant k given the observations y_k. To do that, the particle filtering, which is a popular technique explored by several works in prognostics domain [8], [9], is used. This tool can be applied to systems with non-linear dynamics and non-Gaussian noise. However, contrary to traditional utilization, in this paper a particle is considered as a vector representing the state of health (inoperability) of the system components. Thus, the weight associated to a particle represents the approximation of the inoperability probabilities of all the M components at the same time, as shown in Fig. 2. The process of estimating the inoperability state of a system at time k is explained below.

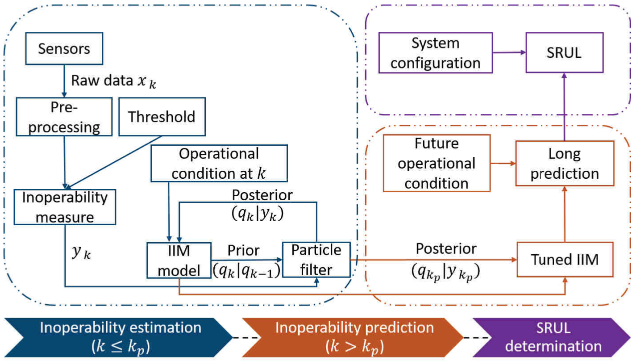

Fig. 1: Uncertainty quantification methodology for system-level prognostics.

Firstly, using the IIM presented in Section II, the prior probability density distributions PDFs of the system components inoperabilities $p(q_k|q_{k-1})$ at time k are predicted based on the ones at the previous time $k-1$:

$$p(q_k|q_{k-1}) \sim IIM(q_{k-1}) \tag{2}$$

Next, given new observations y_{k_i} at time k for a component i, $i \in \{0,1,...,M\}$, the system posterior PDFs inoperabilities are updated by the particle filtering. In detail, considering a set of N particles $\{q^{(l)}\}_{l=1,...,N}$, their associated normalized weights $\{w^{(l)}\}_{l=1,...,N}$ are evaluated by the likelihood functions $p(y_{k_i}|q_{k_i})$ using the importance distribution functions $\pi(q_{k_i}|q_{k-1_i}, y_{1:k_i})$:

$$w_k^{(l)} \propto w_{k-1}^{(l)} \prod_i^M \frac{p(y_{k_i}|q_{k_i}^{(l)})\, p(q_{k_i}^{(l)}|q_{k-1_i}^{(l)})}{\pi(q_{k_i}|q_{k-1_i}, y_{1:k_i})} \tag{3}$$

Finally, to overcome the degeneracy problem, a resampling

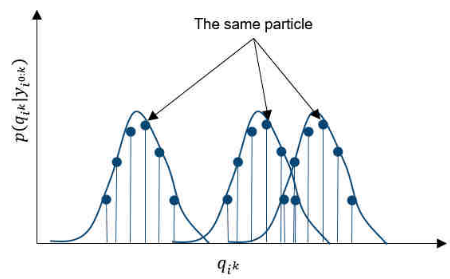

Fig. 2: Inoperability PDFs of a system with 3 components.

process is applied in each time step to replace particles having low importance weights with particles that have higher importance weights.

The posterior PDFs of the system inoperability at time k (Fig. 1) can be approximated before the resampling step by:

$$p(q_k|y_{0:k}) \approx \sum_{l=1}^{N} w_k^{(l)} \delta_{q_k}^{(l)}(q_k) \tag{4}$$

where $\delta(\cdot)$ denotes the Dirac delta function.

The estimation procedure is repeated at every instant k, $k \in \{1,2,...,k_p\}$, where k_p is the starting time of the prediction step presented in the next subsection.

B. Inoperability uncertainty prediction

Prognostics, and thus generation of long-term predictions, is a problem that goes beyond the scope of filtering problem, since it involves future time horizons in which no measurements are available for the Bayesian updating through the equation (3). Thus, the particle filtering, which is more suitable for estimation problems, needs to be adapted to use it for predictions.

In this work, to reduce the computation requirement, we suggest to follow the procedure proposed in [10] and which is based on the assumption that the particle weights are constant from time k_p to time k. According to this procedure, the predicted PDF of the inoperability of the system's components at time k (i.e., $p(q_k|y_{1:k_p})$) can be obtained by applying recursively (2) to $q_{k_p}^{(l)}$.

Once the prediction of the future system inoperability is done, it will be used to determine the system remaining useful life (SRUL), as explained in the next subsection.

C. SRUL determination

The SRUL provides information related to the time when the whole system fails (i.e., when the combined failures of individual components lead to system failure) [2] or when a system reaches performance level that is considered unacceptable. However, the consequence of the degradation of one or more components depends on the considered architecture (e.g. parallel or series). Therefore, the SRUL must be calculated according to system configuration.

Assuming that the system is healthy at time $k_p - th$, moment when the prediction algorithm is launched, the SRUL can be computed as follows:

$$SRUL = \tau_F - k_p \tag{5}$$

with τ_F is the system time-of-failure ToF (or the system end-of-life (EOL)). ToF is chosen in this work because it is a more general concept which can be used in multiple applications [11].

$$\tau_F = inf(k \in N : system\ failure\ at\ k) \tag{6}$$

In practice, and given the complexity of industrial systems, it is important to consider the uncertainty associated with the ToF. To do this, the notations and the new paradigms proposed in [8], [11] are used in the remainder of this paper.

Let's denote a healthy system (with no occurrence of catastrophic failure) and a faulty system (with occurrence of catastrophic failure) at $k-th$ by H_k and F_k, respectively. Let's also consider $H_{k_p:k} = (H_{k_p}, H_{k_p}+1, \cdots, H_k)$ as the sample space that determines all possible sequences where a system has not catastrophically failed until the time k. Then, according to the definition of the conditional probability, the failure probability at $k-th$ is given by:

$$P(F_k) = \frac{P(F_k, H_{k_p:k-1})}{P(H_{k_p:k-1}|F_k)} \tag{7}$$

As the system can only fail once (without maintenance), given that the failure has occurred at time k, the probability of staying healthy until time $k-1$ is $P(H_{k_p:k-1}|F_k) = 1$.

$$P(F_k) = P(F_k, H_{k_p:k-1}) = P(F_k|H_{k_p:k-1})p(H_{k_p:k-1}); \forall k > k_p \tag{8}$$

where $P(F_k|H_{k_p:k-1})$ is given by:

$$P(F_k|H_{k_p:k-1}) = \int_{R^{nq}} p(failure|q_k)p(q_k|y_{1:k_p})d_{q_k} \tag{9}$$

The second term of (8), $p(H_{k_p:k-1})$, stands for the probability that one component is healthy from k_p-th until time $(k-1)-th$, which can be expressed as:

$$p(H_{k_p:k-1}) \quad = \quad \prod_{h=k_p+1}^{k-1} p(H_h|h_{k_p:h-1}) \tag{10}$$

As F_k and H_k are exclusive events, the failure event can be modeled through a Bernoulli stochastic process: $p(H_j|H_{k_p:j-1}) = 1 - p(F_j|H_{k_p:j-1})$. It follows that:

$$p(H_{k_p:j-1}) \quad = \quad \prod_{h=k_p+1}^{k-1} (1 - p(F_h|H_{k_p:h-1})) \tag{11}$$

The expressions presented in (8) and (11) are valid whether for prognostics of a single component or complex systems. However, when considering a multi-components system, the way of characterizing $p(F_k|H_{k_p:k-1})$ will change according to the system configuration.

For exemple, in series configuration of M components, the probability that a system will fail at time k, conditional that it is healthy at $k-1$, is a finite union of the components failure events. As only one component failure can appear at an instantaneous moment, the components failure events can be considered as incompatible. Then, the system failure probability can be written as:

$$p(F_k|H_{k_p:k-1}) = \sum_{i=1}^{M} p(F_{_ik}|H_{k_p:k-1}) \tag{12}$$

where $p(F_{_ik}|H_{k_p:k-1})$ is the probability that component i will fail at time k, conditional that the system is healthy at $k-1$. Then:

$$p(F_k|H_{k_p:k-1}) = \sum_{i=1}^{M} \int_{q_k \in R^{nq}} p(failure_i|q_{ik})p(q_{_ik}|y_{_i1:k_p})d_{q_k} \tag{13}$$

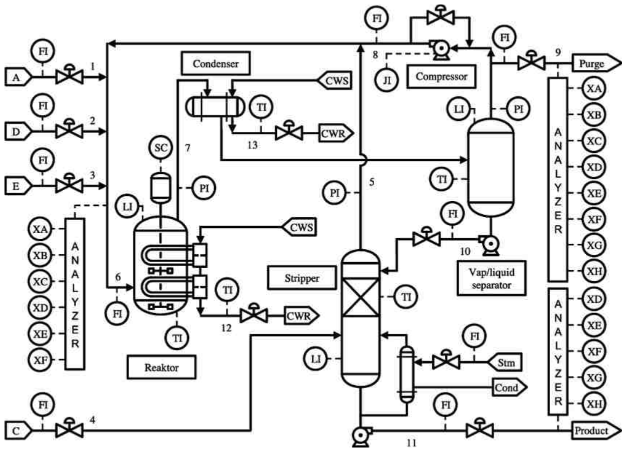

Fig. 3: P&ID of the Tennessee Eastman Process [12].

IV. Tennessee Eastman Process

In this section, the proposed methodology is applied to solve the failure prognostics issue of the Tennessee Eastman Process (TEP).

A. Process description

The Tennessee Eastman Process built by Eastman Chemical Company has been widely used as a realistic benchmark for process control optimization, fault diagnostics and, to a lesser extent, for component-level prognostics. Downs and Vogel [12] described it in detail and provided its simulation where the components, kinetics, and operating conditions have been modified for proprietary reasons. The TEP involves five major units (working in open-loop) including a two-phase reactor, a partial condenser, a separator, a stripper, and a compressor. The schematic piping and instrumentation diagram (P&ID) of the TEP is shown in Fig. 3.

In the TEP, the gaseous reactants $\{A, C, D, E\}$ are fed to the reactor where the liquid products $\{G, H\}$ and the byproduct $\{F\}$ are formed through the following exothermic reactions:

$$\begin{cases} A(g) + C(g) + D(g) \rightarrow G(l), \\ A(g) + C(g) + E(g) \rightarrow H(l), \\ A(g) + E(g) \rightarrow F(l), \\ 3D(g) \rightarrow 2F(g). \end{cases}$$

The process has a total of 53 measured variables, of which 22 variables are continuous process measurements (such as: temperatures, pressures, flow rates, levels), 19 variables are composition measurements and remaining 12 are manipulated variables. 28 faults can be injected in the process [13], which can be related to step changes, drifts, and random variation of variables, etc.

B. Problem Statement

In this case study, we consider a failure as interruption of the operational continuity resulting from violation of the variables shutdown limits. Therefore, only components with shutdown constraints are considered, i.e. reactor, stripper and separator. Each of these components is monitored by a single parameter: pressure for the reactor, and level for the stripper and the separator. Table I lists the specific operational constraints related to the system parameters that the control system should respect.

TABLE I: Process operating constraints [12].

	Normal operating limits		Shut down limits	
Process variables	Low limit	High limit	Low limit	High limit
Reactor pressure	none	2895 kPa	none	3000 kPa
Separator level	3.3 m	9.0 m	1.0 m	12 m
Stripper level	3.5 m	6.6 m	1.0 m	8.0 m

In Matlab® simulations of the TEP [14], three disturbances predefined in [13] were injected. Two disturbances are deviations in the reactor and the stripper parameters that are, respectively: deviation in the reactor cooling water flow ($c_1(t)$) and deviation in the heat transfer of the heat exchanger of the stripper ($c_2(t)$). These deviations $c_1(t)$ and $c_2(t)$ are assumed to follow equations (14) and (15), respectively.

$$c_1(t) = \alpha . c_1(t-1) + \beta \tag{14}$$

$$c_2(t) = \gamma . c_2(t-1) \tag{15}$$

with α, β and γ are the parameter of the two models.

As shown in Fig. 4, when these two faults are injected, the system shuts down at 0.8 hour.

The third injected fault is a random variation of the reactor cooling water inlet temperature and can be considered as process noise.

The purpose, here, is to estimate the SRUL when taking into account the process uncertainty, which is characterized by a random variation of the reactor cooling water inlet temperature. To do this, the proposed IIM, Eq.(1) where $c(t)$ is given by Eq.(14) and (15), is used to model the system degradation. The IIM parameters are estimated and adjusted thanks to data collected from the TEP simulation. In the next subsection, construction of the IIM is detailed.

C. Inoperability input-output model of TEP

The raw data acquired from the TEP simulation are normalized to the initial state and the failure threshold, as

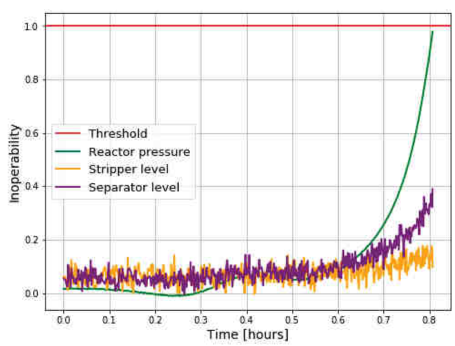

Fig. 4: Components inoperabilities evolution when considering two faults.

shown in Fig. 4. They represent the inoperability of the three components: reactor $q_1(t)$, stripper $q_2(t)$ and separator $q_3(t)$. Their initial states correspond to the parameters base value [12] and the failure thresholds to the shutdown limits values (Table I).

The IIM of the TEP is built from data obtained after injection of these faults. It should be noted here that the IIM model does not model the physical phenomena operating in the process, but is more used to fit the actual data of the process. For this case study, the interdependence matrix (A) is estimated and adjusted to make the proposed model better fit the process data. Then, we obtain the following result:

$$A = \begin{bmatrix} 0 & 0.1 & 0 \\ 10^{-5} & 0 & 0 \\ 0.4 & 0.3 & 0 \end{bmatrix} \tag{16}$$

Regarding the matrix K, it has been assumed that the environmental conditions have no effects on the evolution of the components inoperabilities. Therefore, its diagonal is equal to 1.

D. Inoperability estimation and prediction

After application of the random variation of the reactor cooling water inlet temperature, the data obtained from the process and the built IIM are used in the particle filtering to estimate the components inoperabilities. To evaluate the inoperabilities densities, 200 particles were used with the initial distributions of the components inoperabilities considered as Gaussian. The selection of the particles to be retained after each filtering step was done by using residual resampling.

When the inoperability of one component exceeds the normal operating limits (as indicated in Table I), the prediction step will be launched (at time k_p).

The results of the estimation and prediction of the inoperabilities uncertainty of the system components are shown in Fig. 5. The reactor is the first component to go out of its normal operating limits after 0.34 hour. This time corresponds to the time where the long-term inoperability prediction is launched. Also, it is the reactor pressure, that triggers the system shutdown (system failure) at 0.44 hour.

From these results, we can highlight the power of estimation and prediction of the proposed methodology. Indeed, before the adding of the random variation (the process noise), the system shuts down at 0.8 hour and after the fault injection, it shuts down at 0.44 hour. Despite this high process variability, the particle filtering was able to estimate the actual inoperability of the system (as shown in Fig. 5). Also, one can notice that the uncertainty related to the predicted inoperability increases for $k > k_p$. This is due to the fact that no measurements were received and, therefore, there is neither updating of the particle weights nor resampling.

E. SRUL determination

For this case study, the operability of the studied system depends on the operability of its components since they all contribute to realization of the system function (G and H production). Therefore, one can conclude that the system has a series architecture. Consequently, the system ToF probability will be determined using the equation (13).

Fig. 6 shows the PMF of the SRUL. The mean value of the SRUL is equal to 0.1 hour. The true SRUL, which is equal to 0.102 hour, is within the 95% confidence interval of the predicted SRUL distribution. We can conclude that

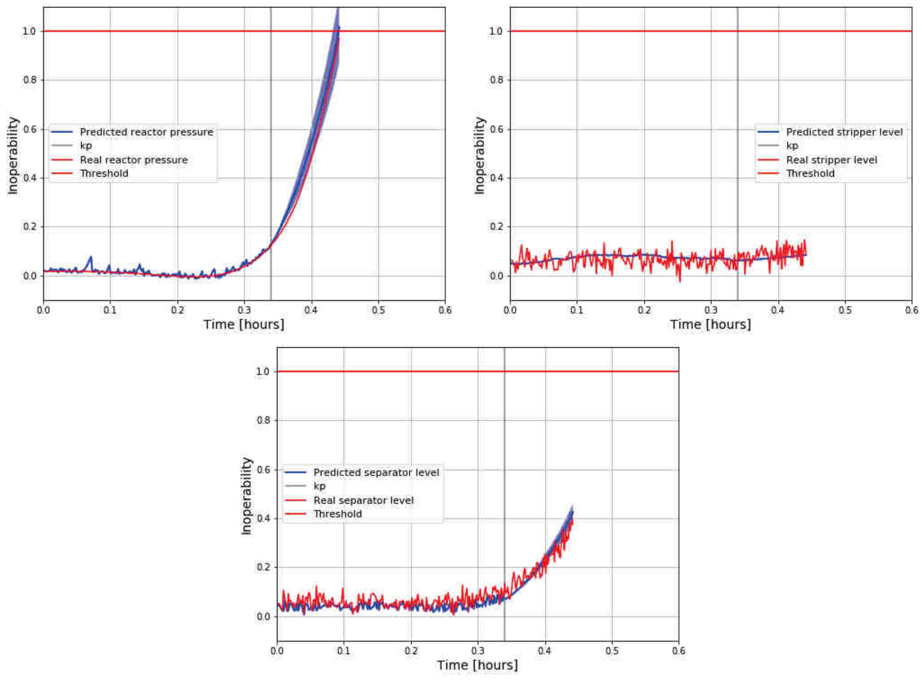

Fig. 5: Components real and predicted inoperabilities evolution when considering three faults.

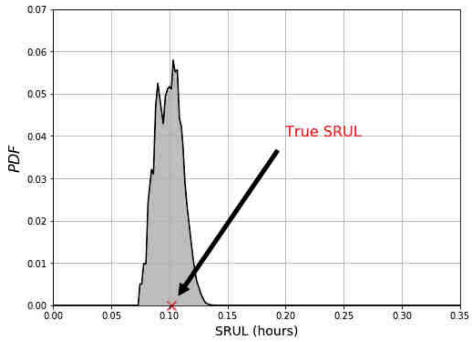

Fig. 6: SRUL probability distribution of the system.

the predicted SRUL is close to reality and is slightly pessimistic, which does not put the system, its operators and the environment in danger.

In order to discuss the results of the proposed methodology, a study of the impact of the prognostics horizon on uncertainty intervals is performed by considering the α-accuracy metric, which determines whether a prediction falls within an α% interval. In fact, α-accuracy is a useful metric to judge if one prognostics algorithm converges to the true value as more information are accumulated over time. Indeed, a faster convergence is desired to achieve a high confidence, keeping the prediction horizon as large as possible. To that end, in this study, the prognostics time has been varied within the interval shown in Fig. 7 and the accuracy is defined with $\alpha = 10\%$. The figure shows the mean values and uncertainties of the predicted SRUL distributions compared with the true SRUL. As it can be seen, the prediction of the SRUL becomes more accurate each time the measurements are obtained. One can also notice that for $t \leq 0.6$, the SRUL remains almost constant because the values of the monitored parameters do not change a lot, as it can be seen in Fig. 4.

V. Conclusion

A methodology for the uncertainty quantification at system-level prognostics is proposed in this paper. This methodology results in three main contributions. The first concerns the modeling of the system degradation by using the inoperability input-output model. The second deals with the health state estimation based on an adapted particle filtering. Finally, the third contribution is related to the calculation of the system remaining useful life, based on the recent developments and achievements proposed at component-level prognostics and generalized in this paper to system-level prognostics.

The proposed methodology was applied on the Tennessee Eastman process in order to predict the shutdown time caused by violation of the process constraints. The obtained results show the effectiveness of the methodology in estimating and predicting the system remaining useful life and characterization of uncertainty.

For future work, a systematic method to determine the IIM parameters from a real data will be proposed. Also, it would be worthwhile to apply the other predefined disturbances in the TEP in order to test the robustness of the methodology.

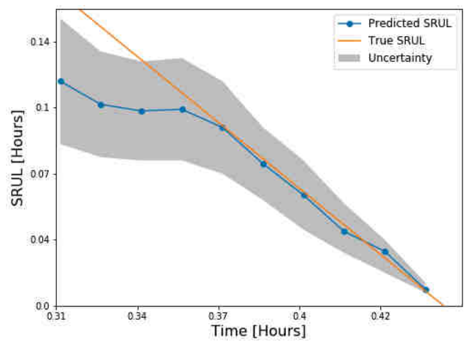

Fig. 7: SRUL prediction performance with $\alpha = 0.1$.

References

[1] V. Atamuradov, K. Medjaher, P. Dersin, B. Lamoureux, and N. Zerhouni, "Prognostics and health management for maintenance practitioners-review, implementation and tools evaluation," *International Journal of Prognostics and Health Management*, vol. 8, no. 060, pp. 1–31, 2017.

[2] L. R. Rodrigues, "Remaining useful life prediction for multiple-component systems based on a system-level performance indicator," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 1, pp. 141–150, 2018.

[3] B. Saha and K. Goebel, "Uncertainty management for diagnostics and prognostics of batteries using bayesian techniques," in *Aerospace Conference, 2008 IEEE*. IEEE, 2008, pp. 1–8.

[4] W. W. Leontief, "Quantitative input and output relations in the economic systems of the united states," *The review of economic statistics*, pp. 105–125, 1936.

[5] Y. Y. Haimes and P. Jiang, "Leontief-based model of risk in complex interconnected infrastructures," *Journal of Infrastructure systems*, vol. 7, no. 1, pp. 1–12, 2001.

[6] J. R. Santos and Y. Y. Haimes, "Modeling the demand reduction input-output (i-o) inoperability due to terrorism of interconnected infrastructures," *Risk Analysis: An International Journal*, vol. 24, no. 6, pp. 1437–1451, 2004.

[7] F. Tamssaouet, T. P. K. Nguyen, and K. Medjaher, "System-level prognostics based on inoperability input-output model," in *Annual conference of the prognostics and health management society*, vol. 10, 2018.

[8] D. E. Acuña and M. E. Orchard, "Particle-filtering-based failure prognosis via sigma-points: Application to lithium-ion battery state-of-charge monitoring," *Mechanical Systems and Signal Processing*, vol. 85, pp. 827 – 848, 2017.

[9] M. Orchard, "A particle filtering-based framework for on-line fault diagnosis and failure prognosis," Ph.D. dissertation, Georgia Institute of Technology, Atlanta, CA, 2006.

[10] A. Doucet, S. Godsill, and C. Andrieu, "On sequential monte carlo sampling methods for bayesian filtering," *Statistics and computing*, vol. 10, no. 3, pp. 197–208, 2000.

[11] D. Acuña and M. Orchard, "A theoretically rigorous approach to failure prognosis," in *Annual conference of the prognostics and health management society*, vol. 10, 2018.

[12] J. J. Downs and E. F. Vogel, "A plant-wide industrial process control problem," *Computers & chemical engineering*, vol. 17, no. 3, pp. 245–255, 1993.

[13] A. Bathelt, N. L. Ricker, and M. Jelali, "Revision of the tennessee eastman process model," *IFAC-PapersOnLine*, vol. 48, no. 8, pp. 309–314, 2015.

[14] N. L. Ricker, "Tennessee eastman challenge archive," http://depts.washington.edu/control/LARRY/TE/download.html, online; accessed 15 December 2018.