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

Data-Driven Fault Prognosis Based on Incomplete Time Slice Dynamic Bayesian Network[1]

Zhengdao Zhang*. Feilong Dong. Linbo Xie*.**

** Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, 214122, P.R.C. (e-mail: wxzzd@ Jiangnan.edu.cn, xielb@126.com).*
***School of IoTs, Jiangnan University, Wuxi, 214122, P.R.C. (e-mail: 478309108@qq.com)*

Abstract: Based on a dynamic Bayesian network with an incomplete time slice and a mixture of the Gaussian outputs, a data-driven fault prognosis method for model-unknown processes is proposed in this article. First, according to the requirement of fault prognosis, an incomplete time slice Bayesian network with unknown future observed node is constructed. Moreover, the future states are described by the current measurements and his historic data in the form of conditional probability. Second, according to the completed part of historical data, a parameter-learning algorithm is used to obtain network parameters and the weight coefficients of distribution components. After that, using such weight coefficients as input-output data, the subspace identification method is employed to build a forecasting model which can predict weight coefficients at next sampling time. To achieve fault prognosis, an inference algorithm is developed to predict hidden faults based on the distribution of the measurements directly. Furthermore, the remaining useful life of process is estimated via iterative one-step ahead prognosis. As an example, the proposed method is applied to a continuous stirred tank reactor system. The results demonstrate that the proposed method can efficiently predict and identify the fault, and estimate the remaining useful life of process, even though the measurements are partly missing.

Keywords: fault prognosis, data missing, dynamic Bayesian network, subspace identification, remaining useful life.

1. INTRODUCTION

With the increasing of reliability requirement in industry, we hope that we cannot only detect the faults after they occur but also predict failures or unacceptable degradations of performance before they occur. Fault prognosis focuses on using the historical and current measurements of system to deal with fault prediction and the estimation of the remaining useful life (RUL) in acceptable operating state (Yu, Wang, Luo, et al., 2011). By now, fault prognosis plays a key role in reliability, low environmental risks, and human safety.

The existing prognosis methods can be divided into three categories: physical-model based methods, knowledge based methods, and data-driven methods (Peng, Dong & Zhou, 2010). The difficulties in obtaining the precise mathematical models obstruct the applications of physical-model based methods. Knowledge based methods rely heavily on whether the domain knowledge can completely describe the normal or faulty conditions of the system. In addition, to obtain the complete knowledge both on the normal states and on faulty situations is difficult, and sometimes the knowledge even contains incorrect rules. Data-driven methods just use the measurements of system input and output variables, which can avoid the difficulties in obtaining the precise mathematical models and the complete knowledge of the process. Based on the training dataset represented the normal or faulty conditions, neural network can forecast the states or distinguish the normal and fault states (Huang, Xi, Li, et al., 2013). Masoud et al. used locally linear neuro-fuzzy model to predict and detect some common faulty conditions in the cement rotary kiln (Masoud & Alireza, 2011). Chen et al. (2011) developed a prognostic method based on adaptive neuro-fuzzy inference systems (ANFISs) and high-order particle filtering. Abdenour et al. combined the hidden Markov models and ANFIS for the prognostics and health management of roller bearings, and the estimation of the remaining time of the current state is predicted (Soualhi, Razik, Clerc, et al., 2014). However, the prediction performance for the aforementioned methods depends on whether the neural network is properly trained. The incomplete train dataset and the unexpected noise can cause the performance degradation.

[1] This work was supported in part by the National Natural Science Foundation of China (No. 61374047, No. 61202473) and the Fundamental Research Funds for Central Universities (JUSRP111A49).

As an important kind of data-driven method, multivariate statistical process monitoring method has been successfully applied in the field of conditional monitoring and fault detection (Qin, 2012; MacGregor & Cinar, 2012), but there are rarely applications in the fault prognosis field. The most important reason is that multivariate statistical process monitoring methods essentially cannot forecast the future states of system, while fault prognosis need combine fault diagnosis with event forecasting. Therefore, integrating model based state estimation into multivariate statistical process monitoring could be a candidate method for fault prognosis. Liu et al.(2012) proposed a novel data-model-fusion prognostic framework integrating the data-driven method and the particle filtering approach, which can improve the accuracy of system state with long-horizon forecasting. Wang & Wang (2012) proposed a prognostic procedure based on expectation maximization and unscented Kalman filter. System states are estimated by an unscented Kalman filter upon sensor measurements, and then, component-specific parameters in a degradation process are identified on the estimation of the degradation process. Mosallam et al. (2013) combined the PCA and AR modelling to machinery health prognosis. Li, Qin, Ji & Zhou (2010) proposed a multivariate fault prognosis approach for continuous processes, in which fault magnitude is estimated via fault reconstruction and is predicted by a vector AR model.

Data with missing values is a common problem in practice (Deng & Huang, 2012). However, traditional data-driven methods, such as PCA, PLS, ICA, consider well-conditioned data sets only. Data imputation approaches are commonly used to avoid detection delays or failures in FDD, which is caused by incomplete data (Imtiaz & Shah, 2008). The main drawback of the imputation approaches is the variances of the data may be considerably changed after imputation (Khatibisepehr, Huang & Khare, 2013). Expectation maximization (EM) imputes the missing values in a statistical way so that there is no change in the statistical properties of the data set. Yin et al. incorporated the partial robust M-regression method in the EM framework, and then, the KPI-related prediction and diagnosis were accomplished with PLS (Yin, Wang & Yang, 2014). However, the efficiency of EM-based imputation depends on the missing rate of the original samples. With the increase of missing rate of the data set, EM-based imputation may cause the misled filled-in values as well as the misled FDD results.

Based on our prior researches, we found that dynamic Bayesian network with mixture of Gaussian output (DBNMG) can deal with the missing data in the manner of the marginal density function and conditional probability density function (Zhang & Dong, 2014). Expanding our prophase research works, a fault prognosis method, which integrates the subspace identification with incomplete DBNMG (IT-DBNMG), is proposed in this article. First, we construct the IT-DBNMG which is adapted for fault prognosis. After the parameter learning dynamic Bayesian network (DBN) is implemented, a Gaussian component coefficients prediction model is established by applying the subspace identification method. Then, a modified inference algorithm is developed to accomplish the fault prognosis and to estimate the RUL of process. This paper is organized as follows. In section 2, the IT-DBNMG for fault prognosis is constructed. After a briefly introduction of the parameter learning algorithm, the prediction modelling and a modified inference algorithm are deduced in section 3. Then, the procedure of proposed fault prognosis method is introduced in section 4. In section5, as an example, we use continuous stirred tank reactor (CSTR) process to show the efficiency of the approach. Finally, the last section concludes the paper.

2 PROBLEM FORMULATION OF IT-DBNMG BASED FAULT PROGNOSIS

In training phase, IT-DBNMG consists of three time slices as shown in Fig. 1 (Zhang & Dong, 2014), where the node R_t represents set of random variables at the tth time instant, which can take on N possible values, $R_t = \{1,...,N\}$, M_t denotes the number of mixture Gaussian components at the tth time instant, and node Y_t denotes the observed variable at the tth instant.

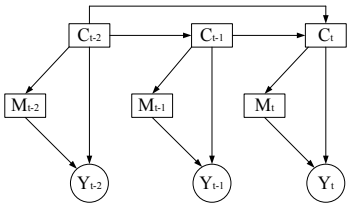


Fig.1. Structure of IT-DBNMG model in training phase.

The main parameters of IT-DBNMG are initial state probability distribution $P(C_0)$, state transition model $P(C_t \mid C_{t-2}, C_{t-2})$ and observation model $P(Y_t \mid C_t)$, which can be rearranged in the form of vector as

$$\theta = \{P(C_t \mid C_{t-1}, C_{t-2}), P(Y_t \mid C_t, M_t), P(M_t \mid C_t)\} \quad (1)$$

At instant t, the measurements Y_{t+1} are unknown. Noticed that Y_{t+1} is used for probabilistic distribution estimating rather than for the fault prediction directly. So, we only need to establish a forecasting model for the weight coefficients of the probabilistic components here. The schematic diagram of fault prognosis via IT-DBNMG is shown in Fig.2.

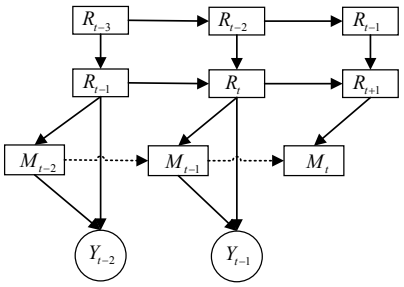


Fig. 2. Schematic diagram of IT-DBNMG in prognosis phase.

3 PREDICTION MODELLING AND PROGNOSIS INFERENCE ALGORITHM

3.1 Parameter learning algorithm and weight coefficients of distribution components

In the training phase, only the well-condition data is chosen. Using EM algorithm iteratively, θ can be calculated by EM algorithm (Zhang & Dong, 2014), which can be summarized as follows, where k means the iteration step.

(i) E-step

$$\begin{aligned} &Q(\theta \mid \theta_k) \\ &= \sum_{a=1}^{S}\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{N}\sum_{m=1}^{M}\{P(R_t=i,R_{t-1}=j,R_{t-2}=n,M_t=m \mid y_{a,t},\theta_k) \\ &\cdot \log P(R_t=i,R_{t-1}=j,R_{t-2}=n,M_t=m \mid \theta)\} \\ &+ \sum_{a=1}^{S}\sum_{t=1}^{T}\sum_{n=1}^{N}\sum_{j=1}^{N}\sum_{i=1}^{N}\sum_{m=1}^{M}\{P(R_t=i,RC_{t-1}=j,R_{t-2}=n,M_t=m \mid y_{a,t},\theta_k) \\ &\cdot \log P(y_{a,t} \mid R_t=i,R_{t-1}=j,R_{t-2}=n,M_t=m,\theta)\} \end{aligned} \tag{2}$$

(ii) M –step

$$\theta_{k+1} = \arg\max_{\theta} Q(\theta \mid \theta_k) \tag{3}$$

If the measurement y_t is complete at time t, the weight coefficients of distribution components can be computed as

$$P(M_t=m \mid R_t=i) = \frac{P(y_t \mid R_t=i,M_t=m)}{\sum_{i=1}^{N}\sum_{m=1}^{M} P(y_t \mid R_t=i,M_t=m)} \tag{4}$$

For the incomplete measurement y_t, we define the missing part of y_t is $y_{t,u}$ and the observable part of y_t is $y_{t,o}$. Since the covariance matrix of $P(y_t \mid R_t=i,M_t=m)$ is a diagonal matrix, we have

$$P(y_{t,o} \mid R_t=i,M_t=m) = \int P(y_{t,o},y_{t,u} \mid R_t=i,M_t=m)dy_{t,u}.$$

Thus, the weight coefficients of distribution components $P(M_t=m \mid R_t=i)$ can be computed as

$$P(M_t=m \mid R_t=i) = \frac{\int P(y_{t,o},y_{t,u} \mid R_t=i,M_t=m)dy_{t,u}}{\sum_{i=1}^{N}\sum_{m=1}^{M}\int P(y_{t,o},y_{t,u} \mid R_t=i,M_t=m)dy_{t,u}} \tag{5}$$

3.2 Model formulation Prediction modelling method

Base on the above results, if we can forecast the $P(M_{t+1} \mid R_{t+1})$, we can predict the fault via modified inference algorithm.

Let $x(t) \in R^l$, $u(t) = [P(M_t \mid R_t)^T,...,P(M_{t-k+1} \mid R_{t-k+1})^T]^T \in R^{kM}$ and $z^i(t+1) = P(M_{t+1} \mid R_{t+1}) \in R^M$, then the prediction model in the output error configuration is:

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) \\ z(t) = Cx(t) \end{cases} \tag{6}$$

where $A \in R^{l\times l}$, $B \in R^{l\times kM}$, $C \in R^{M\times l}$. In this article, the N4SID method (Overschee & Moor, 1994) is used to estimate the matrix A, B and C.

3.3 Prognosis inference algorithm

According to the historical measurement of observed nodes Y_t and Y_{t-1}, and the historical states of hidden nodes R_t and R_{t-1}, the posterior probability of y_t can be obtained as

$$\begin{aligned} P(R_t=i \mid y_t,R_{t-1}=j,R_{t-2}=n) &= \frac{P(R_t=i,y_t,R_{t-1}=j,R_{t-2}=n)}{P(y_t,R_{t-1}=j,R_{t-2}=n)} \\ &= \beta_t P(R_t=i \mid R_{t-1}=j,R_{t-2}=n)\cdot P(y_t \mid R_t=i) \end{aligned} \tag{7}$$

where $\beta_t = \frac{1}{P(y_t)}$ is the normalized factor.

Substituting Eq 2 into Eq 7 gives

$$\begin{aligned} P(R_t=i \mid y_t,R_{t-1}=j,R_{t-2}=n) &= \beta_t P(R_t=i \mid R_{t-1}=j,R_{t-2}=n) \\ &\cdot \sum_{m=1}^{M} P(Y_t \mid R_t=i,M_t=m)P(M_t=m \mid R_t=i) \end{aligned} \tag{8}$$

By replacing $P(M_{t+1}=m \mid R_{t+1}=q)$ with the predicted $\hat{P}(M_{t+1}=m \mid R_{t+1}=q)$, the estimation of the probability density function of the measurements can be obtained as:

$$\begin{aligned} \hat{P}(y_{t+1} \mid R_t=i) &= \sum_{m=1}^{M}\sum_{q=1}^{N}\{\hat{P}(M_{t+1}=m \mid R_{t+1}=q) \\ &\cdot P(R_{t+1}=q \mid R_t=i)P(y_{t+1} \mid M_{t+1}=m,R_{t+1}=q)\} \end{aligned} \tag{9}$$

where $P(y_{t+1} \mid M_{t+1}=m,R_{t+1}=q) = P(y_t \mid M_t=m,R_t=q) = N(\mu_{q,m},\Sigma_{q,m})$ was obtained by parameter learning algorithm.

Combining Eqs (8) and (9), the posterior probability of future state can be expressed as

$$\begin{aligned} \hat{P}(R_{t+1}=q \mid \hat{y}_{t+1},R_t=i,R_{t-1}=j) &= \beta_{t+1}P(R_{t+1}=q \mid R_t=i,R_{t-1}=j) \\ &\cdot \sum_{m=1}^{M}\hat{P}(M_{t+1}=m \mid R_{t+1}=q)P(\hat{y}_{t+1} \mid M_{t+1}=m,R_{t+1}=q) \end{aligned} \tag{10}$$

where $P(R_{t+1}=q \mid R_t=i,R_{t-1}=j)$ which can be approximated by $P(R_t=q \mid R_{t-1}=i,R_{t-2}=j)$.

We can judge the state of R_{t+1} based on the maximum posterior probability of state, namely

$$\begin{aligned} R_{t+1} = \max_q\{&P(R_{t+1}=q \mid R_t=i,R_{t-1}=j) \\ &\cdot \sum_{m=1}^{M}\hat{P}(M_{t+1}=m \mid R_{t+1}=q)P(\hat{y}_{t+1} \mid M_{t+1}=m,R_{t+1}=q)\} \end{aligned} \tag{11}$$

4 FAULT PROGNOSIS BASED ON IT-DBNMG

Consider the nonlinear process system, whose mathematic model is unknown. The output measurement at time instant t is $y(t)$. An indication vector λ_t is assigned to simulate the missing phenomenon of process data, whose elements are a Bernoulli distributed stochastic variable. If the ith element in $y(t)$ is missing, we set $\lambda_t^i = 0$. The indicator $f = 1...N$ is defined to represent the unknown work states of the system, where $f_t = 1$ denotes the normal state and $f_t = i$ $(i \neq 1)$ denotes the different faults, N is the number of all states including normal state and faults.

For simplicity, we assume that only one fault occurs at the same time in this work. Our task is to determine the value of indicator f in the future time instant. The framework of the prognosis procedure is sketched it as follow,

In the training phase:

Step (a): Construct the IT-DBNMG as shown in Fig. 2, where $y(t)$ is expressed by the node Y_t and f is expressed by the node R_t and the initial value of θ_t is selected randomly.

Step (b): According to measurements of system, determine the number of the mixed Gaussian components.

Step (c): Obtain the network parameters according to the Parameters learning algorithm, Eqs (2) and (3).

Step (d): Obtain the weight coefficients of Gaussian components in training process according to Eqs (4) and (5), and use them to establish the subspace forecasting model.

In the prognosis phase:

Step (e): Use the current measurement to obtain a series of posterior distribution according to the inference algorithm of the IT-DBNMG. And then, according to Eq (11), the current state of the system can be determined.

Step (f): Use the current measurement to obtain the current weight coefficients of Gaussian components, then forecast the weight coefficients of future based on the obtained model.

Step (h): Use the prognosis inference algorithm to judge the system prognosis state and to achieve the fault prognosis.

5 CASE STUDY

A case study on nonisothermal CSTR with time varying parameters (Nikravesh, Farell & Stanford, 2000) is used to illustrate the application of the proposed fault prognosis method. The measurements of process include outlet concentration C_A and reacting mixture temperature T. A measurement noise follows $N(0, 0.64)$ is added to the reactor temperature variable, a noise follows $N(0, 0.0016)$ is added to the outlet reagent concentration variable, and an external disturbance follows $N(0,1)$ is added to the reactor feed flow rate variable. The sampling period of process is set to 0.1min.

We set the faults are feed flow rate fault, feed concentration fault and linear fouling fault, which are described by Eqs (12)-(14),

$$q(t) = \begin{cases} 100 & ,t < t_0 \\ q(t_0) + \Gamma_q (1 - \exp(\frac{t - t_0}{80}) & ,t \geq t_0 \end{cases} \tag{12}$$

$$C_{Af}(t) = \begin{cases} 1 & ,t < t_0 \\ C_{Af}(t_0) + \Gamma_C (1 - \exp(\frac{t - t_0}{80}) & ,t \geq t_0 \end{cases} \tag{13}$$

$$\phi_h(t) = \begin{cases} 1 & ,t < t_0 \\ 1 - \alpha_h (t - t_0) & ,t \geq t_0 \end{cases} \tag{14}$$

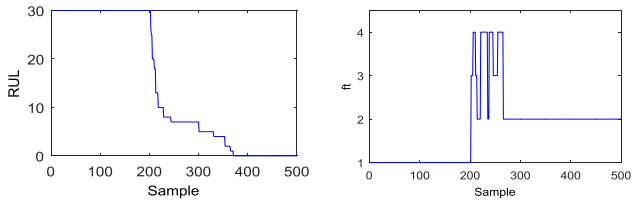
where q is the low rate, C_{Af} is the feed concentration, ϕ_h is the fouling coefficient. Γ_q, Γ_C and α_h denote the fault magnitudes which are assumed to $\Gamma_q = 1$, $\Gamma_C = 0.02$ and $\alpha_h = 0.01$ in the train phase, and assumed to $\Gamma_q = 1.5$, $\Gamma_C = 0.02$ and $\alpha_h = 0.008$ in the test phase. The system is considered as normal when the variables C_A and T fluctuate within the normal ranges $[7.36 \times 10^{-2}, 9.36 \times 10^{-2}](mol/L)$ and $[431, 449](K)$ respectively, otherwise the system is faulty.

According to Fig. 2, let the output node $Y_t = (C_A, T)^T$. The root node R_t denotes the indicator f, where $R_t = \{2, 3, 4\}$ denotes the process is under the feed flow rate fault, the feed concentration fault, and the linear fouling fault, respectively. The number of Gaussian is obtained with a 10-fold cross-validation, and we get $M_t = 3$. Assume that the fault occurs at $t_0 = 100$ sample in the training phase, and it occurs at $t_0 = 200$ in the test phase. The simulation results are evaluated in terms of two performance indexes: the accuracy rate and the advanced time. Let the number of samples from first time the fault is predicted to the time that one of the two measurements exceeds their threshold is TN and the number of fault alarms with correct category is TL, then the accuracy rate is defined as TL/TN. The advance time is defined as the time between the first time the fault is predicted and the time the index exceeds the threshold.

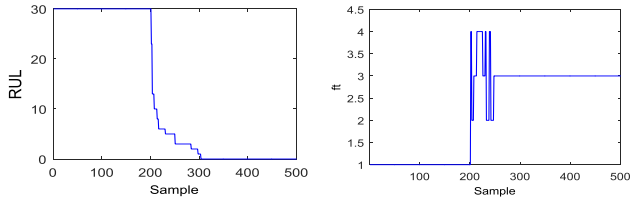
Using the proposed one-step ahead prognosis method iteratively, the RUL can be estimated online. the When we set the maximum of prediction steps to 30, the RUL prediction results are depicted in Figure 3, where the subfigure of RUL shows the RUL prediction results and the subfigure of the indicator f shows the fault identification results for the future faults. With the fault growing, the predicted RUL drops under 30, which means an early warning is triggered. And then, the predicted fault is classified to different fault states. For the feed flow rate fault, the alarm is triggered at 202st instant because the RUL drops to 29. And then, from 365th instant, we judge the fault will happen at next instant. The advance time is 17.5 min. the

increasing of advance time is due to that forecasting modelling amplifies the failure symptom and makes it clear. Therefore, RUL predictions are effective and can help to improve the system safety. Also, due to the forecasting error, the diagnosis results have some mistakes in the early phase. But, we can also obtain the correct diagnosis results before one-step prognosis does.

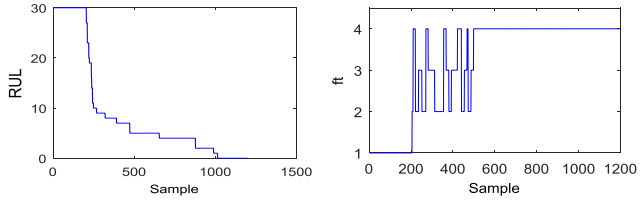
We list the 30-step ahead prognosis results of fault prognosis performance for the feed concentration fault process by applying the approach in (Li, Qin, Ji & Zhou, 2010) and the proposed method in Fig. 4, where the subfigure of RUL shows the RUL prediction results and the subfigure of fault magnitude shows the prediction results of fault magnitude for the future faults. Comparing the Fig.4b with Fig. 3, we can find that the predicted RULs of two methods drop under 30 steps at same time, which both give an early warning effectively. The RUL prediction of the proposed method is more stable, that is valuable to arrange the maintenance beforehand, and improve the system safety significantly. Furthermore, our method can identify the faults, which is not considered by the method in Li, Qin, Ji & Zhou (2010).



(a) Results under the feed flow rate fault process



(b) Results under the feed concentration fault process



(c) Results under the fouling fault process

Fig. 3. 30-step ahead prognosis results with complete data

To evaluate the efficiency of the proposed method when the measurements are partly missing, we increase the data missing rate from 0 to 10%. The part of results are shown in Fig.5. Due to the missing data existing, the distribution information of measurements is imprecise. So, it is hard to identify the correct fault in the beginning phase of prognosis. Also, the first time for prognosis has a little bit of delay.

Monte Carlo simulations are repeated 100 times to evaluate the performance with different missing data rates, and the results are summarized in Table 1. The proposed method can prognosis and identify different faults accurately when the data is complete. With the missing data rate increasing, prognosis accuracy rates are kept at same levels. With the increasing number of prediction step, the advance times of fault alarm are increased because the fault symptoms are amplified further. Meanwhile, at the early stage of prognosis, the errors of fault identification are increasing. So, the accuracy rates for three faults are all decreased. Comparing the ahead steps with the different missing rates, we can find that the increasing of the value of ahead step makes more predicted fault be misclassified, which means that the forecasting modelling method should be studied further.

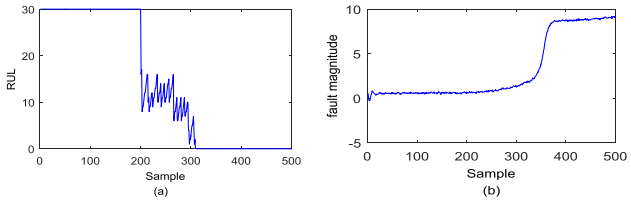
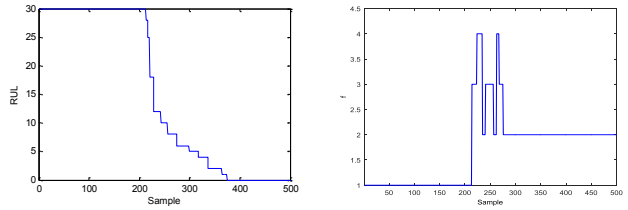
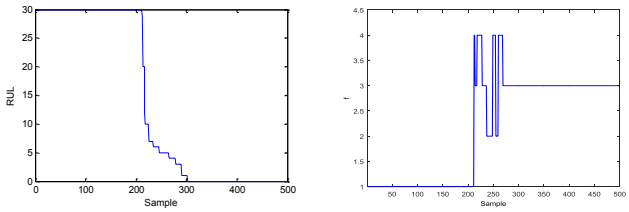


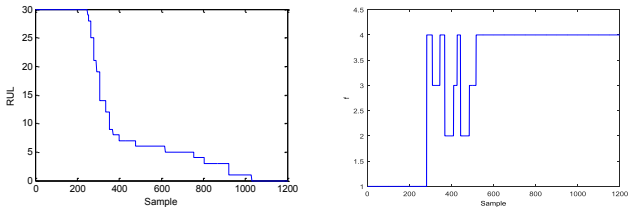
Fig. 4. 30-step ahead prognosis performance of the reported scheme in Li, Qin, Ji & Zhou (2010) under the feed concentration fault process



(a) Results under the feed flow rate fault process



(b) Results under the feed concentration fault process



(c) Results under the fouling fault process

Fig. 5. 30-step ahead prognosis results with 10% missing data

6 CONCLUSIONS

This paper proposes a data-based fault prognosis method for continuous processes with partly missing measurements. The method uses the IT-DBNMG to represent the relationships between the measurements and the hidden faults. The distribution features of process measurements are extracted and be estimated based on a subspace forecasting model. The prediction of the remaining useful life is intergraded with the fault prognosis method in this paper. The effectiveness of the proposed approach is demonstrated on a CSTR process. Also, a comparison with our method and method in (Li, Qin, Ji & Zhou, 2010) is included to evaluate the approach. The prediction model may affect the performance significantly, which should be chosen properly in next.

REFERENCES

Chen C., Zhang B., Vachtsevanos G., et al. (2011). Machine condition prediction based on adaptive neuro–fuzzy and high–order particle filtering, *IEEE Trans. Ind. Electron.*, 58(9),4353-4364.

Deng J., and Huang B. (2012). Identification of nonlinear parameter varying systems with missing output data, *AIChE Journal*, 58(11), 3454-3467.

Huang R., Xi L., Li X., et al. (2007). Residual life prediction for ball bearings based on self–organizing map and BP neural network methods, *Mechanical Syst. Signal Process.*, 21(1), 193-207.

Imtiaz S., and Shah S. (2008). Treatment of missing values in process data analysis, *Can. J. Chem. Eng.*, 86(5), 838-858.

Khatibisepehr S., Huang B., and Khare S. (2013). Design of inferential sensors in the process industry: a review of Bayesian methods, *J. Proc. control*, 23(10), 1575-1596.

Li G., Qin S. J., Ji Y., and Zhou D. (2010). Reconstruction based fault prognosis for continuous processes, *Control Eng. Pract.*, 18(10), 1211-1219.

Liu J., Wang W., Ma F., et al. (2012). A data–model–fusion prognostic framework for dynamic system state forecasting", *Eng. Appl. Artif. Intell.*, 25(4), 814-823.

MacGregor J., and Cinar A. (2012). Monitoring, fault diagnosis, fault–tolerant control and optimization: data driven methods, *Comput. Chem. Eng.*, 47(52), 111-120.

Masoud S., and Alireza F. (2011). Identification, prediction and detection of the process fault in a cement rotary kiln by locally linear neuro-fuzzy technique, *J. Proc. cont.*, 21(2), 302-308.

Mosallam A., Medjaher K., and Zerhouni N. (2013). Nonparametric time series modelling for industrial prognostics and health management, *Int. J. Adv. Manuf. Technol.*, 69(5-8), 1685-1699.

Nikravesh M., Farell A., and Stanford T. (2000). Control of nonisothermal CSTR with time varying parameters via dynamic neural network control (DNNC), *Chem. Eng. J.*, 76(1), 1-16.

Overschee P. V., and Moor B. D. (1994). N4SID: Subspace algorithms for the identification of combined deterministic–stochastic systems, *Automatica*, 30(94), 75-93.

Peng Y., Dong M., and Zhuo M.J. (2010). Current status of machine prognostics in condition–based maintenance: a review, *Int. J. of Adv. Manuf. Technol.*, 50(1-4), 297-313.

Qin S. J. (2012). Survey on data–driven industrial process monitoring and diagnosis, *Annu. Rev. Control*, 36(2): 220-234.

Soualhi A., Razik H., Clerc G., et al. (2014). Prognosis of bearing failures using hidden markov models and the adaptive neuro–fuzzy inference system, *IEEE Trans. Ind. Electron.*, 61(6), 2864-2874.

Wang Y., and Wang W. (2012). Failure prognostic of systems with hidden degradation process, *J. Syst. Eng. Electron*, 23(2), 314-324.

Yin S., Wang G., and Yang X. (2014). Robust PLS approach for KPI–related prediction and diagnosis against outliers and missing data, *Int. J. Syst. Sci.*, 45(7), 1375-1382.

Yu M., Wang D., and Luo M. (2014). An integrated approach to prognosis of hybrid systems with unknown mode changes, *IEEE Trans. Ind. Electron.*, 62(1), 503-515

Zhang Z., and Dong F. (2014). Fault detection and diagnosis for missing data systems with a three time–slice dynamic Bayesian network approach, *Chem. Intell. Lab. Syst.*, 138(11), 30-40.

Table 1: The results of 100 Monte Carlo iterance

Prediction Step value		Feed flow fault			Feed concentration fault			Fouling fault		
	Missing rate (%)	0	5	10	0	5	10	0	5	10
10	Prediction rate (%)	97	94	93	99	96	95	96	93	92
	Accuracy rate (%)	94	94	93	92	92	91	83	82	81
	Advance time (min)	16.30	15.90	15.70	10.10	9.40	9.10	80.50	78.20	76.90
30	Prediction rate (%)	99	97	96	99	97	96	99	97	97
	Accuracy rate (%)	81	81	82	87	85	84	76	74	74
	Advance time (min)	17.50	16.90	16.50	10.40	9.80	9.40	84.50	81.90	80.20