



Online monitoring and control of a cyber-physical manufacturing process under uncertainty

Saideep Nannapaneni¹ · Sankaran Mahadevan² · Abhishek Dubey³ · Yung-Tsun Tina Lee⁴

Received: 17 February 2020 / Accepted: 14 June 2020 / Published online: 24 June 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Recent technological advancements in computing, sensing and communication have led to the development of cyber-physical manufacturing processes, where a computing subsystem monitors the manufacturing process performance in real-time by analyzing sensor data and implements the necessary control to improve the product quality. This paper develops a predictive control framework where control actions are implemented after predicting the state of the manufacturing process or product quality at a future time using process models. In a cyber-physical manufacturing process, the product quality predictions may be affected by uncertainty sources from the computing subsystem (resource and communication uncertainty), manufacturing process (input uncertainty, process variability and modeling errors), and sensors (measurement uncertainty). In addition, due to the continuous interactions between the computing subsystem and the manufacturing process, these uncertainty sources may aggregate and compound over time. In some cases, some process parameters needed for model predictions may not be precisely known and may need to be derived from real time sensor data. This paper develops a dynamic Bayesian network approach, which enables the aggregation of multiple uncertainty sources, parameter estimation and robust prediction for online control. As the number of process parameters increase, their estimation using sensor data in real-time can be computationally expensive. To facilitate real-time analysis, variance-based global sensitivity analysis is used for dimension reduction. The proposed methodology of online monitoring and control under uncertainty, and dimension reduction, are illustrated for a cyber-physical turning process.

Keywords Cyber-manufacturing · Cyber-physical · Monitoring · Control · Uncertainty · Bayesian network

Disclaimer: Certain commercial systems are identified in this paper. Such identification does not imply recommendation or endorsement by NIST; nor does it imply that the products identified are necessarily the best available for the purpose. Further, any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NIST or any other supporting U.S. government or corporate organizations.

✉ Saideep Nannapaneni
saideep.nannapaneni@wichita.edu

¹ Department of Industrial, Systems, and Manufacturing Engineering, Wichita State University, 1845 Fairmount St, Box 0035, Wichita, KS 67260, USA

² Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN 37235, USA

³ Department of Electrical Engineering and Computer Science, Vanderbilt University, Nashville, TN 37235, USA

⁴ Systems Integration Division, Engineering Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA

Introduction

Cyber-physical manufacturing systems (CPMS) refer to the integration of manufacturing processes and computing subsystems to perform several types of data analytics such as process monitoring and control to achieve resilient performance (Lee et al. 2016). A CPMS is a dynamic system, where the computing subsystem continuously monitors the manufacturing process and provides the appropriate actuation to reduce the part quality losses and increase its performance. Following our previous work (Nannapaneni et al. 2017a), we consider a CPMS as being a composition of four subsystems—manufacturing process, sensors, computing, and actuation (control); these subsystems continuously influence each other in a coupled manner as shown in Fig. 1. The performance of each subsystem is affected by different types of uncertainty sources, which affect the overall CPMS performance. The uncertainty sources affecting a manufacturing process include the inherent process variability and the

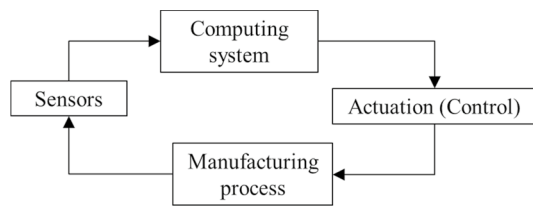


Fig. 1 Interactions between several subsystems in a cyber-physical manufacturing system

mathematical models used to analyze the process. Sensors are often associated with noise and performance uncertainty. Similar to the manufacturing process, actuation systems, which are typically mechanical systems such as a hydraulic or pneumatic pump, are also associated with inherent variability and uncertainty in the mathematical models used to describe them.

The computing nodes perform analytics on the sensor data and provide necessary actuation. To perform analytics, a computing node should have required hardware resources such as battery power and memory (data storage). When networks of computing nodes are utilized, it is possible that the network bandwidth can become clogged due to high data traffic resulting in unsuccessful data transmission. Also, communication uncertainty may also exist between several computing nodes if the computing subsystem has multiple nodes, and between the computing nodes and the actuation system and the sensors. The additional uncertainty sources associated with the computing nodes in a CPMS are the availability of hardware resources, sensor-to-node communication, node-to-node communication, and node-to-actuation communication. Quantification and incorporation of various uncertainty sources can enable a reliable design and an effective operation of CPMS.

Precision machining is a manufacturing strategy that is implemented to produce parts with high precision and low tolerances, and ultra-precision machining, an advancement to precision machining followed when even greater precision and lower tolerances are desired when compared to those in precision machining (Hatefi and Abou-El-Hossein 2020; Lee et al. 2006). In-process sensor-based monitoring and control strategies are often implemented in precision and ultra-precision machining to ensure the produces parts are within the desired requirements. Deep drilling is one machining operation whose performance can be improved by implementing monitoring and control strategies. Deep drilling is a drilling process where the depth of the hole is at least five times the drill diameter (Khan et al. 2017). Drills are subjected to severe machining conditions in terms of high thrust force, poor heat evacuation and chip jamming. To improve the process performance, Kavaratzis and Maiden monitored the drilling thrust and torque, and controlled the machining

parameters such as feed rate, spindle speed and tool position to ensure safety of the tool and work piece under high penetration rates (Kavaratzis and Maiden 1990). Kim et al. used peck drilling and thrust force monitoring during deep-micro-hole drilling of steel to improve the tool life by changing the one-step feed-length (OSPL) (Kim et al. 2009). In reality, there exists several uncertainty sources that impact the sensor data collection, process models used to calculate the control actions, and uncertainty in the implementation of control actions. A monitoring and control system should consider all these uncertainty sources and their interactions for reliable precision and ultra-precision machining.

We review below some uncertainty quantification (UQ) methods that were used in the manufacturing domain. Mehta et al. used Bayesian inference for calibration of the machining force model and estimation of cutting force with limited experimental tests (Mehta et al. 2017). Adnan et al. used fuzzy logic to predict the surface roughness and estimation of cutting force in machining processes (Mohd Adnan et al. 2015). Bhinge et al. used a Gaussian process model to quantify the uncertainty in the energy prediction of milling process (Bhinge et al. 2017). Reza et al. used a fuzzy set approach to characterize the uncertainty in energy synthesis and demonstrated the approach for a paved road system (Reza et al. 2013). Pehlken et al. estimated energy efficiency in the processing of raw materials under various uncertainty sources such as weather and soil conditions using Monte Carlo simulations (Pehlken et al. 2015). Bayesian network approaches were developed for parameter estimation and uncertainty quantification in energy prediction of manufacturing processes (Nannapaneni et al. 2016; Nannapaneni and Mahadevan 2016). Karandikar et al. used a Markov Chain Monte Carlo (MCMC) approach to estimate parameters of a turning model under uncertainty for tool life prediction (Karandikar et al. 2014). Dynamic Bayesian network approaches were used for diagnosis, prognostics, and optimization in maintenance strategies in (Tobon-Mejia et al. 2012; Weber and Jouffe 2006). UQ methods for prediction, parameter estimation, diagnosis and prognosis have been developed for traditional manufacturing process; however, such methods for CPMS are unavailable.

Process monitoring allows us to obtain the process performance in real time and enables us to change any process parameters to improve the part quality. Wu et al. used a fog computing framework for process monitoring and prognosis, and demonstrated the methods for monitoring vibrations of pumps in a power plant and energy consumption of CNC machines (Wu et al. 2017). Rao et al. used a combination of recurrent predictor neural network along with Bayesian parameter estimation using a particle filter for real-time identification of surface morphology variations in ultra-precision machining process (Rao et al. 2014). Arul et al. developed an online process monitoring

mechanism based on acoustic emissions for quality control in the drilling of polymeric composites (Arul et al. 2007). Wang and Yan developed a real-time monitoring framework in chemical processes by analyzing the process data for abnormalities using a principal component analysis (PCA) model (Wang and Yan 2019). Rao et al. developed an online process control framework for additive manufacturing processes through statistical analysis and nonparametric Bayesian modeling approaches (Rao et al. 2015). Gonzaga et al. used an artificial neural network-based soft sensor for online estimates of polyethylene terephthalate (PET) viscosity to control of industrial polymerization process (Gonzaga et al. 2009). Mosallam et al. developed a Bayesian data-driven approach for prognostics and remaining useful life (RUL) prediction and demonstrated the approach for battery and turbofan degradation (Mosallam et al. 2016). Given the growing interest in CPMS, methodologies for UQ and its incorporation in process monitoring and control are increasingly becoming necessary and this paper seeks to address this issue.

Fuzzy modeling approach is one of the commonly-used mathematical approach implemented in advanced machining to predict the process performance metrics given the uncertainty in input parameters. Syn et al. employed fuzzy expert system for the prediction of surface quality and the dross inclusion in a laser cutting process (Syn et al. 2011). Park et al. employed a fuzzy pattern recognition-based system for monitoring weld quality in a laser welding process (Park et al. 2001). Maji et al. used adaptive network-based fuzzy inference system to model input–output relationships if an electrical discharge machining process (Maji and Pratihari 2010). Vundavilli et al. used a fuzzy logic-based expert system for prediction of depth of cut in an abrasive water jet machining process (Vundavilli et al. 2012). Kovac et al. used a combination of fuzzy logic and regression analysis for modeling surface roughness in a face milling process (Kovac et al. 2013). In this paper, we employed the Bayesian approach as it facilitates both performance prediction considering various uncertainty sources and inference for updating process parameters using real-time sensor data.

The interactions between subsystems in a CPMS (Fig. 1) occur in a time sequence with small (but finite) time lags between them. This paper analyzes the coupled interactions between individual subsystems in a sequential manner using a two-level dynamic Bayesian network (DBN) approach. When process parameters are unknown, real-time sensor data can be used to estimate them using Bayesian calibration, and are later used for process control. As the number of uncertain parameters increases, their estimation in real time can become computationally expensive. To reduce the computational complexity, we employ variance-based sensitivity analysis to identify critical parameters and reduce the number of uncertain parameters.

Technological advancements in cloud computing and cloud services have led to a new manufacturing paradigm called cloud manufacturing, which is a new service-oriented manufacturing paradigm that facilitates on-demand access for customers, ranging from individual users to large OEMs, to diversified and distributed manufacturing resources to enhance production efficiency, reduce product life-cycle costs, and allow for optimal resource loading in response to variable customer demands (Wu et al. 2013; Xu 2012; Zhang et al. 2019). Cloud manufacturing may or may not necessarily provide direct interactions to the machine tools and physical devices. Cyber-physical manufacturing systems enable direct interactions between computing, manufacturing, actuation, and sensor systems, and enable online control of the manufacturing processes using real-time sensor data. Recently, Liu et al. proposed a paradigm combining the cloud manufacturing and cyber-physical systems called Cyber-Physical Manufacturing Cloud (CPMC) that combines the principles of cloud manufacturing and cyber-physical systems (Liu et al. 2017). The CPMC paradigm is a service-oriented manufacturing paradigm where various manufacturing processes can be monitored and controlled from the cloud. In this paper, we focus on developing online monitoring and control algorithms considering various uncertainty sources that arise from the manufacturing, sensor, computing, and actuation systems. The proposed methodology is general, and can be applied with different computing environments (edge, mainframe or cloud).

The overall contributions made through this paper are: (1) Quantification of multiple uncertainty sources (including the computing uncertainty) through a Bayesian probabilistic framework; (2) Development of a multi-level DBN for uncertainty propagation; (3) Real-time control of CPMS using the DBN; (4) Dimension reduction to enable real-time analysis; and (5) Illustration of the proposed quality control framework for a cyber-physical turning process.

The rest of the paper is organized as follows. “[Background](#)” section provides a brief background to dynamic Bayesian networks and sensitivity analysis, which are later used in the proposed methodology described in “[Online monitoring and control under uncertainty](#)” section. “[Illustrative example: turning process](#)” section illustrates the proposed control framework and dimension reduction for a cyber-physical turning process, followed by concluding remarks in “[Conclusion](#)” section.

Background

Dynamic Bayesian networks

A dynamic Bayesian network is a probabilistic framework used to model time-dependent systems (Murphy 2002). In

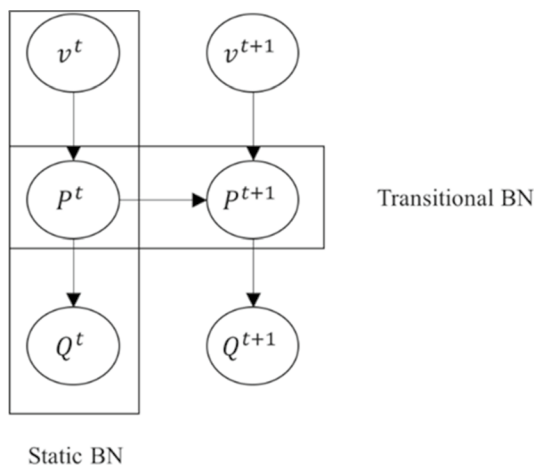


Fig. 2 DBN between two time steps

this framework, the continuous time is discretized into discrete time steps, and dependence between variables is modeled within a single time step and across time steps. The DBNs are typically considered with a Markov assumption, i.e., the variables in any time step are dependent only on the variables within the current time step and the previous time step Fig. 2 shows an illustrative DBN.

A DBN model follows a state-space modeling framework, where the behavior of the system at any time is represented using a set of variables called state variables. When the state variables are unobservable, they are estimated indirectly by observing another set of variables called the observation variables. The dependence between various variables can be given as

$$\mathbf{P}^{t+1} = G(\mathbf{P}^t, \mathbf{v}^{t+1}) + \epsilon_P \quad (1)$$

$$\mathbf{Q}^t = H(\mathbf{P}^t) + \epsilon_Q \quad (2)$$

where \mathbf{P}^t and \mathbf{P}^{t+1} represent the state variables in two time steps. \mathbf{Q}^t represents the observation variable at the current time step t . The estimation of \mathbf{P}^{t+1} from \mathbf{P}^t is through Eq. (1). \mathbf{v}^{t+1} refer to system inputs at time $t + 1$. Equation (2) represents the relationship connecting observation variables \mathbf{Q}^t to the state variables \mathbf{P}^t . G and H represent the models, either physics-based or data-driven, connecting the state variables at consecutive time steps, and connecting the state and observation variables at any given time step respectively. ϵ_P and ϵ_Q represent the noise (error) terms associated with the prediction of \mathbf{P}^{t+1} and \mathbf{Q}^t .

A DBN is typically constructed in two steps; the static BN is constructed in the first step and the transitional BN is learnt in the second step (Murphy 2002). The static BN is constructed using physics-based models, domain expert knowledge, data, or hybrid approaches, i.e., using

a combination of physics, expert knowledge, and data. In a hybrid approach, segments of the static BN, i.e., a partial BN can be obtained using physics models and domain knowledge, while the remaining dependencies are learnt using data (Nannapaneni et al. 2017b). Similar to the static BN, the learning of the transitional BN can be carried out using available physics models, expert knowledge, data, or either combination. In data-driven analysis, learning becomes a variable selection analysis, i.e., the variables at time step t that affect the variables at time $t + 1$ are identified using several variable and feature selection techniques (Saeys et al. 2007).

Several exact and approximate inference techniques are available to estimate the state variables in real-time depending the complexity of the relationships between state and observation variables. In this paper, we use particle filtering methods for inference as they can be used in the presence of complex non-linear relationships between variables. Some commonly used particle filtering algorithms include Sequential Importance Sampling (SIS), Sequential Importance Resampling (SIR) and Rao-Blackwellized Particle Filter (Arulampalam et al. 2002). In this paper, we use the SIR algorithm (Wang et al. 2019); the steps of the algorithm are given below.

1. Generate N samples of the state variables at the current time step, $P_k^t, k = 1, 2 \dots N$.
2. Compute the likelihood of each of the N particles by propagating them through the static BN and by using the observation data.
3. Compute weights for each particle as being proportional to their likelihood measures.
4. Resample the generated N values of the state variables according to their weights and obtain N values, which are used to obtain their posterior distributions.
5. These posterior samples are then used to obtain the prior distributions of the state variables in the next time step by propagating them through the transitional BN.

After a background to DBN, we now review variance-based sensitivity analysis, which can be used to perform dimension reduction in the presence of a high-dimensional state space in a DBN to enable real-time analysis for process control.

Variance-based sensitivity analysis

Consider a model G with n input variables X_1, X_2, \dots, X_n given by

$$Y = G(X_1, X_2, X_3, \dots, X_n) \quad (3)$$

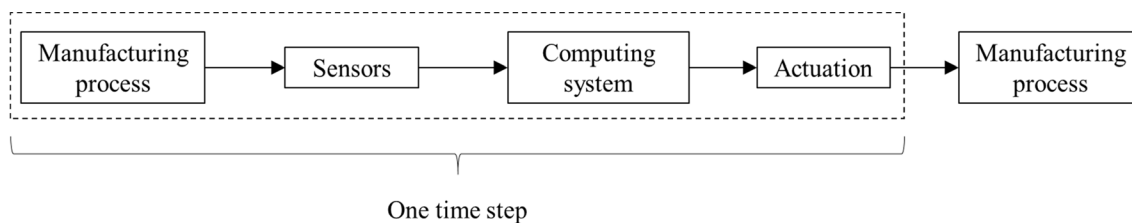


Fig. 3 One time step in the dynamic Bayesian network analysis of a CPMS

Two types of indices are computed for each variable in variance-based sensitivity analysis—main effect (or first-order effect), and total effect. The first-order effect index (S_i^I) quantifies the individual contribution from a variable X_i , without considering its interaction with other variables, to the uncertainty in the output whereas the total effect index quantifies the contribution from X_i including its interactions with all the other variables $X_{\sim i}$. The first-order effect is given as

$$S_i^I = \frac{Var_{X_i}(E_{X_{\sim i}}(Y|X_i))}{Var(Y)} \tag{4}$$

where $E_{X_{\sim i}}(Y|X_i)$ represents the expected value of output Y when X_i is fixed at a specific value, and Var_{X_i} computes the variance of this expected value when the uncertainty in X_i is included. The total effects index is given as

$$S_i^T = 1 - \frac{E_{X_{\sim i}}(Var_{X_i}(Y|X_{\sim i}))}{Var(Y)} \tag{5}$$

where $Var_{X_i}(Y|X_{\sim i})$ represents the variance of Y when all variables other than X_i are fixed at specific values, and $E_{X_{\sim i}}$ calculates the expected value of this variance considering the randomness in $X_{\sim i}$. Some techniques to compute these sensitivity indices are the Sobol’s scheme (Sobol’ 2001), Fourier amplitude sensitivity test (FAST) (Saltelli et al. 1999), improved FAST (Tarantola et al. 2006), importance sampling and kernel regression (Sparkman et al. 2016), and the stratified sample-based approach for sensitivity analysis (Li and Mahadevan 2016). Here, we use the stratified sample-based approach for its computational efficiency. Next, we use the concepts of DBN and sensitivity analysis for UQ, dimension reduction, and process control in a CPMS.

Online monitoring and control under uncertainty

In this section, we detail the construction of a two-level DBN for modeling a CPMS and then use it for process monitoring and control under uncertainty.

Multi-level DBN

As discussed in “Introduction” section, the coupled interactions between various subsystems in a CPMS occur with a time lag. We define a time step (denoted as t) as the time taken for one analysis of various CPMS subsystems (manufacturing process, sensing, computing, and actuation) as shown in Fig. 3. We discuss below the construction of the multi-level DBN in a CPMS.

Since manufacturing processes and actuation systems are physical systems, their associated BNs can be constructed from available physics models, expert knowledge, data or their combination. There has been extensive literature on the construction of BNs for physical systems (Nannapaneni and Mahadevan 2016; Scutari 2010). However, the BN corresponding to a computing subsystem is not straightforward, as it does not have an associated physics-based model. Moreover, the dependence between various computing nodes and associated uncertainty in the computing subsystem depend on the number of computing nodes and the type of interactions between them. Therefore, this paper particularly deals with the uncertainty sources related to the computing subsystem, and their aggregation with the uncertainty sources from the manufacturing process, actuation subsystem, and the sensors.

The interactions (communication) can be between computing nodes, sensors and computing nodes, and computing nodes and the actuation system. There are two types of interactions—(1) one-way or asynchronous interaction, and (2) request–reply or request–response or synchronous interaction (Dubey et al. 2011). A brief introduction to these interactions is given below. Assume a one-way interaction from C_1 to C_2 , where C_1 and C_2 can represent computing nodes, sensors or the actuation subsystem. The data transmission does not occur in a single instance but occurs over a time interval during which the data is transmitted in several data packets in a sequential manner. Since data transmission is a dynamic process, we model it using another DBN. Let us define a time step n in the lower level DBN as the time required for the transmission of one data packet. Let E_{12}^n represent the event of transmitting one data packet. E_{12}^n is a binary outcome event where 0/1 represent successful/unsuccessful transmission.



Fig. 4 DBN for a 2-node asynchronous communication

Table 1 Conditional probabilities of data transfer between two lower-level time steps

	Successful at time step $n + 1$	Unsuccessful at time step $n + 1$
Successful at time step n	R_{00}	R_{01}
Unsuccessful at time step n	R_{10}	R_{11}

A Markov assumption is made for the outcome of the transmission event at time $n + 1$, i.e., the outcome of E_{12}^{n+1} is dependent on the outcome of E_{12}^n . A practical rationale for this assumption is that if a data packet transmission is successful at time step n , then it is likely to be successful at time $n + 1$. On the contrary, if the transmission is unsuccessful at time step n , which could be due to high data traffic, then a successful transmission is unlikely at time $n + 1$. The DBN for the asynchronous interaction is shown in Fig. 4.

Let p data packets be transmitted from nodes C_1 to C_2 , and if r packets get successfully transmitted, then all the data can be reconstructed at C_2 . Since we assumed that one data packet is transmitted per time step n in the lower-level DBN, the value of n goes from $n = 1$ to $n = p$, i.e., there are p time steps. The joint probability of all the events corresponding to the transmission of p data packets is equal to $P(E_{12}^1, E_{12}^2 \dots E_{12}^p)$. This joint probability can be decomposed into a product of marginal and conditional probabilities as

$$P(E_{12}^1, E_{12}^2 \dots E_{12}^p) = P(E_{12}^1) \times P(E_{12}^2 | E_{12}^1) \times \dots \times P(E_{12}^p | E_{12}^1, E_{12}^2 \dots E_{12}^{p-1}) \tag{6}$$

Using the Markov assumption, Eq. (6) can be simplified as

$$P(E_{12}^1, E_{12}^2 \dots E_{12}^p) = P(E_{12}^1) \times P(E_{12}^2 | E_{12}^1) \times \dots \times P(E_{12}^p | E_{12}^{p-1}) \tag{7}$$

Let R^1 represent the probability of a successful data packet transfer at $n = 1$. The conditional dependence for data transmission is given in Table 1, where R_{ij} is the probability of data transmission event in the current time step j conditioned on the data transmission event in the previous time step i ($i, j = 0, 1$). These probabilities can be estimated through an aggregation of historical data, and simulations regarding the communication network.

A request-reply interaction is associated with a sequence of *request* and *reply* messages, i.e., C_2 requests for information and C_1 replies, as shown in Fig. 5. Here, we define a

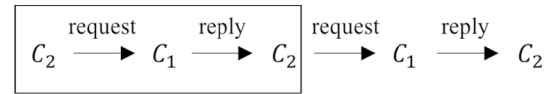


Fig. 5 Request and reply messages in a 2-node synchronous communication

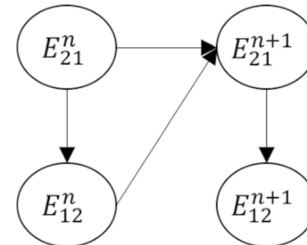


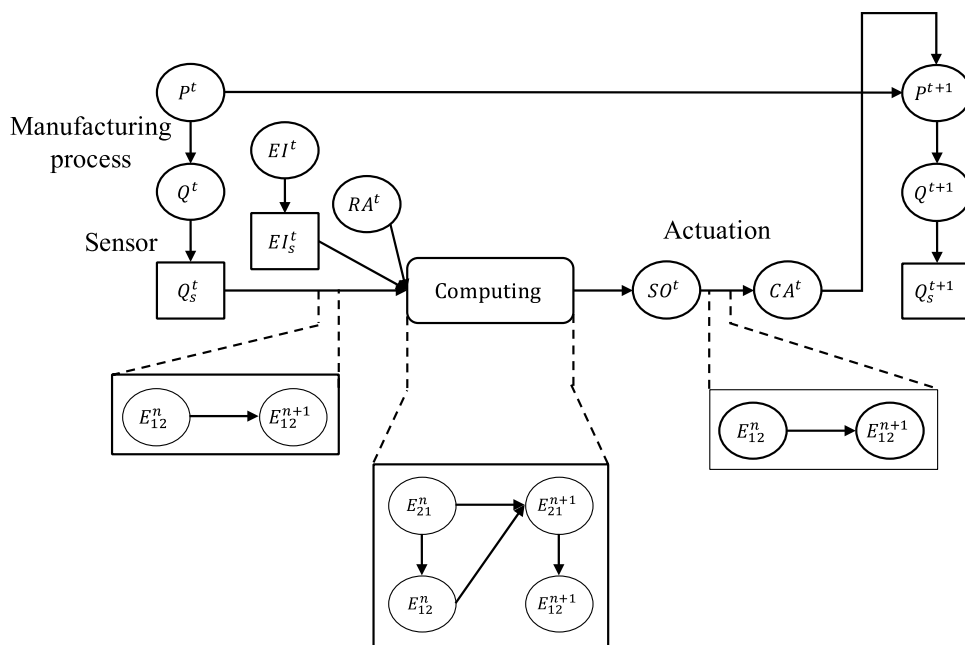
Fig. 6 DBN for a 2-node synchronous communication

time step as the time required for one *request* and one *reply* message transmission. Let E_{12}^n and E_{21}^n represent the *reply* and *request* events accordingly at a lower level time step n . As the request and reply messages occur in a sequential manner, the success of a message (*reply/request*) is dependent on the previous message (*reply/request*). The lower-level DBN for the 2-node synchronous interaction is given in Fig. 6, where if E_{21}^n is successful, E_{21}^{n+1} is dependent on E_{12}^n and if E_{21}^n is not successful (failed request message and this implies no reply message), then E_{21}^{n+1} is assumed to be dependent on E_{21}^n . As opposed to the one-way interaction system, we assume we require r successful request-reply pairs in the request-reply interaction system since a *reply* does not occur unless there is a *request* and *reply* does not always occur for every *request*. The joint probability of p request-reply interactions, assuming that one occurs at each lower-level time step can be computed using Eqs. (6) and (7).

Let R_2 represent the probability of a successful *request* message at lower-level time step $n = 1$. Let R_{12} represent the probability of successful reply message when the request message is successful. Therefore, $R_2 \times R_{12}$ refers to the reliability of a request-reply pair at $n = 1$. For illustration, we assume the same conditional relationships between two requests across two successive time steps as provided in Table 1. Given the dependence relationships across time steps, the probability of r successful pairs out of p can be computed. The same procedure can be extended to the interactions of the computing nodes with the sensors and actuation system. For illustration, this paper considered a Markov model to estimate the success/failure of a data packet transmission. In future, we shall consider other sophisticated performance models of network communication based on queuing theory and Poisson distribution (Ray et al. 2005).

It should be noted that we have two types of discretization of the continuous time in the two-level DBN. The

Fig. 7 A two-level DBN of a conceptual cyber-physical manufacturing system



higher-level discretization, which corresponds to the time step at which sensor data are available is denoted using t and the lower-level discretization, which discretizes the time between two higher-level time steps t and $t + 1$ is denoted using n . The lower-level time step corresponds to the time step at which data packets are transmitted between computing nodes, and with the sensors and the actuation subsystem. The reliability of communication between the computing nodes, and between the sensors and actuation subsystem are not impacted by the higher-level time step t , and only dependent on the lower-level time step n as detailed in Eqs. (6–9).

To perform the required analysis, the computing nodes should have the necessary hardware resources such as power and computing memory. In some cases, power may be available to computing nodes through battery power; this is typically observed in mobile computing nodes, which are computing nodes that can be transported while they are in operation (Hoang et al. 2012). If there are N computing nodes and $E_{i,t}, i = 1, 2, 3 \dots N$ represent the events corresponding to their resource availability at any time step t . The joint probability of resource availability is given as

$$P(E_{1,t}, E_{2,t} \dots E_{N,t}) = P(E_{1,t}) \times P(E_{2,t}|E_{1,t}) \times \dots \times P(E_{N,t}|E_{1,t}, E_{2,t} \dots E_{N-1,t}) \quad (8)$$

Under the assumption that each node has its own hardware resources, the resource availability of one node is independent to that of another node. Thus, Eq. (8) can be simplified as

$$P(E_{1,t}, E_{2,t} \dots E_{N,t}) = P(E_{1,t}) \times P(E_{2,t}) \times \dots \times P(E_{N,t}) \quad (9)$$

Let $S_{r,i,t}$ refer to the probability of resource availability of node i at time step t . The probability that all the events, $E_{i,t}, i = 1, 2, 3 \dots N$, are successful is equal to $\prod_i^N S_{r,i,t}$. The resulting overall two-level DBN for a CPMS is shown in Fig. 7, and the description of variables is given in Table 2. The two-level DBN begins with the variables P^t and Q^t associated with the manufacturing process. We then have the sensor subsystem that collects real-time data on Q^t , denoted as Q_s^t , which is transmitted to the computing subsystem. The sensor data is used to estimate the posterior distributions of the state variables, and calculate the control action in the next time step. In some cases, the process performance depends on the environmental inputs (EI^t) such as temperature or humidity. The data on the environmental inputs (EI_s^t) are also transmitted to the computing subsystem. As discussed earlier in this section, the performance of the computing subsystem depends in the availability of computing resources; this is denoted as RA^t in Fig. 7. The

Table 2 Variables in the two-level DBN model

Parameter	Description
P^t, P^{t+1}	State variable at time $t = t, t + 1$
Q^t, Q^{t+1}	Observation variables at time $t = t, t + 1$
Q_s^t, Q_s^{t+1}	Sensor measurements of observation variables at time $t = t, t = t + 1$
EI_s^{t+1}	Sensor measurements of environmental inputs at time $t = t + 1$
RA^{t+1}	Resource availability at time $t = t + 1$
SO^{t+1}	Computing system output at time $t = t + 1$
CA^{t+1}	Control action at time $t = t + 1$

output of the computing subsystem is denoted as SO^t ; this is transmitted to the actuation subsystem, which implements the control action. It should be noted that the implemented control action may not be the same as the software output due to the communication uncertainty between the computing and actuation subsystems.

The type of interactions within the computing subsystem and with the sensors and actuation subsystem depend on the computing architecture available in the CPMS. For illustration, we represented the communication between the computing subsystem and the sensors, and with the actuation subsystem using an asynchronous interaction, and the communication within the computing subsystem using a synchronous interaction.

The steps for the construction of the multi-level DBN are summarized below.

- (1) Obtain the conditional probability relationships between the manufacturing process variables (P^t and Q^t) using physics-based or data-driven models.
- (2) Model the conditional probability relationships between the observation variables (Q^t) and the sensor measurements (Q_s^t), and between the environmental inputs (ET^t) and their sensor measurements (ET_s^t) using sensor uncertainty.
- (3) Identify the required resources for the computing nodes and estimate the probability of their availability of each computing node at each time step t .
- (4) Identify the type of interaction (asynchronous or synchronous) between the sensors and computing nodes, between the computing nodes, and between the computing nodes and the actuation subsystem; this depends on the computing architecture in the CPMS.
- (5) Identify the number of lower-level time steps in the communication between computing nodes, and with sensors and actuation subsystem.
- (6) Depending on the types of interaction, construct the lower-level DBN as shown in Figs. 4 and 6.
- (7) Obtain the conditional probability relationships of the state variables at time $t + 1$ dependent on the state variables at time t and the implemented control action at time t .

After discussing the construction of the multi-level DBN, we now discuss its application for online process monitoring and control in a CPMS.

Process monitoring and control in a CPMS

The monitoring and control follows a ‘measure-update-optimize’ procedure. The sensor data regarding several variables such as Q^t in Fig. 7 are obtained and communicated to

the computing subsystem. The computing subsystem then updates the uncertain model variables P^t through Bayesian calibration. The updated distributions are used to estimate the control action that minimizes the quality losses and applied through the actuation system in the next time step. All the above analyses are performed by the computing subsystem, and the result is communicated to the actuation subsystem.

The computational effort of Bayesian calibration is dependent on the number of uncertain parameters. As the number of parameters increases, the number of particles required in the SIR algorithm (used for Bayesian calibration and discussed in “Dynamic Bayesian networks” section) increases. Since process monitoring and control need to be performed in real time, high computing time is not affordable. In such cases, we use the variance-based sensitivity analysis for dimension reduction, i.e., obtain a subset of uncertain parameters that have a major influence on the observation variable. At the start of the process monitoring, we have the prior distributions of the all uncertain parameters. These distributions are later updated to obtain the posterior distributions using the sensor data. For dimension reduction, we perform sensitivity analysis using the prior distributions, and if the sensitivity of a parameter is less than a pre-defined threshold, then that parameter is assumed to be deterministic at its nominal value, such as the mean or mode. The remaining parameters after dimension reduction analysis are updated using the sensor data. The steps in the process monitoring and control are summarized below.

- (1) Perform variance-based sensitivity analysis for dimension reduction of state variables (P^t) using their prior distributions.
- (2) Given the computing and communication architecture (asynchronous and/or synchronous), compute the reliability (success probability) values of (i) data transmission between the sensors and computing nodes, (ii) data transmission between the computing nodes to carry out required analysis, and (iii) data transmission between the computing nodes and the actuation subsystem.
- (3) Compute the probability of all the computing nodes to have the necessary resources using Eq. (9).
- (4) Obtain a binomial random sample (using success probability from step 2(i)) to simulate the data transmission between the sensors and the computing nodes.
- (5) If the data transmission in step 4 is successful, then obtain two binomial random samples using the success probabilities from steps 2(ii) and 4, to simulate the data transmission between the computing nodes, and resource availability.
- (6) If the data transmission and resource availability in step 5 are successful, then the posterior distributions

of the state variables and the required control action for the next time step are computed and stored in the computing subsystem. The posterior distributions are computed using the SIR algorithm in “[Dynamic Bayesian networks](#)” section. The control action is calculated to minimize a loss function defined over the system quantity of interest (QoI)

- (7) Obtain a binomial random sample (using the success probability from step 2(iii)) to simulate the data transmission between the computing nodes and the actuation subsystem.
- (8) If the data transmission is successful, then the computed control action is implemented on the manufacturing process else the control action implemented in the previous time step is implemented in the current time step.
- (9) Steps 3–8 are repeated until a part is manufactured

We demonstrate below the construction of the multi-level DBN and online process monitoring and control for a cyber-physical turning process.

Illustrative example: turning process

In this section, we provide a brief introduction to a turning process, discuss a cyber-physical version of the turning process, and use it to demonstrate the proposed process monitoring and control framework.

Turning process

Turning is a machining operation, where the material of a rotating cylindrical part is removed when it moves linearly against a cutting tool, and along the axis of rotation. Let D_o and D_f represent the initial and final (target) diameters of the cylindrical part. Given D_o and D_f , the depth of the cut (d) is calculated as

$$d = \frac{D_o - D_f}{2} \tag{10}$$

The inputs to the turning operation are feed rate (f) and cutting speed (v). Feed rate is the speed at which the cutting tool is fed on to the part, and cutting speed is the relative speed measured along the axis between the part and the cutting tool. Over time, the tip of the cutting tool wears out, which is known as flank wear (w), and this affects the dimensional accuracy of the part.

The dimensional accuracy is affected by several parameters in a turning process: cutting speed, feed rate, depth of cut, coolant, coating type, chip breaker geometry, nose radius, and shape of insert (Yih-Fong 2006). Of these, cutting speed, feed

rate, and depth of cut are process parameters while the remaining are equipment parameters. A coolant is typically used in a machining process to decrease the cutting temperature, decrease the amount of power consumed in a cutting process, and increase tool life (Yildiz and Nalbant 2008). Tools are often coated to improve their hardness, protect against abrasion, and to provide lubrication between them and the parts. There are primarily two categories of coating available: physical vapor deposition (PVD) and chemical vapor deposition (CVD). Each coating type provides different benefits to the tool such as hardness, abrasion protection and lubrication (Nalbant et al. 2009). Chip breakers are often used for chip control method in machining processes (Lotfi et al. 2015). Nose radius is the radius of the tip of the cutting tool that comes in contact with a part and inserts are replaceable components that are attached to the cutting tool. For the sake of illustration, this paper considers only the process parameters (cutting speed, feed rate, and depth of cut) when estimating tool wear.

To compensate for the tool wear and achieve the target diameter, the position of the cutting tool needs to be adjusted; this is referred to as tool wear compensation (δ). As the flank wear increases with time, the tool wear compensation needs to be increased accordingly. When k finished parts are produced by a turning process and $\delta_i, i = 1, 2, \dots, k$ represent the tool wear compensation in each of those operations, the revised depth of cut after considering tool wear is given as

$$d_i = \frac{D_o - D_f}{2} + \delta_i \tag{11}$$

where d_i refers to the depth of the cut of the i^{th} part. The tool wear cannot be precisely estimated using physics models; however, empirical models are commonly adopted (Abdelmaguid and El-hossainy 2012). The tool wear is given as

$$w_i = k_w v^{\alpha_w} f^{\beta_w} d_i^{\gamma_w} (t_{w,i} + t)^{\sigma_w} \tag{12}$$

where $k_w, \alpha_w, \beta_w,$ and σ_w are the model parameters estimated using experimental data. t and w_i refer to the cutting time spent on part i and the tool wear on the i^{th} part after spending time t . $t_{w,i}$ refers to the time that needs to be spent on the i^{th} part to achieve the same tool wear that is achieved after processing $(i - 1)$ parts. If W_{i-1} refers to the tool wear after processing $(i - 1)$ parts, then $t_{w,i}$ is obtained as

$$t_{w,i} = \left(\frac{1}{k_w} v^{-\alpha_w} f^{-\beta_w} d_i^{-\gamma_w} W_{i-1} \right)^{\frac{1}{\sigma_w}} \tag{13}$$

If $\bar{D}_i = D_o - d_i$ and L represent the diameter of the cylindrical part after the turning operation, and length of the part respectively, then the cutting time for the i^{th} part is calculated as

$$t_{c,i} = \frac{\Pi(D_o - d_i)L}{vf} \quad (14)$$

Using Eqs. 12–14, the total wear after processing i parts is given as

$$W_i = k_w v^{\alpha_w} f^{\beta_w} d_i^{\gamma_w} (t_{w,i} + t_{c,i})^{\sigma_w} \quad (15)$$

The continuous wear in the cutting tool causes a drift in the final diameter of a part, which can be estimated as

$$\Delta_i = 2 \times (w_i - W_{i-1}) \tan(\theta) \quad \text{for } 0 \leq t \leq t_{c,i} \quad (16)$$

where θ is the clearance angle associated with the cutting tool. Δ_i is the drift in the i th part. The clearance angle is the angle made by the cutting tool with the axis of rotation of the cylindrical part (Baradie 1996). The drift calculated using Eq. (16) represents the additional variation in the i th part's diameter when compared to the $(i - 1)$ th part. To achieve the target diameter, a tool wear compensation (δ_i) is implemented in response to the drift (Δ_i). The final diameter after considering tool wear compensation and drift can be computed as

$$D_i = D_f - 2 \times \delta_i + \Delta_i \quad (17)$$

The quality loss due to the deviation from the target diameter ($|D_f - D_i|$) can be quantified as

$$Q_w = \int_{T_{i-1}}^{T_i} (2 \times \delta_i - \Delta_i)^2 dt \quad (18)$$

where Q_w represents the quality loss and T_i represents the cumulative machining time over i parts.

Cyber-physical turning process

In a traditional turning process, the diameter of a part is measured after it is produced in an offline manner, and any tool wear compensation is implemented to the next part. This implementation works for parts of shorter lengths, where the variation in diameters at the head and tail ends is not significant, and in longer parts where the acceptable quality tolerance is greater than the deviations due to tool wear. Such offline techniques may not be suitable for ultra-high precision parts where the acceptable tolerances are less than the deviations due to tool wear. In a cyber-physical turning process, the part's diameter is measured while it is being manufactured; these measurements are used to appropriate tool wear compensation in real time.

Several uncertainty sources are present in the above cyber-physical turning process. A source of input uncertainty is the actual clearance angle of the cutting tool, which can be different from the intended angle due to the

Table 3 Part specifications from the cyber-physical turning process

Parameter	Value (mm)
Initial diameter (D_o)	100
Final diameter (D_f)	98
Length of part (L)	100
Lower bound of target diameter ($D_{f,L}$)	97.98
Upper bound of target diameter ($D_{f,U}$)	98.02

Table 4 Probability distributions of the parameters in the tool wear empirical model

Parameter	Value
k_w	$N(8.2961 \times 10^{-5}, 8.2961 \times 10^{-7})$
α_w	$N(2.747, 0.02747)$
β_w	$N(1.473, 0.01473)$
γ_w	$N(1.261, 0.01261)$
σ_w	$N(0.43, 0.0043)$

placement of cutting tool. There exists uncertainty in the diameter measurements obtained the scanning laser beam method. The control action (input) on the turning process is the actual tool wear compensation, which can be different from the calculated tool wear compensation due to inherent process variability. The uncertainty in the process performance prediction is caused by the uncertainty in the model parameters of the tool wear empirical models. Additional uncertainty sources due to the presence of computing subsystem include the communication uncertainty and uncertainty in resource availability.

Monitoring and control in a cyber-physical turning process

Table 3 provides the desired part specifications and quality bounds. The process parameters such as feed rate and cutting speed are assumed at 60 m/min and 0.065 mm/rev respectively. The initial tool wear compensation is assumed to be 0.009 mm. The values of the model parameters in the tool wear empirical model are obtained from (Abdelmaguid and El-hossainy 2012), and provided in Table 4. The clearance angle was assumed at 15° . The uncertainty in the actual tool wear compensation, variability in the sensor measurements, and the variability in the clearance angle are modeled using Gaussian distributions with zero means and standard deviations of 0.0005 mm, 0.0025 mm, and 0.5° respectively. Gaussian distributions are used to represent uncertainties in the paper for illustration purposes only, and can be replaced with the actual probability distributions, if available.

We assume asynchronous communication interaction between the sensors and the computing node, and between

Table 5 Conditional probabilities of data transfer in cyber-physical turning process

	Successful at time step $n + 1$	Unsuccessful at time step $n + 1$
Successful at time step n	0.95	0.05
Unsuccessful at time step n	0.9	0.1

the computing node and actuation subsystem. For illustration, we assume that the sensor data is sent in three data packets and two packets are required for successful data transmission. For successful communication between the computing and actuation subsystems, we assume that one of the two data packets is required. The reliability of the first packet is assumed to be 0.95, while the reliability of the following data packets is obtained using the conditional probability table in Table 5. The reliability values are obtained from (Dulman et al. 2003). The probability that the necessary computing resources are available is assumed as 0.95.

We assume that the diameter measurements are obtained at 0.25 s intervals. The tool wear compensation of calculated using the diameter measurements and by minimizing the loss function in Eq. (19). Due to the uncertainty in the model parameters, clearance angle, tool wear compensation, and sensor measurements, the loss and the constraint functions become stochastic. The optimization formulation for the tool wear compensation is written as

$$MinE \left[\int_{T_{i-1}}^{T_i} \left(\delta_i - \frac{\Delta_i}{2} \right)^2 dt \right]$$

such that

$$Pr (D > D_{f,L} \cap D < D_{f,U}) \geq 0.95 \tag{19}$$

where the objective function refers to the minimization of the expected value of the loss function, provided in Eq. (18), which relates to the discrepancy between the predicted and target diameter values, as shown in Eq. (17).

The measured diameter values (sensor data) are used to calibrate the tool wear model parameters ($k_w, \alpha_w, \beta_w, \gamma_w,$ and σ_w), clearance angle (θ) and tool wear compensation (δ). Sensitivity analysis is used to reduce the number of parameters using the stratified sample-based algorithm (Li and Mahadevan 2016). The sensitivity indices of parameters are given in Table 6, from which α_w and β_w are considered for calibration and the other parameters ($k_w, \gamma_w, \sigma_w, \theta$ and δ) are assumed deterministic at their mean values. However, it should be realized that the true and unknown values of these parameters (denoted using a superscript ‘T’ in Table 7) may be different than the values at which they are fixed.

Table 6 Sensitivity indices of parameters in the tool wear empirical model

Parameter	Sensitivity index
k_w	0.0075
α_w	0.8585
β_w	0.1201
γ_w	0.0045
σ_w	0.0076
θ	0.0072
δ	0.0022

Table 7 Underlying true values of parameters in the tool wear empirical model

Parameter	True value
k_w^T	8.2676×10^{-5}
α_w^T	2.7195
β_w^T	1.4899
γ_w^T	1.2627
σ_w^T	0.4247
θ^T	15.0427
δ^T (initial)	0.0093

The DBN model is shown in Fig. 8, where superscript k refers to the time step. Parameters α_w and β_w are deterministic and do not vary with time. Therefore, $\alpha_w^k = \alpha_w^{k+1}$ and $\beta_w^k = \beta_w^{k+1}$. d^k, w^k, Δ^k refer to the depth of cut, tool wear and drift at the k th time step. D^k and D_S^k are the predicted output diameter and its sensor measurement respectively. R_1^k is the probability of the availability of required computing resources at k th time step. The SIR algorithm (“Dynamic Bayesian networks” section) with 1000 samples was used for parameter estimation using the sensor data. The prior and posterior distributions of α_w and β_w , along with their true values, for a sample part are given in Fig. 9.

In Fig. 10, we compare the diameter profiles in two scenarios: with and without the real-time control. In the absence of real-time control, the tool wear compensation does the change as the part is being produced; however, in real-time control, the tool wear compensation changes at each time step with the diameter measurements.

Figure 11 shows the change in the tool wear compensation with each time step. The computed and the actual values of the tool wear compensation are shown in Fig. 11. The difference between the two plots is due to the presence of uncertainty in implementing the tool wear compensation. The calculation of tool wear compensation requires real-time analysis using the computing resources. When the computing resources are unavailable, then the tool wear compensation implemented in the previous time step is continued in the current time step. This is shown in the region marked by the red circle in Fig. 11. Figure 12 shows the availability of computing resources with time.

Fig. 8 DBN model for the cyber-physical turning process

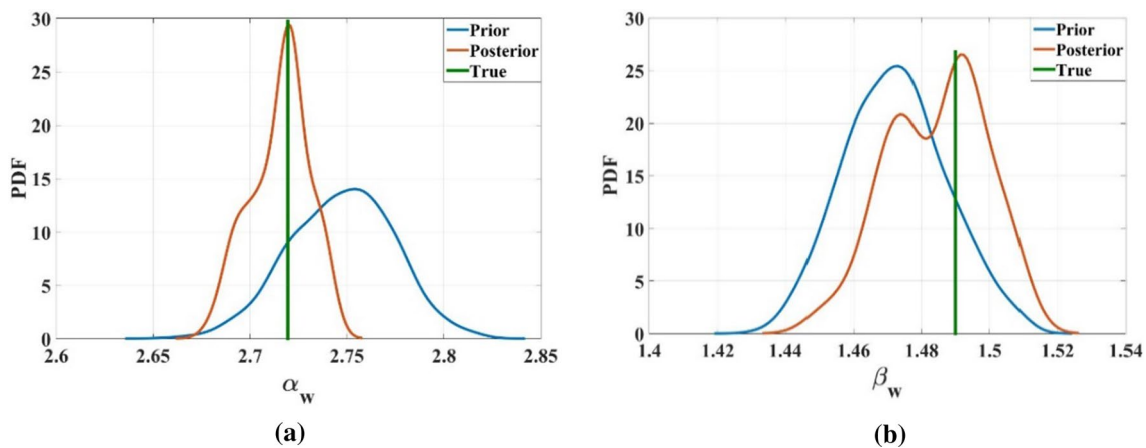
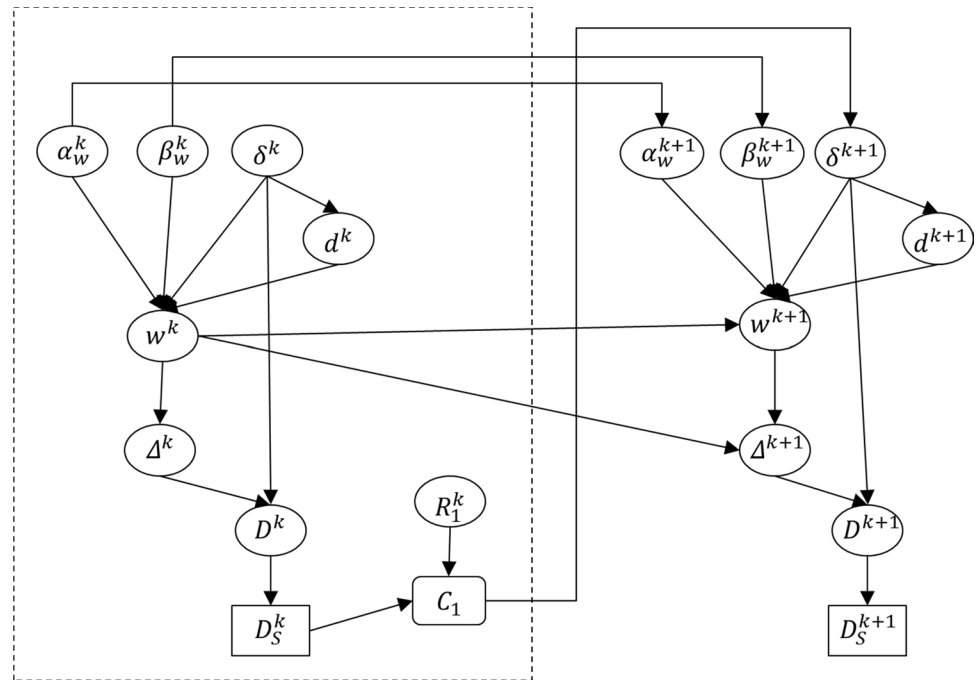


Fig. 9 Prior and posterior distributions of calibration parameters, α_w and β_w

In Fig. 12, ‘0’ and ‘1’ represent success and failure of an analysis, which can be parameter estimation or tool wear compensation estimation. Parameter estimation requires successful communication between the sensors and computing subsystem, and availability of computing resources. However, implementation of tool wear compensation requires successful communication between the computing and actuation systems in addition to successful calibration. In Fig. 12, the implementation of tool wear compensation was unsuccessful at one instant, whereas calibration was successful at all time steps. The unsuccessful implementation of tool wear compensation can be attributed to the loss of communication between the computing and actuation systems. To quantify the effect of various uncertainty sources in

the turning process, the computing subsystem, the actuation subsystem and the sensors, several parts are produced and variations in the diameter profiles can be obtained (Fig. 13).¹

In this paper, we have illustrated the proposed methods for a closed system; however, the proposed methods can be extended to open integration of manufacturing and computing systems. For an open integration and interoperability, the computing system requires a predictive model associated with the manufacturing process that is being monitored, and real-time process sensor data. With the availability of Predictive

¹ All the codes used in this example can be found at <https://github.com/Saideep259/JIM20>.

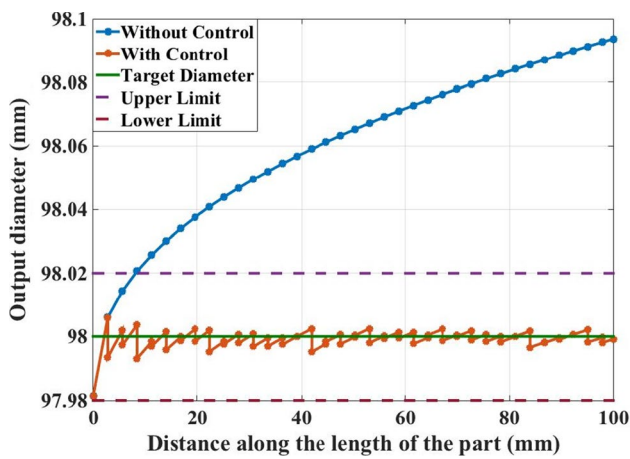


Fig. 10 Comparison of output diameter profiles with and without real-time control

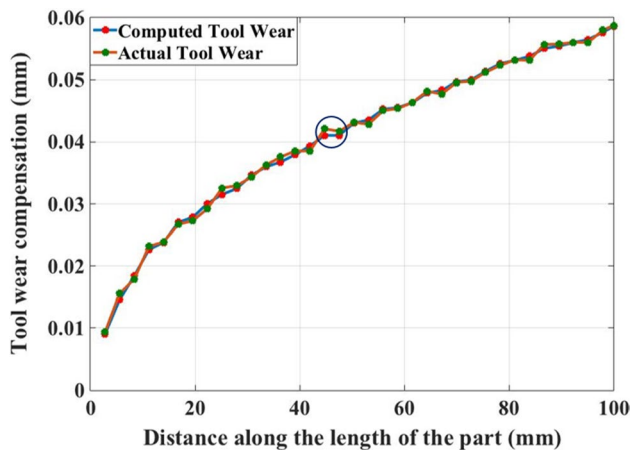


Fig. 11 Comparison of the computed and the actual tool wear compensation

Model Markup Language (PMML) technical standard (Grossman et al. 1999) and MTConnect (Vijayaraghavan et al. 2008), there exists standardized approaches for the representation of predictive models (Nannapaneni et al. 2018; Park et al. 2017) and transmission of real-time sensor data from the production floor equipment to the computing systems (Lynn et al. 2018). The sensor data will be used to update the process parameters, and the predictive model will be used to calculate the optimal control action. Therefore, the proposed methodology can be implemented with either closed or open system.

Conclusion

This paper proposed a two-level dynamic Bayesian network (DBN) framework for online monitoring and control of a cyber-physical manufacturing system (CPMS)

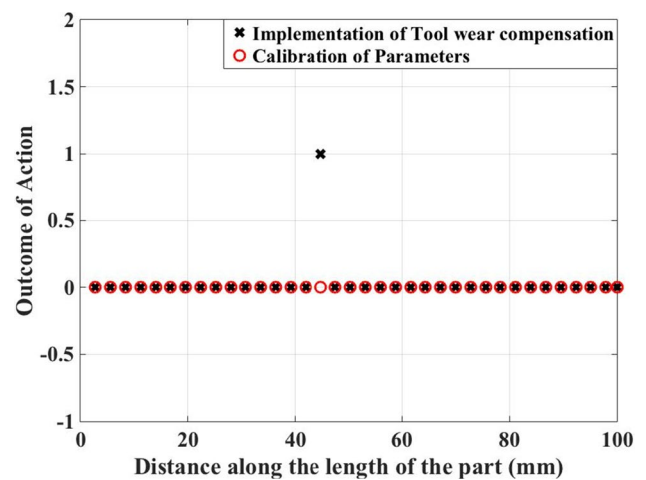


Fig. 12 Completion of tool wear compensation and calibration analyses due to availability of computational resources and successful communication between sensors, computational and actuation systems

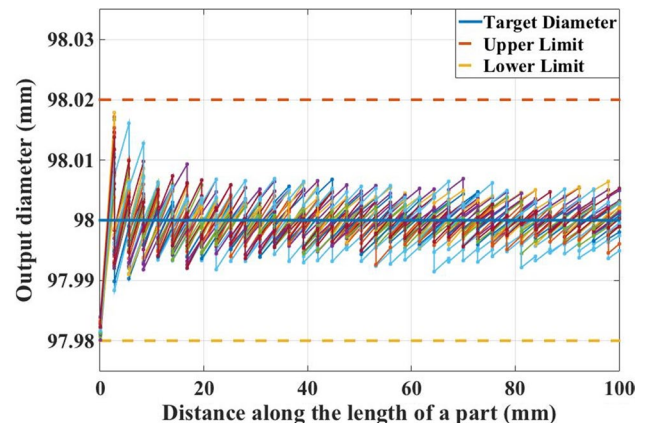


Fig. 13 Diameter profiles of parts considering aleatory and epistemic uncertainty from cyber-physical turning analysis

under uncertainty. A CPMS was assumed to be composed of four interdependent subsystems—manufacturing process, computing and actuation systems, and sensors. In the two-level DBN, the higher level captures the dependence between the individual subsystems while the lower level DBN captures the interactions between various computing nodes in a computing subsystem, and its communication with sensors and the actuation subsystem. The uncertainty sources associated with the computing subsystem include the communication between several computing nodes, and with the sensors and the actuation system, and hardware resource availability. The uncertainty associated with the sensors include the sensor measurement uncertainty. The manufacturing process is associated with uncertainty in the process models, which can be physics-based or data-driven. The actuation system implements the control input,

which can be different from the calculated control due to machine imperfections.

When the model parameters used to predict the system quantity of interest (QoI) are unknown, they are estimated in real-time using the sensor data. As the number of unknown parameters increases, the computational complexity of the calibration process increases. In such cases, variance-based sensitivity analysis is used for dimension reduction. The proposed methods are demonstrated for a cyber-physical turning process, where online monitoring and control are performed in real time while a part is being produced. The output diameter measurement is used to estimate the tool wear and an appropriate tool wear compensation is provided to reduce the deviations of the output part from the desired values.

This paper demonstrated the proposed control framework for a single manufacturing process. Future work should consider application of the proposed methods to a network of CPMS and the Industrial Internet-of-Things (IIoT) systems. Moreover, advanced computing paradigms such as fog and edge computing need to be considered for efficient computing, along with machine-to-machine interactions to minimize the product quality losses. Since human operators are involved on the production floor, their interactions with the CPMS, forming a Human-in-the-loop CPMS or H-CPMS, also need to be investigated.

Acknowledgements The research presented in this paper was supported by funds from the National Institute of Standards and Technology (NIST) under the Smart Manufacturing Data Analytics Project (Cooperative Agreement No. 70NANB16H297). The support is gratefully acknowledged.

References

- Abdelmaguid, T. F., & El-hossainy, T. M. (2012). Optimal cutting parameters for turning operations with costs of quality and tool wear compensation. In: *Proceedings of the 2012 international conference on industrial engineering and operations management*, Istanbul, Turkey, July 3–6 (pp. 924–932).
- Arul, S., Vijayaraghavan, L., & Malhotra, S. K. (2007). Online monitoring of acoustic emission for quality control in drilling of polymeric composites. *Journal of Materials Processing Technology*, 185, 184–190.
- Arulampalam, M., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50, 174–188.
- Bhingre, R., Park, J., Law, K., Dornfeld, D., Helu, M., & Rachuri, S. (2017). Toward a generalized energy prediction model for machine tools. *Journal of Manufacturing Science and Engineering*, 139, 041013.
- Dubey, A., Karsai, G., & Mahadevan, N. (2011). A component model for hard real-time systems: CCM with ARINC-653. *Software: Practice and Experience*, 41, 1517–1550.
- Dulman, S., Nieberg, T., Wu, J., & Havinga, P. (2003). Trade-off between traffic overhead and reliability in multipath routing for wireless sensor networks. In *IEEE wireless communications and networking conference, WCNC* (pp. 1918–1922). <https://doi.org/10.1109/WCNC.2003.1200680>.
- El Baradie, M. (1996). The effect of varying the workpiece diameter on the cutting tool clearance angle in tool-life testing. *Wear*, 1, 201–205.
- Gonzaga, J. C. B., Meleiro, L. A. C., Kiang, C., & Maciel Filho, R. (2009). ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process. *Computers & Chemical Engineering*, 33, 43–49.
- Grossman, R., Bailey, S., Ramu, A., Malhi, B., Hallstrom, P., Pulleyn, I., et al. (1999). The management and mining of multiple predictive models using the predictive modeling markup language. *Information and Software Technology*, 41, 589–595. [https://doi.org/10.1016/S0950-5849\(99\)00022-1](https://doi.org/10.1016/S0950-5849(99)00022-1).
- Hatefi, S., & Abou-El-Hossein, K. (2020). Review of single-point diamond turning process in terms of ultra-precision optical surface roughness. *International Journal of Advanced Manufacturing Technology*, 106, 2167–2187. <https://doi.org/10.1007/s00170-019-04700-3>.
- Hoang, D. T., Niyato, D., & Wang, P. (2012). Optimal admission control policy for mobile cloud computing hotspot with cloudlet. In *IEEE wireless communications and networking conference, WCNC* (pp. 3145–3149). <https://doi.org/10.1109/WCNC.2012.6214347>.
- Karandikar, J., Abbas, A., & Schmitz, T. (2014). Tool life prediction using Bayesian updating. Part 2: Turning tool life using a Markov Chain Monte Carlo approach. *Precision Engineering*, 38, 18–27.
- Kavaratzis, Y., & Maiden, J. D. (1990). Real time process monitoring and adaptive control during CNC deep hole drilling. *International Journal of Production Research*, 28, 2201–2218. <https://doi.org/10.1080/00207549008942862>.
- Khan, S. A., Nazir, A., Mughal, M. P., Saleem, M. Q., Hussain, A., & Ghulam, Z. (2017). Deep hole drilling of AISI 1045 via high-speed steel twist drills: Evaluation of tool wear and hole quality. *International Journal of Advanced Manufacturing Technology*, 93, 1115–1125. <https://doi.org/10.1007/s00170-017-0587-4>.
- Kim, D. W., Lee, Y. S., Park, M. S., & Chu, C. N. (2009). Tool life improvement by peck drilling and thrust force monitoring during deep-micro-hole drilling of steel. *International Journal of Machine Tools and Manufacture*, 49, 246–255. <https://doi.org/10.1016/j.ijmactools.2008.11.005>.
- Kovac, P., Rodic, D., Pucovsky, V., Savkovic, B., & Gostimirovic, M. (2013). Application of fuzzy logic and regression analysis for modeling surface roughness in face milling. *Journal of Intelligent Manufacturing*, 24, 755–762. <https://doi.org/10.1007/s10845-012-0623-z>.
- Lee, J., Bagheri, B., & Jin, C. (2016). Introduction to cyber manufacturing. *Manufacturing Letters*, 8, 11–15. <https://doi.org/10.1016/j.mfglet.2016.05.002>.
- Lee, D., Hwang, I., Valente, C., Oliveira, J., & Dornfeld, D. (2006). Precision manufacturing process monitoring with acoustic emission. In L. Wang & R. X. Gao (Eds.), *Condition monitoring and control for intelligent manufacturing* (pp. 33–54). Berlin: Springer.
- Li, C., & Mahadevan, S. (2016). An efficient modularized sample-based method to estimate the first-order Sobol index. *Reliability Engineering and System Safety*, 153, 110–121. <https://doi.org/10.1016/j.res.2016.04.012>.
- Liu, X. F., Shahriar, M. R., Al Sunny, S. M. N., Leu, M. C., & Hu, L. (2017). Cyber-physical manufacturing cloud: Architecture, virtualization, communication, and testbed. *Journal of Manufacturing Systems*, 43, 352–364. <https://doi.org/10.1016/j.jmsy.2017.04.004>.
- Lotfi, M., Akhavan Farid, A., & Soleimanimehr, H. (2015). The effect of chip breaker geometry on chip shape, bending moment, and cutting force: FE analysis and experimental study.

- International Journal of Advanced Manufacturing Technology*, 78, 917–925. <https://doi.org/10.1007/s00170-014-6676-8>.
- Lynn, R., Wescoat, E., Han, D., & Kurfess, T. (2018). Embedded fog computing for high-frequency MTConnect data analytics. *Manufacturing Letters*, 15, 135–138. <https://doi.org/10.1016/j.mfglet.2017.11.002>.
- Maji, K., & Pratihari, D. K. (2010). Forward and reverse mappings of electrical discharge machining process using adaptive network-based fuzzy inference system. *Expert Systems with Applications*, 37, 8566–8574. <https://doi.org/10.1016/j.eswa.2010.05.019>.
- Mehta, P., Kuttolamadom, M., & Mears, L. (2017). Mechanistic force model for machining process—Theory and application of Bayesian inference. *The International Journal of Advanced Manufacturing Technology*, 91, 3673–3682. <https://doi.org/10.1007/s00170-017-0064-0>.
- Mohd Adnan, M. R. H., Sarkheyli, A., Mohd Zain, A., & Haron, H. (2015). Fuzzy logic for modeling machining process: A review. *Artificial Intelligence Review*, 43, 345–379. <https://doi.org/10.1007/s10462-012-9381-8>.
- Mosallam, A., Medjaher, K., & Zerhouni, N. (2016). Data-driven prognostic method based on Bayesian approaches for direct remaining useful life prediction. *Journal of Intelligent Manufacturing*, 27, 1037–1048. <https://doi.org/10.1007/s10845-014-0933-4>.
- Murphy, K. P. (2002). *Dynamic bayesian networks: Representation, inference and learning*. Berkeley: University of California.
- Nalbant, M., Gökkaya, H., Toktaş, I., & Sur, G. (2009). The experimental investigation of the effects of uncoated, PVD- and CVD-coated cemented carbide inserts and cutting parameters on surface roughness in CNC turning and its prediction using artificial neural networks. *Robotics and Computer-Integrated Manufacturing*, 25, 211–223. <https://doi.org/10.1016/j.rcim.2007.11.004>.
- Nannapaneni, S., Dubey, A., & Mahadevan, S. (2017a). Performance evaluation of smart systems under uncertainty. In *2017 IEEE SmartWorld, ubiquitous intelligence & computing, advanced & trusted computed, scalable computing & communications, cloud & big data computing, internet of people and smart city innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI)* (pp. 1–8).
- Nannapaneni, S., & Mahadevan, S. (2016). Manufacturing process evaluation under uncertainty: A hierarchical bayesian network approach. In *Proceedings of the ASME 2016 international design engineering technical conferences and computers and information in engineering conference* (p. V01BT02A026).
- Nannapaneni, S., Mahadevan, S., Dubey, A., Lechevalier, D., Narayanan, A., & Rachuri, S. (2017b). Automated uncertainty quantification through information fusion in manufacturing processes. *Smart and Sustainable Manufacturing Systems*, 1, 153–177.
- Nannapaneni, S., Mahadevan, S., & Rachuri, S. (2016). Performance evaluation of a manufacturing process under uncertainty using Bayesian networks. *Journal of Cleaner Production*, 113, 947–959.
- Nannapaneni, S., Narayanan, A., Ak, R., Lechevalier, D., Sexton, T., Mahadevan, S., et al. (2018). Predictive model markup language (PMML) representation of Bayesian networks: An application in manufacturing. *Smart and Sustainable Manufacturing Systems*, 2, 87–113.
- Park, J., Lechevalier, D., Ak, R., Ferguson, M., Law, K., Lee, Y., et al. (2017). Gaussian process regression (GPR) representation in predictive model markup language (PMML). *Smart and Sustainable Manufacturing Systems*, 1, 121–141.
- Park, H., Rhee, S., & Kim, D. (2001). A fuzzy pattern recognition based system for monitoring laser weld quality. *Measurement Science & Technology*, 12, 1318. <https://doi.org/10.1088/0957-0233/12/8/345>.
- Pehlken, A., Decker, A., Kottowski, C., & Kirchner, A. (2015). Energy efficiency in processing of natural raw materials under consideration of uncertainties. *Journal of Cleaner Production*, 106, 351–363.
- Rao, P., Bukkapatnam, S., Beyca, O., Kong, Z., & Komanduri, R. (2014). Real-time identification of incipient surface morphology variations in ultraprecision machining process. *Journal of Manufacturing Science and Engineering*, 136, 021008.
- Rao, P., Liu, J., Roberson, D., Kong, Z., & Williams, C. (2015). Online real-time quality monitoring in additive manufacturing processes using heterogeneous sensors. *Journal of Manufacturing Science and Engineering*, 137, 61007.
- Ray, S., Starobinski, D., & Carruthers, J. (2005). Performance evaluation of wireless network in presence of hidden node: A queuing theory approach. *Computer Communications*, 28, 1179–1192.
- Reza, B., Sadiq, R., & Hewage, K. (2013). A fuzzy-based approach for characterization of uncertainties in emergent synthesis: An example of paved road system. *Journal of Cleaner Production*, 59, 99–110.
- Saeyns, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, 23, 2507–2517.
- Saltelli, A., Tarantola, S., & Chan, K. (1999). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41, 39–56.
- Scutari, M. (2010). Learning Bayesian networks with the bnlearn R Package. *Journal of Statistical Software*, 35, 1–22.
- Sobol', I. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55, 271–280. [https://doi.org/10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6).
- Sparkman, D., Garza, J., Millwater Jr, H., & Smarslok, B. (2016). Importance sampling-based post-processing method for global sensitivity analysis. In *Proceedings of the 18th AIAA non-deterministic approaches conference*.
- Syn, C. Z., Mokhtar, M., Feng, C. J., & Manurung, Y. H. P. (2011). Approach to prediction of laser cutting quality by employing fuzzy expert system. *Expert Systems with Applications*, 38, 7558–7568. <https://doi.org/10.1016/j.eswa.2010.12.111>.
- Tarantola, S., Gatelli, D., & Mara, T. (2006). Random balance designs for the estimation of first order global sensitivity indices. *Reliability Engineering & System Safety*, 91, 717–727.
- Tobon-Mejia, D. A., Medjaher, K., & Zerhouni, N. (2012). CNC machine tool's wear diagnostic and prognostic by using dynamic Bayesian networks. *Mechanical Systems and Signal Processing*, 28, 167–182.
- Vijayaraghavan, A., Sobel, W., Fox, A., Dornfeld, D., & Warndorf, P. (2008). Improving machine tool interoperability using standardized interface protocols: MT connect. In *2008 international symposium and flexible automation*.
- Vundavilli, P. R., Parappagoudar, M. B., Kodali, S. P., & Benguluri, S. (2012). Fuzzy logic-based expert system for prediction of depth of cut in abrasive water jet machining process. *Knowledge-Based Systems*, 27, 456–464. <https://doi.org/10.1016/j.knsys.2011.10.002>.
- Wang, B., & Yan, X. (2019). Real-time monitoring of chemical processes based on variation information of principal component analysis model. *Journal of Intelligent Manufacturing*, 30, 795–808.
- Wang, J., Gao, R. X., Yuan, Z., Fan, Z., & Zhang, L. (2019). A joint particle filter and expectation maximization approach to machine condition prognosis. *Journal of Intelligent Manufacturing*, 30, 605–621. <https://doi.org/10.1007/s10845-016-1268-0>.
- Weber, P., & Jouffe, L. (2006). Complex system reliability modelling with dynamic object oriented Bayesian networks (DOBN). *Reliability Engineering & System Safety*, 1, 149–162.
- Wu, D., Greer, M. J., Rosen, D. W., & Schaefer, D. (2013). Cloud manufacturing: Strategic vision and state-of-the-art. *Journal of Manufacturing Systems*, 32, 564–579. <https://doi.org/10.1016/j.jmsy.2013.04.008>.

- Wu, D., Liu, S., Zhang, L., Terpenney, J., Gao, R. X., Kurfess, T., et al. (2017). A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing. *Journal of Manufacturing Systems*, *43*, 25–34.
- Xu, X. (2012). From cloud computing to cloud manufacturing. *Robotics and Computer-Integrated Manufacturing*, *28*, 75–86. <https://doi.org/10.1016/j.rcim.2011.07.002>.
- Yih-Fong, T. (2006). Parameter design optimisation of computerised numerical control turning tool steels for high dimensional precision and accuracy. *Materials and Design*, *27*, 665–675. <https://doi.org/10.1016/j.matdes.2004.09.029>.
- Yildiz, Y., & Nalbant, M. (2008). A review of cryogenic cooling in machining processes. *International Journal of Machine Tools and Manufacture*, *48*, 947–964. <https://doi.org/10.1016/j.ijmactools.2008.01.008>.
- Zhang, Y., Xi, D., Yang, H., Tao, F., & Wang, Z. (2019). Cloud manufacturing based service encapsulation and optimal configuration method for injection molding machine. *Journal of Intelligent Manufacturing*. <https://doi.org/10.1007/s10845-017-1322-6>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.