Computers & Industrial Engineering 101 (2016) 70–80

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

Risk based uncertainty quantification to improve robustness of manufacturing operations

Cinzia Giannetti [a,*], Rajesh S. Ransing [b]

[a] College of Science, Swansea University, Singleton Campus, Swansea SA2 8PP, UK
[b] College of Engineering, Swansea University, Bay Campus, Crymlyn Burrows, Swansea SA1 8EN, UK

ARTICLE INFO

Article history:
Received 2 April 2016
Received in revised form 15 July 2016
Accepted 8 August 2016
Available online 9 August 2016

Keywords:
Tolerance synthesis
ISO9001:2015
7Epsilon
Six Sigma
Industry 4.0
Digital manufacturing

ABSTRACT

The cyber-physical systems of Industry 4.0 are expected to generate vast amount of in-process data and revolutionise the way data, knowledge and wisdom is captured and reused in manufacturing industries. The goal is to increase profits by dramatically reducing the occurrence of unexpected process results and waste. ISO9001:2015 defines risk as effect of uncertainty. In the 7Epsilon context, the risk is defined as effect of uncertainty on expected results. The paper proposes a novel algorithm to embed risk based thinking in quantifying uncertainty in manufacturing operations during the tolerance synthesis process. This method uses penalty functions to mathematically represent deviation from expected results and solves the tolerance synthesis problem by proposing a quantile regression tree approach. The latter involves non parametric estimation of conditional quantiles of a response variable from in-process data and allows process engineers to discover and visualise optimal ranges that are associated with quality improvements. In order to quantify uncertainty and predict process robustness, a probabilistic approach, based on the likelihood ratio test with bootstrapping, is proposed which uses smoothed probability estimation of conditional probabilities. The mathematical formulation presented in this paper will allow organisations to extend Six Sigma process improvement principles in the Industry 4.0 context and implement the 7 steps of 7Epsilon in order to satisfy the requirements of clauses 6.1 and 7.1.6 of the ISO9001:2015 and the aerospace AS9100:2016 quality standard.

1. Introduction

Industry 4.0, also called the fourth industrial revolution, has already started to take place and it will involve a complete digital transformation of many manufacturing activities. This revolution will break the existing boundaries of manufacturing operations to deliver a new generation of intelligent, co-operating and interconnected manufacturing systems capable of monitoring system performance real time to control costs, reduce downtime and prevent faults (Foresight, 2013). The new manufacturing systems will be characterised by cyber-physical systems able to interoperate via networked connections and interact with humans in complex smart factory environments. These systems will make extensive use of data and predictive analytics to manage manufacturing processes more efficiently and allow production of customised products with increased profitability and energy efficiency (Deloitte, 2015; Germany Trade & Invest, 2015; Manyika, 2012; Rockwell Automation, 2014). As new technologies are starting to be deployed as part of the fourth industrial revolution, one of the biggest challenges manufacturing companies are facing is to develop capabilities to timely access and reuse the sheer volume of data and information scattered across diverse business functions to gain new insights and to create knowledge and value for the enterprise (Foresight, 2013). As part of this digital transformation new predictive analytics tools will need to be developed to access, integrate and use the vast, multi-faceted and heterogeneous data sets that will become available, including machine and human generated data collected through sensors and other interconnected IT systems.

In the context of continual improvement, undoubtedly the new generation of manufacturing systems represent an important opportunity for leveraging existing continual improvement capabilities by exploiting the potential to create new knowledge from in-process data and enabling real-time decision making capabilities. Continual improvement is defined by the ISO9001:2015 standard as a "recurring activity to enhance performance", and the one that generally leads to a corrective or preventive action (International Standard Organisation ISO, 2014, p. 16). This

* Corresponding author.
E-mail addresses: c.giannetti@swansea.ac.uk (C. Giannetti), r.s.ransing@swansea.ac.uk (R.S. Ransing).

http://dx.doi.org/10.1016/j.cie.2016.08.002
0360-8352/© 2016 The Authors. Published by Elsevier Ltd.

typically involves reducing variation in production processes to satisfy customer requirements. According to Stricker and Lanza (2014), the robustness of a production system should aim for both a target value of the process outcome and a stable or consistent performance with minimum deviation or variation. In multiprocess manufacturing achieving process robustness is a challenging activity because the quality of the final product is often influenced by hundreds of factors as well as part specific quality constraints (Giannetti et al., 2014, 2015; Ransing & Ransing, 2014; Roshan, Giannetti, Ransing, & Ransing, 2014). Production processes in foundries are a typical example of multiprocess manufacturing as they consists of many sub-processes (i.e. patternmaking, molding, coremaking, melting and pouring, heat treatment, welding and finishing), with their quality determined by the effect and interactions of many process inputs. For these processes, quality of the final product cannot be simply achieved by limiting process variability to predefined thresholds determined according to the customer requirements. In fact, despite working within specifications, a process may still exhibit a large amount of variance in its output target value. Process knowledge is often necessary to implement changes which will lead to enhanced performance and achieve process robustness. Recently a novel methodology, called 7Epsilon (2015), has been developed which promotes the use of risk based analysis of in-process data to create new product specific process knowledge and evaluate opportunities that will lead to improvement of manufacturing processes, as required by the ISO9001:2015 standard. (Giannetti et al., 2015; Ransing, Batbooti, Giannetti, & Ransing, 2016; Roshan et al., 2014). Ransing et al. (2016) have shown that new product specific process knowledge can be created from in-process data by means of tolerance synthesis. In the literature process tolerance synthesis is defined as the problem of allocating tolerances of process variables to achieve a specified quality at a minimum costs (Ding, Jin, Ceglarek, & Shi, 2000). Extending this definition to the context of multiprocess manufacturing, tolerance synthesis is the study of variability in all process inputs (including interactions among process inputs) in order to discover optimal regions that correlate with the occurrences of expected process outputs (results) (Ransing et al., 2016). Owing to its definition, tolerance synthesis involves developing a sound understanding of how variability of process factors (i.e. process input settings) affects the expected target value and the variability of responses (i.e. process outputs). Process robustness is then achieved by selecting optimal tolerance limits of process variables that will reduce variation of responses (Ransing et al., 2016). One approach to solve the tolerance synthesis problem and predict process robustness is to attempt to model the relationships between process factors and responses from in-process data. In the literature data driven predictive methods have been used and applied to several industrial sectors, including manufacture of fabricated metal products, computers and electronic goods (Köksal, Batmaz, & Testik, 2011). The influence of design and process parameters has also been studied via numerical simulation methods (Lewis, Manzari, Ransing, & Gethin, 2000; Lewis & Ransing, 2000; Pao, Ransing, Lewis, & Lin, 2004; Postek, Lewis, Gethin, & Ransing, 2005), decision trees (Bakır et al., 2006) and Bayesian networks (Lewis & Ransing, 1997). Typically these methods attempt to model the complex relationships between process inputs and outputs to characterise or, sometimes, predict process behaviour and find improvement opportunities. However, for complex manufacturing processes, these relationships are not easily captured due to several reasons. First of all, in multiprocess manufacturing operations, the quality of the final product is often influenced by a combination of large number of product and process variables, including both categorical and continuous variables. Secondly, relationships between inputs and quality characteristics are related not only to some physical phenomena but also to interactions of different process settings. Trying to model these relationships can become very cumbersome with the risk of including variables with little effect on the final quality output (Giannetti et al., 2014). Traditional data driven approaches, such as regression analysis, tend to fail due the inability to model complex interactions and overfitting problems due to the presence of noise. Unless some prior knowledge about the underlying model is available, fitting the data with simple models, such as a linear model, would fail to capture the complex interactions (Bakır et al., 2006). On the other hand, using more complex models (e.g. polynomials) would lead to overfitting because of the presence of noise and small amount of observations. Overfitting will then produce a model that performs very well on the available data but has very poor predictive performance. In order for process knowledge to be learnt robustly, there is the need to analyse weak patterns in noisy and heterogeneous datasets. Furthermore, because of the presence of noisy data, uncertainty of the model results need to be quantified to overcome the lack of process knowledge.

In this paper a novel algorithm is proposed to predict the robustness of a process by quantifying uncertainty in manufacturing operations. The main motivation of this work is to develop a robust and general purpose method for tolerance synthesis to quantify the combined effects of process variables on the quality output without making distributional assumptions and overcome the linearity assumption of previous algorithms for risk based tolerance synthesis (Giannetti et al., 2014; Ransing et al., 2016). This is achieved by introducing a novel mathematical formulation of the tolerance synthesis problem in terms of conditional quantiles of response variables and a robust algorithm based on quantile regression to find optimal tolerance limits. The method improves the previous quality correlation algorithm for tolerance synthesis (Ransing et al., 2016) by using the concept of likelihood ratio for probabilistic estimation of the effects of the new tolerance limits on the quality output. Uncertainty quantification of the newly developed hypotheses is performed using the bootstrap method to predict process robustness and aid development of new product specific process knowledge.

The paper is organised as follows. Section 2 reviews regression trees methods and their industrial applications, including traditional least square and quantile regression approaches. Section 3 introduces the tolerance synthesis problem, its mathematical formulation and the proposed algorithm. The latter includes a probabilistic approach for hypotheses validation based on calculation of likelihood ratio with bootstrap method. The method is illustrated using test data from the UCI machine learning repository. In Section 4 the proposed algorithm is applied to an industrial case study to show its application for uncertainty quantification in multiprocess manufacturing systems. The paper is concluded in Section 5.

2. Related methods: regression trees and quantile regression

Decision tree learning is a common method used for classification and regression problems, owing its popularity to easiness of interpretation and the ability to visually and explicitly represent decision making rules (Bakır et al., 2006). The general method builds a tree shaped structure to predict or classify a dependent variable (often called response variable) by recursive partitioning the data set into groups of observations with similar values of the dependent variable (Breiman, Friedman, Stone, & Olshen, 1984). One main advantage of decision tree learning is that it can deal simultaneously with continuous and categorical predictor variables, without the need of further transformations and making distributional assumptions (Francke, López-Tarazón, & Schroder, 2008). Regression trees are particular types of decision tree designed to work with continuous response variables, while