Engineering Applications of Artificial Intelligence 54 (2016) 17–25

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Trend modelling with artificial neural networks. Case study: Operating zones identification for higher SO_3 incorporation in cement clinker

R.N. Lima [a], G.M. de Almeida [b,*], A.P. Braga [c], M. Cardoso [b]

[a] Holcim, Brazil
[b] Department of Chemical Engineering, Federal University of Minas Gerais, Brazil
[c] Department of Electronic Engineering, Federal University of Minas Gerais, Brazil

A R T I C L E I N F O

Article history:
Received 6 July 2015
Received in revised form
23 February 2016
Accepted 2 May 2016
Available online 12 May 2016

Keywords:
Trend modelling
ANN
Instantaneous measurements
Process data dispersion
Time alignment
Variable relationship

A B S T R A C T

Instantaneous measurements of process variables are usually not representative of the process effects as a whole when defining the condition of an output sample mainly in case of laboratory analysis. Moreover, process data have considerable dispersion. This leads to uncertainty in input–output time alignment and in variable relationship. This work employs a trend data-based approach to overcome the negative effects of these uncertainties in both tasks variable selection commonly supported by correlation analysis and model identification. Two real case studies using a clinker rotary kiln from a cement plant and a chemical recovery boiler from a pulp mill were used for illustration purposes. More reliable data-driven system representation enhances the comprehension of the underlying system phenomena supporting a more rational basis for decision making.

1. Introduction

Transformation of data into useful information for supporting decision making has been a great challenge to modern organizations worldwide (Tata, 2013; Manyika et al., 2011). Currently, only small portions of data are really exploited for information extraction. The reason is the lack of ready-made recipes for handling, exploring and analyzing the increasingly massive and also complex amount of data. This new data rich-world demands new approaches for performing data analysis including in thinking (NRC, 2013). On the other hand, a continuous knowledge generation to answer the growing tighter government and society regulatory constraints in addition to market competitiveness is mandatory; data-oriented decision making plays a key role in it (Mayer-Schonberger and Cukier, 2013).

Continuous processing systems such as steel, petrochemicals, and pulp and paper mills, have also experienced the data rich-word paradigm (Qin, 2014; Venkatasubramanian, 2009). Nowadays, a multitude of plant sensors registers hundreds or thousands of measurements of process variables even in a fraction of a second. On the other hand, descriptions of continuous industrial systems are more and more challenging as a result of their inherent growing complexity, namely multivariable, non-linear, noisy and of partial knowledge, which restricts the use of purely mathematical models. Alternatively, the data-driven concept using process historical data has been employed for understanding the physics and chemistry of processes, or at least for their emulation (Venkatasubramanian et al., 2003; Chiang et al., 2001; Wang, 1999). Nowadays, extracting useful information directly from data plays a key role in making decisions for obtaining safer, cleaner, more economical and more efficient industrial operations.

Apart from continuous process representation, dealing with raw data sets is also a challenging task once they are composed of e.g. redundant and irrelevant information, asynchronous sample times, discrete and continuous variables, and laboratory and field sources. Therefore, to some degree, a pre-processing step is necessary before using it. Nevertheless, making use of punctual information given by individual measurements of process variables may be not appropriate. This resides on the fact that *instantaneous measurements* are commonly not representative of process effects as a whole defining the condition of a given output sample. This is more critical when such condition derives from laboratory analysis. In addition, the time a portion of material spends inside an industrial equipment depends on several factors e.g. feeding load and fuel type to mention a few. However, this *time interval* (*or residence time*) usually assumed fix changes continuously in practice. This uncertainty regarding time alignment also hampers obtaining realistic input–output correspondence mainly in case of having laboratory data. Furthermore, an inherent characteristic of process data is considerable *dispersion*. This arises from the natural stochastic uncertainty given by automation devices, systematic

* Corresponding author.

http://dx.doi.org/10.1016/j.engappai.2016.05.002
0952-1976/

uncertainty mainly due to loss of sensor calibration over time, and mainly the extensive operating ranges accepted in continuous processing systems. Multiple modes of operation (or operating states) and the frequent interchanges among them also contribute to a higher dispersion.

This way, instantaneous measurements, fixed residence time and considerable data dispersion result in masking of the real input–output time alignment and of the form and intensity of variable relationship. This resulting uncertainty negatively compromises the performance of both tasks variable selection commonly based on correlation analysis, and model identification regarding input–output mapping. The following studies illustrate the necessity of taking these uncertainties into account when working with industrial process data. Most of the applications concern data reconciliation. Hodouin and Everell (1980) investigate the weighted sum of squared residuals to satisfy mass conservation equations. Hodouin et al. (1998) material conservation equations to deal with the integration error caused by dynamic variations. Kongsjahju et al. (2000) use a derivation from the unbiased estimation technique to detect gross error under serial correlation. Abuelzeet et al. (2002) study a global dynamic data reconciliation strategy with the aim of coping with systematic bias caused by miscalibrated instruments and outliers caused by process peaks fluctuations. Chen et al. (2013) reduce the effect of random and gross errors in process data by using an entropy based-estimator for satisfying balance equations in a data reconciliation study. Zhu et al. (2015a, 2015b) use, respectively, probabilistic principal component analysis and probabilistic principal component regression, to treat outliers with focus on soft sensors development. An alternative approach to treat process data uncertainty is to make use of trend rather than individual measurements.

The present paper proposes a trend data-based approach to overcome uncertainties present in input–output time alignment and variable relationship. After a pre-processing step transforming instantaneous data to trend series, the Multi-Layer Perceptron artificial neural network is employed for variable selection and model identification. The proposed methodology is illustrated using two real case studies namely a clinker rotary kiln of a cement mill and a chemical recovery boiler from a pulp mill.

The remainder of the paper is organized as follows. Section 2 depicts the proposed methodology. The results followed by discussion are presented in Section 3, and Section 4 gives final considerations.

2. Material and methods

The proposed methodology is depicted in Fig. 1. The first three steps including a trend data generation procedure constitutes the Data Pre-Processing Phase.

Step 1: *Data Set Preparation.* This initial step deals with the removal of samples in the raw data set that present some type of inconsistency, e.g. registry errors and extreme, missing and anomalous data. A visual analysis using graphical representations supports this task. In case of unequal sample times, a synchronization procedure needs to be accomplished. This point often arises in case of having several data sources mainly between field and laboratory data.

Step 2: *Variable Selection.* For a better parameter estimation and lesser time computation, redundant and irrelevant information should be not considered for modelling purposes. The previous graphical analysis in conjunction with a measure of degree of association using original and transformed (x^2, $\sqrt{x}$, and $\log_{10}x$) variables are employed (Chatterjee and Hadi, 2006).

Step 3: *Trend Data Set Preparation.* This step employs a trend-cycle filter to generate trend series to be used as model inputs

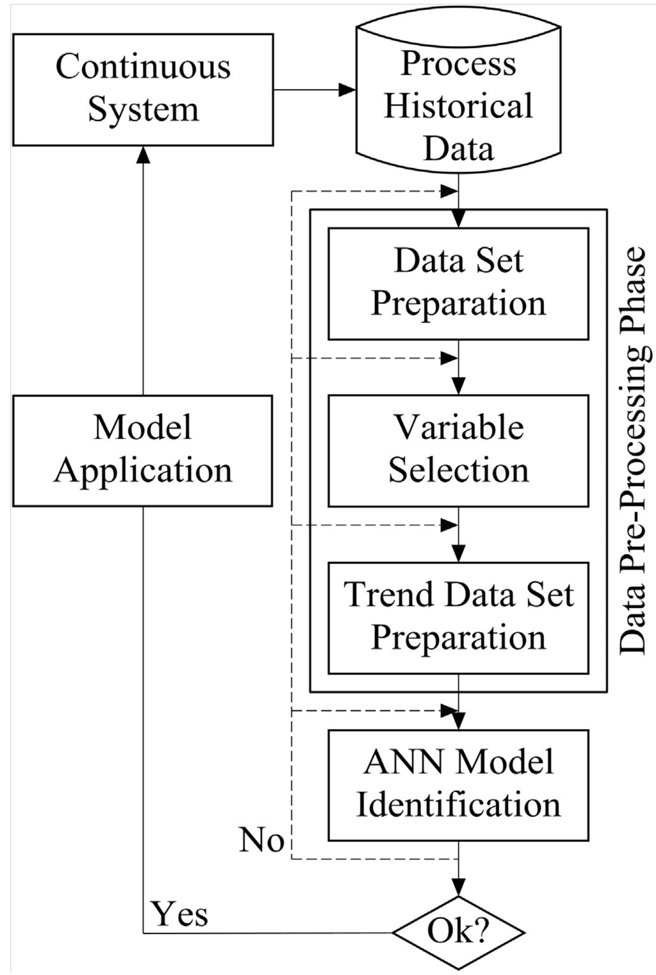

Fig. 1. Sequence of steps of the proposed methodology.

instead of individual measurements. Trend-cycle decomposition is addressed e.g. in Dupasquier et al. (1997) and Chagny and Dopke (2002). This work makes use of the Mohr filter (Mohr, 2005). It consists in an univariate approach to decompose a times series into a trend, a cyclical and a seasonal component. The Mohr filter can be seen as an extension of the well known Hodrick–Prescott filter (HP filter) (Hodrick and Prescott, 1997) by considering explicit stochastic models for the cyclical, assumed to follow a stationary ARMA-process, and the seasonal components. Furthermore, the stochastic trend model, restricted to a second order random walk process in the HP filter, which may not be always appropriate depending on the properties of the series to be filtered, gives place to a stochastic trend of arbitrary order in the Mohr filter. This work applies its particular form called Trend-Cycle filter (TC filter). The HP filter is obtained by minimising the following objective function, in matrix form: $(X-X^T)'(X-X^T)+\lambda X^{T'}\nabla^{2'}\nabla^2 X^T$, for x_t^T (where X and X^T are $[T\times 1]$ original and trend series, λ is a smoothing parameter, and ∇^2 is a second-difference matrix). Its solution follows from the first order conditions: $X^T=(I+\lambda\nabla^{2'}\nabla^2)^{-1}X$, and $X^C=X-X^T$. In contrast, the optimization problem formulation of the TC filter is as follows: $(X-X^C-X^T)'(X-X^C-X^T)+(\nabla^{d-1}(\nabla X^T-Ub))'(\nabla^{d-1}(\nabla X^T-Ub))+X^{C'}A'(BB')^{-1}AX^C$, for X^T and X^C (the $[T\times 1]$ cyclical series), and b if the trend is assumed to follow a first-order random walk process with drift ($d=1$) (where A and B are $[N-2c\times N]$ matrices representing the *AR* and the *MA* process, respectively, U is a $[T\times 1]$ vector $[0, 1, \ldots, 1]'$, $M_C\equiv(I+A'(BB')^{-1}A)^{-1}$, and $M_T\equiv(I+\nabla'(I-UU'(N-1)^{-1})\nabla)^{-1}$, if