Fault Detection and Replacement of a Temperature Sensor in a Cement Rotary Kiln

Tiago Matias[1], Dulce Gabriel[1,2], Francisco Souza[1], Rui Araújo[1] and J. Costa Pereira[2]
[1]Institute of Systems and Robotics (ISR-UC), and
Department of Electrical and Computer Engineering (DEEC-UC),
University of Coimbra, Pólo II, PT-3030-290 Coimbra, Portugal.
[2]Department of Chemistry (DQ-UC),
University of Coimbra, PT-3004-535 Coimbra, Portugal.
tmatias@isr.uc.pt, dracag@isr.uc.pt, fasouza@isr.uc.pt, rui@isr.uc.pt, jcpereira@qui.uc.pt

Abstract

This paper proposes a method for fault detection and replacement of the sensor responsible by the measurement of the burning zone temperature in a rotary cement kiln. The control of the burning zone temperature is crucial for the control of kiln temperature and therefore for the control of the cement quality, pollutant emissions, and consumed energy. However the flying dust within the kiln can block the pyrometer sensor, causing faults in the temperature sensor. Exploring the analytical redundancy that usually exist in industrial processes, the proposed methodology uses a neural network trained using an online sequential extreme learning machine to online construct a model to estimate the burning zone temperature. Using the error between the measured and estimated temperatures, faults in the measurement can be detected and thus the replacement of the measured temperatures by the estimated output is made.

1 Introduction

Cement manufacture involves grinding a mixture of limestone, shale, clay, sand, and smaller quantities of other substances in controlled proportions to a fine powder, and heating the resulting raw cement in a large rotary kiln. Inside the kiln a complex exothermic chemical reaction takes place at temperatures in the range of 1200-1700°C. Such temperatures bring the cement to a semi-liquid state and converts it to small black nodules called clinkers, which are cooled and ground again to make the finished product [18].

During the burning process, the control of the temperature inside the kiln is crucial: insufficiently high maximum temperatures in the kiln result in incompletely reacted products and poor-quality cement, while excessive maximum temperatures waste energy and propitiate the formation of NO_x pollutant compounds that have several environmental impacts [21]. NO_x is formed by the reaction of nitrogen in air and fuel with oxygen at the high temperatures reached during the clinker production process. In addition to contributing to the formation of ground-level ozone, acid rain, and fine particle pollution, NO_x is linked with a number of adverse effects on the respiratory system [4].

The correct and online measure of the temperature in the burning zone is crucial in order to ensure a good-quality clinker and low level of pollutant emissions. As the contact temperature measurement is impossible, the measurement of the temperature is made using a pyrometer. However, due to flying dust within the kiln system that blocks the sensor, many measurements have faults. When the measured burning zone temperature is too low (based on the operator knowledge), the sensor is removed and cleaned, which can take some minutes. Furthermore, since the sensor is giving erroneous measurements until the operator becomes aware of this, it may take a long time.

Methodologies for abnormal/faulty working condition detection in cement rotary kiln are an active topic of research. In [17] and [22] a model for normal working conditions was constructed by a locally linear neuro-fuzzy model trained by Locally Linear Model Tree algorithm. Then, the outputs of the model are compared with the plant outputs and abnormal working conditions can be detected. A methodology that uses a fuzzy logic fault diagnoser for fault detection in a cement rotary kiln is proposed in [2]. In [13] an algorithm for fault detection in a cement kiln is also proposed. This method is based on a binary ant colony and support vector machines. All of these methods focus in helping the operators in insuring the normal operation of the kiln and in the diagnosis of faults. However, to the best of the authors' knowledge, no method for fault detection in the burning zone temperature sensor can be found in the literature. As previously mentioned, a tool that allows a correct measurement of the burning zone temperature is crucial to ensure a good-quality clinker, and to control the temperature inside of the kiln and, therefore, reducing the pollutant emissions.

The aim of this work is to create a methodology to detect failures in the burning zone temperature sensor. The proposed methodology explores the redundancy of the information available from all the sensors in the kiln to construct a model to estimate the temperature in burning zone. Comparing these estimates with the temperatures measured by physical sensor it is possible to detect faults in the sensor. When a fault is detected, the estimates obtained by the proposed methodology replace the measures obtained from the sensor and warns the operator to the necessity of cleaning the sensor. This work was made with the ultimate objective of attaining a correct control of the cement kiln. The control of the kiln can be improved using correct measures of the temperatures in the burning zone, improving the cement quality, and reducing the pollutant emissions and the waste of energy.

The paper is organized as follows. In Section 2 it is presented a brief description of the cement production process. The sensor failure detection and replacement methodology proposed in this paper is presented in Section 3. Section 4 presents the obtained results. Finally, Section 5 gives concluding remarks.

2 Process Description

The cement is made mostly of a mixture of limestone (supplies the bulk of calcium oxide or lime (CaO)) and clay, marl or shale (supplies the bulk of silica (SiO_2), alumina (Al_2O_3), and iron oxide (Fe_2O_3)).

These materials first pass through a series of crushing, stockpiling, milling and blending stages, which yield an intimately mixed and dry raw meal. The raw meal then passes through a preheater and frequently also a precalciner (combustion chamber which is able to drive the dominant endothermic process of calcination), to initiate the dissociation of calcium carbonate to calcium oxide and carbon dioxide (calcination), before entering a rotary kiln. The more carbon dioxide (CO_2) is released from the raw material, the less work needs to be performed on the feed in the kiln, which increases the efficiency of the process greatly.

A preheater is a heat exchanger tower consisting of several suspension cyclone stages in which the moving powder is dispersed in a stream of hot gas coming from the kiln. The feed temperature at the lowest cyclone stage reaches at least 800°C. The feed flows down the kiln and the gas is drawn upwards by a ventilator at the exhaust. In Fig. 1 the layout of a kiln with preheater and calciner is shown.

In the kiln, the meal is heated and it occurs the reactions between calcium oxide and other elements to form new compounds (tricalcium silicate, dicalcium silicate, tricalcium aluminate and tetracalcium aluminoferrite) at a temperature up to 1450°C. Primary fuel is used to keep the temperature high enough in the burning zone for the chemical reactions to take place. For each ton of material that goes into the feed end of the kiln, two thirds of a ton then comes out the discharge end as a nodular material called clinker. The clinker then drops onto the cooler where it is rapidly cooled to approximately 150°C. In the final stage, the clinker will then be intergrounded with gypsum, limestone and/or ashes, producing a fine product called cement.

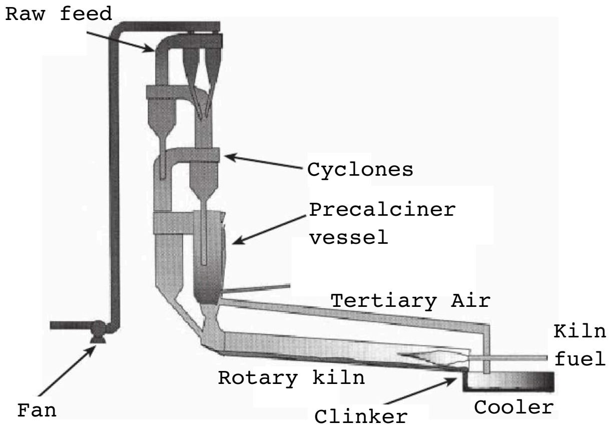

Figure 1: A diagram of a cement rotary kiln plant, adopted from [19].

3 Adopted Failure Detection and Replacement Strategy

Sensor failure can be defined as an unpermitted deviation from the acceptable/usual/standard condition of at least one characteristic property or parameter of the system obtained by the sensor [1]. Fault detection techniques for sensors and actuators is an active topic of research [3, 11, 14, 24]. In the construction of a fault tolerant system, the factor that needs to be explored is the redundancy.

The most direct form of redundancy is the physical redundancy. It consists in the use of more than one physical sensor to measure the same variable. When physical and mathematical relationships among the system inputs and measured outputs can be known, the physical redundancy can be substituted by the presence of another redundant source of information, which is called analytical redundancy (AR) [5].

The methodology proposed in this work is an online learning-based sensor failure detection approach that explores the possible AR of the system to detect faults and replace an optical piezoelectric sensor responsible for the measurement of the burning zone temperature in a cement kiln process. Exploring the AR, a model that can be used to detect erroneous measurements from the physical sensor is constructed online. In the proposed methodology, a single hidden-layer feedforward neural network (SLFN) trained using a online sequential extreme learning machine (OS-ELM) is used to describe the relation between the input variables and the variable measured by physical sensor. The error between the predicted output of the model and the measure of the physical sensor will be used to detect possible abnormal measurements obtained by the physical sensor. In case of abnormal measurements, the real-sensor measurements are replaced by the measurement obtained by the SLFN.

3.1 Online Data Pre-processing

The proposed methodology have as objective his real implementation, so in order to achieve a successful model construction, online methods for outlier detection and scaling of the variables are required.

3.1.1 Online Outlier Detection

An outlier is an observation that is inconsistent with the remaining data. Outliers may be generated by abnormal operating conditions of the plant, failure in measurement, in hardware, or in transmission. If outlier observations are used to update the model, the predictions from the updated model may deteriorate.

For the detection of the outliers it was assumed that the data is identically and independently distributed (i.i.d), and the Hampel identifier was used [7]. Among the existing statistical outlier identifier methods, the Hampel identifier is one of the most robust and efficient [15], both for static operation and for on-line operation. One-class classifiers like density estimators or reconstruction methods can also be used for on-line outlier detection; however these methods are computational more expensive, and a training data set or some knowledge about the outliers is required [23]. Furthermore, due to the time-variant behavior of the process, the collection of a training data set representative of the process dynamics is difficult. Therefore these methods are not appropriate for the application discussed in this paper. In the Hampel algorithm, an incoming data sample χ at time-instant k can be identified as outlier if the following condition is fulfilled [15]:

$$|\chi(k) - median_{ws}(\chi)| > 3 \times MAD(\chi), \tag{1}$$

where $median_{ws}(\chi) = median([\chi(k - ws + 1), \ldots, \chi(k)]^T)$ is the median of χ in a window with size ws, and MAD is the median absolute deviation from the median that is defined as:

$$MAD(\chi) = 1.4826 \times median(|\chi(k) - median_{ws}(\chi)|). \tag{2}$$

This identifier is similar to the 3σ rule [5], however the mean and standard deviation operators were replaced by the median and MAD operators. Using these last operators, the scale estimation of the variable is much more robust against outliers [15].

In this work, after the outlier identification procedure, if an outlier is identified in a sample $\chi(k)$, then the sample is replaced by $median_{ws}(\chi)$.

3.1.2 Online Data Scaling

In a industrial process the measured variables may have different magnitudes. In order to perform an accurate model building, all input variables should have the same magnitude and therefore they should be scaled. In the present work, each new incoming data point $\chi(k)$ of variable χ was scaled to zero mean and unit variance using the following equation:

$$\tilde{\chi}(k) = \frac{\chi(k) - \overline{\chi}(k)}{\sigma^2_{\chi}(k)}, \tag{3}$$

where $\tilde{\chi}(k)$ is the scaled version of $\chi(k)$, and $\overline{\chi}(k)$ and $\sigma^2_{\chi}(k)$ are the mean and standard deviation of variable χ at instant k. Many industrial processes have time-variant behaviors, so the mean and standard deviations of the variables change over the time. Using the following equations the mean and standard deviation of χ at instant k can be recursively obtained [12]:

$$\overline{\chi}(k) = \frac{k-1}{k}\overline{\chi}(k-1) + \frac{1}{k}\chi(k), \tag{4}$$

$$\sigma^2_{\chi}(k) = \frac{k-2}{k-1}\sigma^2_{\chi}(k-1) + \frac{1}{k-1}(\chi(k) - \overline{\chi}(k))^2. \tag{5}$$

The data can be also unscaled using:

$$\chi(k) = \tilde{\chi}(k)\sigma^2_{\chi}(k) + \overline{\chi}(k). \tag{6}$$

After variable scaling, equal conditions are ensured for each input variable to influence the model. The model will be constructed using a single hidden-layer feedforward neural network trained with a OS-ELM approach and the online scaled variables.

3.2 Online Modeling using an Online Sequential Extreme Learning Machine

For the model construction an online sequential extreme leaning machine proposed in [9] was used. This methodology is a sequential approach of the batch extreme learning machine (ELM) proposed in [10]. The sequential approach was used because, in the presence of processes with time-variant behaviors, it is crucial the adaptation of the parameters of the network. Using the traditional backpropagation learning techniques, the training of the network is very slow and computationally expensive. Using the principle of the ELM, the weights of the first level can be randomly generated and the output weights of the network can be obtained by least squares, which is very fast when compared with traditional backpropagation methods. Furthermore, the ELM algorithm tends to provide better generalization performance compared with conventional learning algorithms [9, 10].

Consider a single hidden-layer feedforward neural network (SLFN) with n_i input variables, n_j hidden-layer neurons, and one neuron in the output layer with linear activation function. The output of a neural network at time-instant k is given by:

$$\tilde{y}(k) = \mathbf{v}^T\mathbf{s}(k), \tag{7}$$

where $\mathbf{v} = [v_1, \ldots, v_{n_j}]^T$ is the vector of output weights v_j $(j = 1, \ldots, n_j)$, and $\mathbf{s}(k) = [s_1(k), \ldots, s_{n_j}(k)]^T$ is the vector of the outputs of the n_j neurons given by:

$$\mathbf{s}(k) = \mathbf{\Gamma}\left(\mathbf{W}^T\tilde{\mathbf{x}}(k)\right), \tag{8}$$

where

$$\mathbf{W} = \begin{bmatrix} w_{01} & w_{02} & \dots & w_{0n_j} \\ w_{11} & w_{12} & \dots & w_{1n_j} \\ \vdots & \vdots \ddots \vdots & & \\ w_{n_i1} & w_{n_i2} & \dots & w_{n_in_j} \end{bmatrix}. \qquad (9)$$

w_{0j} is the bias of hidden neuron j, and w_{ij} is the weight between the i-th input variable and the j-th hidden-layer neuron. $\tilde{\mathbf{x}}(k) = [1, \tilde{x}_1(k), \tilde{x}_2(k), \dots \tilde{x}_{n_i}(k)]^T$ is the vector of scaled inputs where the first element is for the bias of each hidden neuron, and $\mathbf{\Gamma}(\boldsymbol{\eta}) = [g(\eta_1), \dots, g(\eta_{n_j})]^T$ for $\boldsymbol{\eta} = [\eta_1, \dots, \eta_{n_j}]^T$. It is assumed that the activation function of each neuron is a sigmoidal function, $g(\eta_j) = 1/(1 + \exp(-\eta_j))$ $(j = 1, \dots, n_j)$.

In [8] it is proved that in the SLFNs the weights between the input layer and the hidden layer, and the bias of the neurons of the hidden layer do not need to be adjusted during the training, and can be randomly assigned. In the batch ELM with N available samples, the input weights and bias matrix $\mathbf{W}$ are randomly assigned and, considering an output neuron with a linear activation function, the output weights vector $\mathbf{v}$ is estimated as:

$$\hat{\mathbf{v}} = \mathbf{S}_N^* \tilde{\mathbf{y}}_N^d, \qquad (10)$$

where $\mathbf{S}_N^*$ is the Moore-Penrose generalized inverse of the hidden layer output matrix

$$\mathbf{S}_N = [\mathbf{s}(1), \dots, \mathbf{s}(N)], \qquad (11)$$

and $\tilde{\mathbf{y}}_N^d = [\tilde{y}^d(1), \dots, \tilde{y}^d(N)]^T$, is the vector of the scaled target outputs.

Considering that $\mathbf{S}_N \in \mathbb{R}^{N \times n_j}$ with $N \geq n_j$ and $rank(\mathbf{S}_N) = n_j$ the Moore-Penrose generalized inverse of $\mathbf{S}_N$ can be given by:

$$\mathbf{S}_N^* = (\mathbf{S}_N^T \mathbf{S}_N)^{-1} \mathbf{S}_N^T. \qquad (12)$$

So substituting (12) into (10), the estimation $\hat{\mathbf{v}}$ of $\mathbf{v}$ can be obtained by the following least-squares solution:

$$\hat{\mathbf{v}} = (\mathbf{S}_N^T \mathbf{S}_N)^{-1} \mathbf{S}_N^T \tilde{\mathbf{y}}_N^d. \qquad (13)$$

The sequential implementation of the ELM results in the application of recursive least-squares (RLS) to estimate the output weights vector. Considering that N_0 ($N_0 \geq n_j$) initial data samples are available and that $rank(\mathbf{S}_{N_0}) = n_j$. So the estimate $\hat{\mathbf{v}}$ can be obtained by the following algorithm:

1. Initialize the covariance matrix $\mathbf{M}_0 = (\mathbf{S}_{N_0}^T \mathbf{S}_{N_0})^{-1}$ and the output weights estimation $\hat{\mathbf{v}}_0 = \mathbf{M}_0 \mathbf{S}_{N_0}^T \tilde{\mathbf{y}}_{N_0}^d$.
2. In each newly available data sample k, the output weights estimation and the covariance matrix can be recursively obtained by [16]:

$$\hat{\mathbf{v}}_k = \hat{\mathbf{v}}_{k-1} + \mathbf{M}_{k-1}\mathbf{s}(k) \times \frac{\tilde{y}^d(k) - \mathbf{S}^T(k)\hat{\mathbf{v}}_{k-1}}{\lambda + \mathbf{S}^T(k)\mathbf{M}_{k-1}\mathbf{s}(k)}, \qquad (14)$$

$$\mathbf{M}_k = \frac{1}{\lambda}\left(\mathbf{M}_{k-1} - \frac{\mathbf{M}_{k-1}\mathbf{s}(k)\mathbf{S}^T(k)\mathbf{M}_{k-1}}{\lambda + \mathbf{S}^T(k)\mathbf{M}_{k-1}\mathbf{s}(k)}\right), \qquad (15)$$

where λ is a forgetting factor.

3.3 Methodology

The proposed methodology for sensor failure detection and replacement is described in the following algorithm:

1. Collect a small initial training set $[\mathbf{X}_{N_0}, \mathbf{y}_{N_0}^d]$ to boost the learning algorithm, where $\mathbf{X}_{N_0} = [\mathbf{x}(1), \dots, \mathbf{x}(N_0)]$. The number of the samples N_0 can be equal to the number of neurons of hidden layer n_j. However, in order to obtain a good prediction and avoid false sensor faults in initial samples of on-line operation stage, a larger number of initial samples can be considered;
2. Define a windows size ws for the outlier detection. For each training sample k with $k \geq ws$: verify the existence of outliers in input variables x_i ($i = 1, \dots, n_i$) and output variable y^d using (1). If any sample was identified as an outlier, replace it by the median of the respective variable in the present window of the most recent ws samples;
3. Update the mean and the variance of the input and output variables using (4) and (5);
4. Scale the data using (3), obtaining $[\tilde{\mathbf{X}}_{N_0}, \tilde{\mathbf{y}}_{N_0}^d]$;
5. Assign an arbitrary input weights and bias matrix $\mathbf{W}$;
6. Calculate the matrix of the outputs of the neurons of hidden layer $\mathbf{S}_{N_0}$ using (11) and (8);
7. Estimate the initial output weights vector $\hat{\mathbf{v}}$ using (13);
8. For each newly available data $[\mathbf{x}(k), y^d(k)]$ at sample k:
   (a) If $k \geq ws$: verify the existence of outliers in each newly available sample using (1). If any sample was identified as an outlier, replace it by the median of the respective variable in the present window of the most recent ws samples;
   (b) Update the mean and the variance of the input and output variables using (4) and (5);
   (c) Scale the data using (3), obtaining $[\tilde{\mathbf{x}}(k), \tilde{y}^d(k)]$;
   (d) Calculate the vector of the outputs of the neurons of hidden layer $\mathbf{s}(k)$ using (8);
   (e) Obtain the estimated output $\tilde{y}(k)$ using (7);
   (f) Unscale $\tilde{y}(k)$ using (6), obtaining $y(k)$;
   (g) Compute the error $e(k) = y(k) - y^d(k)$;
   (h) If the error $e(k)$ is inside a desired interval, update the output weights vector $\hat{\mathbf{v}}$ using (14) and (15). Else, signalize a sensor fault and replace the sensor measurement $y^d(k)$ by $y(k)$;

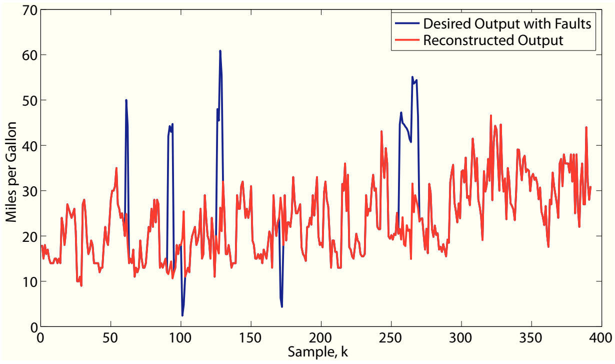

Figure 2: Desired output with injected faults and the output obtained with the proposed methodology in the modified Automobile MPG dataset.

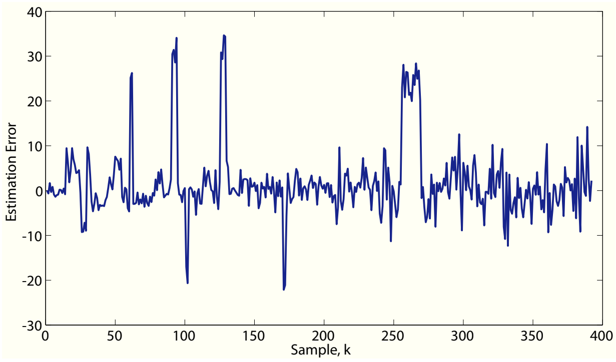

Figure 3: Estimation error in miles per gallon between the desired output with injected faults and the unscaled neural network output in the modified Automobile MPG dataset.

3.4 Methodology Validation

For the validation of the proposed methodology, it was applied to Automobile MPG dataset [6]. Faults were injected in the target output to evaluate the sensor fault detection capability and, as the original output is know, the replacement capability of the proposed methodology can be also evaluated.

This dataset corresponds to a problem of predicting the city-cycle fuel consumption of a vehicle in number of miles per gallon (MPG). It is a six input, single output regression problem. The six input variables are the number of cylinders, displacement, horsepower, weight, acceleration, and model year. The number of available samples is 398. It was considered that at sample k the measured output is $y^d(k) + \nu(k)$ where:

$$\nu(k) = \begin{cases} 30 + rand(-2,2), & \text{if } k = 61, 62, 91\text{-}94, 126\text{-}129, \\ -20 + rand(-2,2), & \text{if } k = 101, 102, 171, 172, \\ 25 + rand(-2,2), & \text{if } k = 256\text{-}269, \\ 0 & \text{otherwise,} \end{cases} \tag{16}$$

where $rand(-2, 2)$ are uniformly distributed generated random numbers on the interval $(-2, 2)$.

The results of the application of the proposed methodology are presented in Figures 2-4. From the analysis of Fig. 2 can be seen that all the injected faults are detected by the proposed methodology. In Fig. 3 is presented the error between the output with injected faults and the unscaled model output. As can be seen, in normal operation (when don't are faults in the sensor) the estimation error e is small. However, in the samples where faults were injected the estimation error increases and the faults are detected. The first 15 samples were used to initialize the method. In the experiment it was used a window size of 25 samples in the online outlier detection methodology, a neural network with 5 neurons in the hidden-layer, and a forgetting factor $\lambda = 0.92$. It was considered that the target sensor was working well if $|e(k)| < 15$ miles per gallon, where $e(k)$ is the error in sample k between the unscaled neural network output and the desired output. In the detected faults the measured outputs are replaced by the estimated output that, as can be seen in Fig. 4, is a good estimation of the original output without the injected faults.

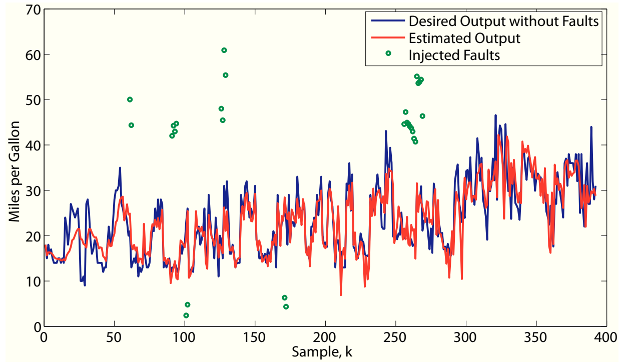

Figure 4: Original output without injected faults and the unscaled neural network output in the modified Automobile MPG dataset.

After the validation of the methodology, it was applied in the detection and replacement of faults in the temperature sensor in a cement kiln process.

4 Case Study

As previously mentioned, the objective of this work is the detection of the faults in the measurement of the temperature in the burning zone of a cement kiln, and the replacement is the faulty measurement by an adequate estimate. A correct measurement of this temperature is crucial for the control of the pollutant emissions, cement quality, and waste of energy. In this work, the experiments are made in simulation environment using a real-world dataset[1] obtained in a cement kiln plant. This dataset is composed by 194 monitorized variables, recorded with a sampling interval of $T = 1$ [min]. The monitored variables refer to several system variables from the preheater (cyclone) tower until the chimney and cement mill, and include, for example, temperatures, pressures, and manual and laboratorial entries. The used data has a total of 10000 samples which represent approximately one week.

The proposed sensor failure detection approach for the

[1]Provided by "Acontrol - Automação e Controle Industrial, Lda", Coimbra, Portugal.

burning zone temperatures of a cement kiln and the replacement of faulty readings by estimates is based on analytical redundancy that usually exists in the processes. Therefore, due to large number of input variables, the selection of the most relevant variables for the estimation of the burning zone temperature is needed. Moreover, as a neural network model was used, the presence of redundant or irrelevant variables slow down the learning process and causes the over-fitting of the model [25]. At a approach the selection of the initial set of input variables was based on the knowledge of the process. From the 193 available variables, 17 variables were selected. Some of the variables represent certain temperatures and pressures in the input and output of the kiln, fuel flows (coal and alternative fuels), temperatures in the coal mill and in the cooler, etc. In a second step, the set of input variables was refined using the sequential backward search (SBS) approach proposed in [20]. After this procedure, the following set of input variables was obtained:

- Temperature of the clinker in output of the kiln;
- Pressure of the air inside of the kiln in the burner area;
- Speed of the fan responsible for the return from the kiln to cyclone (tertiary air);
- Flow of alternative fuels;
- Flow of the fuel in the central burner;
- Flow of the fuel in the radial burner;
- Temperature in the input of the kiln.

With this set of selected input variables, the sensor fault detection and replacement methodology presented in Section 3.3 was applied. The first 30 samples were used to initialize the method. The following parameters were used in the method: a window size of 25 samples in the online outlier detection methodology, a neural network with 5 neurons in the hidden-layer, and a forgetting factor $\lambda = 0.998$. It was considered that the target sensor was working well if $|e(k)| < 100°\text{C}$.

Figures 5 and 6 present the results of the application of the proposed methodology. The first figure presents the temperatures in the burning zone measured using a pyrometer, the temperatures in the burning zone obtained after the application of the proposed fault detection and replacement methodology, and the unscaled outputs of the neural network. As previously explained, the proposed methodology uses the estimation error of Fig. 6 to detect faults in the hardware sensor and, when a fault is detected, the measured sample is replaced by the unscaled estimation of the neural network.

As previously mentioned the flying dust within the kiln system blocks the pyrometer, and this can be seen as a sensor faults. Such faults make that the measured temperatures will be lower than the real temperatures. From the

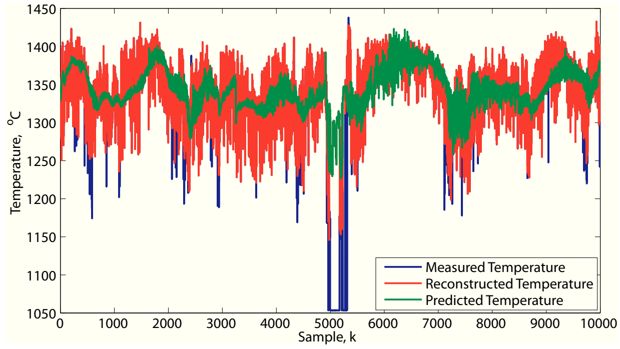

Figure 5: Measured burning zone temperature, reconstructed burning zone temperature using the proposed methodology and the unscaled neural network output in the cement kiln process.

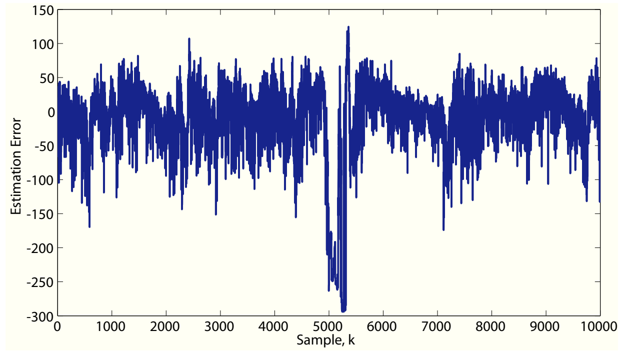

Figure 6: Estimation error in degree Celsius between the measured burning zone temperature and the unscaled neural network output in the cement kiln process.

analyses of the results it is possible to see that the proposed methodology detects some of these faults in the pyrometer sensor. Furthermore, it can be seen that usually the faults have a short duration and in some minutes the sensor is clean. However, during the week period under consideration, one fault of approximately 400 samples has occurred.

Using the proposed methodology the operator can be warned about the abnormal operation conditions, contributing to a correct measurement of the burning zone temperature and a correct control of the kiln temperature.

5 Conclusion

The paper proposed a new methodology to detect and replace faults in the sensor that measures the burning zone temperature in a rotary cement kiln. In the studied setup, the temperature was measured using a pyrometer. However, due to the flying dust within the rotary kiln, frequently there exist faults in the measurements. The correct control of the burning zone temperature is crucial for the correct control of the kiln temperature, and thereby allowing an improvement in the quality of the cement, and a reduction of the pollutant emissions and energy wasting.

Using the redundancy of information available from the sensors of the process (analytical redundancy), a model is constructed online and its output is compared with the measured temperatures in order to detect abnormal operation conditions in the hardware sensor. In order to achieve a successful model construction, online data-preprocessing methodologies were used. The proposed methodology was validated in a regression dataset. Several faults were injected in the output variable and therefore the capabilities of fault detection and replacement of the proposed methodology could be evaluated. The results show that all faults were detected and the replaced measurements were closer to the original measurements without the injected faults, when compared to the faulty measurements.

The results obtained in the fault detection and sensor replacement of the temperature sensor in the burning area of the cement kiln shows that the proposed methodology was successfully used. Several measurements, mainly where the estimated temperatures are higher that the measured temperatures, were identified as faults and therefore were replaced by the unscaled model output. Using the fault detection methodology the operator can be quickly warned of the need to clean the pyrometer sensor. This fast clean of the sensor and the replacement capability of the proposed methodology can improve the control of the temperature of the rotary cement kiln.

Acknowledgment

This work was supported by Project FAir-Control "Factory Air Pollution Control" (reference: E!6498), supported by the Eurostars Programme of the EUREKA network, financed by "Fundação para a Ciência e a Tecnologia" (FCT), do Ministério da Educação e Ciência, "Agência de Inovação" (AdI), and the Seventh Framework Programme for Research and Technological Development (FP7) of the European Union.

Francisco Souza has been supported by "Fundação para a Ciência e a Tecnologia" (FCT) under grant SFRH/BD/63454/2009.

References

[1] H. Alwi, C. Edwards, and C. Tan. Fault tolerant control and fault detection and isolation. In *Fault Detection and Fault-Tolerant Control Using Sliding Modes*, Advances in Industrial Control, pages 7–27. Springer London, 2011.

[2] A. Correcher, F. Morant, E. Garcia, R. Blasco-Gimenez, and E. Quiles. Failure diagnosis of a cement kiln using expert systems. In *IEEE 2002 28th Annual Conference of the Industrial Electronics Society (IECON 2002)*, volume 3, pages 1881–1886, 2002.

[3] X. fa Hu, Z. Zhao, S. Wang, F. li Wang, D. kuo He, and S. kang Wu. Multi-stage extreme learning machine for fault diagnosis on hydraulic tube tester. *Neural Comput. Appl.*, 17(4):399–403, June 2008.

[4] W. B. C. for Sustainable Development. *CSI - Guidelines for Emissions Monitoring and Reporting in the Cement Industry*, March 2012.

[5] L. Fortuna, S. Graziani, A. Rizzo, and M. G. Xibilia. *Soft Sensors for Monitoring and Control of Industrial Processes*. Springer, 2007.

[6] A. Frank and A. Asuncion. UCI machine learning repository, 2010.

[7] F. R. Hampel. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):383–393, 1974.

[8] G.-B. Huang, L. Chen, and C.-K. Siew. Universal approximation using incremental constructive feedforward networks with random hidden nodes. *IEEE Transactions on Neural Networks*, 17(4):879–892, July 2006.

[9] G.-B. Huang, N.-Y. Liang, H.-J. Rong, P. Saratchandran, and N. Sundararajan. On-line sequential extreme learning machine. In M. H. Hamza, editor, *IASTED International Conference on Computational Intelligence*, pages 232–237. IASTED/ACTA Press, July 2005.

[10] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew. Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1–3):489 – 501, 2006.

[11] I. Hwang, S. Kim, Y. Kim, and C. Seah. A survey of fault detection, isolation, and reconfiguration methods. *Control Systems Technology, IEEE Transactions on*, 18(3):636–653, May 2010.

[12] P. Kadlec, R. Grbić, and B. Gabrys. Review of adaptation mechanisms for data-driven soft sensors. *Computers & Chemical Engineering*, 35(1):1 – 24, 2011.

[13] O. Kadri, L. Mouss, and M. D. Mouss. Fault diagnosis of rotary kiln using svm and binary aco. *Journal of Mechanical Science and Technology*, 26(2):601–608, 2012.

[14] C. Laurentys, G. Ronacher, R. Palhares, and W. Caminhas. Design of an artificial immune system for fault detection: A negative selection approach. *Expert Systems with Applications*, 37(7):5507 – 5513, 2010.

[15] H. Liu, S. Shah, and W. Jiang. On-line outlier detection and data cleaning. *Computers & Chemical Engineering*, 28(9):1635–1647, 2004.

[16] L. Ljung, editor. *System Identification: Theory for the User*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 2nd edition, 1999.

[17] I. Makaremi, A. Fatehi, B. N. Araabi, M. Azizi, and A. Cheloeian. Abnormal condition detection in a cement rotary kiln with system identification methods. *Journal of Process Control*, 19(9):1538–1545, 2009.

[18] A. A.-R. M.M. Shoaib, M.M. Balahab. Influence of cement kiln dust substitution on the mechanical properties of concrete. *Cement and Concrete Research*, 30:371–377, 2000.

[19] J. Newman and B. Choo. *Advanced Concrete Technology–part 1–Constituent Materials*. Butterworth-Heinemann, 2003.

[20] E. Romero and J. M. Sopena. Performing feature selection with multilayer perceptrons. *IEEE Transactions on Neural Networks*, 19(3):431–441, March 2008.

[21] M. Sadeghian and A. Fatehi. Identification of nonlinear predictor and simulator models of a cement rotary kiln by locally linear neuro-fuzzy technique. *World Academy of Science, Engineering and Technology*, 58:1121–1127, 2009.

[22] M. Sadeghian and A. Fatehi. Identification, prediction and detection of the process fault in a cement rotary kiln by locally linear neuro-fuzzy technique. *Journal of Process Control*, 21(2):302–308, 2011.

[23] D. M. J. Tax. *One-class classification: Concept learning in the absence of counter-examples*. PhD thesis, Delft University of Technology, 2001.

[24] L. Wu and D. W. C. Ho. Fuzzy filter design for itô stochastic systems with application to sensor fault detection. *IEEE Transactions on Fuzzy Systems*, 17(1):233–242, Feb. 2009.

[25] G. P. Zhang. Avoiding pitfalls in neural network research. *IEEE Transactions on Systems, Man, and Cybernetic - Part C: Applications and Reviews*, 37(1):3–16, January 2007.