Fault Detection of an Industrial Heat-Exchanger: A Model-Based Approach

Dejan Dragan*
University of Maribor, Faculty of Logistics, Slovenia

One of the key issues in modelling for fault detection is how to accommodate the level of detail of the model description to suit the diagnostic requirements. The paper addresses a two-stage modelling concept to an industrial heat exchanger, which is located in a tyre factory. Modelling relies on both, prior knowledge and recorded data. During the identification procedure, the estimates of continuous model parameters are calculated by the least squares method and the state variable filters (SVF). It is shown that the estimates are largely invariant of the bandwidth of the SVFs. This greatly reduces the overall modelling effort and makes the whole concept applicable even for less experienced users. The main issues of the modelling procedure are emphasized. Based on the process model, a simple detection system is derived. An excerpt of the results obtained on operating records is given.

©2010 Journal of Mechanical Engineering. All rights reserved.

Keywords: industrial heat exchanger, fault detection, condition monitoring, model-based detection, modelling, identification

0 INTRODUCTION

Model-based condition monitoring of industrial processes aims at early revelation of degradations in process equipment and instrumentation. A sensible process model acts as an additional virtual instrument, which contributes to a higher quality of production and better safety. There are many papers and books dealing with model-based techniques for detecting, isolating and identifying faults [1] to [6]. However, in many real applications deriving a proper model still takes a bulk of overall design effort. Moreover, proper shaping of the model precision with respect to the diagnostic requirements remains to be rather an art.

This paper deals with the design of a fault detector for a heat exchanger as a possible alternative to some other approaches [7] to [11]. The work represents part of the prototyping design of a condition monitoring system for the process of incineration of vulcanisation gasses located in a tyre factory [6].

A brief idea of the process can be grasped from Fig. 1. Vulcanisation gas (VU gas), which is one of hand products of vulcanisation, is generated on vulcanisation lines. Prior to the emission into the atmosphere, all the carbon particles contained in the gas need to be destroyed by incineration in a combustion chamber.

The entire system consists of three major parts [6]:
• pre-heating of the VU gas in a gas-gas heat exchanger;
• incineration of the VU gas in a combustion chamber;
• transportation of the flue gasses to the chimney.

The focus of this paper is on model-based condition monitoring of the cold part of heat exchanger. The goal is to improve the support to the maintenance team through permanent monitoring of the condition of sensors and detection of fouling. The project aims at designing a detector with the highest precision possible. For that purpose, the extra redundancy is achieved by means of an analytical model of the plant.

The modelling concept consists of two stages in which prior knowledge and recorded data are combined (grey-box modelling concept). In the first stage, the model structure is derived up to unknown parameters by strongly relying on reasoning from first principles. After taking the available instrumentation into consideration, the set of prior assumptions and the diagnostic requirements the modelling procedure ended up with a continuous-time model linear in parameters.

For the purpose of parameter estimation, the least squares method (LSM) combined with state variable filters (SVF) is adopted in the second
stage. It is shown that the use of LSM leads to the identification results, which are largely invariant of the bandwidth of the SVFs. This observation deserves attention since the choice of bandwidth, as a design parameter, turns out to be quite an easy task. This is believed to be one of the contributions of the paper.

Fig. 1. The incineration system

The derived model of the cold part of an exchanger employs the temperatures \( T_{co} \), \( T_{ho} \) and flow of the VU gas \( \Phi_{VU} \). The measured signals (c.f. Fig. 1) are collected, displayed and stored by the FactoryLink\textsuperscript{TM} SCADA system (USDATA Corp.).

Finally, the model-based diagnostic algorithm is designed to run on-line as an external C module of SCADA.

The proposed design of the condition monitoring system is very simple and the suggested detector is a helpful indicator for the operator to take corrective action.

The derivation of the model structure and the parameter estimation approach is described in the first section. An excerpt of the experimental results is given in the second section. Finally, the diagnostic procedure is overviewed in the third section.

1 SYSTEM IDENTIFICATION

1.1 Determination of Model Structure

The purpose of this section is to emphasize the importance of prior knowledge in deriving the model structure. This knowledge is essential in early modelling steps, in particular in defining causal relationships between process variables. Depending on the diagnostic requirements, these relationships can range from loose (qualitative) to precise (quantitative) descriptions. In case of poor prior knowledge and/or loose diagnostic requirements, model precision can be restricted to qualitative relationships defined on sign, interval or fuzzy sets. Perfect prior knowledge allows for precise expressions that are fully defined on the set of real numbers. Since in many practical cases prior knowledge is incomplete, process data that represent the carrier of additional information needed to complete the model description need to be employed. Here, the finest level of detailed description is observed.

The process of the heat exchange between (cool) VU gas and (hot) flue-gas is illustrated in Fig. 2.

The modelling procedure starts with setting the energy balance equation for an infinitesimally small piece of the VU gas channel and pipe wall. This results in the following partial differential equations respectively [1]:

\[
c_{VU} \frac{\partial T_C(x,t)}{\partial t} + c_{VU} \cdot \Phi_{VU}(t) \cdot \frac{\partial T_C(x,t)}{\partial x} = 0
\]

\[
\frac{\partial T_H(x,t)}{\partial t} = \frac{o_2}{\rho_{VU} \cdot A_{cl}} \cdot \alpha_2(t) \cdot \left[ T_W(x,t) - T_H(x,t) \right]
\]

where \( T_C(x,t), T_H(x,t) \) and \( T_W(x,t) \) represent space-time behaviour of the temperatures of the cold part of the exchanger, hot part and pipe wall, respectively, while \( o_1 = 2\pi r_1, o_2 = 2\pi r_2 \).

In order to identify the relationship between the measured variables \( T_{co}, T_{ho}, \Phi_{VU} \), the distributed parameter models in Eqs. (1) and (2) need to be converted into the lumped models.

The lumping procedure is based on the following set of assumptions:
1. specific mass \( \rho_{VU} \) and specific heat \( c_{VU} \) of the VU gas are assumed to be constant,
2. flow of the VU gas is space independent \( \Phi_{VU}(x,t) = \Phi_{VU}(t) \),
convective heat exchange coefficients are independent of time and space \((\alpha_1(x,t) = \alpha_1, \alpha_2(x,t) = \alpha_2)\), as the wall is only 4 mm thin \(A_w \approx 0\) and \(\partial T_w(x,t) / \partial t \approx 0\) can be assumed.

5. The retention time of vulcanisation gasses in the heat exchanger is fairly low, as the speed of VU gasses is relatively high while the length of the heat exchange channel is quite short. This implies an almost momentary formation of temperature profiles along the heat exchange channel.

Assumption 4 implies short time constant for the dynamics of the wall temperature \(T_w(x,t)\) in Eq. (2). Transients can be neglected so that the following static relation emerges:

\[
T(x,t) = \frac{\alpha_1 \cdot \alpha_2}{\alpha_1 + \alpha_2} \cdot T_c(x,t).
\]

In the next step let us combine Eqs. (3) and (1) at \(x = L\), which results in:

\[
\frac{\partial T_c(x,t)}{\partial t} \bigg|_{x=L} = \frac{1}{A_{c1} \cdot \Phi_{vu}(t)} \cdot \frac{\partial T_c(x,t)}{\partial x} \bigg|_{x=L} = \frac{\alpha_2 \cdot m}{b_1} \cdot T_c(L,t) - \frac{\rho \cdot c_v}{\alpha_1} \cdot \frac{T_v(L,t) - T_c(L,t)}{\Phi_{vu}(t)}.
\]

Then, according to assumption 5, the temperature profiles along the heat exchange channels exhibit almost static behaviour. From the steady-state condition at \(x = L\), Eq. (4) results in the following expression:

\[
\frac{\partial T_v(x)}{\partial x} \bigg|_{x=L} = \frac{\alpha_2 \cdot m_1}{b_1} \cdot \left( T_v(L,t) - T_c(L,t) \right) - \frac{\rho \cdot c_v}{\alpha_1} \cdot \left( T_v(L,t) - T_c(L,t) \right).
\]

where bars over symbols denote stationary values.

However, a number of identification runs, carried out on model structures based on Eq. (5), turned to produce one-step-ahead predictor with relatively poor performance. Careful examination of the unmodelled effects related to the derivation of Eq. (5), suggests the approximation of the gradient \(\partial T_c(x,t) / \partial x\) at \(x = L\) by a richer structure in order to improve the predictive power of the final process model. It can be shown that the gradient is related with the measured input and output temperatures and flow in a very complicated manner. Instead of an exact solution a black-box structure is sought so that experimental data are fitted as well as possible while keeping the number of unknown parameters at a minimum. Among many candidates (polynomials, neural networks) it turned out that already a simple structure with only two free parameters can significantly raise the quality of fit of the model Eq. (4). The approximation reads very similarly to Eq. (5) i.e.

\[
\frac{\partial T_c(x,t)}{\partial x} \bigg|_{x=L} \approx \epsilon_1 \cdot T_{ho}(t) - \epsilon_2 \cdot T_{co}(t).
\]

If Eq. (6) is entered to Eq. (4), the following expression is obtained:

\[
\dot{T}_{co}(t) + b_1 \cdot \Phi_{vu}(t) \left( \epsilon_1 \cdot T_{ho}(t) - \epsilon_2 \cdot T_{co}(t) \right) = a_1 \cdot \left( T_{ho}(t) - T_{co}(t) \right),
\]

i.e.
\[
\dot{T}_{co}(t) + \frac{b_1}{b_2} \cdot \varepsilon_2 \cdot \left[ \Phi_{Fy}(t) \cdot T_{ho}(t) \right] + \frac{b_1}{b_2} \cdot \varepsilon_2 \cdot \left[ -\Phi_{Fy}(t) \cdot T_{co}(t) \right] = a_1 \cdot \left[ T_{ho}(t) - T_{co}(t) \right].
\]

Eq. (8) represents the lumped model of the cold part of the heat exchanger. The same modelling procedure can be applied to derive a model of the hot part. Unfortunately, the task is not that easy due to a lack of a flow sensor in the flue-gas channel (hot part). Nevertheless, a careful analysis indicates that the unmeasurable flow mostly depends on flow of the vulcanisation gasses, roughly in a linear way. With this in mind a model of the hot part being entirely similar to that of the cold part is achieved. Due to a lack of space the hot part will not be treated here. In the sequel emphasis is on the cold part only.

1.2 Parameter Estimation

In order to identify the model parameters batch the least squares method is chosen.

To estimate the parameters of the continuous time model in Eq. (8), the derivative of \( T_{co} \) is needed. As direct differentiation is prone to significant errors due to measurement noise, the problem is alleviated by using state variable filters. In this case, any stable filter with relative degree \( \geq 1 \) and appropriate bandwidth would suit. A simple transfer function is employed to filter the signals in Eq. (8):

\[
G_f(S) = \frac{1}{\tau \cdot S + 1},
\]

where \( 1/\tau \) is the bandwidth of filter (\( \tau \) is the time constant). If Eq. (9) is applied to both sides of Eq. (8), filtering preserves the original model structure.

To improve the numerical properties of the algorithm, the filtered signals from Eq. (8) are transformed in the next step as for example:

\[
\Delta T_{co_{mf}}(t) = T_{co_{mf}}(t) - T_{co_{mf}}(t-1),
\]

where index “mf” refers to the value of the filtered signal. Differentiated data preserve the original model structure while eliminating the problem of offsets in prediction error.

After filtering all the signals in Eq. (8) and after transformation of filtered data, an appropriate form for identification procedure is obtained. For the data collected by the SCADA system, an over-determined system of equations, which can be represented in the following form, is obtained:

\[
\begin{align*}
Y &= \Psi \cdot \theta + \Delta n_{f1}, \\
\end{align*}
\]

where:

\[
\Psi = 
\begin{bmatrix}
\Delta T_{co_{mf}}(1) - \Delta T_{ho_{mf}}(1) \\
\vdots \\
\Delta T_{co_{mf}}(N) - \Delta T_{ho_{mf}}(N)
\end{bmatrix},
\]

\[
\begin{bmatrix}
-\Delta \dot{T}_{co_{mf}}(1) & \Delta F_{1_{mf}}(1) & \Delta F_{2_{mf}}(1) \\
\vdots & \vdots & \vdots \\
-\Delta \dot{T}_{co_{mf}}(N) & \Delta F_{1_{mf}}(N) & \Delta F_{2_{mf}}(N)
\end{bmatrix},
\]

\[
\theta = \begin{bmatrix}
1 \\
\frac{b_1}{a_1} \\
\frac{b_2}{a_1} \\
\frac{K}{a_1}
\end{bmatrix},
\]

\[
\Delta n_{f1} = 
\begin{bmatrix}
\Delta n_{f1}(1) \\
\vdots \\
\Delta n_{f1}(N)
\end{bmatrix}.
\]

The unknown constant \( K \) is added in order to take into account the bias in the prediction error. In case of perfect model structure (i.e. free of modelling errors) the estimated \( K \) should be zero. Vector \( \Delta n_{f1} \) is added to encounter the noise effects.

The unknown parameters of the system Eq. (10) result as follows [12] and [13]:

\[
\hat{\theta} = \left( \Psi^T \cdot \Psi \right)^{-1} \cdot \Psi^T \cdot Y.
\]

2 PRACTICAL IDENTIFICATION RESULTS

Process identification was carried out on a batch of 14000 samples taken during normal operating regime (interval [1, 14000] [min]). Outliers and intervals with non-informative data are carefully eliminated from further processing.

The results of the estimation achieved at various bandwidths of SVF’s (1/\( \tau \)) are shown in Fig. 3. It can be seen that the results of estimation do not vary significantly with respect to SVF bandwidth changing over two decades. This statement includes the variations of parameter \( \tilde{K}(\tau)/\tilde{a}_1(\tau) \), which are small compared to the elements of vector \( Y \) (Eq. (11)) and can be
neglected. Identification results largely invariant to pre-filtering imply great freedom in choosing the SVF bandwidth. This is an advantage for every practitioner.

The purpose of the designed model is to predict the difference between temperatures of cold and hot part at the end of exchanger.

Thus, by taking into account Eq. (8) process output can be represented as follows:

\[ y_f(t) = T_{co,mf}(t) - T_{ho,mf}(t), \quad (13) \]

and the predicted output as:

\[ \hat{y}_f(t) = \left( \frac{1}{a_1} \right) \cdot \left( \hat{T}_{co,mf}(t) + \hat{b}_1 \right) + \left( \frac{b_2}{a_1} \right) \cdot F_{1,mf}(t) + \left( \frac{b_2}{a_1} \right) \cdot F_{2,mf}(t) + \hat{K}_{y_f}, \quad (14) \]

respectively. Similarly to Eq. (11), estimated constant \( \hat{K}_{y_f} \) is added to avoid bias in modelling error.

A comparison between process output Eq. (13) and predicted output Eq. (14) on validation set containing 5000 samples (interval \([14000, 19000]\) [min]) is shown in Fig. 4. The model fits the process reasonably well so that the prediction error does not exceed 10% of the dynamic range of the signal Eq. (13). In other words, the underlying mathematical model can be viewed as an additional (virtual) instrument, which obviously brings extra redundancy into the system.

3 DIAGNOSTIC RESULTS

Inference about faults is made on the basis of the residual signal defined as follows:

![Fig. 3. The estimated parameters in dependence of different bandwidths (1/τ) of the SVF](image)

![Fig. 4. Model validation: measured output, predicted output and prediction error; a) y_f(t); b) prediction error](image)
\[ r(t) = y_f(t) - \hat{y}_f(t), \quad (15) \]

where \( y_f(t) \) and \( \hat{y}_f(t) \) are process output Eq. (13) and predicted output Eq. (14), respectively.

Based on the residual Eq. (15) the presence of a fault can be detected. This means that if the residual is near zero, there is no evidence that a fault is present. On the contrary, if the residual departs significantly far from zero, the presence of a fault can be inferred. Generally, purely on the basis of the residual Eq. (15) it is not possible to determine the location of the fault. The exception is discussed in the literature [3] and [5]. Indeed, provided the parameters \( \theta \) of the process model bear physical meaning, the regressor form of the model Eq. (10) reads:

\[ Y = \Psi \cdot \theta + \Delta n_{f1} \quad \text{and} \quad \hat{Y} = \Psi \cdot \hat{\theta}, \]

for the true process and the mathematical model. Symbols \( Y \) and \( \hat{Y} \) denote true and predicted outputs, while \( \theta \) and \( \hat{\theta} \) denote true and estimated process parameters, respectively. By differentiating the two equations the following is obtained:

\[ e = Y - \hat{Y} = \Psi \cdot \theta + \Delta n_{f1} - \Psi \cdot \hat{\theta} = \]
\[ = \Psi \cdot (\theta - \hat{\theta}) + \Delta n_{f1} = \Psi \cdot \Delta \theta + \Delta n_{f1}, \]

where \( e \) denotes the vector of residuals. Since \( e \) results in a straightforward manner from the recorded data and the nominal process model, the estimate of the vector \( \Delta \theta \) can easily be calculated. Any non-zero term in \( \Delta \theta \) indicates a variation in the process parameters due to the fault. Knowing the physical origin of such a term in the vector \( \theta \) the position of the fault can be inferred. An important technical requirement is that the data matrix \( \Psi \) is full rank (persistent excitation).

However, the idea is not applicable in the present case for three reasons:

1. The temperature signals \( T_{ho} \) and \( T_{co} \) are generally too poor from the point of view of information content, which means that with recursive parameter estimation it would not be possible to unambiguously estimate all the parameters in Eq. (14);
2. Signal to noise ratio in certain intervals of process operation can be rather low.
3. Not all the parameters in Eq. (14) reflect the physical properties of the system (the model Eq. (8) is semi-physical); consequently such parameteres would be of little use in fault localisation.

Having residual Eq. (15) it is necessary to draw the decision about alarm. If pure Boolean logic were applied, then frequent, even small deviations in residuals in the vicinity of the threshold value would lead to large variations in the diagnostic results [14] and [15]. In order to smooth the diagnostic output, approximate reasoning techniques seem to be a better alternative. A residual is no longer qualified as zero (0) or non-zero (1) but is associated a degree of being zero, which is a number between 0 and 1. In this way, incremental changes in the residual Eq. (15) result in incremental changes in the diagnostic results.

For the sake of detection the belief mass (Fig. 5) is introduced, which can be represented in the following form [14]:

\[ bel(r) = \frac{1}{1 + \frac{1 - \delta_d}{\delta_d} (\frac{h}{r})^{2 \gamma_d}}, \quad (16) \]

\( h \) and \( \gamma_d \) are threshold and the adjustable smoothing parameter respectively, while \( \frac{\gamma_d}{\delta_d} = bel(r=h) \).

Fig. 6 shows the response of the detector, when two consecutive faults are injected into the system. Firstly, the temperature sensor related to \( T_{ho} \) is stuck high. This fault is emulated by fixing the sensor output to 560 °C for the period of 700 minutes. The next fault is emulated drift in the sensor related to \( T_{co} \). Obviously, the detector is quite sensitive to the occurrence of both faults.

Since belief in a non-zero residual also significantly increases, the detector provides clear evidence of the presence of both faults (c.f. Fig. 6). Though obvious, the first fault remains undetected.
by the classical alarm system. Indeed, the alarm is set at 600 °C as lower values can be easily reached during normal operating conditions (not presented in Fig. 6). On the contrary, the model-based detector reacts quite quickly and accurately. This detector reacts relatively promptly in the case of a drift type of fault as well. For example, it reports the presence of fault at approx. \( t = 4100 \) min with 70% belief. The classical alarm system would react, but much later (at approx. 5200 min).

To sum up, by running the suggested detector on the recorded data sets it has been possible to reveal several temporary erroneous sensor readings during operation of the real plant. These were completely overlooked by the existing alarm system (missed alarms).

![Graph showing classical alarm vs. model-based detection](image)

**Fig. 6. Detection of sticking high of sensor \( T_{ho} \) and detection of drift in sensor \( T_{co} \)**

### 4 CONCLUSIONS

A two-stage modelling procedure is presented and applied to a heat exchanger in an incineration unit. It relies on blending prior knowledge with the information contained in data records. Prior knowledge serves to derive the model structure, while data are used to identify unknown model parameters.

The paper makes a significant contribution in two ways. Firstly, a set of prior (heuristic) assumptions provides a means for determining the model structure. Secondly, it is shown that the results of estimation, obtained using the least squares method, are largely invariant of the bandwidth of the SVF. This greatly reduces the overall modelling effort and is an advantage for every practitioner.

The suggested detection algorithm is very simple for execution in real time. The diagnostic results show that the module is able to accurately indicate the presence of incipient faults and thus facilitate timely on-condition maintenance. Traditional alarm systems based on thresholding are insensitive to such faults, i.e. do not react until large deviations and failures occur.

### 5 REFERENCES


