18th IMEKO TC10 Conference
"Measurement for Diagnostics, Optimisation and Control to Support Sustainability and Resilience"
Warsaw, Poland, September 26–27, 2022

Adaptive Methods for Fault Detection on Research Engine Test Beds

Doris Schadler[1], Michael Wohlthan[2], Andreas Wimmer[3]

[1]LEC GmbH, Graz, Austria, doris.schadler@lec.tugraz.at, +43 316 873 30136
[2]LEC GmbH, Graz, Austria, michael.wohlthan@lec.tugraz.at
[3]Graz University of Technology, Graz, Austria

Abstract **– In the case of test beds for research engines, fault detection methods that use models based on historical data face a particular challenge. Due to the experimental design of the test bed, offline training of statistical models with a data set containing all possible variations is simply not possible. The methods must adapt to the current data situation directly on-site. But this involves risks. First, computational time and memory requirements can become extremely large with high data volumes. Second, the data may be faulty and thus negatively affecting the models. To avoid both, a selection of data is made before it is used to build the fault-free reference model. For this purpose, a new statistic is presented as the combination of the Mahalanobis distance and the forecast residual. With it, it is possible to reduce the update frequency and to increase the rate of detected faulty points, since the models are no longer manipulated by faulty data points and thus the residuals provide a better structure for fault detection.**

Keywords – Fault detection; engine test beds; adaptive methods; data selection; Industry, Innovation and Infrastructure.

I. INTRODUCTION

Engine test beds are an essential tool in the experimental investigation of combustion concepts and in engine development. The complexity of these test beds, combined with the large number of sensors and data, makes an automated method of monitoring essential. As with many other research topics, various approaches to fault detection already exist. For model-based fault detection, physical models such as energy balances and mass balances can be used. The disadvantage is that such models require a lot of expert knowledge and it is often not possible to monitor all variables with this limited number of models. Therefore, models based on a historical data set provide a useful complement.

But even for such statistical models, there exists a critical challenge, at least on research engine test beds. In particular, the engine and its configuration change frequently. This means that it is not possible to collect data points from all possible data situations for an offline modelling phase. Instead, the models must learn online directly on the test bed using the current data. Unfortunately, continuously adapting the models to each new data point can become very time-consuming and memory-intensive. Additionally, there is also the risk of adding faulty data to the data set used for model building. This can lead to manipulated models that subsequently make it impossible to detect further faults. This means that it is not only the models themselves that are important for fault detection, but also the quality and quantity of the data used for fitting them.

II. RELATED RESULTS IN THE LITERATURE

The modeling of engines is closely related to the modeling of engine test benches, as various engines, engine settings and operating points are analyzed there, for example. Fault detection has been investigated in both cases. [1] uses, for example, process models of the subsystems of a Diesel engine such as intake, injection/combustion and exhaust systems for fault detection. In [2], physical and statistical methods are combined by using a combination of knowledge-based residual generation and statistical residual evaluation for automotive Diesel engines. Online monitoring can also be applied to electric motors, as described in [3]. Engine test beds are well explained and a fault detection approach using historical process data is shown in [4]. [5] describes the specific physical relationships applied for fault detection at engine test beds in detail and also explains the important step of fault isolation.

Thus, fault detection on test beds is a well-known topic, and also adaptive methods have already been investigated by several authors. [6] gives a good overview of various approaches. For example, the process data can be sorted by operating mode and different models can be used, as in [7]. In just-in-time learning technique, the computation and relationship building is done online using similar data in the database as in [8]. Recursive weighting can also be used to make well-known fault detection methods such as principal component analysis adaptive, as in [9]. Further, it is also possible to update the system with information about faults [10]. The selection of data before updating the model represents a special approach for such an adaptive procedure. The necessity of such a preliminary analysis is also shown in [11].

III. DESCRIPTION OF THE METHOD

A. Model building and fault detection

Assume that we have already observed n measurements of $(p+1)$ variables. One variable is the monitored one and thus the response variable of the model, $\boldsymbol{y}$. The remaining variables are used as predictors and are summarized to the matrix of predictors $\boldsymbol{X}$. With the given data set, the unknown parameters of the model can be estimated. Complex methods such as machine learning approaches can be used for modelling as in [12]. Since we are dealing with small data sets and strong linear correlations between the variables, the models are built using multiple linear regression. Previous studies have shown that despite the simplicity of these models, their quality was still sufficiently high to be suitable for fault diagnosis on engine test beds. In multiple linear regression, we simply need to estimate the parameter vector $\boldsymbol{\beta}$ of the model represented in Equation (1):

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\boldsymbol{\varepsilon}$ is the model error term.

With this model, the new data point given by the value of the response variable y_+ and the corresponding values of the predictor variables $\boldsymbol{x}_+$ can be evaluated. This is done by analyzing the forecast residual given in Equation (2):

$$e_+ = y_+ - \boldsymbol{x}_+\widehat{\boldsymbol{\beta}}. \tag{2}$$

This forecast residual should be close to zero in the fault-free case. An error alarm is raised if a sequence of residuals falls outside of a defined control region.

B. Data selection

For data selection, two measures are to be used. To assess whether the new data point is in a new region of the data space, a distance measure is used which is one factor of the statistic. Various distance measures were investigated in this regard. In order to take into account the different scaling of the variables, but also the covariance between them, the Mahalanobis distance from equation (3) is used for data selection:

$$M^2 = (\boldsymbol{x}_+ - \overline{\boldsymbol{x}})^T \boldsymbol{S}^{-1} (\boldsymbol{x}_+ - \overline{\boldsymbol{x}}), \tag{3}$$

where $\bar{x}$ is the sample mean vector and S is the sample covariance matrix.

Furthermore, the value of the response variable is considered by analyzing the standardized forecast residual as shown in equation (4):

$$r = \frac{y_+ - x_+\widehat{\boldsymbol{\beta}}}{\hat{\sigma}\sqrt{1 + x_+ (X^T X)^{-1} x_+^T}}, \tag{4}$$

where is $\hat{\sigma}$ is the estimated model standard deviation.

The statistic used for data selection is a combination of these two measures, given in equation (5):

$$T = \frac{\widetilde{M}^2}{r^2}, \tag{5}$$

where $\widetilde{M}^2 = \frac{n(n-p)}{p(n+1)(n-1)}$ is the standardized version of the Mahalanobis distance.

This statistic is large when either the distance is large or the residual is small. Both conditions indicate that the model should be updated. If the distance is small, the data point is close to those already observed. Since this indicates only a small amount of new information, the point is excluded. As with fault detection, large residuals are an indicator of a fault in the data, at least if the distance is relatively small and the model quality is sufficiently large. Therefore, the corresponding data points are also excluded from the modeling. In summary, the ratio of the two measures must be appropriate for the point to be added to the data set.

The criterion of data selection requires the determination of a threshold value. This threshold can be determined with theoretical considerations. It is shown in the literature that both statistics follow an F-distribution under certain assumptions. Furthermore, similar to the analysis of Cook's distance, one can show the independence of these two factors. Combining both results, it follows that the new statistic also follows an F-distribution, at least approximately. Thus, the criterion for the selection of the data can be set up as given in equation (6). Accordingly, a point is excluded from the modeling if the corresponding statistic is sufficiently small:

$$T < q_\alpha, \tag{6}$$

where $q_\alpha$is theoretical α-quantile of the $F_{p,1}$ -distribution. In this way, α percent of all data points are excluded from the modeling.

IV. RESULTS AND DISCUSSIONS

The data originate from a single-cylinder research engine test bed. The schematic of the test bed is shown in Figure 1. There, the air-gas exhaust path is illustrated and the coolant pipe is shown. Various engines, but also combustion concepts, are investigated on such test beds.

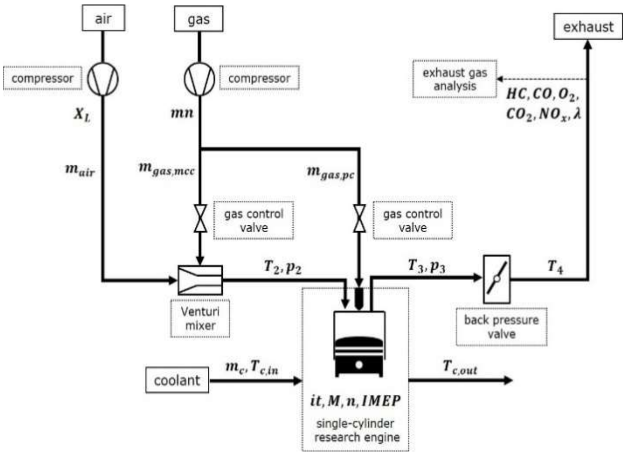

Fig. 1. Schematic of a single-cylinder research engine test bed

Process faults occur frequently, but sensor faults are also an issue. The fault detection scheme shown here aims to detect sensor faults of two different types, abrupt faults and drifts.

To prove the advantage of data selection as a step before model building, we compare the results of the method with continuous updating, i.e., without any data selection, with the method in which data points are selected and thus updating is performed only when the data selection criterion is not fulfilled. For this comparison, we used four data sets, each containing 500 measurements from the research engine test bed. Each data set contains observations of the 23 variables. Only one of these variables was used in the analysis as a response and thus monitoring variable. The data selection criterion parameter was defined as α=0.1.

We compare the methods in two specific settings. The first is the fault-free setup. The goal is that despite the fact that less data is used for model fitting, the model quality does not decrease. The performance of the model is assessed using two criteria: the mean absolute percentage error and the mean percentage error. The results for the fault-free case are shown in Table 1. We find that the model performance is not negatively affected by the data selection procedure and both measures remain almost at the same level.

Table 1. Evaluation measures for the fault-free case.

	No selection	Selection
Fraction of points excluded	0.000	0.127
Mean absolute percentage error	0.345	0.369
Mean percentage error	-0.044	-0.054

In the faulty case, we analyzed different fault scenarios. In such a scenario, a fault of a certain type and with a defined fault intensity was simulated on the data of the response variable beginning at a certain onset and lasting over 100 data points. The fault intensity f describes the relative amount of the fault compared to the measured value:

$$y_f = y + f \cdot y. \tag{7}$$

Faults with a low intensity are the most difficult to detect. Therefore, we analyzed intensities in the range between 0.01 and 0.1.

For the analysis, we count the number of faulty data points that were actually detected as such and calculate the true-positive rate (TPR). For each fault type and intensity, multiple scenarios at different onsets and data sets were examined and summed to produce one result. For abrupt faults, the results for the two different data selection criteria are shown in Figure 2. The results for drifts are shown in Figure 3. In both figures, the increase in the true-positive rate due to the use of the data selection criterion can be clearly observed. The greater the fault intensity, the larger is the improvement. In the best case, 50% more faulty data points were actually detected as such during data selection than during continuous updating.

In summary, although less data is used for model fitting (about 10% is excluded from the data set), neither the performance in the fault-free case nor the rate of true positives decreases.

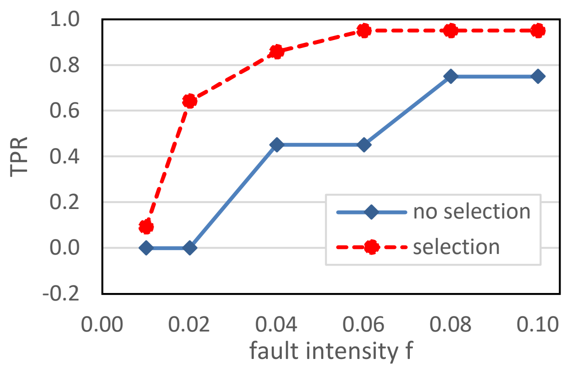

Fig. 2. True positive rate for abrupt faults

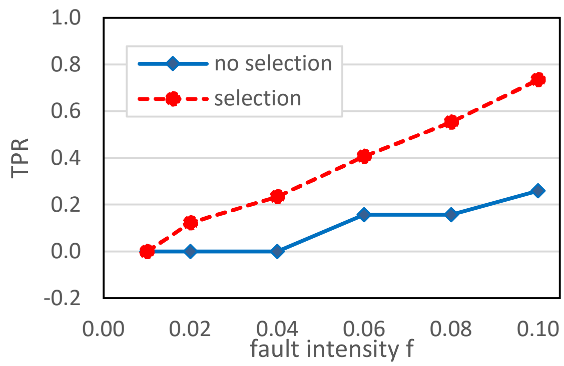

Fig. 3. True positive rate for drifts

V. CONCLUSIONS AND OUTLOOK

Online adaptation of models used for fault detection is inevitable on research engine test beds. The selection of data with the reformulated statistic as a criterion leads to a significant increase in the performance of the fault detection method. Furthermore, fewer updates of the model are required. The result is an adaptive method that has more stability due to the fact that only selected data are used for model building.

Nevertheless, linear regression models are the simplest way to describe the relationship between variables. Non-linear dependencies and dynamic processes are more common in practice. The models and the distance measure must account for both cases. Consequently, the new statistic must be defined for a broader range of data situations and thus models. A key assumption of the shown methodology concerns the distribution of the data. In practice, the Gaussian distribution is not the normal case. It must be analyzed how this fact affects the distribution of the new statistic and thus the data selection procedure. A new definition of the threshold for the criterion may become necessary.

VI. ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support of the "COMET - Competence Centers for Excellent Technologies Programme" of the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT), the Austrian Federal Ministry of Science, Research and Economy (BMWFW) and the Provinces of Styria, Tyrol and Vienna for the K1-Center LEC EvoLET. The COMET Programme is managed by the Austrian Research Promotion Agency (FFG).

REFERENCES

[1] **Kimmich, F.; Schwarte, A.; Isermann, R.:** Fault detection for modern Diesel engines using signal- and process model-based methods, in: *Control Engineering Practice* 13.2 (2005), pp. 189 –203.

[2] **Svärd, C.; Nyberg, M.; Frisk, E.; Krysander, M.:** Automotive engine FDI by application of an automated model-based and data-driven design methodology, in: *Control Engineering Practice* 21.4 (2013), pp. 455 –472.

[3] **Vass, J., Cristalli, C.:** Bearing fault detection for on-line quality control of electric motors, in *10th IMEKO TC10 International Conference on Technical Diagnostics*, (2005)

[4] **Haghani, A.; Jeinsch, T.; Roepke, M.s; Ding, S. X.; Weinhold, N.:** Data-driven monitoring and validation of experiments on automotive engine test beds, in: *Control Engineering Practice* 54 (2016), pp. 27 –33.

[5] **Wohlthan, M.; Schadler, D.; Pirker, G.; Wimmer, A.:** A multi-stage geometric approach for sensor fault isolation on engine test beds, in: *Measurement* 168 (2021), p. 108313.

[6] **Kadlec, P.; Grbic, R.; Gabrys, B.:** Review of adaptation mechanisms for data-driven soft sensors, in: *Computers & Chemical Engineering* 35.1 (2011), pp. 1–24.

[7] **Yong, G.; Xin, W.; Zhenlei, W.:** Fault detection for a class of industrial processes based on recursive multiple models, in: *Neurocomputing* 169 (2015), pp. 430–438.

[8] **Yu, H.; Yin, S.; Luo, H.:** Robust Just-in-time Learning Approach and Its Application on Fault Detection, in: *IFAC-PapersOnLine* 50.1 (2017), pp. 15277 –15282.

[9] **Portnoy, I.; Melendez, K.; Pinzon, H.; Sanjuan, M.:** An improved weighted recursive PCA algorithm for adaptive fault detection, in: *Control Engineering Practice* 50 (2016), pp. 69 –83.

[10] **Benndorf, G. A.; Wystrcil, D.; Réhault, N.:** A fault detection system based on two complementary methods and continuous updates, in: *IFAC-Papers-OnLine* 51.24 (2018), pp. 353 –358.

[11] **Frye, M., & Schmitt, R. H.:** Structured Data Preparation Pipeline for Machine Learning-Applications in Production, in: *17th IMEKO TC10 Conference* (2020), pp. 241-246.

[12] **Zhao, W., Egusquiza Estévez, E., Valero Ferrando, M. D. C., Egusquiza Montagut, M., Valentín Ruiz, D., & Presas Batlló, A.:** A novel condition monitoring methodology based on neural network of pump-turbines with extended operating range, in_ *16th IMEKO TC10 Conference*, (2019), pp. 154-159.