Robust Functional Regression for Outlier Detection

Harjit Hullait[1], David S. Leslie[1], Nicos G. Pavlidis[1], and Steve King[2]

[1] Lancaster University, Lancaster, UK
[2] Rolls Royce PLC, Derby, UK

Abstract. In this paper we propose an outlier detection algorithm for temperature sensor data from jet engine tests. Effective identification of outliers would enable engine problems to be examined and resolved efficiently. Outlier detection in this data is challenging because a human controller determines the speed of the engine during each manoeuvre. This introduces variability which can mask abnormal behaviour in the engine response. We therefore suggest modelling the dependency between speed and temperature in the process of identifying abnormalities. The engine temperature has a delayed response with respect to the the engine speed, which we will model using robust functional regression. We then apply functional depth with respect to the residuals to rank the samples and identify the outliers. The effectiveness of the outlier detection algorithm is shown in a simulation study. The algorithm is also applied to real engine data, and identifies samples that warrant further investigation.

Keywords: Robust Functional Data Analysis · Robust Model Selection · Outlier Detection.

Before a jet engine is delivered it must complete a Pass-Off test. In a Pass-Off test a controller performs manoeuvres, which can be defined as various engine accelerations and decelerations starting and ending at a set idle speed. The purpose of this test is to ensure the engine complies with set standards. During the test, data is captured by sensors measuring engine speed, pressure, temperature and vibration in different parts of the engine. This high-frequency measurement data offers the ability to automate the detection of engine problems. By building statistical models for the Pass-Off test data we can aid the engineers in identifying engine issues efficiently.

One of the key manoeuvres in a Pass-Off test is the Vibration Survey (VS). In this manoeuvre the engine is accelerated slowly to a certain speed then slowly decelerated. We have 199 VS datasets, which include the turbine pressure ratio (TPR) that measures the engine speed, and the turbine gas temperature (TGT) which is a key temperature feature. In Fig. 1 we have plots of the TPR and TGT for the 30 VS manoeuvres. We have transformed the time index to the interval $[0, 1]$ and the range of sensor measurements to $[0, 100]$.

Automated detection of abnormal engine behaviour has been studied before [14], [9]. Both approaches require a training set of "normal" samples to build a normality model. They then apply novelty detection using an appropriate distance measure and threshold. We will instead use Functional Data Analysis (FDA) methods to identify VS manoeuvres that display unusual temperature behaviour in response to the variable (human-controlled) TPR time series. We will robustly build a normality model without requiring a set of "normal" samples. FDA techniques have been used effectively to model sensor data [13], as they combine information across samples and exploit the underlying behavioural structure. However this is to the best of our knowledge the first time these techniques are being used for modelling jet engine data.

We will use robust Functional Linear Regression (FLR) to build a model of "normal" behaviour. We shall then use the residuals from this model to identify outlying behaviour. The residuals are time series therefore using metrics such as the mean-square error means we lose a lot of information. Instead we will apply functional depth [6], which is capable of identifying various types of outlier behaviour.

There are a number of functional outlier detection methods, including the threshold approach [8], the Functional Boxplot [23] and the Outliergram [2], which use functional depth [16] to rank the curves. Alternative approaches use Directional Outlyingness measures, such as MS-plots [7] and Functional Outlier Maps [20]. There are also approaches for multivariate functional data [10]. These methods do not model the dependency between the functional response and functional input, and may therefore miss important outliers. Robust FLR can model this dependency structure, which can improve the detection of outliers.

This paper is organised as follows. In Section 1 we summarise the FDA methods, which will be used in the outlier detection algorithm. In Section 2, we will develop robust FDA techniques to obtain a robust regression model. In Section 3, we show how the robust regression model can be used to identify outliers. In Section 4 we give simulation results comparing the robust model with a classical model. Finally in Section 5 we apply the robust model on the engine data and highlight the outliers identified.

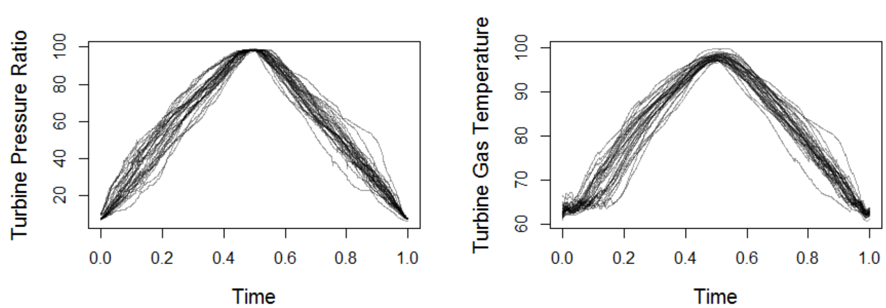


Fig. 1. Plots of 30 TPR and TGT time series.

1 Classical Functional Data Analysis

In this section we give a brief summary of the FDA tools that we will later apply in our model. In the following sections we will use the vector space $L^2(I)$ which is the Hilbert space of square integrable functions on the compact interval I with the inner product $\langle f, g\rangle = \int_I f(t)g(t)dt$ for functions $f, g \in L^2(I)$.

We will define $X(t), Y(t)$ to be univariate stochastic processes defined on I, with mean functions $\mu^X(t)$ and $\mu^Y(t)$, and covariance functions $C_X(s,t) = cov\{X(s), X(t)\}$ and $C_Y(s,t) = cov\{Y(s), Y(t)\}$ for all $s, t \in I$. We shall define $x(t) = [x_1(t), ..., x_n(t)]$ and $y(t) = [y_1(t), ..., y_n(t)]$ be n samples from $X(t)$ and $Y(t)$ respectively.

In practice we observe $x_i(t)$ and $y_i(t)$ at discrete time points. We shall assume for simplicity of exposition that observations are made at equally spaced time points $t_1, ..., t_N$. We will outline Functional Linear Regression and Functional Principal Component Analysis with respect to the underlying functions. In Section 1.3 we need to use the discretely observed data to define a suitable model selection criterion.

1.1 Functional Linear Regression

In this section we will introduce the FLR model [17], which we will use to model the relationship between TGT and TPR for the VS manoeuvre. In FLR we model the relationship between predictor $x_i(t)$ and response $y_i(t)$ as:

$$y_i(t) = \alpha(t) + \int_I x_i(s)\beta(s,t)ds + \epsilon_i(t), \tag{1}$$

where $\alpha(t)$ is the intercept function, $\beta(s,t)$ is the regression function and $\epsilon_i(t)$ is the error process. For a fixed t, we can think of $\beta(s,t)$ as the relative weight placed on $x_i(s)$ to predict $y_i(t)$. For simplicity we will assume the mean functions $\mu^X(t) = 0$ and $\mu^Y(t) = 0$ which thereby means $\alpha(t) = 0$. This is a reasonable assumption as in practice we can calculate the mean functions $\mu^X(t)$ and $\mu^Y(t)$ efficiently for dense data and then pre-process the data by subtracting $\mu^X(t)$ and $\mu^Y(t)$ from the observed curves.

FLR in the function-on-function case is a well studied model. There are typically two approaches taken: basis methods [24], [5] and grid based methods [11], [21]. The basis approach will be used as it is computationally efficient.

We will represent $x_i(t)$ and $y_i(t)$ in terms of M pre-chosen basis functions $\phi_j^X(t), \phi_j^Y(t)$ respectively:

$$x_i(t) = \sum_{j=1}^{M} z_{ij}\phi_j^X(t) \text{ and } y_i(t) = \sum_{j=1}^{M} w_{ij}\phi_j^Y(t).$$

For notational simplicity we have assumed that $x_i(t)$ and $y_i(t)$ can be represented by the same number of functions M, however this assumption can be easily relaxed.

We define $\phi^X(t) = [\phi_1^X(t), ..., \phi_M^X(t)]$, $\phi^Y(s) = [\phi_1^Y(s), ..., \phi_M^Y(s)]$, $z_i = [z_{i1}, ..., z_{iM}]$ and $w_i = [w_{i1}, ..., w_{iM}]$. We will then model the regression surface using a double basis expansion [18]:

$$\beta(s,t) = \sum_{l=1}^{M} \sum_{m=1}^{M} b_{ml} \phi_m^X(s) \phi_l^Y(t) = \phi^X(s)^T B \phi^Y(t), \tag{2}$$

for an $M \times M$ regression matrix B. We can then write:

$$y_i(t) = z_i B \phi^Y(t) + \epsilon_i(t). \tag{3}$$

Letting $\epsilon_i(t) = q_i \phi^Y(t)$ [5] we can reduce Equation (3) to:

$$w_i = z_i B + q_i. \tag{4}$$

This simplification enables us to estimate B using standard multivariate regression methods.

1.2 Functional Principal Component Analysis

In this section we describe Functional Principal Component Analysis (FPCA), which we will use to build data-driven basis functions $\phi^X(t)$ and $\phi^Y(t)$ for $x_i(t)$ and $y_i(t)$, respectively. These basis functions give effective, low-dimensional representations and will be used in the Functional Linear Regression model described in Section 1.1.

Functional Principal Component Analysis (FPCA) is a method of finding dominant modes of variance for functional data. These dominant modes of variance are called the Functional Principal Components (FPCs). FPCA is also used as a dimensionality reduction tool, as a set of observed curves can be effectively approximated by a linear combination of a small set of FPCs. These FPCs form an orthonormal basis over $L^2(I)$ [22].

The FPCs, $\phi_k^X(t)$ for $k = 1, 2, ...$, are the eigenfunctions of the covariance function $C_X(s,t)$ with eigenvalues λ_k^X. Note that the eigenfunctions are ordered by the respective eigenvalues. The Karhunen-Loéve theorem shows that $x_i(t)$ can be decomposed as $x_i(t) = \sum_{k=1}^{\infty} z_{ik} \phi_k^X(t)$ where the principal component score $z_{ik} = \int_I x_i(t) \phi_k^X(t) dt$.

We can define the M-truncation as:

$$\hat{x}_i(t) = \sum_{k=1}^{M} z_{ik} \phi_k^X(t), \tag{5}$$

which gives the minimal residual error:

$$\frac{1}{n} \sum_{i=1}^{n} ||x_i - \hat{x}_i||^2 = \frac{1}{n} \sum_{i=1}^{n} \int_I [x_i(t) - \hat{x}_i(t)]^2 dt, \tag{6}$$

over all possible M functions. To choose M we will use an information criterion outlined in Section 1.3.

1.3 Bayesian Information Criterion for FLR

In this section we formulate a Bayesian Information Criterion (BIC) to determine the basis size M, similarly to Matsui [12]. A component of the BIC is the log likelihood, often expressed as a squared error term. It is tempting to use the squared error resulting from Equation (4). However the objective is to fit the data y_i so we must use a likelihood of this data instead of a squared error term of basis coefficients.

Letting $y = [y_1, ..., y_n]$, where $y_i = [y_i(t_1), ..., y_i(t_N)]$, and $\phi^Y = [\phi^Y(t_1), ..., \phi^Y(t_N)]$ we obtain the discrete version of Equation (3):

$$y_i = z_i B\phi^Y + \epsilon_i, \tag{7}$$

where the error $\epsilon_i = [\epsilon_i(t_1), ..., \epsilon_i(t_N)]$ is assumed for simplicity to be sampled from $N(0, v^2 I_N)$, where I_N is the identity matrix of size N.

Using Equation (7) we can define the likelihood for sample i under parameters $\theta = (B, v^2, M)$ as:

$$f(y_i|\theta) = \frac{1}{(2\pi)^{\frac{N}{2}} v} \exp\left\{ -\frac{1}{2v^2} ||y_i - z_i B\phi^Y||_2^2 \right\}. \tag{8}$$

Then the log-likelihood is $l(\theta) = \sum_{i=1}^{n} \log(f(y_i|\theta))$. We can then write:

$$BIC(M) = -l(\theta) + \frac{1}{2}[M^2 + 1]\log(n). \tag{9}$$

where the number of free parameters is equal to $M^2 + 1$ [12], where M^2 comes from the regression terms and the 1 comes from the estimation of v^2 in the covariance matrix of the residuals.

To summarise, we estimate the FPCs for X and Y and solve the FLR model for multiple M values. We then choose the value of M that minimises the BIC criterion. The robust equivalent of this procedure is given in Algorithm 1.

2 Robust FLR model

In Section 1 we have defined the FLR model and have outlined the use of FPCA bases to estimate parameters of the model. In this section we will introduce robust versions of the FDA techniques outlined in Section 1. This will allow us to fit a normality model even in the presence of outliers. We shall also propose a robust BIC procedure for model selection.

We will replace classical FPCA, which can be shown to be heavily affected by outliers [3], with robust FPCA estimates by Boente and Salibian-Barrera [4]. Analogous to Equation (5), the robust FPCs $\tilde{\phi}_k^X(t)$, $\tilde{\phi}_k^Y(t)$, $k = 1, ..., M$ are orthonormal functions such that:

$$x_i(t) \approx \sum_{k=1}^{M} \tilde{z}_{ik} \tilde{\phi}_k^X(t) \quad y_i(t) \approx \sum_{k=1}^{M} \tilde{w}_{ik} \tilde{\phi}_k^Y(t).$$

are good approximations for $x_i(t)$ and $y_i(t)$.

We define $\tilde{y}_i(t) = \tilde{w}_i \tilde{\phi}^Y(t)$ and assume as in Equation (4) that $\epsilon_i = \tilde{q}_i \tilde{\phi}^Y(s)$. We can now write the robust counterpart of Equation (4) as:

$$\tilde{w}_i = \tilde{z}_i \tilde{B} + \tilde{q}_i. \tag{10}$$

To obtain a robust estimate of the regression matrix $\tilde{B}$, we will use the Multivariate Least Trimmed Squares (MLTS) [1] estimator. The objective of MLTS is to find a subset of our data of some pre-chosen size k, which gives minimal L^2 error over all possible subsets of size k. This is robust as outliers will not be in the subset by definition so shall not affect the model estimation. We will choose a subset of size $k = \lceil 0.8n \rceil$.

In Section 1.3 we defined a BIC procedure to choose the number of basis functions M. In our case we wish to estimate M in the presence of outliers, we therefore propose a robust BIC (RBIC) procedure. As in the BIC procedure in Section 1.3 we shall use the observed data y_i. We can then define the Mahalanobis distance for sample i under model M as:

$$\Delta_i(M) = \frac{1}{v^2}(y_i - \tilde{z}_i \tilde{B} \tilde{\phi}^Y)^T (y_i - \tilde{z}_i \tilde{B} \tilde{\phi}^Y). \tag{11}$$

To make the information criteria robust, we shall take a subset S of size k with the smallest Mahalanobis distances analogous to MLTS. We obtain the log-likelihood $\tilde{l}(\theta) = \sum_{i \in S} \log(f(y_i|\theta))$, which we use to obtain:

$$RBIC(M) = -\tilde{l}(\theta) + \frac{1}{2}(M^2 + 1)\log(k). \tag{12}$$

In Algorithm 1 we outline the calculation of the robust FLR model. In the algorithm we estimate the model for different values of M and choose the model with the minimum RBIC value. We consider $M = 1, \ldots, P$ where P is chosen to ensure that 99.99% of the variance in the raw data is captured by the first P FPCs.

Data: Let (x_i, y_i) be time series of length N for $i = 1, ..., n$, and P be the number of models.

1. Estimate mean functions $\tilde{\mu}_X(t)$ and $\tilde{\mu}_Y(t)$ [19]
2. Centre the time series (x_i, y_i)
3. Estimate $\{\tilde{\phi}^X_1(t), ..., \tilde{\phi}^X_P(t)\}$, $\{\tilde{\phi}^Y_1(t), ..., \tilde{\phi}^Y_P(t)\}$ [4].

for $M = 1$ to P **do**

  Estimate the regression matrix B using MLTS [1].

  Obtain the RBIC value using Equation (12)

end for

4. Select model M^* with smallest RBIC value

return Regression matrix $\tilde{B}$ from model M^* and $\{\tilde{\phi}^X_1(t), ..., \tilde{\phi}^X_{M^*}(t)\}$, $\{\tilde{\phi}^Y_1(t), ..., \tilde{\phi}^Y_{M^*}(t)\}$

Algorithm 1: Robust FLR procedure

3 Outlier Detection

In this section we will describe an outlier detection procedure using the robust FLR model estimated in Algorithm 1. We will use the robust FLR model to obtain estimates $\tilde{y}_i = \tilde{z}_i \tilde{B} \tilde{\phi}^Y$ for $i = 1, ..., n$.

For an outlier we expect the residual $r_i = y_i - \tilde{y}_i$ to deviate in behaviour from the other residuals. Traditionally, we would use the L^2 error to identify outliers. However we have found using a functional depth approach [8] is more effective in identifying outliers. The approach assigns a depth value to samples r_i. Samples with small depth values lie far away from the other samples.

We will use the h-modal depth [6] to rank samples r_i. For a given kernel K_h (typically Gaussian with bandwidth h), the h-modal depth of r_i with respect to $r = \{r_1, ..., r_n\}$ is given by:

$$D(r_i|r, h) = E(K_h||z - r||) \approx \frac{1}{n} \sum_{l=1}^{n} K \left(\frac{||r_i - r_l||_2}{h} \right). \quad (13)$$

with the bandwidth h taken to be the 15th percentile of the empirical distribution of $\{||r_i - r_j||_2^2, i, j = 1, ..., n\}$ [8].

In the algorithm we need a threshold to identify outliers, which is chosen such that $P(D(r_i|r, h) \leq C) = \delta$, where δ is a pre-chosen percentile. We take a sample of size m from the set of samples excluding the $\alpha\%$ of samples with the smallest depth. Then we bootstrap to obtain an estimate of the threshold C. We describe the outlier detection algorithm in Algorithm 2.

Data: Centred time series (x_i, y_i) for $i = 1, ..., n$ and percentile δ.
1. Use Algorithm 1 to obtain $\tilde{\phi}_k^Y(t)$, $\tilde{z}_k$ and $\tilde{B}$.
2. Take $\tilde{\phi}^Y = (\tilde{\phi}^Y(t_1), ..., \tilde{\phi}^Y(t_N))$
3. Estimate residual vectors r_i as follows:
for $i = 1$ to n **do**
  Estimate $\tilde{y}_i = \tilde{z}_i \tilde{B} \tilde{\phi}^Y$
  Obtain residual vectors $r_i = y_i - \tilde{y}_i$
end for
4. Estimate bandwidth h [8]
5. For each r_i calculate $D(r_i|r, h)$.
6. Estimate C [8] for given percentile δ
7. If $D(r_i|r, h) < C$ sample i is an outlier.

Algorithm 2: Outlier Detection

4 Simulations

In this section we will outline simulation results comparing the effectiveness of BIC and RBIC in identifying the true model. Using the true model we shall also compare our robust FLR model to the classical FLR approach in terms of outlier detection.

4.1 Scenario

We will generate samples $x(t)$ using a FPCA based model with mean function $\mu_X(t) = -10(t-0.5)^2 + 2$ for $t \in [0,1]$ and eigenfunctions:

$$\phi_1^X = \sqrt{2}\sin(\pi t),\ \ \phi_2^X = \sqrt{2}\sin(7\pi t),\ \ \phi_3^X = \sqrt{2}\cos(7\pi t).$$

The principal scores are sampled from Gaussian distributions with mean 0 and variances 20, 10 and 5 for the eigenfunctions respectively. We generate 100 samples, $x_1(t), ..., x_{100}(t)$. Note that we do not create any outliers in the predictors.

The samples $y(t)$ will have eigenfunctions:

$$\phi_1^Y = \sqrt{2}\sin(12\pi t),\ \ \phi_2^Y = \sqrt{2}\sin(5\pi t),\ \ \phi_3^Y = \sqrt{2}\cos(2\pi t).$$

Using these eigenfunctions we will generate $\beta(s,t) = \phi^X(s)^T B \phi^Y(t)$ where B will have random entries between $[-1,1]$. Outliers will be generated by replacing B with $B' = B + R$ where R has random entries sampled from $N(0, 0.5)$. We will then add a mean function $\mu^Y(t) = 80\exp(-(t-1)^2)$. The residual function $\epsilon_i(t)$ will be a linear combination of $\phi^Y(t)$ with coefficients sampled from $N(0, 0.1)$. We will consider three cases when the number of outliers are $a = 1, 5$ and 10.

4.2 Results

We have used $M = 3$ principal components to generate $x(t)$ and $y(t)$ and $\beta(s,t)$. To test BIC and RBIC we will make 200 repetitions and calculate the proportion of times the true model is chosen. The results are given in Table 1. BIC and RBIC have similar levels of accuracy for small number of outliers. However for large number of outliers $a = 10$ RBIC outperforms BIC considerably.

Next we will compare the outlier detection capabilities of the Robust FLR (RFLR) model and the classical FLR (CFLR) model. To compare the models we will use the sensitivity (sens) defined as the the proportion of actual outliers detected and specificity (spec) defined as the proportion of non-outlying samples correctly identified as non-outliers. We want a model that has high sensitivity and specificity. The outlier detection algorithm relies on choosing a suitable percentile δ. In our simulated data we choose δ equal to the true proportion of outliers in the simulated data.

In Table 2 we have the sensitivity and specificity for the robust and non-robust model. We can see the robust model and non-robust model have the same level of sensitivity however the robust model has higher specificity. This shows the robust model is more effective in differentiating between non-outliers and outliers.

5 Engine data

In Fig. 1 we have a plot of 30 VS TPR and TGT time series. The TPR is controlled during the test, and therefore the variability comes from the controller

Table 1. Proportion of times true model chosen using BIC and RBIC over 200 replications.

	a=0	a=1	a=5	a=10
BIC	1	0.995	0.99	0.965
RBIC	1	0.995	0.995	0.995

Table 2. Sensitivity and Specificity over 200 replications.

	a=1		a=5		a=10	
	sens	spec	sens	spec	sens	spec
Classical FLR	1	0.956	0.999	0.959	1	0.957
Robust FLR	1	0.964	1	0.970	1	0.964

performing the manoeuvre. This causes variability in the TGT time series. We therefore want to model the TGT and TPR relationship to account for this controller induced variability.

Applying Algorithm 1 we select model $M = 9$, which we use in Algorithm 2 to obtain the residual curves in Fig. 2. Using a percentile $\delta = 0.01$ we obtain four outliers, which seem plausible from a visual inspection. We know that the test with the smallest depth value corresponds to an engine with damaged hardware. The other three outliers obtained for this value of δ are currently being investigated. Instead of using a fixed value for δ, the user can apply the proposed algorithm to sort the residual time series in increasing order of depth and thus obtain a priority list of tests to be investigated by engineers.

Standard outlier detection methods for functional data do not model the dependency between the predictor and response functions, and may therefore miss outliers. To show this explicitly we applied the outlier detection algorithm [8] on the TPR and TGT time series. The outlier detection algorithm outputs the same three outliers for the TPR and TGT, as shown in green in Figure 3. This supports our argument that abnormal speed profiles cause abnormal temperature profiles, which in turn can mask real outliers. In comparison our approach models the relationship between the TPR and TGT time series and is therefore able to identify four different outliers shown in red in Figure 3.

6 Conclusion

We have proposed a new outlier detection method to identify anomalous samples with respect to the engine temperature. We have built a robust functional regression model to capture the temporal relationship between speed and temperature

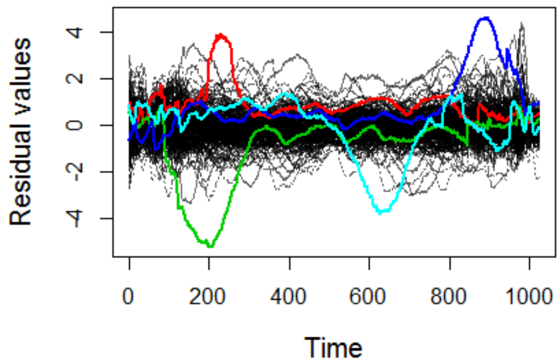
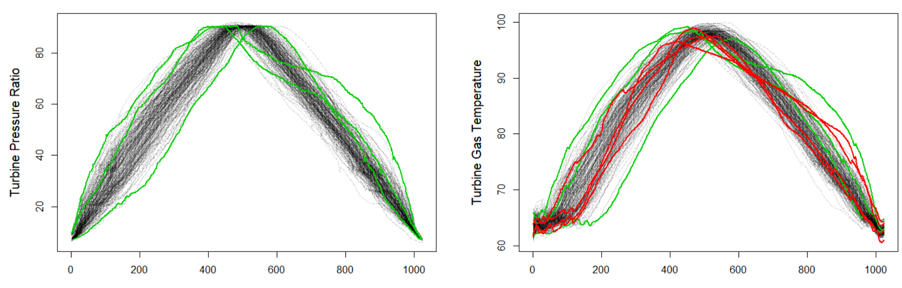


Fig. 2. Residuals using robust FLR model on TGT response and TPR predictor functions. The four outliers identified using Algorithm 2 are coloured.



(a) TPR (b) TGT

Fig. 3. Plots of the TPR and TGT time series with outliers using RFLR in red and those using a standard outlier detection approach [8] on the TPR and TGT curves directly in green.

and have used this model to identify outliers. In our experimental results we have shown that this model is capable of detecting outliers effectively, and identified samples of interest in the jet engine data.

In future work we aim to investigation the connection between the outliers identified and possible engine issues. We shall also develop theoretical results for the robust FLR model and BIC procedure.

References

1. Agulló, J., Croux, C., Van Aelst, S.: The multivariate least-trimmed squares estimator. J. Multivar. Anal. **99**(3), 311–338 (2008)
2. Arribas-Gil, A., Romo, J.: Shape outlier detection and visualization for functional data: the outliergram. Biostatistics **15 4**, 603–19 (2014)
3. Bali, J.L., Boente, G., Tyler, D.E., Wang, J.L.: Robust functional principal components: A projection-pursuit approach. Ann. Statist. **39**(6), 2852–2882 (2011)
4. Boente, G., Salibian-Barrera, M.: S-estimators for functional principal component analysis. Journal of the American Statistical Association **110**(511), 1100–1111 (2015)
5. Chiou, J.M., Yang, Y.F., Chen, Y.T.: Multivariate functional linear regression and prediction. Journal of Multivariate Analysis **146**, 301 – 312 (2016), special Issue on Statistical Models and Methods for High or Infinite Dimensional Spaces
6. Cuevas, A., Febrero, M., Fraiman, R.: Robust estimation and classification for functional data via projection-based depth notions. Comput. Stat. **22**(3), 481–496 (Sep 2007)
7. Dai, W., Genton, M.G.: Multivariate functional data visualization and outlier detection. Journal of Computational and Graphical Statistics **27**(4), 923–934 (2018)
8. Febrero-Bande, M., Galeano, P., Gonzãlez-Manteiga, W.: Outlier detection in functional data by depth measures, with application to identify abnormal nox levels. Environmetrics **19**, 331 – 345 (06 2008)
9. Hayton, P.M., Schölkopf, B., Tarassenko, L., Anuzis, P.: Support vector novelty detection applied to jet engine vibration spectra pp. 946–952 (2001)
10. Hubert, M., Rousseeuw, P.J., Segaert, P.: Multivariate functional outlier detection. Statistical Methods & Applications **24**(2), 177–202 (2015)
11. Ivanescu, A.E., Staicu, A.M., Scheipl, F., Greven, S.: Penalized function-on-function regression. Computational Statistics **30**(2), 539–568 (2015)
12. Matsui, H.: Quadratic regression for functional response models. ArXiv e-prints (2017)
13. Morris, J.S.: Functional Regression. Annual Review of Statistics and Its Application **2**, 321–359 (2015)
14. Nairac, A., Townsend, N.W., Carr, R., King, S., Cowley, P., Tarassenko, L.: A system for the analysis of jet engine vibration data. Integrated Computer-Aided Engineering **6**, 53–66 (1999)
15. Narisetty, N.N., Nair, V.N.: Extremal depth for functional data and applications. Journal of the American Statistical Association **111**(516), 1705–1714 (2016)
16. Nieto-Reyes, A., Battey, H.: A topologically valid definition of depth for functional data. Statist. Sci. **31**(1), 61–79 (02 2016)
17. Ramsay, J.O., Dalzell, C.J.: Some tools for functional data analysis. Journal of the Royal Statistical Society. Series B (Methodological) **53**(3), 539–572 (1991)

18. Ramsay, J., Silverman, B.W.: Functional Data Analysis (Springer Series in Statistics). Springer-Verlag. 2nd ed. (2005)
19. Rousseeuw, P.J., Driessen, K.V.: A fast algorithm for the minimum covariance determinant estimator. Technometrics **41**(3), 212–223 (1999)
20. Rousseeuw, P.J., Raymaekers, J., Hubert, M.: A measure of directional outlyingness with applications to image data and video. Journal of Computational and Graphical Statistics **27**(2), 345–359 (2018)
21. Scheipl, F., Staicu, A.M., Greven, S.: Functional additive mixed models. Journal of Computational and Graphical Statistics **24**(2), 477–501 (2015)
22. Shang, H.L.: A survey of functional principal component analysis. AStA Advances in Statistical Analysis **98**(2), 121–142 (2014)
23. Sun, Y., Genton, M.G.: Functional boxplots. Journal of Computational and Graphical Statistics **20**(2), 316–334 (2011)
24. Yao, F., Müller, H.G., Wang, J.L., et al.: Functional linear regression analysis for longitudinal data. The Annals of Statistics **33**(6), 2873–2903 (2005)