International Journal of Applied Information Systems (IJAIS) ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12 - No. 33, September 2020 - www.ijais.org

Outliers Detection in Sensor Time Series using Robust moving Least Squares

Crislânio de Souza Macêdo, José Everardo Bessa Maia
State University of Ceará – UECE
60714-903 - Fortaleza – Ceará – Brazil

ABSTRACT

Sensors are ubiquitous elements, whether through smart phones and other personal devices, or via wireless sensor networks, body area networks or IoT in general. However, due to noise, intermittent operation or message loss, sensor time series often arrive with outliers at processing centers. In this work, the problem of detecting isolated outliers in sensor time series is addressed using Robust Moving Least Square prediction (RMLS). The performance of RMLS is compared against that of the Sequentially Discounting Autoregressive (SDAR), which is a well-established state of the art method. The results show that RMLS has performance compatible with SDAR in all tests, with the advantage that RMLS is less sensitive to outliers present in the predictors window.

General Terms:

Outlier Detection, Robust Moving Least Square, Signal Processing

Keywords

Outlier Detection, Sensor Time Series, Robust Moving Least Square, Sequentially Discounting Autoregressive, Linear Prediction

1. INTRODUCTION

A Sensor Time Series (STS) is a series of data numerical points indexed in time order. Often, STS is the sequence of samples for a continuous variable, for example, temperature, taken at successive equally spaced points in time. In other applications, it can represent stock prices, sales volume, stock volume in a company or even the sequence of the numbers of the evolution of an epidemic. Thus it is a sequence of discrete-time ordered date which can be univariate or vector.

The acquired data is normally not directly usable, as they suffer from problems such as noise, data loss and outliers. In addition to the noise inherent in any measurement, phenomena in the local context may temporarily aggravate the noise. Missing data usually occurs due to packet loss and node failure, and outliers can originate from transmission errors or intense interferences in physical quantity being measured [22].

Extracting useful knowledge from raw sensor data is not a trivial task. The context of the sensor data makes the design of an appropriate outlier detection technique more challenging. Given this scenario, even though many techniques for detecting outliers have been developed in the past and several strategies have been applied in sensor techniques to increase reliability, detect outliers and isolate defective readings, there is still a vast field of research to mitigate the effects of these outleirs in the sensors' time series [20, 15].

Two Sensor Time Series (STS) processing modes are found. In offline mode, all or part of the STS is received and stored and must be processed without processing or storage time restrictions. In this mode, processing at one point in the series can benefit from knowledge of past and future values at that point. This mode is typical of applications such as financial series [23], retail sales series [16] and some climate monitoring applications [25].

In data stream mode, it is generally desired to detect events as they occur, with no time between samples for intensive processing, such as the construction of sophisticated models. This mode is typical of applications for continuous environmental monitoring [6, 12], health care monitoring [8] and industrial [11, 19] and IoT [3, 27, 5] monitoring in general. In this context, outliers must be detected (novelty, anomaly) without, however, the outliers present in the STS bias the detection of the following outliers.

Depending on the problem at hand, there are two approaches to tackling outliers in sensor time series and in data in general [7, 4]: detecting and treating outliers, or using robust processing methods for outliers. Robust methods are useful when, for example, through domain knowledge, outliers are not a source of information but only undesirable interference.

However, in many applications, such as those looking for novelty or anomaly, outlier detection is necessary and processing of the outlier is imposed. In this situation, robust methods may not be adequate as they can hide new information coming.

An outlier detector can fail in two ways: fail to detect an outlier (false negative) or classify a normal sample as outlier (false positive). In a detector based on a sliding window, when a wrongly classified sample enters the window, it biases the detector and may lead to new detection errors. In this scenario, a detection method that is robust to outliers within the data window is welcome. This is because, in the case of a false positive, if any treatment for outliers is being used, it will replace that real sample with another value when it should not be done. And in the case of a false negative, no action is taken as the outlier is unknown and the outlier will be part of the window and bias the procedure.

Autoregressive (AR) [13] or sliding window (MA) [24] prediction techniques are often used to detect outliers in this context. SDAR (Sequentially Discounting Autoregressive) [26] prediction is a well-established technique in the class of autoregressive models and MLS (Moving Least Square) [28] prediction uses a sliding window.

The research questions faced in this paper are to evaluate the effect of past undetected or untreated outliers on the accuracy of these algorithms to detect new isolated outliers, and to compare the performances against that of robust moving least squares linear prediction. The authors do not know of any other work with a specific focus on this problem. Prediction based on robust regression over a data sliding window has other applications in ecology cite and economics cite . In this work it is associated with an error threshold to detect outliers.

In what follows, this article is organized like this. Section 2 presents the theoretical foundations of the investigated methods and Section 3 analyzes some works directly related to this. In Section 4 the data used, the experiments and the results are described and analyzed. Section 4 concludes this work.

International Journal of Applied Information Systems (IJAIS) ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12 - No. 33, September 2020 - www.ijais.org

2. METHODS

Time series outlier detection methods based on prediction build generative models of the time series using past data and analyze the error between the value predicted by the model and the data arriving to classify it as outlier or not outlier. This section reviews the fundamentals of the techniques used in this work.

2.1 Robust OLS Regression

A multiple linear regression model seeks to approximate the relationship between a dependent or output variable y, and a number of independent or input variables $\{x_i\}$ by a linear function. In the experimental context, when n observations are available, this relationship can be represented, in matrix notation, by Equation 1:

$$\mathbf{y} = \mathbf{X}\beta + \epsilon, \tag{1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observed response values, $\mathbf{X}$ is the $n \times p$ matrix of the predictor variables, β is the $p \times 1$, and ϵ is the $n \times 1$ vector of random error terms.

The purpose of the model is to find the estimates of unknown parameters vector β and it describes what is called the regression surface [18]. In practice it occurs that due to measurement inaccuracies Equation 1 almost never has an exact solution and thus OLS (Ordinary Least Squares) is used to find the best estimate of β's with the least squares criterion which minimizes the sum of squared distances of all of the points from the actual observation to the regression surface.

The solution of Equation 1 by the OLS method is well known and is given by

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{X}^{+}\mathbf{y}, \tag{2}$$

where $\hat{\beta}$ is the estimate of β for this data set, and $\mathbf{X}^{+}$ is called the MoorePenrose inverse of the data matrix $\mathbf{X}$.

A feature of the OLS method is that it gives equal importance (weight) to all the examples in the data set. One weighted least squares (WLS) method can be obtained from OLS by taking a diagonal matrix of $\mathbf{W}$ weights of dimension $n \times n$ with the diagonal elements considering the importance of each corresponding example. In this case, Equation 1 becomes

$$\mathbf{y} = \mathbf{W}\mathbf{X}\beta + \epsilon, \tag{3}$$

and its weighted least squares (WLS) solution is given by

$$\hat{\beta} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}. \tag{4}$$

One way for a outlier robust regression method would be to use WLS Equation 4 by assigning lower weights to outlier data. The difficulty is that it is not known in advance which data are outliers. This is an appropriate scenario for applying iterative methods giving rise to the iteratively re-weighted least squares (IRLS) method.

In the general method of robust regression M-estimator, introduced by Huber [18], rather than minimize the sum of squared errors as the objective, the M-estimate minimizes a function of the errors. Thus, the diagonal matrix W is written as

$$\mathbf{W} = diag(f(\frac{\mathbf{y} - \mathbf{X}\hat{\beta}}{\sigma})), \tag{5}$$

where the objective is the least absolute deviation and σ is a (scale) width parameter of the function $f()$ that makes it reduce the weight of the most distant examples of the regression surface of the current iteration. The IRLS method is an iterative solution consisting of the following steps [18]:

(1) Initialize β^0 using the OLS method,
(2) At the each iteration t, calculate residuals and the weight matrix by Equation 5,
(3) Solve for new weighted least squares estimates with $\mathbf{W}$ calculated in step 2.

Repeat Step2 and Step3 until the estimated coefficient converges by some convergence criterion. Various functions are published and tested for $f()$ and the Gaussian function was used. To complete the specification, the estimate of scale σ may be updated after initial estimate based on the variance of the residuals.

Figure 1 illustrates the effect of the IRLS algorithm when compared to OLS in the presence of outliers in the data window. It shows a 20 point STS and the application of OLS and IRLS to estimate the expected value of the next STS measurement. A data window with a size of 3 was used. The two models were built in two moments: one around time 4, free from outliers, and another around time 12, with the presence of outliers.

In the first application, the models, and consequently the predictions, almost coincide. With the presence of outliers, the models differ significantly. Note that in the figure a threshold range was included around the prediction outside which a sample will be classified as outlier. Clearly, the OLS model would fail to detect the next outlier while the IRLS model would be effective in this example built for demonstration.

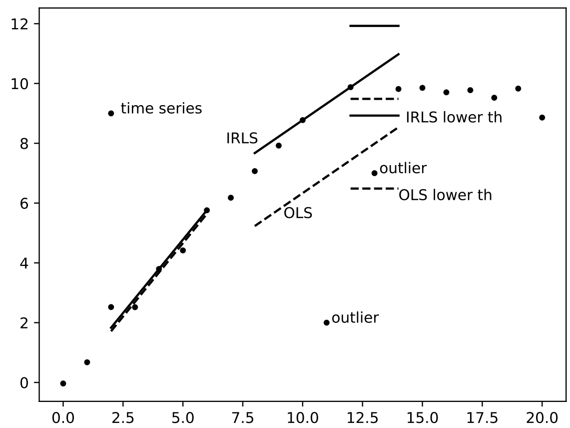

Fig. 1. Effect of an outlier in the design data window on the accuracy of OLS and IRLS outlier detection.

2.2 Sequentially Discounting AutoRegressive (SDAR)

SDAR [21] uses density estimation and the fact that a sum of Gaussian random variables is also a Gaussian random variable. To estimate the value density of a sample x_t arriving at time t, let ϵ_t be a white Gaussian noise with zero mean and variance σ^2. SDAR assumes that the monovariate time series $\{x_t, t = 1, 2, 3, ...\}$ was generated by the model

$$x_t = \sum_{i=1}^{p} a_i x_{t-i} + \epsilon_t, \tag{6}$$

where x_{t-i} is the value of the time series at time $t-i$, $a_i, i = 1...p$ are the model coefficients and p is the model order. In these conditions, $P(x_t|x_{t-1}, ..., x_{t-p}, \theta)$ has normal distribution $N(w, \sigma^2)$ with probability density given by [26]:

$$P(x_t|x_{t-1}, ..., x_{t-p}, \theta) = \frac{1}{\sigma\sqrt{2\pi}}\epsilon^{-(x_t-w)^2/2\sigma^2}, \tag{7}$$

where $\theta = (a_1, ..., a_p, \sigma^2)$ is the vector of model parameters and $w = \sum_{i=1}^{p} a_i x_{t-i}$.

Normal Yule-Walker equations and maximum likelihood estimation are two procedures by which the model parameters can be estimated. In any case, the parameters are obtained by minimizing the sum of the discounted squared errors given by $\sum_{i=p+1}^{n}(1-r)^{t-i}(x_i - A^T\tilde{x}_i)^2$, where $r \in (0,1)$ is a discount factor that reduces the weight of the samples more distant in time in relation to the current time t, x_i is the true value read from the series, $\tilde{x}_t = (x_{t-1}, ..., x_{t-p})^T$ and $A = (a_1, ..., a_p)^T$.
Using the normal equations the algorithm to update the estimates of the mean and variance of the x_t probability density is given in Algorithm 1 [26].

Algorithm 1: Pseudocode for the SDAR algorithm.

Data: A Sensor Time Series $(x_t, t = 1, 2, 3, ..., T)$, $r \in (0, 1)$
Result: A vector $\mathbf{score}()$ of the x_t samples scores.
initialization $\hat{\mu}_0, C_j, \hat{\Sigma}, \hat{\omega}_j (j = 1, ..., k)$;
for $t \in (1, 2, 3, ..., T)$ **do**
  read x_t;
  $\hat{\mu} = (1-r)\hat{\mu} + rx_t$;
  $C_j = (1-r)C_j + r(x_t - \hat{\mu})(x_{t-j} - \hat{\mu})^T$;
  solve for ω: $C_j = \sum_{i=1}^{k}\omega_i C_{j-i} (j = 1, ..., k)$;
  $\hat{x}_t = \hat{\omega}(x_{t-k}^{t-1} - \hat{\mu}) + \hat{\mu}$;
  $\hat{\Sigma} = (1-r)\hat{\Sigma} + r(x_t - \hat{x}_t)(x_t - \hat{x}_t)^T$;
  $\mathbf{score}(t) = -log(p_{t-1}(x_t))$;
end

The $\mathbf{score}(x_t)$ value assesses the change in the probability density function when going from p_{t-1} to p_t. Therefore, the higher the $\mathbf{score}(x_t)$, the greater the possibility that x_t is an outlier [26].

3. RELATED WORKS

Research and publication on outlier detection and treatment in Sensor Time Series is extensive. To contextualize this paper, this section reviews some works directly related to this research [10], [1], [29], [14].

The authors in [10] present the Temporal Outlier Discovery (TOD) framework for detecting temporal outliers in vehicle traffic data. What distinguishes the approach is that it does not look at the behavior of individual objects, but instead seeks to detect disparate behavior in road segment traffic data. Outlier scores are calculated to detect drastic changes in trends in road segment traffic in relation to its history and that of its neighbors. However, the authors say nothing about how data with outliers is incorporated into the formation of traffic history and its effects.

A two-stage probabilistic method for detecting anomalies in natural gas time series data is presented in the work [1]. In the first phase, the probability of a data point being anomalous is determined, using an OLS linear regression model and a geometric probability distribution of the residuals. The second step is to train a Bayesian maximum likelihood classifier based on the types of anomalies identified in the first phase. The method's contribution is its ability to incorporate domain knowledge in the techniques developed in the first phase for the detection of anomalies via Machine Learning techniques in the operation phase. Missing data and outliers are assumed to be absent from the design data.

In [29] a method is proposed to detect and repair data quality problems in STS generated from seafloor observatories. The proposed method consists of three parts: a general pre-test to preprocess data and provide a route for further processing, data outlier detection methods to label suspect data points, and a data interpolation method to fill in missing data and suspects. An integrated autoregressive moving average model (ARIMA) within a sliding window is used to compute the prediction interval, so that the parameters are readjusted each time the window moves a step forward. Processing is offline and the design data in each sliding window is assumed to be free of outliers.

The authors in [14] study the detection and elimination of outliers in self-weighing time series data obtained from connected weight scales. For these, three techniques are analyzed: (1) a method based on autoregressive integrated moving average (ARIMA) time series modelling, (2) another based on median absolute deviation (MAD) scale estimate, and (3) a method based on Rosner statistics [17]. Are applied these methods to both a data set with real outliers and a clean data set corrupted with simulated outliers. The results suggest that the simple MAD algorithm and ARIMA performed well with both test sets while the Rosner statistics was significantly less effective. In addition, the ARIMA approach appeared to be significantly less sensitive to long periods of missing data than MAD and Rosner statistics. The work assumes that the data are approximately normal and the outliers are not taken into account in the design of the models.

The reviewed papers are representative of four applications with practical relevance. Note that they do not consider a configuration in which outliers are present in the design data of the models updated online.

4. METRICS AND RESULTS

Two noise STS were used in the experiments: a synthetic STS and a natural STS. A number of random amplitude and polarity outliers have been inserted in random positions. In a specified number of these positions, occurrences of two outliers were created separated only by a correct sensor reading to stress the methods. It has been provided for the four polarity combinations to occur.

Preprocessing was carried out to guarantee time series without seasonality and trend when applying SDAR. The performance of the OLS, IRLS and SDAR methods were compared using the False Negative (FN) and False Positive (FP) metrics. False Positive accumulates the number of regular measurements that the method classified as outliers and False Negative accumulates the number of outliers that were not detected by the algorithm.

4.1 Result for Synthetic Time Series

For this test an STS of 200 points was generated with the same recursive equation used in [26] given by $x_t = 0.6x_{t-1} - 0.5x_{t-2} + n_t$ where n_t is a white Gaussian additive noise (AWGN) with zero mean and unit variance. The graph of this STS is shown in Figure 2. In this STS, 20 isolated outlier points were inserted, corresponding to 10 % of the series. It was taken care that the STS contains 8 occurrences of outliers separated by just a correct measurement.

The outputs of the executions of the three methods were inspected to generate the results shown in Table 1. For each detection method, total FN and FP (T in Table 1) are shown, and of these, in how many cases there was an outlier within the sliding window of the modeling data for the OLS and IRLS cases. Column w for SDAR aims to verify the behavior of SDAR in situations where OLS and IRLS are more sensitive.

Table 1 shows that the OLS model can only be applied when it is guaranteed that the STS is free from an outlier. IRLS and SDAR performed well with SDAR slightly better in FN and with equal performance in FP. It was observed that some FPs generated by SDAR are not the same as those that appear for IRLS. For SDAR, r = 0.02 was used and, after tests for OLS and IRLS, a size 3 sliding window was used with the parameters given in Section 2.

International Journal of Applied Information Systems (IJAIS) ISSN : 2249-0868
Foundation of Computer Science FCS, New York, USA
Volume 12 - No. 33, September 2020 - www.ijais.org

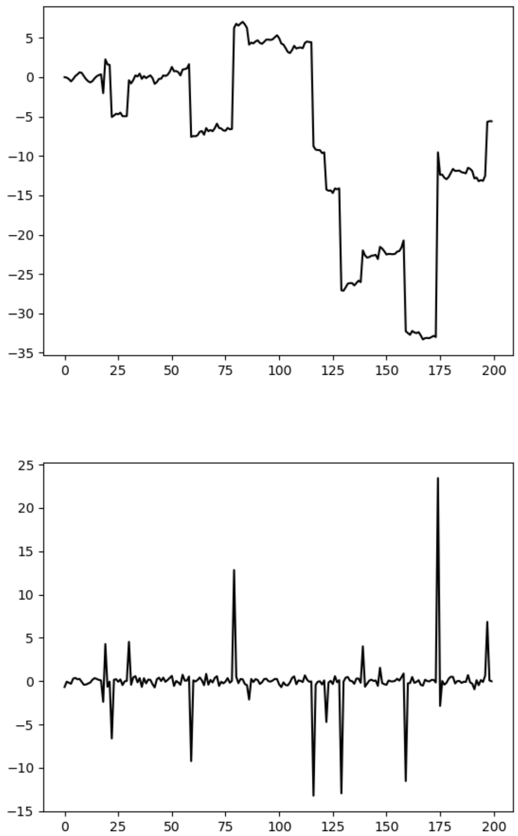

Fig. 2. A synthetic STS with 200 points of which 20 points (10%) are isolated outliers. In these, there are 8 occurrences of outliers that remain in the modeling window used to detect the next outlier.UP: original series, DOWN: difference series with outliers.

Table 1. Outlier detection performances obtained using the SDAR, OLS and IRLS algorithms in the synthetic STS of Figure 2 (FN stands for false negative, FP stands for false positive, T stands for FN and FP totals, and w stands for the FP and FN numbers when an outlier is present in the sliding window).

Method	OLS		SDAR		IRLS	
Metrics	T	w	T	w	T	w
FP	21	03	09	07	10	07
FN	11	04	02	01	03	01

4.2 Result for Natural Time Series

For the experiment with natural STS, a time series of sensor from the Intel-Lab-Data dataset [9] was taken. This dataset is derived from the Intel-Lab-Data dataset [2] with some preprocessing. It contains 14400 temperature measurements from 52 sensors from the original data corresponding to five days of measurement with an interval of 30 s between measurements. For this test, sensor 1 was chosen. In the natural STS of sensor 1, 144 isolated outlier points were inserted, corresponding to 10% of the series. It was provided to guarantee 16 occurrences in which two ouliers are separated by a correct measurement with different polarities. The resulting STS is shown in Figure 3.

A detailed inspection of the results generated Table 2 which summarizes the performance of the algorithms for this STS. For SDAR, r = 0.02 was used and for OLS and IRLS, a sliding window of size 3 was used with the parameters given in Section 2. Other sizes of sliding window were tested and size 3 performed well. The Table shows IRLS and SDAR again with similar performances with IRLS slightly better in FP and slightly worse in FN. As in the previous case, it was observed that some FP and FN are not the same for SDAR and IRLS. The table also confirms the previous understanding that the OLS model can only be applied when the STS is free of outlier.

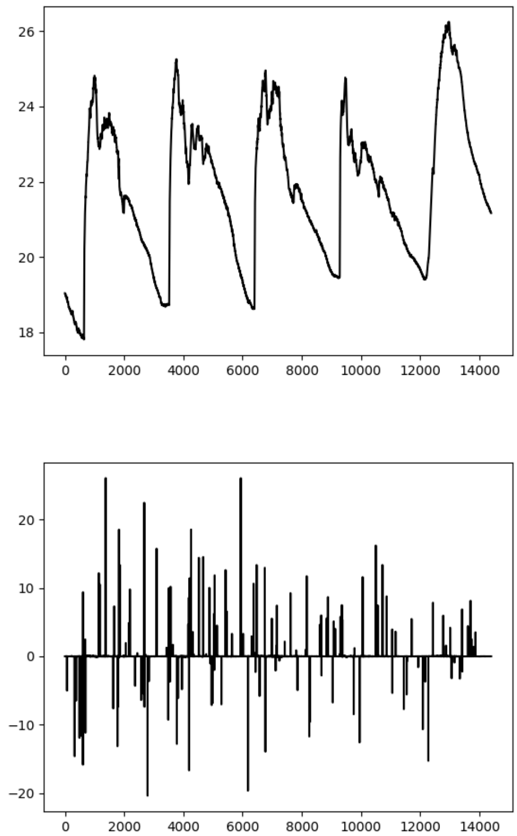

Fig. 3. The natural STS-20 with 14400 points of which 144 points (1%) are isolated outliers. To the original 24 outliers, 120 more challenging artificial outliers were added. In these, there are 16 occurrences of outliers that remain in the modeling window used to detect the next outlier. UP: original series, DOWN: difference series with outliers.

Table 2. Outlier detection performances obtained using the SDAR, OLS and IRLS algorithms in the natural STS of Figure 3 (FN stands for false negative, FP stands for false positive, T stands for FN and FP totals, and w stands for the FP and FN numbers when an outlier is present in the sliding window).

Method	OLS		SDAR		IRLS	
Metrics	T	w	T	w	T	w
FP	37	06	17	04	18	05
FN	41	07	11	03	13	02

5. CONCLUSION

The work investigated the performance of sliding window-based strategies using the OLS and IRLS algorithms against that of SDAR in detecting isolated outliers in Sensor Time Series in the specific context in which the detected outliers are not treated. The motivation is that in some data stream applications there will be no time or interest in treating outliers although it is important to identify them. Specifically, the effect of outliers not removed was investigated when they enter the data window used to update the prediction models.

The results in Tables 1 and 2, for synthetic and natural STS, respectively, show that sliding window with OLS cannot be used in this scenario. They also show that SDAR and sliding window with IRLS have comparable performance although in some situations the IRLS does better. It should be noted that the algorithms have the same number of adjustable parameters and that the SDAR model requires a time series without trend and seasonality while the sliding window can be adjusted locally in some contexts.

This study was not exhaustive in analyzing the entire parameter space of the compared algorithms. Thus, in future work, it is planned to survey the configuration space to identify the regions with the best applicability for each algorithm.

6. REFERENCES

[1] Hermine N Akouemo and Richard J Povinelli. Probabilistic anomaly detection in natural gas time series data. *International Journal of Forecasting*, 32(3):948–956, 2016.

[2] Peter Bodik, Wei Hong, Carlos Guestrin, Sam Madden, Mark Paskin, and Romain Thibaux. Intel lab data. *Online dataset*, 2004.

[3] Andrew Cook, Göksel Mısırlı, and Zhong Fan. Anomaly detection for iot time-series data: A survey. *IEEE Internet of Things Journal*, 2019.

[4] Pedro Galeano and Daniel Peña. Finding outliers in linear and nonlinear time series. In *Robustness and Complex Data Structures*, pages 243–260. Springer, 2013.

[5] Federico Giannoni, Marco Mancini, and Federico Marinelli. Anomaly detection models for iot time series data. *arXiv preprint arXiv:1812.00890*, 2018.

[6] Mustafa Gul and F Necati Catbas. Statistical pattern recognition for structural health monitoring using time series modeling: Theory and experimental verifications. *Mechanical Systems and Signal Processing*, 23(7):2192–2204, 2009.

[7] Manish Gupta, Jing Gao, Charu C Aggarwal, and Jiawei Han. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and data Engineering*, 26(9):2250–2267, 2013.

[8] Milos Hauskrecht, Iyad Batal, Michal Valko, Shyam Visweswaran, Gregory F Cooper, and Gilles Clermont. Outlier detection for patient monitoring and alerting. *Journal of biomedical informatics*, 46(1):47–55, 2013.

[9] Yann-Aël Le Borgne, Jean-Michel Dricot, and Gianluca Bontempi. Principal component aggregation for energy efficient information extraction in wireless sensor networks. *Knowledge Discovery from Sensor Data*, 2007.

[10] Xiaolei Li, Zhenhui Li, Jiawei Han, and Jae-Gil Lee. Temporal outlier detection in vehicle traffic data. In *2009 IEEE 25th International Conference on Data Engineering*, pages 1319–1322. IEEE, 2009.

[11] Alberto Luceño. Detecting possibly non-consecutive outliers in industrial time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(2):295–310, 1998.

[12] Jose E Bessa Maia, Angelo Brayner, and Fernando Rodrigues. A framework for processing complex queries in wireless sensor networks. *ACM SIGAPP Applied Computing Review*, 13(2):30–41, 2013.

[13] Allan D McQuarrie and Chih-Ling Tsai. Outlier detections in autoregressive models. *Journal of Computational and Graphical Statistics*, 12(2):450–471, 2003.

[14] Saeed Mehrang, Elina Helander, Misha Pavel, Angela Chieh, and Ilkka Korhonen. Outlier detection in weight time series of connected scales. In *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1489–1496. IEEE, 2015.

[15] Flávio Nunes and José Maia. Continuous monitoring in wireless sensor networks: A fuzzy-probabilistic approach. In *Anais do XVI Encontro Nacional de Inteligência Artificial e Computacional*, pages 96–107. SBC, 2019.

[16] Giuseppe Nunnari and Valeria Nunnari. Forecasting monthly sales retail time series: A case study. In *2017 IEEE 19th Conference on Business Informatics (CBI)*, volume 1, pages 1–6. IEEE, 2017.

[17] Bernard Rosner et al. Fundamentals of biostatistics, 2011.

[18] Peter J Rousseeuw and Annick M Leroy. *Robust regression and outlier detection*, volume 589. John wiley & sons, 2005.

[19] Maximilian Schmidt and Marko Simic. Normalizing flows for novelty detection in industrial time series data. *arXiv preprint arXiv:1906.06904*, 2019.

[20] Jethro Shell, Simon Coupland, and Eric Goodyer. Fuzzy data fusion for fault detection in wireless sensor networks. In *2010 UK Workshop on Computational Intelligence (UKCI)*, pages 1–6. IEEE, 2010.

[21] Jun-ichi Takeuchi and Kenji Yamanishi. A unifying framework for detecting outliers and change points from time series. *IEEE transactions on Knowledge and Data Engineering*, 18(4):482–492, 2006.

[22] Yee Lin Tan, Vivek Sehgal, and Hamid Haidarian Shahri. Sensoclean: Handling noisy and incomplete data in sensor networks using modeling. *Main*, pages 1–18, 2005.

[23] Jussi Tolvi et al. Outliers in eleven finnish macroeconomic time series. *Finnish Economic Papers*, 14(1):14–32, 2001.

[24] Bin Wang, Xiao-Chun Yang, Guo-Ren Wang, and Ge Yu. Outlier detection over sliding windows for probabilistic data streams. *Journal of Computer Science and Technology*, 25(3):389–400, 2010.

[25] Christine Wright and David Booth. Water treatment control using the joint estimation outlier detection method. *Environmental Modeling & Assessment*, 6(1):77–82, 2001.

[26] Kenji Yamanishi and Jun-ichi Takeuchi. A unifying framework for detecting outliers and change points from nonstationary time series data. In *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 676–681, 2002.

[27] Chunyong Yin, Sun Zhang, Jin Wang, and Neal N Xiong. Anomaly detection based on convolutional recurrent autoencoder for iot time series. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2020.

[28] Yufeng Yu, Yuelong Zhu, Shijin Li, and Dingsheng Wan. Time series outlier detection based on sliding window prediction. *Mathematical problems in Engineering*, 2014, 2014.

[29] Yusheng Zhou, Rufu Qin, Huiping Xu, Shazia Sadiq, and Yang Yu. A data quality control method for seafloor observatories: the application of observed time series data in the east china sea. *Sensors*, 18(8):2628, 2018.