Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2019) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Industry 4.0 and Smart Manufacturing

Decay-parameter Diagnosis in Industrial Domains by Robustness through Isotonic Regression

Salma Mahmoud[a], Florian Sobieczky[a,*], Jorge Martinez-Gil[a], Patrick Praher[a], Bernhard Freudenthaler[a]

[a] *Software Competence Center Hagenberg (SCCH) GmbH, Softwarepark 21, 4232 Hagenberg, Austria*

Abstract

In various industrial production environments a burn-in phase of a specific settling-length precedes a stable (i.e. steady-state, stationary) process mode. The identification of the corresponding physical parameters may be difficult to perform in the presence of strong noise. We propose a method using isotonic regression which circumvents the negative effects of heteroscedasticity related to naive estimation procedures adding robustness against different occurrences of scale on which the run-in effect is observed.

Keywords: Industry 4.0; Decision Support Systems; Failure Detection; Failure Analysis; Isotonic Regression

1. Introduction

Industry 4.0 focuses on digital technology and the connection between physical objects like sensors or machines and the Internet. One of the fields of application for industry 4.0 is predictive maintenance. For the idea of Industry 4.0 to become a reality, much work remains to be done in developing methods for identifying and predicting failures so that industrial process improves their reliability [13]. In this context, our approach to diagnose anomalies associated with the initial, non-stationary run-in phenomena in production processes will shed light on what time series analysis can contribute to computerization of manufacturing. We propose a method determining the corresponding settling-time which is specifically robust with respect to varying time-scales at which these effects appear.

1.1. Problem Description

Initial phenomena of a stochastic process related to the gradual assimilation of a stationary working mode may for example be related to transients in oscillations [16], warm-up phases in temperature sensitive dynamics [40],

* Corresp. Author Tel.: +43-50-343-839, +43-50-343-849.
E-mail address: florian.sobieczky@scch.at, salma.mahmoud@scch.at

1877-0509

initial concentration levels in chemical reactions [36], or transitions into mechanical equilibrium of hydrological lubrication of friction pairs [21] in production machinery. Trimmed mean, as well as quantile regression methods have successfully been applied for rendering decay parameter estimations more robust [9]. More recently, isotonic regression has been proposed for exponential parameter estimation methods overcoming the usual order restrictions [31]. The reciprocal of this parameter then acts as a characteristic measure of the order of magnitude of the settling-time. E.g. in control theory, it is defined to be proportional to the reciprocal of the product of natural frequency and damping factor [19]. In particular, for reliability estimation [4], the isotonic maximum likelyhood estimators have been shown to be superior to standard MLE of 'time-until-failure'. In the present approach, however, the applications are concerned with a monotonic approach of equilibrium, i.e. monotonic functions of time with the addition of noise and outliers, to which monotonic regression is reasonably applied. We consider robustness problems against (**A**) simple outlier-noise, and (**B**) robustness against the ignorance of the scale on which the noisy run-in process is to appear.

1.2. Role of robustness of our approach

Different approaches have been proposed to handle outliers in the training data [33]. Another difficulty involving the lack of robustness is the lack of reliability of parameter estimations in the presence of noise [25]. In particular, small deviations of the parameter probability distribution can cause disproportionately large changes in the estimates of these parameters [42]. In this work, we aim at designing a robust parameter estimation algorithm of run-in anomalies. Specifically, for exponentially decaying anomalies in the initial phase of an asymptotically stationary process, we present a method that reliably estimates the decay parameter on much larger regimes of possible values than the conventional 'naive' exponential regression technique. Our method uses isotonic regression, which, even though not robust to outliers (which we discuss), presents an improvement in terms of the size of the possibly occurring settling-times, before the process assumes a steady state. Such an algorithm makes it possible to eliminate the effects of the run-in period and its interference with other anomalies which may have to be anticipated for maintenance interventions. Our contribution is a method for estimating the anomalies associated with the initial start-up (burn-in, run-in, transients, warm-up etc.) processes that may precede the stationary mode of operation. Since it is usually not known in advance on what time-scale these preceding anomalies act, it is desirable to have a method diagnosing them which is invariant against strongly changing lengths of this characteristic 'settling- time'. This specific form of robustness is the main purpose of this work.

The paper is structured as follows: Section 2 presents the state-of-the-art in relation to the use of the related regression analysis to cope with anomaly diagnosis scenarios in the industrial domain. Section 3 presents the mathematical definition of our approach, i.e. we formulate the related regression problems as linear programs and analyze their associated robustness properties in comparison with the naive approach. Section 4 contains the experiments on artificial and real data associated with tasks **(A)** and **(B)**, followed by the discussion of the results in Section 5.

2. Parameter Estimation for Production Processes

2.1. Frameworks for applying robust parameter estimation

In recent years, data-monitoring in the production industry has made great advances in terms of the ability to document production steps with an arbitrarily high number of sensors. While in the early days, only malfunctioning production history was 'stored' and analyzed, today the observation of also the well-working production modes are performed, as a matter of a standard routine. So, since the data is available on a large scale, its availability has revolutionized manufacturing by offering companies the ability to implement machine-to-machine (M2M) or machine-to-human (M2H) communication and analytical technologies to increase productivity to the highest levels ever known [13]. However, in spite of this enormous progress, there are still several challenges in the way these data are analyzed and used in a prescriptive way that needs to be faced by factories of the future [30].

One of the great challenges of current industrial research is the inability to predict failures before they occur [29]. Moreover, these questions of so-called predictive maintenance [23] are strongly intertwined with the strong increase in the complexity of cyber-physical systems. While the anticipation of failures has always been a very eminent area of study, the emergence of high performing machine learning predictive tools (such as deep learning) has lifted the strategy of maintenance from the case-by-case basis to that of a prediction problem. Both, model predictive control approaches [39], as well as survival analysis techniques [26]. In both of these approaches, robustness plays an eminent role for the potentially positive impact that they can have to detect the malfunctioning of machinery before it fails and causes damage [1].

In this regard, the statistical research community has been working repeatedly on the design of robust methods that can properly anticipate failures [18, 27, 28]. In principle, this may seem like an easy task, but due to robustness problems related to parameter estimation in a noise-affected environment [33], it may be difficult to be put into practice. In particular, standard regression schemes (such as exponential regression [26]) may fail because of the difficulty to obtain reliable estimates of parameters describing the specific types of failures that are going to happen.

2.2. Process decomposition: Systematic part, noise, anomaly

First, it is necessary to define what is commonly understood by failure. Moreover, although there is no broad consensus on its definition, most authors agree that failure is anything deemed to occur when the system experiences an abnormal condition, such as a malfunction in the actuators or sensors. In the last years, there has been some research in the direction of investigating new robust statistical techniques to estimate the dependence of a failure of several factors associated with the system and that can be monitored in real-time, with robust regression being one of the most promising techniques. In robust statistics, robust regression is a form of regression analysis designed to circumvent some traditional limitations of parametric and non-parametric methods. Simple least squares regression schemes are not robust to outliers. Following [43], one possible definition of an outlier is observations that do not follow the pattern of the other observations, even though this does not provide a prescription of its detection.

We think that a good solution for this type of problem can be made through regression techniques [2] along with a definition of an 'anomalous observation', or just 'anomaly'. Given a real-valued discrete-time signal, which we view as a standard [6],[15] stochastic process $X_t : \Omega \to \mathbb{R}, t \in \mathbb{N}$ on a suitable probability space of paths $(\Omega = \mathbb{R}^{\mathbb{N}}, \mathcal{F}, \mu)$ with the product-sigma algebra $\mathcal{F} = \otimes_{i\in\mathbb{N}}\mathcal{B}(\mathbb{R})$ of the Borel σ-algebras of the real numbers $\mathbb{R}$ and a probability measure determined by the finite dimensional distributions, i.e. probabilities $\mu(\{X_{t_i} \in B_i\}_{i\in I})$ with $B_i \in \mathcal{B}(\mathbb{R})$, $|I| < \infty$), of which we assume that it possesses the decomposition into a systematic (non-random) part $f : \mathbb{N} \to \mathbb{R}$, some additive 'noise' $W_t : \Omega \to \mathbb{R}$, and some random process $I_t : \Omega \to \mathbb{R}$

$$X_t(\omega) \ = \ f(t) \ + \ W_t(\omega) \ + \ I_t(\omega). \tag{1}$$

By choosing $W_t(\omega)$ to fulfill $\mathbb{E}[W_t] = 0$, this lets us define the anomaly to be $I_t(\omega)$, i.e. the process describing the deviation of X_t of the systematic part $f(t)$, 'in expectation'.

2.3. Regression Models

Regression models allow predicting the value of a continuous variable based on other variables [22]. As in the previous point, the rationale behind regression is to learn from the relationships of variables in the past by inferring that those relationships will be preserved in the future. This group would include information as useful as the analysis of the remaining lifetime of a mechanical component. Our algorithm here depends on Isotonic regression rather than exponential one.

Exponential regression assumes the data includes a non-random component (trend) which follows an exponential curve. In its simplest form a linear model is fitted to the logarithm of the given data [26]. This transformation linearises the data, but also introduces a non-homogeneous scaling of the variance leading to heteroscedasticity[22]. Moreover, the logarithm only taking positive values as an argument for the (inverse) data-transformation makes it impossible to consider data which by the shift of the noise isn't within this domain. Naturally, smoothing procedures regularising the time series can solve this problem, if the scale parameter ('span' in the case of LOESS; 'f' in [12]) is known. Robustness against the occurrence of a large variety of scales involves comparatively complicated adaptive scale parameter estimation schemes [41].

Isotonic regression [37] or monotonic regression is the technique of fitting a free-form line to a sequence of observations under some ordering constraints of the fitted values [5]. In its simplest form, the fitted free-form line has to be non-decreasing everywhere, and fulfill a least-squares condition of the residuals [32].

In the case of run-in anomalies, it is often necessary to react to the non-linear shape of the underlying systematic trend by a suitably chosen data-transformation in order to linearise the data and make standard regression techniques applicable (as in exponential regression). Isotonic regression, however, allows direct application of minimizing the sum of squared residuals. The specific constraints, if used in a suitable way, already deliver much of the solution of the task imposed by the occurring non-linearities. The a-priori knowledge about where the systematic part of the signal is monotone can be used in an essential advantageous manner to make isotonic regression the method of choice.

3. 'Run-in' anomaly diagnosis in industrial domains

We now define our approach to estimate the parameter characterising a burn-in anomaly. In the first part, we assume an exponential decay anomaly with characteristic exponential decay behaviour. In the second section we generalise the approach to arbitrary anomaly shapes with a scaling parameter (λ) describing the extent of the run-in period.

The estimation procedure is defined by transforming the data with run-in anomalies into a form allowing parameter estimation with linear regression. It is based on a model of the general functional form $I_\lambda(t) = I(\lambda t)$ of the anomaly. By modeling the inverse image of the given data under $f(t) = f_1(t)$ by linear regression and with special care to minimize the bias introduced by strong additive noise, it is possible to retrieve a robust image of the scaling parameter λ.

We start by considering a signal of type

$$X_t(\omega) \quad = \quad f(t) \ + \ W_t(\omega) \ + \ I_t \tag{2}$$

where $W_t \sim \mathcal{N}(0, \sigma)$ is Gaussian noise, and $I_t = a\exp(-\lambda t)$ is chosen non-random, here, of which we assume to know only the amplitude a. (Note the strictly decreasing form of the considered anomaly.)

Let $\mathbf{X} = (t_i, X_{t_i})_{i=1}^N$ be a sample of size N. The following protocol defines our **isotonic regression method**:

- [1] Perform isotonic regression on $\mathbf{X}$ under the constraints of non-increasing values Y_t, and initial value Y_0.
- [2] Using the interval $[0, T]$ of times t for which $Y_t > 0$, define $Z_t = \log(Y_t/Y_0)$.
- [3] Apply a linear model for $Z_t, t \in [0, T]$ to receive the estimate $\widehat{\lambda}$ as the linear regression coefficient.

Note that the set of points $[0, T]$ on which the inverse of $f(t)$ exists is an interval, due to the decreasing nature of the solution Y_t to the isotonic regression problem. For comparison of this specific approach with a more straightforward, but naive solution to this parameter estimation problem, we consider a standard method using a smoothing filter, before applying the inverse transformation.

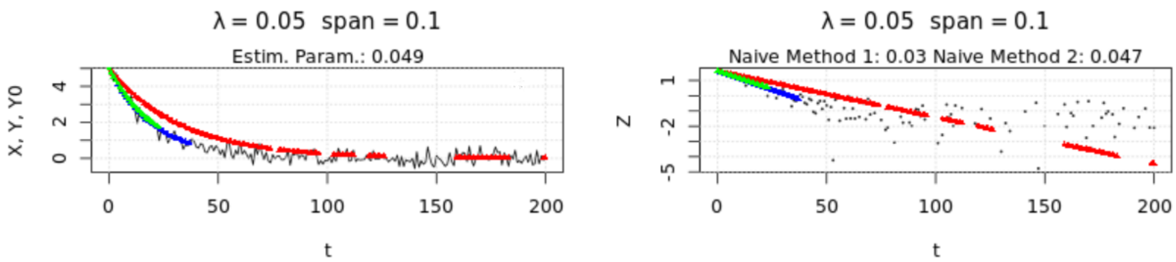

Fig. 1: LEFT: Linear Scale. The blue curve is the function $f(\widehat{\lambda} t) = ae^{-\widehat{\lambda} t}$ for the estimated value of the true $\lambda = 0.05$, while the magenta and green curves represent the naive solution using LOESS-filtering. Red: All points at t with $Y^{(0)} > 0$ have been considered. Green: Only points up until the first occurrence of a negative argument $Y_t^{(0)}$ of the log is included in the linear model. It is seen that the major influence for bias of the naive solutions is the large amount of data excluded with negative values of $Y_t^{(0)}$ (Red), even for the relatively small size of the interval of considered points (Green). RIGHT: Same result on logarithmic scale, only with data $X_t > 0$ shown.

Namely, the LOESS filter, as a local finite degree polynomial regression is applied to the initial data X_t, with a hitherto unknown 'span'-parameter for the smoother's 'stiffness' (the extent of the support of the convolution-kernel).

- (1) Calculate the LOESS-smoothing filtered version $Y_t^{(0)}$ of X_t.

- (2) Using the set $\mathcal{T}$ of time values for which $Y_t^{(0)} > 0$, define $Z_t^{(0)} = \log(Y_t^{(0)}/Y_0^{(0)})$, with $t \in \mathcal{T}$.
- (3) Apply a linear model for $Z_t^{(0)}, t \in \mathcal{T}$ to receive the estimate $\widehat{\lambda_0}$ as the linear regression coefficient.

We will refer to this as the **Naive Method 1**. An improved modification of it, called **Naive Method 2**, differs only in the choice of $\mathcal{T}$ in step (2) not including all data with $X_t > 0$, but only the initial uninterrupted interval of time-values of such (positive) data points.

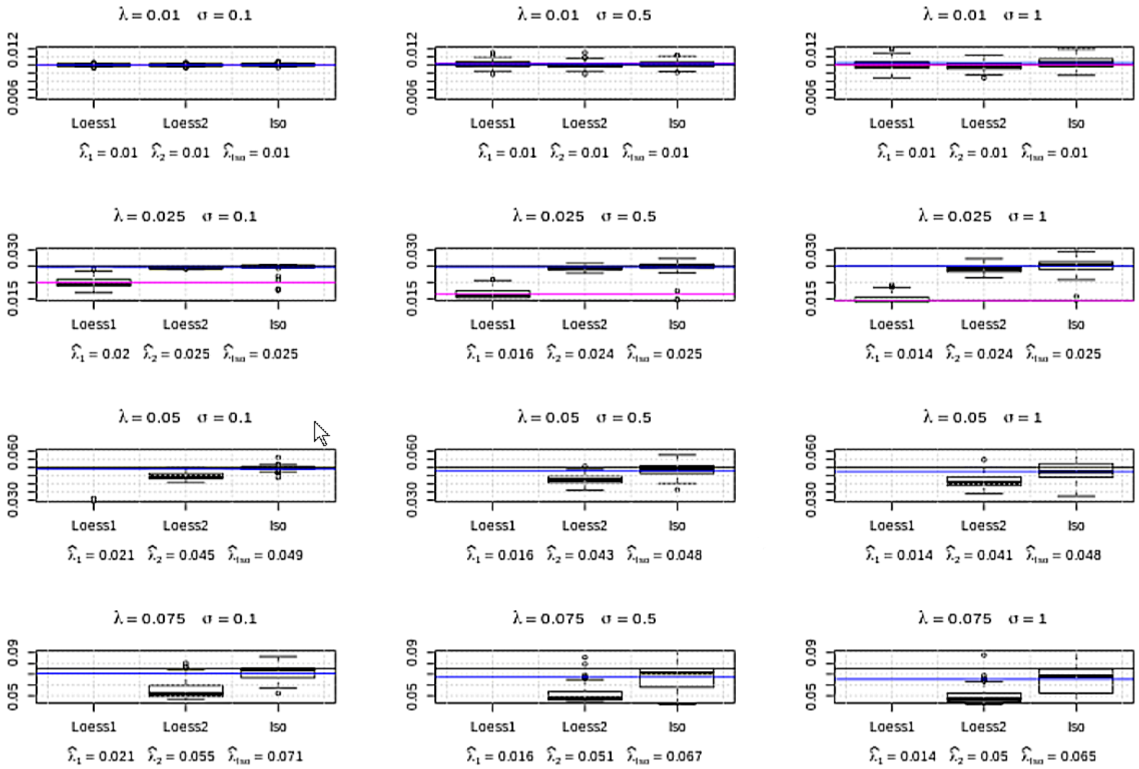

Fig. 2: Twelve runs of the experiment through 3 degrees of noise σ (**columns**) and 4 different scaling parameters λ (**lines**), with a constant span-parameter for the Loess filter of 0.1, for each of which hundred independent noise functions have been sampled. In each instance, the first two boxplots represent **Naive Method 1** (mean shown by **magenta line**) and **Naive Method 2**, while the third boxplot represents our method using isotonic regression. Comparison of the parameter estimates shown underneath the boxplots documents the isotonic regression method to be mostly superior, and to deliver reasonable estimates (the confidence interval containing the target λ value represented by the **blue line**) on a much wider range of parameters λ and noise levels σ than either of the naive methods. (The vertical axes show the scale of the scaling parameter λ.)

It is seen from Figure 1 that the choice of data points $\mathcal{T}$ (which excludes those not in the domain of the logarithm) has a crucial bias. Any parameter estimation method using the (second) data-transformation step has to cope with the values which noise has shifted outside of the range of values accessible by the linearising transformation (the inverse function of $t \mapsto f_\lambda(t)$). In these applications, it therefore has shown to be advantageous to be more strict about the choice of which t to include in $\mathcal{T}$ (see 'Results', end of Section 4.1). Whatever additional constraint for this set has been used, however, the isotonic regression method has always shown to be more accurate than the two smoothed exponential regression methods.

Most importantly regarding the robustness of the estimation method, choosing a parameter (as 'span' in the LOESS-Filter) is not required for our technique. In particular, if the same quality was to be achieved with the LOESS-methods, a considerable amount of additional parameter estimation would have to be invested to pick the right value for span (cmp [41] for smoothing parameter estimation in the framework of generalized additive models).

The fact that the (blue) isotonic regression solution is correctly mapping the decay has to do with the lack of 'stiffness' present in the LOESS-approach, even though the anomaly only consumes a small part of the total considered time interval. This independence of a geometric parameter characterising the scale is the main reason for the robustness of the isotonic regression method.

4. Experiments

The advantage of the application of the isotonic regression method to monotonic anomalies is documented here in the form of two use cases. Please note, that we are particularly interested in accuracy, and robustness.

4.1. Use Case **(A)** *- Observation of a calibrated counting process*

The data shown is taken from a project in which integer-valued counting events are estimated by the difference of a sensor's performance in action and the sensor's performance away from the environment of use. The sensor issues a non-zero noise-related default value, which has to be subtracted from the values obtained while actually measuring the number of counting events for calibration. Even though both of these two estimated counting events are increasing functions in time, the difference is not strictly increasing, as they are independently measured data sets. So, even though the process is expected to be an increasing integer number, the measured value maybe not strictly increasing due to this calibration process. This calibrated counting process is an application for the fault detection topic where these counts can be used to detect trends or outliers which may indicate a deficiency in the machines producing them. The project partner remains anonymous, as are the specific details of the nature of the process producing these events.

In these counting events experiment, we fit three different algorithms. Linear regression, isotonic regression, and quantile regression are compared in order to document how they perform with this data. Figure 3 shows the results that we have obtained for the baseline case. The data shows a particular incidence of run-in anomaly of exponentially decreasing intensity, merging into the steady state mode in which the registered counting events are linearly increasing.

In order to demonstrate their effect, we then proceeded by adding outliers, before we fitted the regression lines again, and measured the RMSE for these new lines' predictions of the original data without the outliers. The RMSE is used to measure the differences between values predicted by a model and the values observed [8]. In Figure 4,

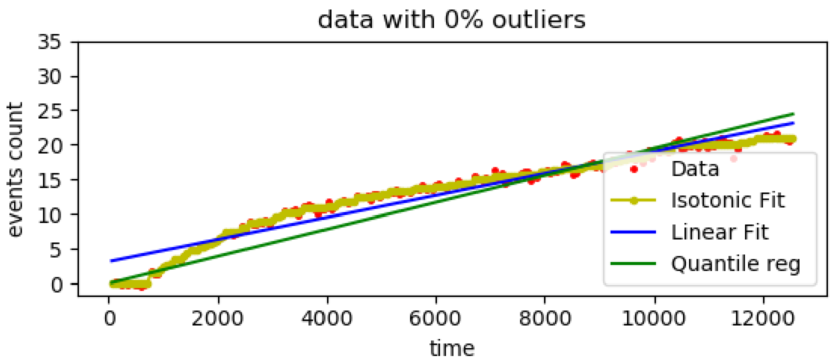

Fig. 3: Use Case **(A)**: This figure shows how the different kinds of regression fit the data points from an experiment with noisy data with a (negatively) curved systematic part.

we have added ten percent of outliers in the upper part. It is seen how the different regression methods react to this modification. We have repeated the same experiment but this time we have added the outliers in the bottom part so that we can observe the behavior of the three different kinds of regression. Figure 4 shows the repetition of this experiment with twenty percent outliers added to the data.

Table 1 shows the results of root-mean-square error (RMSE) for the three methods under study. As it becomes obvious, isotonic regression is best in the low noise regime. It remains to be the least affected in the case of the outliers added *below* the increasing curve with *negative* curvature.

(a) (b)

Fig. 4: On 4a we have 10% of outliers on the upper direction while on 4b the outliers are presented downwards

Method	RMSE	RMSE 10% upwards	RMSE 10% downwards	RMSE 20% upwards	RMSE 20% downwards
Isotonic	0.01	0.08	0.18	0.12	0.25
Linear	0.07	0.09	0.19	0.12	0.27
Quantile	0.10	0.07	0.20	0.08	0.29

Table 1: Summary of the quality of the considered regression types under different degrees of outliers for noisy data with a curved trend. It is seen how isotonic regression looses its advantage over standard linear regression, and quantile regression with increasing level of outlier-concentration, while a slight advantage of its use remains if the outlier is towards the 'inner' (here: lower) side of the curved data. Even though the RMSE is much higher for data with outliers (showing the principal non-robustness of isotonic regression), its slight advantage in the case of outlyingness towards the 'inner side of the curve' is visible.

It seems clear that the overall increasing nature of the data points in this data-set makes it reasonable to apply isotonic regression as the solution to the isotonic regression method is characterized by being a monotonic function. Quantile regression estimates the conditional quantile or median instead of the mean which is used in Linear regression. Moreover, in order to verify the robustness of the three kinds of regression, we have proceeded to introduce extreme outliers to our data by sampling some data points and changing their values (here 1.5*maximum. value) and examine the effect of these outliers on the fitted function.

4.2. Use Case **(B)** *- Power absorbed during drilling*

Data from a production process involving drilling is considered for four different cycles over a period of time including the initial run-in process together with some length of steady-state operation (see Fig. 5). The data shows the momentarily applied power, controlled such as to keep approximately constant drilling speed. It is seen that for the case of span parameters large enough to guarantee a monotonic solution of the naive approach (**Naive Method 2** used here), the isotonic regression method is approximating the true data *significantly* better.

As the span parameter becomes smaller, the error of the LOESS-smoother becomes smaller, but the approximating solution of the naive method progressively develops more fluctuations, which the isotonic method inhibits due to its order restrictions. The result of the experiment is shown in Fig. 2 and compares two applications of Naive Method 2 with our isotonic regression approach. The first LOESS-method uses all positive smoothed values, while the second LOESS-method is restricted to *the time before a third of the first time the smoothed curve becomes negative*. For the isotonic regression, this (1/3) rule has also been applied (to the first time the isotonic regression solution becomes negative). The LOESS-methods are better in the range of small λ. The second LOESS-method remains comparable in the other regimes, but is clearly outperformed by the isotonic regression method in high noise and quickly decaying exponential anomalies.

5. Results and Discussion

A model based parameter estimation method has been suggested to register exponentially decaying run-in anomalies of real-valued time series production process data. The method outperforms standard smoothed exponential regression techniques both in accuracy and in terms of robustness against scale variability.

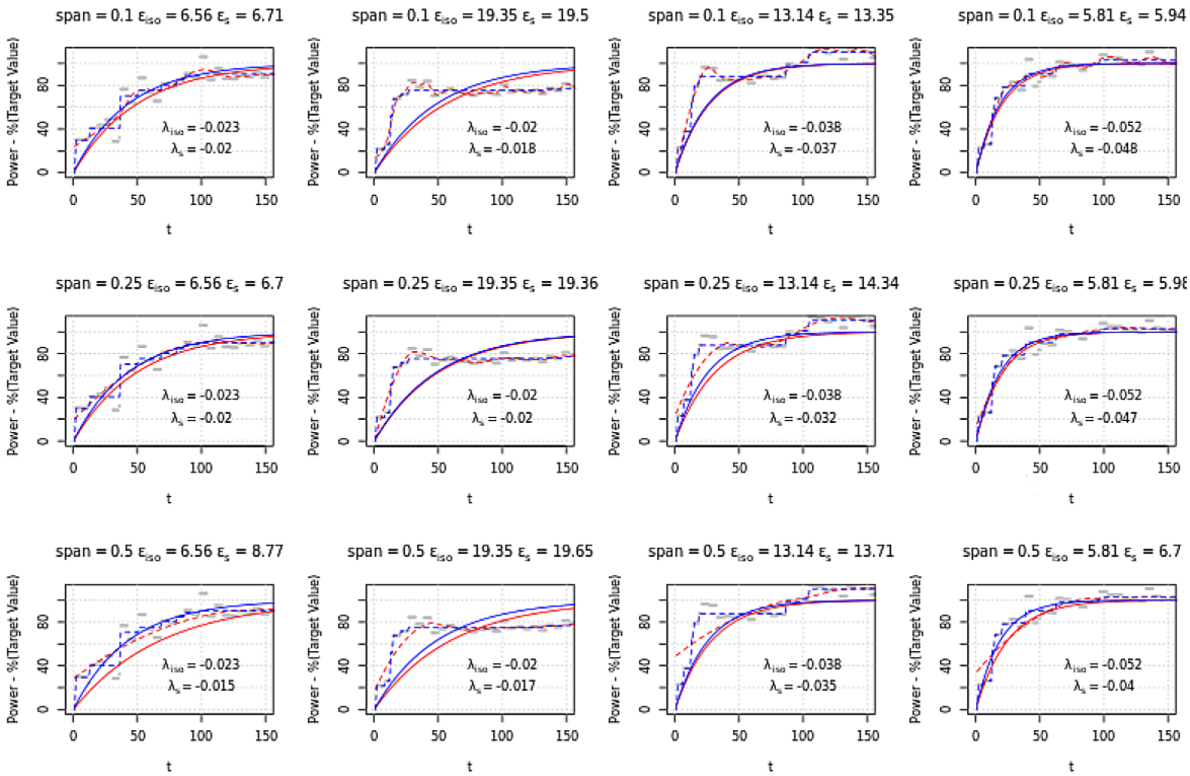

Fig. 5: Use Case **(B)**: Application of the LOESS- and Isotonic Regression method to measured power absorption data of an industrial drilling process (including the run-in effect) controlled to eventually have constant rotational drilling speed. Apart from the data (grey points), the solution of the Naive Method 2 is shown in solid **red**, and the solution of the isotonic method is shown in solid **blue**. The dashed red line is the LOESS-smoothed approximation of the data of step (2), while the dashed blue line represents the isotonic regression of step [1]. Four different measurements (**columns**) are shown, to which three different span-parameters of the LOESS method are applied (**lines**). The RMS-error from the original data is always smaller for isotonic regression (ε_{iso} vs. ε_s). It is particularly superior to the situation in which large non-monotonous fluctuations occur (2nd and 3rd columns). Note, that these can be interpreted as 'outliers' as considered in Use Case (**A**). Moreover, a larger stiffness of the smoother of the naive method, represented by a larger span-parameter λ, generally strongly increases the discrepancy between the results.

The discussion of Use Case **(A)** involving outliers particularly shows that the sensitivity of isotonic regression to outliers is high. In spite of this lack of robustness (comparable to that of simpler regression techniques), we were able to show its advantage for the special case of monotone anomalies in the presence of strong Gaussian noise. This even remains, if outliers remain on the 'inner' side of the curved underlying systematic part of the data. This demonstrates the advantage of knowledge about monotonic and convex behaviour of the systematic (non-noisy) part underlying the input data, to gain an improvement over conventional - if more generally applicable - regression techniques.

Another aspect is the non-parametric nature of the isotonic regression method, proposed here. While the theoretical results 1 show a clear advantage over the smoothed exponential regression technique for a fixed parameter of the LOESS-smoothing range (here: span=0.1), an adaption of this parameter may yield an improvement. However, even though this may be achieved through linear computational cost, the sensitivity to non-independent noise may increase the lack of robustness against outliers, significantly (see [17], Sect. 6).

Typically, robust regression works by assigning weights to data points: The weighting is usually performed automatically and by iteration through a process called iterative re-weighted least squares [22]. In the first iteration, each point is assigned the same weight and model coefficients are estimated using ordinary least squares. In subsequent iterations, the weights are recalculated so that the points furthest from the model predictions in the previous iteration have a lower weight. The model coefficients are then recalculated using weighted least squares. The process continues until the values of the coefficient estimates converge within a specified tolerance. As an example, consider the specific method of the Sawatzky-Golay Filter ([12], Sect. 2, p. 831), which is used here in the form of LOESS [35]. So, even though these filters are designed inherently with features enhancing its robustness, for 'monotone-before-noise' forms of process-data (such as the run-in anomalies, considered, here), the results of Use Case (**B**) show that the isotonic regression method surpasses the performance of these techniques in terms of robustness against changes of scale of the decay parameter, and even makes its estimation feasible in ranges of noise in which more naive methods fail, entirely. This and the discussion of Use Case (**A**) show, that in spite of the general lack of robustness against outliers, knowledge about the qualitative geometric form of the systematic part of the input data may be used advantageously.

6. Conclusions and Future Work

For monitoring of technical systems and collecting control information about sensor data, it is nowadays often necessary to use methods with the highest feasibility and versatility in terms of scale in-variance of the given data anomalies. For the example of exponentially decaying run-in effects we showed the advantages of isotonic regression over standard exponential regression and smoothing methods appearing by the use of knowledge of the input data's monotonicity properties. The results of Fig. 5 are real world data from an industrial drilling process. The increase in applicability to a wider range of scales (as in Fig. 5) makes our approach competitive in comparison with advanced scale sensitive methods [41, 9]. Our estimates of the run-in period in the form of the scaling parameter λ allow for more even control of a constant drilling performance, and should be interesting for applications such as [36, 3, 42]. As future work, we want to lay the foundation for the development of software applications of learning-based fault detection, diagnosis and prediction i.e. information systems for monitoring and analyzing technical systems involving robust statistical parameter estimation as an essential ingredient.

Acknowledgment

This work has been supported by the project **AutoDetect** (Project No. 862019; Innovative Upper Austria 2020 call Digitalization) has been funded by the Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK), the Federal Ministry for Digital and Economic Affairs (BMDW), and the Province of Upper Austria in the frame of the COMET - Competence Centers for Excellent Technologies Programme managed by Austrian Research Promotion Agency FFG.

References

[1] Adlemo, A., Andréasson, S., 1992. Models for fault tolerance in manufacturing systems. J. Intelligent Manufacturing 3, 1–10. URL: https://doi.org/10.1007/BF01471747, doi:10.1007/BF01471747.

[2] Alasalmi, T., Koskimäki, H., Suutala, J., Röning, J., 2018. Getting more out of small data sets - improving the calibration performance of isotonic regression by generating more data, in: Rocha, A.P., van den Herik, H.J. (Eds.), Proceedings of the 10th International Conference on Agents and Artificial Intelligence, ICAART 2018, Volume 2, Funchal, Madeira, Portugal, January 16-18, 2018, SciTePress. pp. 379–386.

[3] Antsalo, P., Makivirta, A., Valimaki, V., Peltonen, T., Karjalainen, M., 2001. Estimation of modal decay parameters from noisy response measurements, in: Audio Engineering Society Conv. 110, Audio Eng. Soc.

[4] Balakrishnan, N.; Beutner, E.K.M., 2009. Order restricted inference for exponential step-stress models. IEEE Transactions on Reliability 58. URL: http://gen.lib.rus.ec/scimag/index.php?s=10.1109/tr.2008.2011686, doi:10.1109/tr.2008.2011686.

[5] Barlow, R.E., Bartholomew, D.J., Bremner, J.M., Brunk, H.D., 1972. Statistical Inderence under Order Restrictions.

[6] Bauer, H., 2002. Measure and Integration Theory. De Gruyter.

[7] Best, M.J., Chakravarti, N., 1990. Active set algorithms for isotonic regression; A unifying framework. Math. Program. 47, 425–439. URL: https://doi.org/10.1007/BF01580873, doi:10.1007/BF01580873.

[8] Botchkarev, A., 2019. A new typology design of performance metrics to measure errors in machine learning regression algorithms, volume 14.

[9] Brazauskas, V., Serfling, R., 2000. Robust estimation of tail parameters for two-parameter pareto and exponential models via generalized quantile statistics. Extremes 3, 231–249.

[10] Bretthorst, G.L., Hutton, W.C., Garbow, J.R., Ackerman, J.J., 2005. Exponential parameter estimation (in nmr) using bayesian probability theory. Concepts in Magnetic Resonance Part A: An Educational Journal 27, 55–63.

[11] Chakravarti, N., 1989. Isotonic median regression: A linear programming approach. Math. Oper. Res. 14, 303–308. URL: https://doi.org/10.1287/moor.14.2.303, doi:10.1287/moor.14.2.303.

[12] Cleveland, W.S., 1981. Lowess: A program for smoothing scatterplots by robust locally weighted regression. The American Statistician 35, 54–54. URL: http://www.jstor.org/stable/2683591.

[13] Doltsinis, S., Ferreira, P., Mabkhot, M.M., Lohse, N., 2020. A decision support system for rapid ramp-up of industry 4.0 enabled production systems. Comput. Ind. 116, 103190. doi:10.1016/j.compind.2020.103190.
[14] Elsheikh, A., Yacout, S., Ouali, M., Shaban, Y., 2020. Failure time prediction using adaptive logical analysis of survival curves and multiple machining signals. J. Intelligent Manufacturing 31, 403–415. doi:10.1007/s10845-018-1453-4.
[15] Feller, W., 1949. On the theory of stochastic processes, with particular reference to applications, in: Proceedings of the [First] Berkeley Symposium on Mathematical Statistics and Probability, Univ. of Calif. Press, Berkeley. pp. 403–432. URL: https://projecteuclid.org/euclid.bsmsp/1166219215.
[16] Frank, K.T., Petrie, B., Fisher, J.A., Leggett, W.C., 2011. Transient dynamics of an altered large marine ecosystem. Nature 477, 86–91.
[17] Friedman, J., 1984. A Variable Span Smoother. Report from the Laboratory for Computational Statistics 5, SLAC PUB-3466. Stanford University. URL: https://www.slac.stanford.edu/cgi-bin/getdoc/slac-pub-3477.pdf.
[18] Fries, T.P., Graham, J.H., 2003. Fusion of soft and hard computing for fault diagnosis in manufacturing systems, in: Proceedings of the IEEE International Conference on Systems, Man & Cybernetics: Washington, D.C., USA, 5-8 October 2003, IEEE. pp. 114–119. URL: https://doi.org/10.1109/ICSMC.2003.1243801, doi:10.1109/ICSMC.2003.1243801.
[19] Gajic, Z., Lelic, M., 1972. Modern Control System Engineering. Prentice Hall.
[20] Garimella, R.V., 2017. A Simple Introduction to Moving Least Squares and Local Regression Estimation. Technical Report. URL: https://doi.org/10.2172/1367799, doi:10.2172/1367799.
[21] Glock, A.C., Sobieczky, F., Jech, M., 2019. Detection of anomalous events in the wear-behaviour of continuously recorded sliding friction pairs, in: Proc. of the ÖTG (Austrian Tribological Society). URL: https://www.oetg.at/publikationen/. ISBN 978-3-901657-6-2-7.
[22] Golberg, M., Cho, H., 2010. Introduction to Regression Analysis.
[23] Hashemian, H.M., 2011. State-of-the-art predictive maintenance techniques. IEEE Trans. Instrumentation and Measurement 60, 226–236. URL: https://doi.org/10.1109/TIM.2010.2047662, doi:10.1109/TIM.2010.2047662.
[24] Kotlowski, W., 2016. Online isotonic regression, in: Feldman, V., Rakhlin, A., Shamir, O. (Eds.), Proceedings of the 29th Conference on Learning Theory, COLT 2016, New York, USA, June 23-26, 2016, JMLR.org. pp. 1165–1189.
[25] Lazzerini, B., Volpi, S.L., 2013. Classif. ensembles to improve the robust. to noise of bearing fault diagnosis. Pattern Anal. Appl. 16, 235–251. doi:10.1007/s10044-011-0209-y.
[26] Liu, L., Gu, H.y., Xi, Y.g., 2007. Robust and stable scheduling of a single machine with random machine breakdowns. The International Journal of Advanced Manufacturing Technology 31, 645–654. doi:10.1007/s00170-005-0237-0.
[27] Locatelli, I., Marazzi, A., Yohai, V.J., 2011. Robust accelerated failure time regression. Comput. Stat. Data Anal. 55, 874–887. doi:10.1016/j.csda.2010.07.017.
[28] Marcos-Jorquera, D., Pérez, F.M., Gilart-Iglesias, V., Capella-D'Alton, A., 2006. Fault tolerance for manufacturing components, in: Proceedings of 11th IEEE International Conference on Emerging Technologies and Factory Automation, ETFA 2006, September 20-22, 2006, Diplomat Hotel Prague, Czech Republic, IEEE. pp. 1013–1016. URL: https://doi.org/10.1109/ETFA.2006.355396, doi:10.1109/ETFA.2006.355396.
[29] Marrison, N.A., 1992. Real time fault detection and diagnosis in dynamic engineering systems using constraint analysis. Ph.D. thesis. University of Glasgow, UK. URL: http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.309807.
[30] Mavrikios, D., Papakostas, N., Mourtzis, D., Chryssolouris, G., 2013. On industrial learning and training for the factories of the future: a conceptual, cognitive and technology framework. J. Intelligent Manufacturing 24, 473–485. URL: https://doi.org/10.1007/s10845-011-0590-9, doi:10.1007/s10845-011-0590-9.
[31] Nagatsuka, H., Uchino, M., Yamamoto, H., 2012. Parameter estimation of multivariate distributions under order restrictions of the parameters: an extension of isotonic regression. Quality Technology & Quantitative Management 9, 283–293.
[32] Rote, G., 2019. Isotonic regression by dynamic programming, in: Fineman, J.T., Mitzenmacher, M. (Eds.), 2nd Symposium on Simplicity in Algorithms, SOSA@SODA 2019, January 8-9, 2019 - San Diego, CA, USA, Schloss Dagstuhl - Leibniz-Zentrum für Informatik. pp. 1:1–1:18. doi:10.4230/OASIcs.SOSA.2019.1.
[33] Rousseeuw, P.J., Leroy, A.M., 2005. Robust regression and outlier detection. volume 589. John Wiley & sons.
[34] Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36, 1627–1639.
[35] Scrucca, L., 2011. Model-based sir for dimension reduction. Computational Statistics & Data Analysis 5, 3010–3026.
[36] Shen, Y., Zhu, Y., Ge, D., Shen, X., . Effects of Initial Concentration on Flocculation Size and Settling Velocity of Marine Hydraulic Fill Clay.
[37] Simone, V.D., Marino, M., Toraldo, G., 2009. Isotonic regression problems, in: Encyclopedia of Optimization, Second Edition, pp. 1774–1777. URL: https://doi.org/10.1007/978-0-387-74759-0_311, doi:10.1007/978-0-387-74759-0_311.
[38] Sobieczky, F., Shahriari, M., Freudenthaler, B., 2020. A data transformation for the estimation of decay types of multivariate distributions. URL: https://www.sciencedirect.com/science/article/pii/S2351978920305874.
[39] Su, Z., Núnez, A., Jamshidi, A., Baldi, S., Li, Z., Dollevoet, R., Schutter, B.D., 2015. Model predict. control for maint. operations planning of railway infrastructures. Technical Report 15-031. Delft Center for Systems and Control, NL.
[40] Takahashi, Y., Tomizuka, M., Auslander, D.M., 1974. Simple discrete control of indust. proc. (fin. time settling control algorithm for single-loop digital contr.). Journal of Dynamic Systems Measurement and Control 97, 354. doi:10.1115/1.3426950.
[41] Wood, S., Pya, N., Säfken, B., 1981. Smoothing parameter and model selection for general smooth models. Journal of the American Statistical Association 111, 1548–1575.
[42] Xu, C., Zhao, S., Liu, F., 2019. Sensor fault detection and diagnosis in the presence of outliers. Neurocomputing 349, 156–163. URL: https://doi.org/10.1016/j.neucom.2019.01.025, doi:10.1016/j.neucom.2019.01.025.
[43] Yu, C.W., Clarke, B., 2010. Asymptotics of bayesian median loss estimation. J. Multivariate Analysis 101, 1950–1958. URL: https://doi.org/10.1016/j.jmva.2010.04.013, doi:10.1016/j.jmva.2010.04.013.