Modeling and Generating Synthetic Anomalies for Energy and Power Time Series

Marian Turowski
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
marian.turowski@kit.edu

Moritz Weber
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
moritz.weber@kit.edu

Oliver Neumann
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
oliver.neumann@kit.edu

Benedikt Heidrich
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
benedikt.heidrich@kit.edu

Kaleb Phipps
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
kaleb.phipps@kit.edu

Hüseyin K. Çakmak
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
hueseyin.cakmak@kit.edu

Ralf Mikut
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
ralf.mikut@kit.edu

Veit Hagenmeyer
Institute for Automation and Applied Informatics
Karlsruhe Institute of Technology
Germany
veit.hagenmeyer@kit.edu

ABSTRACT

With the development of the smart grid, the number of recorded energy and power times series increases noticeably. This increase allows for the automation of smart grid applications such as load forecasting and load management. This automation, however, requires data that only represents the typical behavior of the system. To ensure that such data is available, detecting the anomalies often present in recorded data is important. As a result, anomaly detection methods are a recent research topic. However, their development is often limited by undefined anomaly characteristics and a lack of labeled anomalous data. To overcome this challenge, we propose a method that generates synthetic anomalies based on real-world anomalies that can be inserted into energy and power time series. For this, we analyze real energy and power time series to identify four types of commonly occurring anomalies. Given the identified anomaly types, we formally model each type and use these models to insert synthetic anomalies of each type into arbitrary energy or power time series. We show that our method is not only capable of generating synthetic anomalies with real-world properties, but also beneficial for training supervised anomaly detection methods.

CCS CONCEPTS

• **Computing methodologies** → **Anomaly detection**.

e-Energy '22, June 28–July 1, 2022, Virtual Event, USA

ACM ISBN 978-1-4503-9397-3/22/06.
https://doi.org/10.1145/3538637.3539760

KEYWORDS

anomaly, synthetic anomalies, anomaly detection, energy time series, power time series

ACM Reference Format:
Marian Turowski, Moritz Weber, Oliver Neumann, Benedikt Heidrich, Kaleb Phipps, Hüseyin K. Çakmak, Ralf Mikut, and Veit Hagenmeyer. 2022. Modeling and Generating Synthetic Anomalies for Energy and Power Time Series. In *The Thirteenth ACM International Conference on Future Energy Systems (e-Energy '22), June 28–July 1, 2022, Virtual Event, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3538637.3539760

1 INTRODUCTION

The increasing share of renewable energy sources in energy supply is accompanied by the implementation of smart grids [24]. A key element of smart grids are smart meters that record power or energy consumption and generation as time series [1]. Given the growing availability of these recorded time series, the goal of automating smart grid applications using these time series, e.g., load analysis, load forecasting, load management [44], becomes feasible. As these applications' performance depends on the quality of the input data, they rely on the assumption that the input data only reflects the normal behavior of the underlying system.

However, recorded time series typically contain anomalies [44], i.e., patterns that differ from "a well defined notion of normal behavior" [3, p. 15:2]. While anomalies occur due to various reasons such as theft [21], unusual consumption [37], and technical faults [28], they can all comprise data points or patterns representing wrong or misleading information, which can be particularly problematic for down-stream applications [44]. For example, spikes that violate the underlying distribution that corresponds to normal behavior

could be problematic. Using data containing these spikes can lead to inappropriate forecasting models, operational optimizations, and scheduling, which ultimately may affect energy system stability.

Especially in fully automated smart grid settings, reliably and effectively detecting anomalies is thus important, making anomaly detection in energy time series a recent research topic [17]. While a portion of existing works develop detection methods with an exploratory approach to discover anomalies contained in time series [e.g., 6, 9, 25], several works base their method development on knowledge about the anomalies to be detected [e.g., 15, 20, 21, 33]. However, gaining precise knowledge of the anomalies to be detected is associated with several challenges. By definition, anomalies are scarce and thus available datasets are imbalanced and there are comparatively few instances available in the time series that can be used for developing anomaly detection methods [46]. Furthermore, this scarcity is particularly problematic for promising deep learning anomaly detection methods that require large training datasets to perform well [30]. Additionally, a precise and comprehensive definition of relevant anomalies is missing [17]. Moreover, there is a lack of openly available energy time series with labeled anomalies or at least energy time series known to contain anomalies [16, 17].

In order to meet these challenges, different strategies can be applied for time series. Obviously, one can manually label anomalies in energy time series [33, 36, 47]. This strategy provides potentially very accurately labeled anomalies. However, it is limited to the anomalies contained in the time series, requires knowledge of the underlying system and typical patterns, often involves third parties such as facility managers or users, is time-consuming and costly, and potentially raises privacy concerns [12].

Alternatively, one can apply a means to define the majority of the time series as non-anomalous and the rest as anomalous, including selection [36], rules [15], statistical methods [12], or pattern recognition methods [47]. This strategy depends less on experts. However, it can also be limited to the anomalies contained in the time series, it requires a strong notion of non-anomalous time series, anomalies may remain hidden in the time series, and a time-consuming and costly verification by an expert could still be necessary.

Another way is to increase the number of available time series through generation, augmentation, or sampling methods, either assuming the used time series to be anomaly-free or reproducing time series with labeled anomalies [14, 22, 26, 46, 49]. This strategy allows one to control the number of available time series containing or not containing anomalies, while also requiring a strong notion of non-anomalous time series or time series with labeled anomalies.

Lastly, as a special case of augmentation, one can insert synthetic anomalies into existing time series [12, 36, 47]. This strategy requires anomalies that well resemble real-world anomalies [12, 36]. Being able to control the number and location of specified anomalies provides a properly defined object of investigation for anomaly detection methods [36] and turns an unsupervised into a supervised learning task [39]. Since this strategy can be applied to various datasets from a domain, it can also help increase the use of available, currently underutilized unlabeled datasets to develop anomaly detection methods [35].

Similar to other domains like intrusion detection [10, 34], security [31], and performance monitoring [42], synthetic anomalies are used to develop anomaly detection methods for energy time series [e.g., 8, 20, 21, 27, 41]. However, the inserted synthetic anomalies and their related parameters such as amplitude and quantity are generally not derived from real-world data and do not cover both energy and power, the typically recorded physical quantities. Furthermore, although [23] considers the insertion of anomalies for time series in general, the implementation of methods to generate the considered synthetic anomalies are not openly available and thus cannot yet be directly applied.

Therefore, the present paper proposes a method for generating four types of synthetic anomalies derived from real-world energy and power time series for developing anomaly detection methods. As a first step towards well-defined anomalies derived from real-world data, we identify anomalies in real-world time series containing energy and power measurements that are likely to be technical faults caused by the metering infrastructure and that may violate the underlying distribution corresponding to normal behavior. We then model the identified anomalies with parameters according to their characteristics observed in the considered real-world time series. Given the modeled anomalies, we are able to insert them as synthetic anomalies in an arbitrary time series containing energy and power measurements. We evaluate the identified and modeled anomalies in two ways. First, we examine whether inserted synthetic anomalies resemble the anomalies identified in real-world time series. Second, we show the benefit of inserted anomalies for training supervised anomaly detection methods.

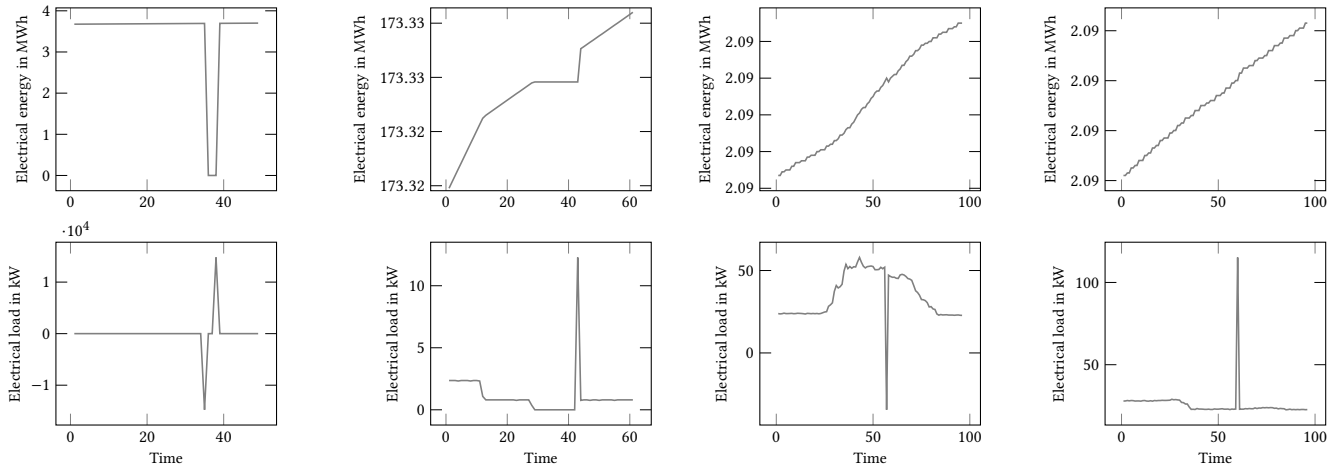
The remainder of the present paper is structured as follows. Section 2 introduces the anomalies identified in real-world time series, before Section 3 describes how the identified anomalies are modeled and inserted as synthetic anomalies in arbitrary time series. In Section 4, we evaluate the generated synthetic anomalies. In Section 5, we discuss the results and our method, before Section 6 concludes the paper.

2 IDENTIFYING REAL-WORLD ANOMALIES

In order to be able to generate realistic synthetic anomalies, we first derive anomalies from real-world time series containing energy and power measurements.

For this purpose, we consider electrical energy and power data collected at the KIT Campus North, which is a subset of the data described in [43]. This subset contains approximately 600 smart meter readings with a quarter-hourly resolution over a period of several years. Since these smart meters are installed in a variety of locations such as office buildings, industrial facilities, gas motors, and photovoltaic panels, their recorded data presents typical patterns and anomalies of consumers and producers found in an ordinary district. For each smart meter, a time series containing energy measurements and a time series containing power measurements is available.

By carefully visually examining these time series, we are able to (i) find typical patterns for each smart meter and (ii) identify unusual patterns across all smart meters. In addition to both, we make use of knowledge about the facilities recorded by the smart meters to distinguish between anomalies and normal behavior in each time series. For example, we expect the power consumption of an office building not to drop to zero while an automatic lighting system may have no power consumption at all during the day.



(a) Anomaly type 1: The energy time series drops to zero for at least one time step and then jumps back to a plausible new value, corresponding to a negative spike potentially followed by zero values and finally a positive spike in the power time series.

(b) Anomaly type 2: The gradient of the energy time series decreases and can fully stagnate for several time steps, before returning to the correct value. This corresponds to a drop to potentially zero followed by a positive spike in the power time series.

(c) Anomaly type 3: The energy time series dips suddenly, which corresponds to a sudden negative spike in the power time series.

(d) Anomaly type 4: The energy time series contains a sudden increase in gradient, corresponding to a sudden positive spike in the power time series.

Figure 1: Examples of the four different anomaly types that are identified in the selected real-world data.

During this examination, we focus on anomalies that are likely to be technical faults caused by the metering infrastructure and that may violate the underlying distribution corresponding to normal behavior with very low or high values.

Across all smart meters, we find that many of the observed anomalies can be assigned to one of four anomaly types (see Figure 1) that also match general classes of anomalies described in the literature (see Table 2 in Appendix B). For each identified anomaly type, we shortly describe its characteristics in time series containing energy and power measurements and provide a potential explanation considering the metering infrastructure in the following.

Anomaly type 1. Anomalies of type 1 are characterized by a drop to zero for at least one time step in the energy time series. After the zero values, the energy time series jumps back to a plausible new value (see Figure 1a). The corresponding power time series is characterized by a negative spike potentially followed by multiple zero power values and finally a positive spike. Anomalies of this type are likely to be caused by missing values in the recorded energy time series that are filled with zeros.

Anomaly type 2. Anomalies of type 2 are characterized by a noticeable decrease in the gradient of the energy time series. For one time step, there can be a decrease in the gradient that can be followed by constant energy values and that ends with a sharp increase in the gradient until a plausible new value is reached. Alternatively, there may be immediately constant energy values ending with an increased gradient (see Figure 1b). In the corresponding power time series, there is a drop followed by a positive spike. If the energy time series contains constant energy values, the power values drop to zero. In most of the observed cases, the height of the power spike is closely related to the length of the anomaly, suggesting that the power values of the constant sequence accumulate at one time step and thus form the spike. Anomalies of this type could be due to an interruption in the transmission of smart meter readings.

Anomaly type 3. Anomalies of type 3 are characterized by a sudden dip in the energy time series (see Figure 1c). In the corresponding power time series, there is a negative power spike at one time step. Since this spike is rather small in some occurrences and rather strong in others, we observe two cases, i.e., a slight and an extreme negative power spike. Anomalies of this type could occur due to an external adjustment of a smart meter such as a recalibration that aims to match the readings from multiple smart meters with a specific amount of energy. Anomalies with an extreme negative power spike could be caused by a reset of the respective smart meter.

Anomaly type 4. Anomalies of type 4 are characterized by a sudden increase in gradient of the energy time series relative to the quarter-hourly resolution (see Figure 1d). In the corresponding power time series, there is a positive power spike at one time step. Since this spike is also rather small in some occurrences and rather strong in others, we again observe two cases, i.e., a slight and an extreme positive power spike. Anomalies of this type can be caused by, for example, the change from daylight saving time to standard time. Because of this clock change by one hour, the consumption or generation within that hour is allocated to a single time step. This type of anomaly can also be observed in combination with anomalies of type 3, indicating an external adjustment of the smart meter.

We label anomalies of these four identified types in 50 one-year energy and one-year power time series. Although it is theoretically possible to derive energy time series to obtain the power time series, we simultaneously label anomalies in both time series to eliminate possible sources of error and guarantee reliable labels. To obtain the 50 time series, we randomly select 23 smart meters from the considered data for 2016, 21 for 2017, and three smart meters that are present in both 2016 and 2017. As shown in Table 1 in Appendix A, the 50 related energy and power time series of the selected smart meters are reasonably diverse, which is consistent with the fact that the used data set comprises smart meters at various locations.

3 MODELING AND GENERATING THE IDENTIFIED ANOMALY TYPES

To be able to generate the identified anomalies as synthetic anomalies, we need to model them and to design an insertion method. Modeling anomalies of the different types requires several parameters. Before describing the concrete modeling of each anomaly type, we briefly introduce the used parameters and how to set them.

For each previously identified anomaly type, we describe the necessary manipulation of values — despite their proportional physical relationship — independently of each other in a given arbitrary time series $E = e_1, e_2, \ldots, e_N$ containing energy measurements and a given arbitrary time series $P = p_1, p_2, \ldots, p_N$ containing power measurements. With the described manipulation, we replicate an anomaly $\hat{e}_{j,i}$ or $\hat{p}_{j,i}$ of type j with start index i.

While anomalies of types 1 and 2 have a length l, anomalies of types 3 and 4 affect all entries after the time series entry i in an energy time series and only the entry i itself in a power time series. More precisely, for anomalies of types 1 and 2, we assume the length $l \sim \mathcal{U}_{[l_{min}, l_{max}]}$ to be from a uniform distribution in an interval $[l_{min}, l_{max}]$. For anomalies of types 1 and 3 for power time series, we additionally consider the fact that the amount of energy at a given time step in the power time series in terms of the constant offset k is lost when deriving a power time series from an energy time series. For anomalies of these types, we thus explicitly consider the constant k, which has to be identical for all anomalies inserted into the same power time series, to better represent the characteristics of power time series. Moreover, anomalies of types 3 and 4 comprise the random value r that determines the amplitude of their spike. We assume to be from a uniform distribution in an interval $[r_{min}, r_{max}]$, i.e., $r \sim \mathcal{U}_{[r_{min}, r_{max}]}$.

To generate anomalies of the modeled types, all these described parameters need to be set. For this, they can either be determined from available labeled data (as, for example, done in Section 4.2) or from values reported in literature (e.g., in Table 3 in Appendix C).

Anomaly type 1. We reproduce anomalies of type 1 with length l in an energy time series E by setting the time series entries $\hat{e}_i$ to $\hat{e}_{i+l-1}$ to zero. We model anomalies of this type as

$$\hat{e}_{1,i+n} = 0, \quad 0 \leq n < l, \tag{1}$$

where l is the length of the anomaly.

In order to insert anomalies of type 1 with length l into a power time series P, we set the first anomalous entry $\hat{p}_i$ to the negative value of the power aggregated up to this time step i. The next $l-2$ entries are set to zero and the last entry of the anomaly $\hat{p}_{i+l-1}$ to the sum of the power aggregated up to time step $i+l-1$ corresponding to the jump in the energy time series. Formally, we describe this as

$$\hat{p}_{1,i+n} = \begin{cases} -1 \cdot (\sum_{t=1}^{i-1} p_t) - k, & n = 0 \\ 0, & 0 < n < l-1, \\ (\sum_{t=1}^{i+l-1} p_t) + k, & n = l-1 \end{cases} \tag{2}$$

where $l \geq 2$ is the anomaly's length and k is the constant offset.

Anomaly type 2. To replicate anomalies of type 2 with length l in an energy time series E, we determine the first anomalous value $\hat{e}_i$ as the average of the observed value at index i, e_i, and the previous value e_{i-1} weighted by the random number $r \sim \mathcal{U}_{[0,1)}$. All following $l-1$ anomalous entries are then set to this first anomalous value $\hat{e}_i$. Anomalies of this type can be described by

$$\hat{e}_{2,i+n} = r \cdot e_i + (1-r) \cdot e_{i-1}, \quad 0 \leq n < l, \tag{3}$$

where l is the length of the anomaly and $r \sim \mathcal{U}_{[0,1)}$ can be assumed from a uniform distribution and the same for all entries. Note that, in the special case of $r = 0$, the anomaly directly starts with the value of the previous time step.

To insert anomalies of type 2 with length l into a power time series P, we scale down the first anomalous entry $\hat{p}_i$ using a random number $r \sim \mathcal{U}_{[0,1)}$ and set the subsequent $l-2$ entries to zero. In order to form the observed peak at the last entry of the anomaly, we set the last entry $\hat{p}_{i+l-1}$ to the sum of the original values of the previously manipulated entries and subtract the first manipulated value $\hat{p}_i$. Formally, the anomaly can be described as

$$\hat{p}_{2,i+n} = \begin{cases} r \cdot p_i, & n = 0 \\ 0, & 0 < n < l-1, \\ (1-r) \cdot p_i + (\sum_{t=i+1}^{i+l-1} p_t), & n = l-1 \end{cases} \tag{4}$$

where $l \geq 2$ is the length of the anomaly and $r \sim \mathcal{U}_{[0,1)}$. Analogously to the energy time series, in the special case of $r = 0$, the manipulated entries directly start with a zero.

Anomaly type 3. We reproduce anomalies of type 3 in an energy time series E by subtracting a certain amount of energy from every time series entry with an index greater than or equal to i. The amount of energy to be subtracted depends on the observed case, i.e., the slight and the extreme negative spike. For the slight negative power spike, we insert anomalies of type 3 by subtracting a value based on the energy difference between the anomalous entry e_i and its predecessor e_{i-1} multiplied by a random value r. The anomaly can formally be described by

$$\hat{e}_{3,i+n} = e_{i+n} - r \cdot |e_i - e_{i-1}|, \quad n \geq 0, \tag{5}$$

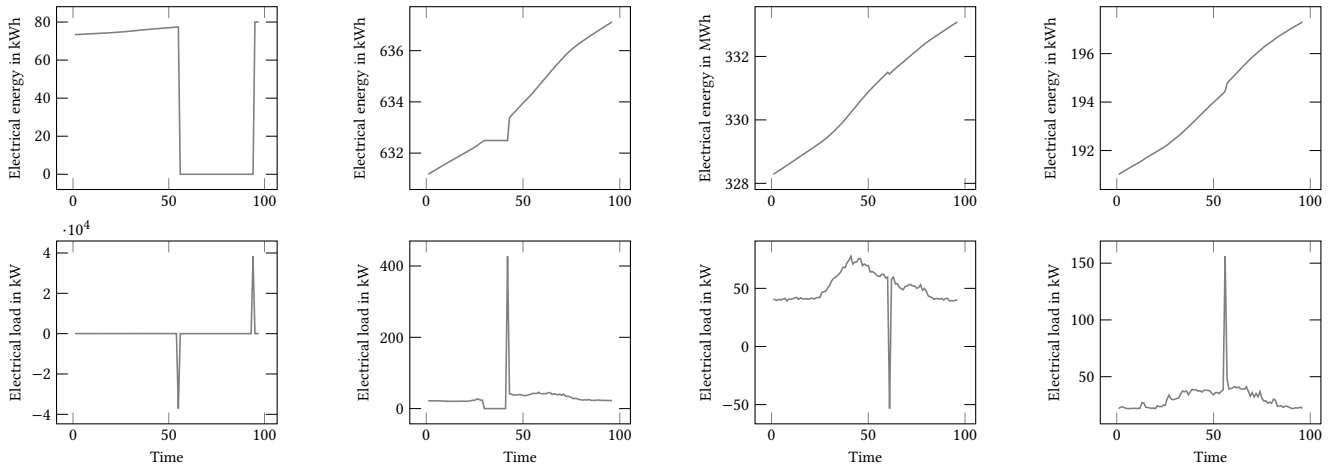
where r is the random value.

For the extreme negative power spike that is likely caused by a smart meter reset to zero, we subtract the value of the entry e_i from all subsequent time series entries, i.e.,

$$\hat{e}_{3,i+n} = e_{i+n} - e_i, \quad n \geq 0. \tag{6}$$

In order to insert anomalies of type 3 into a power time series P, we generate anomalies for the slight negative peak, by setting the anomalous entry $\hat{p}_i$ to the previous value p_{i-1} multiplied by a random value r. Formally, we describe this as

$$\hat{p}_{3,i} = -r \cdot p_{i-1}, \tag{7}$$



(a) Synthetic anomalies of type 1: The characteristic drop to zero in the energy time series and the first negative, then positive spikes in the power time series are clearly visible.

(b) Synthetic anomalies of type 2: We observe the characteristic stagnation in the gradient and following spring in the energy time series, as well as the drop and subsequent positive spike in the power time series.

(c) Synthetic anomalies of type 3: The characteristic decrease in the gradient of the energy time series and corresponding negative spike in the power time series are clearly observable.

(d) Synthetic anomalies of type 4: We observe the characteristic sudden increase in gradient of the energy time series and corresponding positive spike in the power time series.

Figure 2: Examples of generated synthetic anomalies for each the four modeled anomaly types inserted into exemplary data. The generated synthetic anomalies show similar characteristics as the identified real-world anomalies.

where r is the random value.

For the extreme negative spike corresponding to a drop of the energy time series to zero, we use

$$\hat{p}_{3,i} = -1 \cdot (\sum_{t=1}^{i-1} p_t) - k, \tag{8}$$

where k is the previously defined constant offset.

Anomaly type 4. To replicate anomalies of type 4, we apply a similar manipulation as for anomalies of type 3. To cover both the observed slight and extreme cases, we use two different sampling intervals for r. The anomaly is thus defined as

$$\hat{e}_{4,i+n} = e_{i+n} + r \cdot |e_i - e_{i-1}|, \quad n \geq 0, \tag{9}$$

where r is the random value sampled twice to represent the slight and the extreme case

To insert anomalies of type 4 into a power time series P, we model the observed positive spike $\hat{p}_i$ by multiplying its predecessor p_{i-1} with a random value r sampled from two different intervals. Again, to cover both the observed slight and the extreme case, we use two different sampling intervals for r. Formally, it is defined as

$$\hat{p}_{4,i} = r \cdot p_{i-1}, \tag{10}$$

where r is the random value sampled twice to represent the slight and the extreme case.

When generating anomalies of these four types, we need to consider the potential interaction between the modeled anomaly types. With regard to energy time series, one first has to insert anomalies of types 3 and 4 before anomalies of types 1 and 2 because anomalies of types 3 and 4 affect all values after their occurrence and thus potentially influence anomalies of the other types. Concerning power time series, one can, however, insert anomalies in the ascending order of the type. To avoid overlapping anomalies, we use a sequential approach. For each anomaly to be generated, we firstly search for an anomaly-free sequence $(x_i, ..., x_{i+l})$ in the time series X before we insert the anomaly. Figure 2 shows a synthetic anomaly of each anomaly type generated with this implementation.

To be able to reproducibly generate anomalies of the four types for an energy or power time series, we implement an openly available pipeline[1] using pyWATTS[2] [13]. It allows to control the types, the quantities, and the parameters of the synthetic anomalies that are inserted into an arbitrary energy or power time series.

4 EVALUATION

To evaluate the modeled anomalies, we perform a twofold evaluation. First, we examine whether generated synthetic anomalies resemble the anomalies identified in real-world time series. Second, we evaluate the benefit of synthetic anomalies for training supervised anomaly detection methods. Before describing both evaluations in detail, we introduce the selected data, the calculation of the parameters used for the evaluated generation method, the applied evaluation methods, and the experimental setting.

[1] https://github.com/KIT-IAI/GeneratingSyntheticEnergyPowerAnomalies
[2] https://github.com/KIT-IAI/pyWATTS

4.1 Used Data

For the evaluation, we also use the previously introduced electrical energy and power data collected at the KIT Campus North. More specifically, we again consider the previously labeled 50 time series for the evaluation because of the available labels for the related time series containing energy and power measurements. Since time series containing energy measurements are typically monotonically rising and thus non-stationary, one would usually apply differencing to make it stationary and thus useful for time series analyses [19]. As the already available time series containing power measurements are exactly the result of such a differencing due to the proportional physical relationship between energy and power, we focus on them in the following.

To obtain anomaly-free time series containing power measurements for the following analyses, we first use the corresponding manually labeled 50 one-year time series containing energy measurements. More precisely, we mark the labeled anomalies in these time series as missing values and apply the Copy-Paste Imputation (CPI) method [45]. The CPI method has shown a strong performance in imputing missing values with realistic patterns while preserving the amount of energy associated with the missing values. After imputing the anomalies marked as missing values in these energy time series, we calculate their derivative to obtain the corresponding anomaly-free power time series. We use the resulting anomaly-free power time series as the basis for inserting the generated synthetic anomalies used in the evaluation.

For the application of the selected evaluation methods, we finally create overlapping samples with a size of 96 from all considered power time series, namely the power time series containing identified anomalies, the power time series reproducing the identified anomalies with synthetic anomalies, and the power time series containing additional synthetic anomalies.

4.2 Used Anomaly Generation Parameters

We determine the parameters required for the generation of the desired synthetic anomalies from the selected labeled power time series. From these time series, we can directly determine the number of anomalies of all four types as well as the minimum and maximum length l_{min} and l_{max} of type 1 and 2 anomalies. For k, we use the first value in the corresponding available energy times for the comparison between synthetic and identified anomalies or set it to zero when evaluating the benefit of synthetic anomalies for the training of detection methods[3].

Lastly, we determine r_{min} and r_{max} for the slight and the extreme case for anomaly type 3 and 4 using DBSCAN [7]. For both anomaly types, we calculate $r = p_i / \overline{p}$ for all labeled anomalies of this type in the power time series[4], where $\overline{p} = (\sum_{t=i-5}^{i+5} p_t - p_i)/10$ is the local average with an arbitrarily selected range of 10. We cluster the result into two classes. For both types, we assume that the class with the majority of the considered anomalies represents the slight power spike case and the other the extreme spike case. For anomaly type 3, we thus select the smallest and the largest value in the majority class as r_{min} and r_{max}. For anomaly type 4, we select the smallest and the largest value from each class as r_{min} and r_{max} for the corresponding case.

Using this calculation, we aim to reproduce the anomalies contained in the original power time series with the parameters reported in Table 4 in Appendix C to examine whether the synthetic anomalies resemble the anomalies identified in the real-world time series. To evaluate the benefit of synthetic anomalies for training supervised anomaly detection methods, we additionally increase the number of anomalies compared to the original power time series by doubling the number of anomalies (see Table 5 in Appendix C). Note that, for both evaluations, we limit the minimum length of type 1 anomalies to 92 and type 2 anomalies to 44 to consider the imbalanced distributions of these lengths. Additionally, we insert only anomalies of the extreme case of types 3 and 4 as soon as one exists in the corresponding labeled power time series.

4.3 Applied methods

In the evaluation, we apply four different methods, which we describe in the following.

To examine whether the synthetic anomalies resemble the identified anomalies, we apply a statistical visualization and a discriminator method. As visualization method, we use the t-distributed stochastic neighbor embedding (t-SNE) [40]. The t-SNE visualizes high-dimensional data in a two-dimensional map such that similar data points are likely to appear close together and dissimilar data points far apart. As discriminator method, we implement a simple three-layered fully-connected Neural Network (NN) with ten neurons in the hidden layer and ReLU as activation function. In this NN, all neurons are interconnected across the layers and the neuron in the output layer determines whether the input data belong to the original data or not. For training the NN, we use the binary cross-entropy as loss and RMSprop [18] as optimizer.

To evaluate the benefit of synthetic anomalies for their training, we apply two supervised anomaly detection methods. More precisely, we select a kNN classifier and a decision tree classifier. The kNN classifier uses a proximity measure to classify a test sample based on the similarity of training instances [5]. In comparison, as a non-parametric method, the decision tree learns simple decision rules inferred from data features [2].

4.4 Experimental Setting

The experimental setting for the evaluation comprises the selected metrics and the used hard- and software.

Metrics. The evaluation is based on the two following metrics.

For the aforementioned discriminator method, we use the discriminative score. It is defined as

$$\text{Discriminative score} = |\text{Accuracy} - \text{Dummy}|, \tag{11}$$

where Accuracy is the result from the applied discriminator method.

For the supervised anomaly detection methods, we apply the commonly used F1-Score. It is the harmonic mean between precision and recall and is defined as

$$\text{F1-Score} = \frac{\text{TP}}{\text{TP} + \frac{1}{2} \cdot (\text{FP} + \text{FN})}, \tag{12}$$

[3] If energy time series are not available, one could sum the power over a year of data and multiply it by the presumed number of years the smart meter has been in service.

[4] Given energy time series, one could analogously calculate $r = \frac{e_{i-1} - e_i}{e_{i-1} - e_{i-2}}$ for anomalies of type 3 and $r = \frac{e_i - e_{i-1}}{e_{i-1} - e_{i-2}}$ for anomalies of type 4.

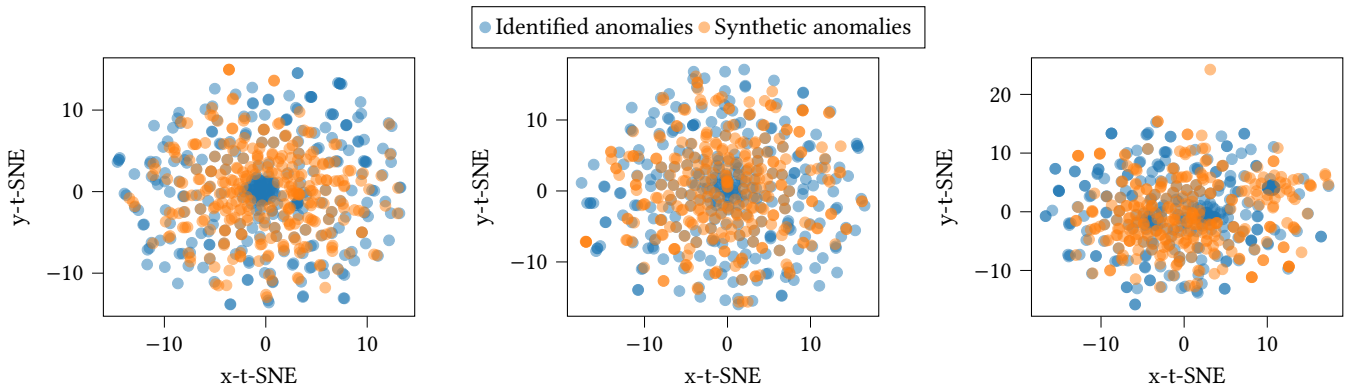


Figure 3: A t-SNE visualization of an identical number of samples containing identified anomalies and synthetic anomalies from three exemplary time series. For all three time series, the majority of the identified and inserted synthetic anomalies overlap, indicating that they have similar properties.

where *TP* are the true positives, *FP* the false positives, and *FN* the false negatives in the considered classification.

Hardware and software. Throughout the evaluation, we apply a standard computer with a four core i7 CPU and 16 GB of RAM. Moreover, all applied methods are implemented in Python. The t-SNE, the decision tree, and the kNN are implemented with SKLearn [32] and the fully-connected NN with Keras [4]. The evaluation is automated using these implementations with pyWATTS[5] [13].

4.5 Comparing Identified and Synthetic Anomalies

For the synthetic anomalies to be useful, they must resemble the identified real-world anomalies and ideally be indistinguishable from them. In this section, we first qualitatively compare identified and synthetic anomalies with the help of t-SNE visualizations, before quantitatively comparing them with the discriminator method.

Firstly, we examine the t-SNE visualizations of an identical number of samples containing identified and synthetic anomalies from three exemplary time series in two versions, with identified anomalies and with synthetic anomalies reproducing the identified anomalies.

As show in Figure 3, we observe that, for all three time series, the samples with identified anomalies and the samples with inserted synthetic anomalies overlap in most cases. The overlap indicates that the synthetic anomalies exhibit properties similar to the identified anomalies.

Secondly, we consider the discriminative score of the discriminator method in detecting the difference between identified anomalies and inserted synthetic anomalies. For this, we consider samples containing anomalies from all 50 considered time series in two versions. The first version comprises the power time series with identified anomalies, whereas the second version contains the anomaly-free power time series with inserted synthetic anomalies.

A histogram of the discriminative score is shown in Figure 4. The discriminative score rounded to one decimal digit, whose maximum is 0.5, is plotted on the x-axis to provide bins for the histogram and the number of occurrences in each bin is shown on the y-axis. We observe that the discriminative score is 0.25 or smaller for a large majority of the samples and higher for only few samples. This result indicates that the discriminator is mostly unable to differentiate between identified and synthetic anomalies.

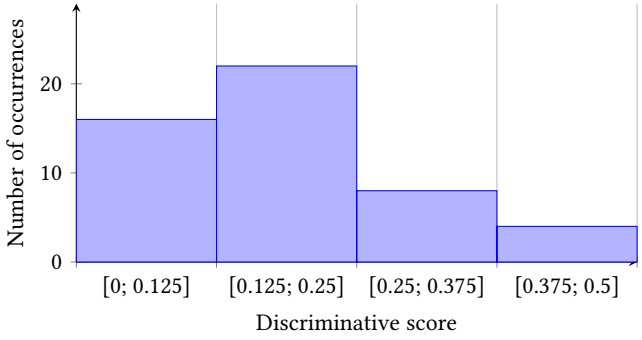


Figure 4: A histogram of the discriminative score of all samples containing identified or synthetic anomalies from the 50 considered time series. The x-axis shows the histogram bins for the discriminative score in steps of 0.125, whereas the number of occurrences in each bin is plotted on the y-axis. The majority of the samples has a discriminative score of 0.25 or smaller, indicating that the discriminator cannot distinguish between identified and synthetic anomalies.

4.6 Benefit of Synthetic Anomalies for Anomaly Detection

To evaluate whether synthetic anomalies exhibiting real-world characteristics are beneficial, we exemplarily analyze this benefit for training supervised anomaly detection methods. For this, we compare the detection performance of two different training strategies based on the F1-Score. The first strategy is Train Real Test Real

[5]https://github.com/KIT-IAI/pyWATTS

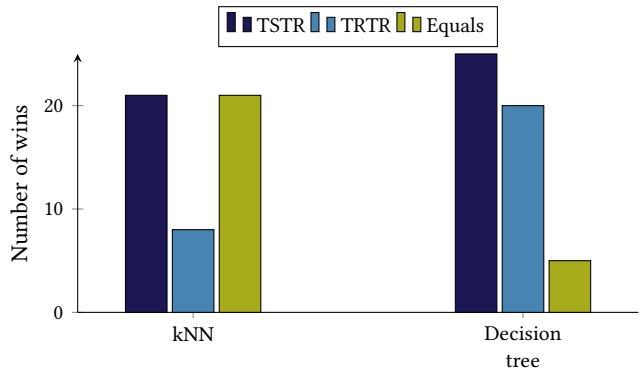


Figure 5: A comparison of the detection performance of the two training strategies Train Real Test Real (TRTR) and Train Synthetic Test Real (TSTR) based on the F1-Score for the two supervised detection methods kNN and decision tree. A strategy *wins*, if the resulting detection performs better than that of the other strategy and is considered *equal* if it performs equally.

(TRTR), where we use original data with identified anomalies for both the training and testing. The second strategy is Train Synthetic Test Real (TSTR), where we train the anomaly detection method on data containing more synthetic anomalies than the original data and test its performance on original data with identified anomalies.

Figure 5 shows a comparison of the detection performance of these two strategies for the kNN and the decision tree. This comparison comprises the number of wins for each strategy, whereby a strategy is considered to win when its detection performance is better than that of the other strategy. If both strategies provide an identical detefection performance, they are considered equal. Independent of the considered detection method, data containing more synthetic anomalies than the original data wins far more often than data only containing identified anomalies. For the kNN, the number of wins with synthetic anomalies is 2.6 times larger (21 vs. 8), while both strategies perform equally in 21 cases. For the decision tree, data containing more synthetic anomalies than the original data wins 25 times, whereas data with identified anomalies only wins 20 times, and in 5 cases the strategies perform equally.

5 DISCUSSION

This section discusses the results, the modeled anomalies, and the proposed method for modeling and generating synthetic anomalies.

In the results, the t-SNE visualizations of synthetic and identified anomalies illustrate that the generated synthetic anomalies mostly overlap with the identified anomalies. Similarly, the histogram of the discriminative scores shows that synthetic and identified anomalies are difficult to distinguish with the discriminator method. Both results confirm that our synthetic anomalies accurately replicate the identified anomalies. From this observation, we conclude that the proposed method is capable of generating synthetic anomalies with real-world properties. Furthermore, since the TSTR strategy performs better or as well as TRTR in most cases, considering these synthetic anomalies in the training of an unsupervised anomaly detection method is beneficial for its detection performance. Given this observation, the proposed anomaly generation method can be used to improve anomaly detection methods in the future.

Despite these promising initial results, we note that our experiments are limited to the considered data, the associated production and consumption, and the anomalies identified in this data. Specifically for the extreme cases of anomaly types 3 and 4, the number of occurrences in our data set are small. Therefore, the parameters selected for the synthetic anomalies are based on a small sample size and we expect that more accurate results could be achieved with more data. Furthermore, the identified anomaly types are likely to be the result of technical failures in the metering infrastructure that cause unusual values such as extreme positive or negative spikes or a series of zeros. These types of anomalies have clearly defined, often extreme characteristics and are therefore relatively easy to detect. We expect that anomalies characterized by typical patterns at uncommon levels – such as unusual consumption – are more difficult to detect and, therefore, synthetic anomalies that reflect these characteristics could further improve anomaly detection methods.

We also note that our method currently inserts synthetic anomalies for energy and power time series separately. Since most applications only consider either energy or power time series separately, we believe this limitation to be not critical. However, due to the physical relationship between energy and power, simultaneously inserting multivariate synthetic anomalies for both energy and power time series could be beneficial in some cases.

6 CONCLUSION

The present paper introduces a method for generating four types of synthetic anomalies derived from real-world anomalies that can be inserted into arbitrary energy and power time series. To develop this generation method, we firstly analyze real-world energy and power time series to identify four commonly occurring anomaly types. Given this identified anomaly types, we formally model each type and then use our generation method to insert a chosen number of each synthetic anomaly type into arbitrary energy and power time series. We show that our method is capable of generating realistic synthetic anomalies and that these anomalies are beneficial for training supervised anomaly detection methods.

In future work, we plan to consider further energy and power time series, especially those that contain anomalies characterized by abnormal patterns such as anomalies caused by unusual consumption. Furthermore, to model the physical relationship between energy and power, future work should consider the simultaneous multivariate generation of synthetic anomalies for energy and power time series. Lastly, future work could include fuzziness into the generation to increase the variation of the synthetic anomalies.

ACKNOWLEDGMENTS

This project is funded by the Helmholtz Association's Initiative and Networking Fund through Helmholtz AI, the Helmholtz Association under the Program "Energy System Design", and the German Research Foundation (DFG) as part of the Research Training Group 2153 "Energy Status Data: Informatics Methods for its Collection, Analysis and Exploitation".

REFERENCES

[1] Damminda Alahakoon and Xinghuo Yu. 2016. Smart Electricity Meter Data Intelligence for Future Energy Systems: A Survey. *IEEE Transactions on Industrial Informatics* 12, 1 (2016), 425–436. https://doi.org/10.1109/TII.2015.2414355

[2] Leo Breiman, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone. 1984. *Classification And Regression Trees* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315139470

[3] Varun Chandola, Arindam Banerjee, and Vipin Kumar. 2009. Anomaly Detection: A Survey. *Comput. Surveys* 41, 3 (2009), 15:1–15:58. https://doi.org/10.1145/1541880.1541882

[4] François Chollet et al. 2015. Keras. https://keras.io.

[5] T. M. Cover and P. E. Hart. 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13, 1 (1967), 21–27. https://doi.org/10.1109/TIT.1967.1053964

[6] Marco De Nadai and Maarten van Someren. 2015. Short-term anomaly detection in gas consumption through ARIMA and Artificial Neural Network forecast. In *2015 IEEE Workshop on Environmental, Energy, and Structural Monitoring Systems (EESMS) Proceedings*. IEEE. https://doi.org/10.1109/EESMS.2015.7175886

[7] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD '96)*. AAAI Press, 226–231. https://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf

[8] Muhammad Fahim, Khadija Fraz, and Alberto Sillitti. 2020. TSI: Time series to imaging based model for detecting anomalous energy consumption in smart buildings. *Information Sciences* 523 (2020), 1–13. https://doi.org/10.1016/j.ins.2020.02.069

[9] Cheng Fan, Fu Xiao, Yang Zhao, and Jiayuan Wang. 2018. Analytical investigation of autoencoder-based methods for unsupervised anomaly detection in building energy data. *Applied Energy* 211 (2018), 1123–1135. https://doi.org/10.1016/j.apenergy.2017.12.005

[10] Wei Fan, Matthew Miller, Salvatore J. Stolfo, Wenke Lee, and Philip K. Chan. 2001. Using artificial anomalies to detect unknown and known network intrusions. *Knowledge and Information Systems* 6 (2001), 507–527. https://doi.org/10.1007/s10115-003-0132-7

[11] Ralph Foorthuis. 2021. On the nature and types of anomalies: a review of deviations in data. *International Journal of Data Science and Analytics* 12, 4 (2021), 297–331. https://doi.org/10.1007/s41060-021-00265-1 arXiv:2007.15634

[12] Megha Gaur, Stephen Makonin, Ivan V. Bajić, and Angshul Majumdar. 2019. Performance Evaluation of Techniques for Identifying Abnormal Energy Consumption in Buildings. *IEEE Access* 7 (2019), 62721–62733. https://doi.org/10.1109/ACCESS.2019.2915641

[13] Benedikt Heidrich, Andreas Bartschat, Marian Turowski, Oliver Neumann, Kaleb Phipps, Stefan Meisenbacher, Kai Schmieder, Nicole Ludwig, Ralf Mikut, and Veit Hagenmeyer. 2021. pyWATTS: Python Workflow Automation Tool for Time Series. *arXiv:2106.10157* (2021).

[14] Benedikt Heidrich, Marian Turowski, Kaleb Phipps, Kai Schmieder, Wolfgang Süß, Ralf Mikut, and Veit Hagenmeyer. under Review. Controlling Non-Stationarity and Periodicities in Time Series Generation. *Applied Intelligence* (under Review).

[15] Yassine Himeur, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira. 2020. A Novel Approach for Detecting Anomalous Energy Consumption Based on Micro-Moments and Deep Neural Networks. *Cognitive Computation* 12, 6 (2020), 1381–1401. https://doi.org/10.1007/s12559-020-09764-y

[16] Yassine Himeur, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira. 2020. Building power consumption datasets: Survey, taxonomy and future directions. *Energy and Buildings* 227 (2020), 110404. https://doi.org/10.1016/j.enbuild.2020.110404 arXiv:2009.08192

[17] Yassine Himeur, Khalida Ghanem, Abdullah Alsalemi, Faycal Bensaali, and Abbes Amira. 2021. Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives. *Applied Energy* 287 (2021), 116601. https://doi.org/10.1016/j.apenergy.2021.116601

[18] Geoffrey Hinton, Nitish Srivastava, and Kevin Swersky. 2012. Neural Networks for Machine Learning Lecture: Lecture 6a Overview of mini-batch gradient descent. http://www.cs.toronto.edu/{~}tijmen/csc321/slides/lecture{_}slides{_}lec6.pdf

[19] Rob J. Hyndman and George Athanasopoulos. 2021. *Forecasting: Principles and Practice* (third ed.). OTexts Melbourne, Australia. https://otexts.com/fpp3/

[20] Vikramaditya Jakkula and Diane Cook. 2010. Outlier Detection in Smart Environment Structured Power Datasets. In *2010 Sixth International Conference on Intelligent Environments*. IEEE, 29–33. https://doi.org/10.1109/IE.2010.13

[21] Paria Jokar, Nasim Arianpoo, and Victor C. M. Leung. 2016. Electricity Theft Detection in AMI Using Customers' Consumption Patterns. *IEEE Transactions on Smart Grid* 7, 1 (2016), 216–226. https://doi.org/10.1109/TSG.2015.2425222

[22] Bartosz Krawczyk. 2016. Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence* 5, 4 (2016), 221–232. https://doi.org/10.1007/s13748-016-0094-0

[23] Nikolay Laptev. 2018. AnoGen: Deep Anomaly Generator. In *Outlier Detection De-constructed (ODD) Workshop (ODD v5.0)*. https://doi.org/10.475/123

[24] Fangxing Li, Wei Qiao, Hongbin Sun, Hui Wan, Jianhui Wang, Yan Xia, Zhao Xu, and Pei Zhang. 2010. Smart Transmission Grid: Vision and Framework. *IEEE Transactions on Smart Grid* 1, 2 (2010), 168–177. https://doi.org/10.1109/TSG.2010.2053726

[25] Xiaoli Li, Chris P. Bowers, and Thorsten Schnier. 2010. Classification of Energy Consumption in Buildings With Outlier Detection. *IEEE Transactions on Industrial Electronics* 57, 11 (2010), 3639–3644. https://doi.org/10.1109/TIE.2009.2027926

[26] Haodong Lu, Miao Du, Kai Qian, Xiaoming He, and Kun Wang. 2021. GAN-based Data Augmentation Strategy for Sensor Anomaly Detection in Industrial Robots. *IEEE Sensors Journal* (2021). https://doi.org/10.1109/JSEN.2021.3069452

[27] Jian Luo, Tao Hong, and Meng Yue. 2018. Real-time anomaly detection for very short-term load forecasting. *Journal of Modern Power Systems and Clean Energy* 6, 2 (2018), 235–243. https://doi.org/10.1007/s40565-017-0351-7

[28] Ramin Moghaddass and Jianhui Wang. 2018. A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data. *IEEE Transactions on Smart Grid* 9, 6 (2018), 5820–5830. https://doi.org/10.1109/TSG.2017.2697440

[29] Kevin Ni, Nithya Ramanathan, Mohamed Nabil Hajj Chehade, Laura Balzano, Sheela Nair, Sadaf Zahedi, Eddie Kohler, Greg Pottie, Mark Hansen, and Mani Srivastava. 2009. Sensor network data fault types. *ACM Transactions on Sensor Networks* 5, 3 (2009). https://doi.org/10.1145/1525856.1525863

[30] Guansong Pang, Chunhua Shen, Longbing Cao, and Anton van den Hengel. 2021. Deep Learning for Anomaly Detection: A Review. *Comput. Surveys* 54, 2 (2021). https://doi.org/10.1145/3439950

[31] Youngja Park, Ian M. Molloy, Suresh N. Chari, Zenglin Xu, Chris Gates, and Ninghi Li. 2015. Learning from Others: User Anomaly Detection Using Anomalous Samples from Other Users. In *Computer Security – ESORICS 2015*, Günther Pernul, Peter Y. A. Ryan, and Edgar Weippl (Eds.). Springer Cham, 396–414. https://doi.org/10.1007/978-3-319-24177-7_20

[32] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. 2011. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 12, 85 (2011), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html

[33] João Pereira and Margarida Silveira. 2018. Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*. IEEE, 1275–1282. https://doi.org/10.1109/ICMLA.2018.00207

[34] Truong Son Pham, Quang Uy Nguyen, and Xuan Hoai Nguyen. 2014. Generating Artificial Attack Data for Intrusion Detection Using Machine Learning. In *Proceedings of the Fifth Symposium on Information and Communication Technology (SoICT '14)*. ACM, 286–291. https://doi.org/10.1145/2676585.2676618

[35] Bruno Rossi and Stanislav Chren. 2020. Smart Grids Data Analysis: A Systematic Mapping Study. *IEEE Transactions on Industrial Informatics* 16, 6 (2020), 3619–3639. https://doi.org/10.1109/TII.2019.2954098

[36] Lukas Ruff, Jacob R. Kauffmann, Robert A. Vandermeulen, Grégoire Montavon, Wojciech Samek, Marius Kloft, Thomas G. Dietterich, and Klaus-Robert Müller. 2021. A unifying review of deep and shallow anomaly detection. *Proc. IEEE* 109, 5 (2021), 756–795. https://doi.org/10.1109/JPROC.2021.3052449

[37] John E. Seem. 2007. Using intelligent data analysis to detect abnormal energy consumption in buildings. *Energy and Buildings* 39, 1 (2007), 52–58. https://doi.org/10.1016/j.enbuild.2006.03.033

[38] Abhishek B. Sharma, Leana Golubchik, and Ramesh Govindan. 2010. Sensor Faults: Detection Methods and Prevalence in Real-World Datasets. *ACM Transactions on Sensor Networks* 6, 3 (2010). https://doi.org/10.1145/1754414.1754419

[39] Georg Steinbuss and Klemens Böhm. 2021. Generating Artificial Outliers in the Absence of Genuine Ones — A Survey. *ACM Transactions on Knowledge Discovery from Data* 15, 2 (2021), 30. https://doi.org/10.1145/3447822

[40] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 86 (2008), 2579–2605. http://jmlr.org/papers/v9/vandermaaten08a.html

[41] Esther Villar-Rodriguez, Javier Del Ser, Izaskun Oregi, Miren Nekane Bilbao, and Sergio Gil-Lopez. 2017. Detection of non-technical losses in smart meter data based on load curve profiling and time series analysis. *Energy* 137 (2017), 118–128. https://doi.org/10.1016/j.energy.2017.07.008

[42] Chengyu Wang, Kui Wu, Tongqing Zhou, Guang Yu, and Zhiping Cai. 2021. TSAGen: Synthetic Time Series Generation for KPI Anomaly Detection. *IEEE Transactions on Network and Service Management* (2021). https://doi.org/10.1109/TNSM.2021.3098784

[43] Long Wang, Yong Ding, Till Riedel, Andrei Miclaus, and Michael Beigl. 2017. Data Analysis on Building Load Profiles: a Stepping Stone to Future Campus. In *2017 International Smart Cities Conference (ISC2)*. IEEE. https://doi.org/10.1109/ISC2.2017.8090823

[44] Yi Wang, Qixin Chen, Tao Hong, and Chongqing Kang. 2019. Review of Smart Meter Data Analytics: Applications, Methodologies, and Challenges. *IEEE Transactions on Smart Grid* 10, 3 (2019), 3125–3148. https://doi.org/10.1109/TSG.2018.

2818167

[45] Moritz Weber, Marian Turowski, Huseyin K. Çakmak, Ralf Mikut, Uwe Kuhnapfel, and Veit Hagenmeyer. 2021. Data-Driven Copy-Paste Imputation for Energy Time Series. *IEEE Transactions on Smart Grid* 12, 6 (2021), 5409–5419. https://doi.org/10.1109/TSG.2021.3101831

[46] Qingsong Wen, Liang Sun, Fan Yang, Xiaomin Song, Jingkun Gao, Xue Wang, and Huan Xu. 2021. Time Series Data Augmentation for Deep Learning: A Survey. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence (IJCAI-21)*. 4653–4660. https://doi.org/10.24963/ijcai.2021/631

[47] Jiuqi Elise Zhang, Di Wu, and Benoit Boulet. 2021. Time Series Anomaly Detection for Smart Grids: A Survey. In *2021 IEEE Electrical Power and Energy Conference (EPEC)*. IEEE, 125–130. https://doi.org/10.1109/epec52095.2021.9621752

[48] Jian Qiu Zhang and Yong Yan. 2001. A Wavelet-Based Approach to Abrupt Fault Detection and Diagnosis of Sensors. *IEEE Transactions on Instrumentation and Measurement* 50, 5 (2001), 1389–1396. https://doi.org/10.1109/19.963215

[49] Bin Zhou, Shenghua Liu, Bryan Hooi, Xueqi Cheng, and Jing Ye. 2019. BeatGAN: Anomalous Rhythm Detection using Adversarially Generated Time Series. In *Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19)*. 4433–4439. https://doi.org/10.24963/ijcai.2019/616

A STATISTICS

Table 1: Overview of the 50 one-year time series from the selected smart meters that are used to label the four identified types of anomalies. For each time series, the overall average power and energy consumption as well as the number, minimum length, maximum length, average power consumption, and average energy consumption of the labeled anomalies of all four types are reported. Note that anomalies of types (3) and (4) always have a length of one and that these types comprise two cases.

	Overall		Type 1					Type 2					Type 3			Type 4		
Time series	kW	kWh	#	Min	Max	kW	kWh	#	Min	Max	kW	kWh	#	kW	kWh	#	kW	kWh
1	11.8	36136.4	8	2	208	16.9	9839.4	3	2	3	15.5	36334.9	0	-	-	1	36779.9	27.1
2	28.6	6308.5	7	2	208	22.9	1590.8	3	2	3	36.8	5310.9	1	6740.1	14.0	1	6776.9	62.7
3	121.3	16508.4	8	2	208	144.3	3905.3	14	2	27	144.0	9231.7	2	6255.1	-10008059.8	4	10786.9	5006955.5
4	35.7	51037.7	7	2	208	34.6	15653.9	3	2	3	36.9	50612.8	1	53132.4	15.5	1	53290.0	80.7
5	91.6	67761.1	9	2	208	-13389.9	18031.2	15	2	27	98.5	34858.4	2	12797.3	-25481.8	3	8718.7	2740483.8
6	1.7	58368.5	8	2	8931	0.9	12781.4	0	-	-	-	-	0	-	-	0	-	-
7	301.7	28417.3	7	2	208	250.8	8649.4	3	2	3	345.5	27697.7	1	30783.5	178.1	1	30923.4	480.5
8	58.1	1247.5	9	2	208	11662.8	375.7	13	2	27	124.6	688.8	2	490.4	-23582.7	2	331.9	456993.9
9	15.5	15227.5	7	2	208	12.3	4678.7	3	2	332	16.8	15105.7	1	16234.2	8.9	1	16311.2	37.1
10	4.4	65001.2	9	2	208	-2523.3	12794.6	7	2	27	0.8	22416.6	0	-	-	0	-	-
11	69.2	29795.1	7	2	208	64.4	9191.6	3	2	332	74.7	29882.1	1	30572.6	37.0	1	30622.0	144.9
12	11.9	49030.3	7	2	208	14.8	15179.6	3	2	332	9.0	49333.9	1	49416.9	4.6	2	49375.3	29.8
13	27.8	7366.5	8	2	208	63.3	2381.6	7	2	270	35.4	7299.3	1	8021.2	21.2	1	8077.8	50.2
14	12.0	25829.4	9	2	208	20.2	8235.4	5	3	270	11.3	25985.9	1	26075.4	5.2	1	26081.9	26.3
15	13.9	62656.2	7	2	208	16.5	19379.0	4	2	5	17.8	63026.7	1	63571.0	8.0	2	63322.9	25.0
16	0.7	6757.4	7	2	208	0.6	2091.7	4	2	6	0.7	6798.9	1	6822.1	0.3	2	6811.0	1.4
17	56.6	72354.2	7	2	208	66.3	22360.6	4	2	5	72.3	72769.3	1	73590.1	33.1	1	73651.3	102.7
18	19.6	39898.6	7	2	208	22.1	12328.3	2	3	16	23.1	40021.8	1	40593.7	8.7	2	40667.8	63.4
19	0.3	8018.0	7	2	208	0.4	2480.6	2	3	16	0.2	8058.0	1	8089.7	0.1	1	8091.9	0.5
20	2.3	34096.7	7	2	208	3.2	10546.5	2	3	16	2.2	34301.2	1	34398.8	1.2	1	34412.7	8.1
21	1.7	173336.8	7	2	208	1.1	53519.8	3	3	16	0.4	171582.1	1	178587.3	0.7	2	174707.7	3.9
22	34.0	24674.7	7	2	208	36.8	7574.9	2	3	16	44.0	24409.0	1	26082.4	21.5	1	26165.2	53.7
23	1.0	8222.8	7	2	208	1.3	2135.3	2	3	16	1.1	8763.0	1	9449.6	0.1	1	9474.5	0.3
24	1.1	79733.7	7	2	208	0.4	24468.9	0	-	-	-	-	0	-	-	1	83031.0	4.9
25	1.7	25918.2	7	2	208	1.7	7405.4	2	3	16	2.0	26573.7	1	28404.9	0.3	1	28593.6	2.4
26	25.3	47407.2	7	2	208	28.6	14622.6	2	3	16	30.3	47470.3	1	48514.9	13.6	1	48585.8	53.9
27	28.2	151172.1	12	2	6	28.8	85601.1	0	-	-	-	-	0	-	-	1	235908.3	41.3
28	7.1	58558.8	5	2	9	8.3	13350.9	2	10	17	6.1	58462.0	0	-	-	3	58827.1	279.6
29	13.1	38328.4	5	2	9	18.6	8748.0	1	17	17	6.5	38388.1	0	-	-	1	38717.3	22.4
30	0.7	570.7	18	2	9	0.5	219.4	1	17	17	0.7	575.1	0	-	-	0	-	-
31	0.2	183.4	16	2	9	0.3	67.0	1	17	17	0.2	183.6	0	-	-	1	187.4	0.5
32	83.3	91543.3	5	2	9	99.4	20869.8	1	17	17	113.4	91660.3	0	-	-	1	92873.7	127.0
33	1.0	56481.2	5	2	9	0.9	12980.3	1	17	17	1.2	56629.4	0	-	-	1	59574.2	2.3
34	1.1	163877.9	6	2	1363	1.9	32573.8	2	1731	6452	0.6	171591.2	0	-	-	0	-	-
35	2.0	170468.5	5	2	9	3.4	39103.9	2	17	3535	2.0	170194.1	0	-	-	2	176801.9	9.2
36	54.6	455610.1	7	2	9	64.9	126438.8	2	17	4425	13.4	461428.7	0	-	-	0	-	-
37	0.4	224231.4	5	2	9	0.3	51062.6	3	17	2290	0.3	224305.7	0	-	-	1	225476.6	1.5
38	1.1	52122.9	5	2	9	1.6	11861.1	2	10	17	1.9	52123.0	0	-	-	2	52182.2	3.2
39	0.6	7139.3	5	2	9	0.6	1627.7	2	7	17	0.9	7086.7	0	-	-	2	7176.6	1.0
40	3.1	11695.8	5	2	9	4.6	2685.1	2	10	17	1.9	11667.7	0	-	-	2	11754.8	165.1
41	19.3	31659.7	6	2	386	36.1	6086.3	5	17	3516	12.4	32017.5	0	-	-	0	-	-
42	0.7	103539.5	6	2	452	0.7	19983.6	1	17	17	0.8	105146.9	0	-	-	2	105784.1	195.0
43	4.6	19312.0	5	2	9	6.8	4425.2	2	2	17	5.6	19260.8	0	-	-	1	19401.5	25.7
44	96.3	93851.6	5	2	9	90.1	21437.6	1	17	17	166.7	94007.3	0	-	-	1	96294.1	188.9
45	85.4	45205.9	5	2	9	67.0	10358.4	1	17	17	193.9	45306.2	0	-	-	1	47441.0	203.2
46	18.8	17973.7	6	2	488	33.1	3497.5	2	727	1957	33.6	18307.7	0	-	-	1	18924.6	19.9
47	0.2	199217.2	5	2	9	0.1	45346.7	0	-	-	-	-	0	-	-	1	199777.8	0.9
48	1.6	43087.4	6	2	9	3.3	9424.6	4	8	17	2.4	40338.4	0	-	-	1	44925.4	5.2
49	1.2	8411.2	7	2	9	1.2	2724.2	3	8	17	1.1	8273.2	0	-	-	1	8594.0	2.6
50	56.4	28194.2	6	2	9	0.0	6050.3	5	8	14869	52.8	28081.2	0	-	-	0	-	-

B ANOMALY TYPES IN THE LITERATURE

Table 2: Overview of the anomaly types identified in the power and energy time series of the considered data and exemplary matching classes in the literature.

	Time series	Matching classes in literature
Anomaly type (1)	Energy	"temporary change (ST-VIIb)" and "variation change (ST-VIIf)" [11], "CONSTANT fault" [38]
	Power	"temporary change (ST-VIIb)" and "variation change (ST-VIIf)" [11]
Anomaly type (2)	Energy	"temporary change (ST-VIIb)" and "variation change (ST-VIIf)" [11], "stuck-at fault" [29], "stuck fault" [48]
	Power	"temporary change (ST-VIIb)" and "variation change (ST-VIIf)" [11]
Anomaly type (3)	Energy	"level shift (ST-VIIc)" [11]
	Power	"local additive (ST-IVe)" [11], "outlier fault" [29], "SHORT fault" [38], "spike fault" [48]
Anomaly type (4)	Energy	"level shift (ST-VIIc)" [11]
	Power	"local additive (ST-IVe)" [11], "outlier fault" [29], "SHORT fault" [38], "spike fault" [48]

C PARAMETERS

Table 3: Summary of the values determined from the 50 one-year power time series of the selected smart meters for the offset k, number, minimum length, maximum length, r_{min}, and r_{max} as presented in Table 4. These values can be used as parameters to generate synthetic anomalies of the four modeled types for power time series. Note that anomalies of types (3) and (4) always have a length of one and that these types comprise two cases.

	Type 1			Type 2			Type 3			Type 4		
k	#	Min	Max	#	Min	Max	#	r_{min}	r_{max}	#	r_{min}	r_{max}
$[177, 431796]$	$[5, 18]$	3	$[3, 4465]$	$[0, 15]$	$[2, 1731]$	$[2, 7434]$	$[0, 2]$	0.61 -	1.62 -	$[0, 4]$	1.15 11.01	8.1 13

Table 4: Overview of the number, minimum length, maximum length, r_{min}, r_{max}, and k used as parameters to generate synthetic anomalies for the evaluated 50 one-year power time series from the selected smart meters using the t-SNE and the discriminative method. Note that anomalies of types (3) and (4) always have a length of one and that these types comprise two cases.

Time		Type 1			Type 2			Type 3			Type 4		
series	k	#	Min	Max	#	Min	Max	#	r_{min}	r_{max}	#	r_{min}	r_{max}
1	35787	8	3	96	3	2	3	0	-	-	1	1.15	8.1
2	5730	7	3	96	3	2	3	1	0.61	1.62	1	1.15	8.1
3	17649	8	3	96	14	2	27	2	-	-	4	11.01	13
4	48127	7	3	96	3	2	3	1	0.61	1.62	1	1.15	8.1
5	68477	9	3	96	15	2	27	2	-	-	3	11.01	13
6	80207	8	3	96	0	-	-	0	-	-	0	1.15	8.1
7	25239	7	3	96	3	2	3	1	0.61	1.62	1	1.15	8.1
8	731	9	3	96	13	2	27	2	-	-	2	11.01	13
9	14104	7	3	96	3	2	48	1	0.61	1.62	1	1.15	8.1
10	49387	9	3	96	7	2	27	0	-	-	0	-	-
11	29056	7	3	96	3	2	48	1	0.61	1.62	1	1.15	8.1
12	49172	7	3	96	3	2	48	1	0.61	1.62	2	1.15	8.1
13	6393	8	3	96	7	2	48	1	0.61	1.62	1	1.15	8.1
14	25862	9	3	96	5	3	48	1	0.61	1.62	1	1.15	8.1
15	62272	7	3	96	4	2	5	1	0.61	1.62	2	1.15	8.1
16	6764	7	3	96	4	2	6	1	0.61	1.62	2	1.15	8.1
17	71565	7	3	96	4	2	5	1	0.61	1.62	1	1.15	8.1
18	39421	7	3	96	2	3	16	1	0.61	1.62	2	11.01	13
19	8020	7	3	96	2	3	16	1	-	-	1	1.15	8.1
20	34042	7	3	96	2	3	16	1	-	-	1	1.15	8.1
21	168614	7	3	96	3	3	16	1	-	-	2	11.01	13
22	23011	7	3	96	2	3	16	1	0.61	1.62	1	1.15	8.1
23	2653	7	3	96	2	3	16	1	-	-	1	11.01	13
24	75229	7	3	96	0	-	-	0	-	-	1	11.01	13
25	17937	7	3	96	2	3	16	1	-	-	1	1.15	8.1
26	46301	7	3	96	2	3	16	1	0.61	1.62	1	1.15	8.1
27	28114	12	3	6	0	-	-	0	-	-	1	1.15	8.1
28	58016	5	3	9	2	10	17	0	-	-	3	11.01	13
29	37605	5	3	9	1	17	17	0	-	-	1	1.15	8.1
30	509	18	3	9	1	17	17	0	-	-	0	-	-
31	177	16	3	9	1	17	17	0	-	-	1	1.15	8.1
32	89750	5	3	9	1	17	17	0	-	-	1	1.15	8.1
33	51978	5	3	9	1	17	17	0	-	-	1	1.15	8.1
34	165403	6	3	96	2	44	48	0	-	-	0	-	-
35	161244	5	3	9	2	17	48	0	-	-	2	11.01	13
36	431796	7	3	9	2	17	48	0	-	-	0	-	-
37	222477	5	3	9	3	17	48	0	-	-	1	1.15	8.1
38	52017	5	3	9	2	10	17	0	-	-	2	11.01	13
39	7001	5	3	9	2	7	17	0	-	-	2	1.15	8.1
40	11188	5	3	9	2	10	17	0	-	-	2	11.01	13
41	31823	6	3	96	5	17	48	0	-	-	0	-	-
42	102079	6	3	96	1	17	17	0	-	-	2	11.01	13
43	18806	5	3	9	2	2	17	0	-	-	1	11.01	13
44	90338	5	3	9	1	17	17	0	-	-	1	1.15	8.1
45	42124	5	3	9	1	17	17	0	-	-	1	1.15	8.1
46	17234	6	3	96	2	44	48	0	-	-	1	1.15	8.1
47	198393	5	3	9	0	-	-	0	-	-	1	11.01	13
48	32838	6	3	9	4	8	17	0	-	-	1	11.01	13
49	8159	7	3	9	3	8	17	0	-	-	1	1.15	8.1
50	26747	6	3	9	5	8	48	0	-	-	0	-	-

Table 5: Overview of the number, minimum length, maximum length, r_{min}, r_{max}, and k used as parameters to generate synthetic anomalies for the evaluated 50 one-year power time series from the selected smart meters regarding the training of the evaluated supervised anomaly detection methods. Note that anomalies of types (3) and (4) always have a length of one and that these types comprise two cases.

Time		Type 1			Type 2			Type 3			Type 4		
series	k	#	Min	Max	#	Min	Max	#	r_{min}	r_{max}	#	r_{min}	r_{max}
1	0	16	3	96	6	2	3	0	-	-	2	1.15	8.1
2	0	14	3	96	6	2	3	2	0.61	1.62	2	1.15	8.1
3	0	16	3	96	28	2	27	4	-	-	8	11.01	13
4	0	14	3	96	6	2	3	2	0.61	1.62	2	1.15	8.1
5	0	18	3	96	30	2	27	4	-	-	6	11.01	13
6	0	16	3	96	0	-	-	0	-	-	0	-	-
7	0	14	3	96	6	2	3	2	0.61	1.62	2	1.15	8.1
8	0	18	3	96	26	2	27	4	-	-	4	11.01	13
9	0	14	3	96	6	2	48	2	0.61	1.62	2	1.15	8.1
10	0	18	3	96	14	2	27	0	-	-	0	-	-
11	0	14	3	96	6	2	48	2	0.61	1.62	2	1.15	8.1
12	0	14	3	96	6	2	48	2	0.61	1.62	4	1.15	8.1
13	0	16	3	96	14	2	48	2	0.61	1.62	2	1.15	8.1
14	0	18	3	96	10	3	48	2	0.61	1.62	2	1.15	8.1
15	0	14	3	96	8	2	5	2	0.61	1.62	4	1.15	8.1
16	0	14	3	96	8	2	6	2	0.61	1.62	4	1.15	8.1
17	0	14	3	96	8	2	5	2	0.61	1.62	2	1.15	8.1
18	0	14	3	96	4	3	16	2	0.61	1.62	4	11.01	13
19	0	14	3	96	4	3	16	2	-	-	2	1.15	8.1
20	0	14	3	96	4	3	16	2	-	-	2	1.15	8.1
21	0	14	3	96	6	3	16	2	-	-	4	11.01	13
22	0	14	3	96	4	3	16	2	0.61	1.62	2	1.15	8.1
23	0	14	3	96	4	3	16	2	-	-	2	11.01	13
24	0	14	3	96	0	-	-	0	-	-	2	11.01	13
25	0	14	3	96	4	3	16	2	-	-	2	1.15	8.1
26	0	14	3	96	4	3	16	2	0.61	1.62	2	1.15	8.1
27	0	24	3	6	0	-	-	0	-	-	2	1.15	8.1
28	0	10	3	9	4	10	17	0	-	-	6	11.01	13
29	0	10	3	9	2	17	17	0	-	-	2	1.15	8.1
30	0	36	3	9	2	17	17	0	-	-	0	-	-
31	0	32	3	9	2	17	17	0	-	-	2	1.15	8.1
32	0	10	3	9	2	17	17	0	-	-	2	1.15	8.1
33	0	10	3	9	2	17	17	0	-	-	2	1.15	8.1
34	0	12	3	96	4	44	48	0	-	-	0	1.15	8.1
35	0	10	3	9	4	17	48	0	-	-	4	11.01	13
36	0	14	3	9	4	17	48	0	-	-	0	-	-
37	0	10	3	9	6	17	48	0	-	-	2	1.15	8.1
38	0	10	3	9	4	10	17	0	-	-	4	11.01	13
39	0	10	3	9	4	7	17	0	-	-	4	1.15	8.1
40	0	10	3	9	4	10	17	0	-	-	4	11.01	13
41	0	12	3	96	10	17	48	0	-	-	0	-	-
42	0	12	3	96	2	17	17	0	-	-	4	11.01	13
43	0	10	3	9	4	2	17	0	-	-	2	11.01	13
44	0	10	3	9	2	17	17	0	-	-	2	1.15	8.1
45	0	10	3	9	2	17	17	0	-	-	2	1.15	8.1
46	0	12	3	96	4	44	48	0	-	-	2	1.15	8.1
47	0	10	3	9	0	-	-	0	-	-	2	11.01	13
48	0	12	3	9	8	8	17	0	-	-	2	11.01	13
49	0	14	3	9	6	8	17	0	-	-	2	1.15	8.1
50	0	12	3	9	10	8	48	0	-	-	0	-	-