DOI: 10.1111/cgf.70000

COMPUTER GRAPHICS *forum*
*Volume 44 (2025), number 1, e70000*

# MANDALA—Visual Exploration of Anomalies in Industrial Multivariate Time Series Data

J. Suschnigg,[1,2] B. Mutlu,[1] G. Koutroulis,[3] H. Hussain[4] and T. Schreck[2]

[1]Pro2Future GmbH, Graz, Austria
{josef.suschnigg, belgin.mutlu}@pro2future.at
[2]Institute of Visual Computing, Graz University of Technology, Austria
tobias.schreck@tugraz.at
[3]AVL List GmbH, Graz, Austria
georgios.koutroulis@avl.com
[4]Institute of Interactive Systems and Data Science, Graz University of Technology, Austria
hussain@tugraz.at

**Abstract**

*The detection, description and understanding of anomalies in multivariate time series data is an important task in several industrial domains. Automated data analysis provides many tools and algorithms to detect anomalies, while visual interfaces enable domain experts to explore and analyze data interactively to gain insights using their expertise. Anomalies in multivariate time series can be diverse with respect to the dimensions, temporal occurrence and length within a dataset. Their detection and description depend on the analyst's domain, task and background knowledge. Therefore, anomaly analysis is often an underspecified problem. We propose a visual analytics tool called MANDALA (**M**ultivariate **AN**omaly **D**etection **A**nd exp**L**or**A**tion), which uses kernel density estimation to detect anomalies and provides users with visual means to explore and explain them. To assess our algorithm's effectiveness, we evaluate its ability to identify different types of anomalies using a synthetic dataset generated with the GutenTAG anomaly and time series generator. Our approach allows users to define normal data interactively first. Next, they can explore anomaly candidates, their related dimensions and their temporal scope. Our carefully designed visual analytics components include a tailored scatterplot matrix with semantic zooming features that visualize normal data through hexagonal binning plots and overlay candidate anomaly data as scatterplots. In addition, the system supports the analysis on a broader scope involving all dimensions simultaneously or on a smaller scope involving dimension pairs only. We define a taxonomy of important types of anomaly patterns, which can guide the interactive analysis process. The effectiveness of our system is demonstrated through a use case scenario on industrial data conducted with domain experts from the automotive domain and a user study utilizing a public dataset from the aviation domain.*

**Keywords:** anomaly detection, interactive data exploration, kernel density estimation, multivariate time series analysis, visual analytics

**CCS Concepts:** • Computing methodologies → Anomaly detection; Visual analytics; • Mathematics of computing → Time series analysis

## 1. Introduction

The increasing amount of large and complex datasets in various industrial domains leads to a growing demand for advanced data analytics methods that provide actionable insights to support decision-making. Data-driven technologies have the potential to optimize processes, reduce costs and improve the quality of products. However, industrial data's volume, variety and velocity pose significant challenges in data management, processing, analysis and interpretation [THS*19].

Yet, out-of-the-box approaches are typically insufficient to analyze industrial data effectively. Instead, it is necessary to integrate background knowledge from fields such as engineering, operations research and business. This is because industrial data is often complex and specific to the particular domain and requires a deep understanding of the underlying processes and factors that influence the data [SMS23]. Thus, visual analytics provides ways to visualize and interact with data, allowing for a more intuitive understanding. Visual analytics also facilitates collaboration and communication among stakeholders [MHK*19] and supports informed decision-making based on data-driven insights.

Anomaly detection in time series data has been widely used in various industrial applications to identify unexpected or abnormal events, for example, for predictive maintenance [GFS*19], cyber security [FFV*12], condition monitoring [WZC*18] or quality control [SZB*20]. Moreover, visual analytics can be a powerful tool for anomaly detection because it allows analysts to use their visual perception to identify relevant patterns through interactive exploration and their domain knowledge.

This research is challenging because visual analytics tools must effectively bring together human intuition with automated data analysis to interpret complex industrial datasets, often dealing with unspecified issues within several domains. Creating these tools necessitates the integration of domain-specific knowledge with advanced analytics to identify patterns and anomalies, particularly contextual anomalies, which have not been extensively covered in previous research. Additionally, the lack of ground truth data makes it difficult to validate models and solutions in great depth. Due to these complexities, exploration through visualizations and collaboration with domain experts is essential for developing effective, interactive and reliable analytics solutions that can tackle the unique challenges of industrial time series data.

A broad spectrum of algorithms exists for detecting anomalies within industrial processes. Our literature review shows that most visual analytics approaches apply anomaly detection on multivariate time series data by handling each data dimension individually [SFS*22] and aggregating the results into an overall anomaly score [SMK*21]. However, this methodology can overlook the complex interdependence between variables that are relevant to industrial applications. On the other hand, approaches that involve all dimensions simultaneously for anomaly calculation often rely on difficult-to-interpret models, such as dimensionality reduction techniques [FSS*21b, GD21]. Our first research challenge involves the creation of an interactive user interface that allows users to investigate anomalies within multivariate datasets, including the associated dimensions. We aim to ensure that the automated data analysis underpinning the user interface remains comprehensible to humans in the loop. Further, domain experts should be able to identify and accurately interpret anomalies, particularly ambiguous ones, via visualization. Users should be able to explore a variety of types of anomalies that can occur in time series data [BGCML21], including contextual anomalies that have received limited attention in previous research. The second research challenge we address is the visualization of the kernel density estimation (KDE) anomaly detector that we utilize in this research work. Researchers raised the issue of exploring density estimates in multidimensional datasets [Che17] and state that graphical illustration or interpretation for more than two-variate data is at least difficult if not impossible [Weg18]. Motivated by these challenges, we make the following contributions:

The first contribution of this research work addresses these challenges and proposes the visual interactive tool MANDALA (**M**ultivariate **AN**omaly **D**etection **A**nd exp**L**or**A**tion). It is designed for industrial applications that generate multivariate cyclic time series data. This data typically exhibits cyclic patterns due to the repetitive nature of industrial processes, such as a single cycle representing a production cycle in manufacturing or a test cycle in validation and verification [SMK*21]. In many industrial tasks, the processes often work as expected and produce similar data. However, even in static environments, occasional deviations or anomalies can occur. These anomalies can be due to unexpected changes in process conditions, equipment malfunction or degradation. Hence, the exploration problem is twofold: (1) defining what is expected and considered normal, (2) understanding factors contributing to differences in data and how to characterize them. This is typically an open and underspecified problem, which requires the human expert in the loop. We propose a visual analytics approach to achieve this. MANDALA uses an approach that allows users to label normal data as reference data. Anomalies can then be identified using KDE, and the results can be examined using scatterplot matrices (SPLOMs). Additionally, hexagonal binning plots are used to visualize the reference data that the user has selected. The SPLOM integrates semantic zooming capabilities to enhance its scalability, while line plots have been added to visualize the time dimension of the data, which is not visible in 2D scatterplots. In addition, the system implements colour-encoding strategies for scatterplots that support the analysis of two analysis scopes: multivariate anomaly detection and bivariate anomaly detection. Line plot views allow for exploration of the temporal scope of time series and enable brushing and linking techniques. In addition, the system calculates and ranks suggestion views based on a user's selection. As a second contribution, we present the results of a use case scenario that analyses an industrial dataset from the automotive domain with a subject matter expert (SME). Based on the use case scenario, we summarize interesting patterns in our prototype through a taxonomy.

As our third contribution, we recruited 15 researchers, predominantly postgraduates, and conducted a user study with a dataset from the aviation domain using NASA-TLX [HS88] to measure workload and the system usability scale (SUS) [SL09] for the assessment of the usability of MANDALA. The fourth contribution is lessons learned, design implications and research directions for future work.

The paper is organized as follows. In the next section, we review related work that has influenced this research, followed by a detailed discussion of the design requirements. The following Sections 4 and 5 describe MANDALA's automated data analysis and visual analytics methodology. Then, we present the results of our evaluation through a use case scenario and user studies and conclude the article by discussing our findings and future research directions.

## 2. Related Work

In this section, we review the research most relevant to our work, organized by the research subjects (i) anomaly detection in automated data analysis, (ii) anomaly detection in visual analytics and (iii) information visualization.

### 2.1. Anomaly detection algorithms

The research topic of anomaly detection has gained significant attention in recent decades, leading to a large body of literature on the subject. As a result, several comprehensive surveys have investigated and reviewed this research area. In their survey, Chandola et al. provide an overview of anomaly detection techniques and connect them to multiple application domains [CBK09a]. Gupta et al.'s survey focuses on various methods proposed for detecting anomalies in temporal data [GGAH14]. More recently, Blázquez-García et al. reviewed anomaly detection in time series data based on a proposed taxonomy of anomaly detection techniques [BGCML21]. Shaukat, et al.'s survey reviews time series anomaly detection techniques and highlights challenges in anomaly detection [SAL*21]. Cook et al. conducted a comprehensive study specifically for IoT time series data [CMF20], although most of the study results can be generalized to other application domains. The study by Schmidl et al. evaluates 71 state-of-the-art anomaly detection algorithms from different domains and their effectiveness, efficiency and robustness in 976 time series datasets [SWP22]. The results provide an overview of the techniques, their strengths and weaknesses, support the facilitation of algorithm selection and open up new research directions.

The non-parametric statistical anomaly detection techniques applied in this research work are discussed in the literature [CBK09a, LLP07]. These techniques are selected because they do not require data distribution assumptions, which makes them suitable for a generic anomaly detection approach. Kernel-based techniques, a type of non-parametric anomaly detection algorithm, have been widely investigated in numerous studies on time series data. For example, Das et al. applied kernel-based techniques to flight operation quality assurance data in the context of aviation safety [DMSO10], while Subramaniam et al. used a similar approach in a sensor network dataset [SPP*06].

Only a few research papers address contextual anomalies. The context of anomalies can be other dataset dimensions, including the time dimension. Viewing data isolated from the context suggests normal data while adding the context dimension to the anomaly reveals abnormal behaviour [CBK09a]. Stojanovic et al. provide an example, proposing an algorithm that can detect anomaly contexts by computing excessive correlations in multidimensional space. The approach focuses on scalability and has been applied to various industrial use cases and time series data [SDS17]. Hayes et al. have developed an anomaly detection framework that separates contextual anomaly detection into two stages (1) content detection and (2) context detection using clustering techniques. The authors applied their framework to sensor data and a publicly available UCI dataset [HC14]. The paper by Golmohammadi et al. focuses on using contextual anomaly detection (CAD) to identify stock market manipulation. The authors emphasize the importance of contextual information and propose a novel prediction-based approach to analyze complex time series data, such as financial market data.

## 2.2. Visual analytics for anomaly detection

Visual analytics systems are used to explore and analyze complex datasets, including those for detecting anomalies. This section reviews some of the most related work in this area.

Xu et al. recently published a comprehensive survey that examined visual analysis to detect anomalous user behaviour [SLT*22]. The survey discusses four distinct application fields: social interaction, travel, network communication and financial transactions. Data types, visualization techniques and interactive analysis methods are reviewed for each field. The authors also identify trends and preferences and suggest potential research directions. Zhou et al.'s survey reviews the literature on visualization technologies for smart manufacturing in industrial applications [ZLL*19]. The authors examine examples throughout the entire life cycle of industrial products [THS*19] and highlight the main challenges that require further investigation.

The MTSAD system proposed by Pham et al. visualizes anomalies in multivariate time series data using radar charts and several interaction techniques, including filtering, ranking, lensing and brushing and linking [PNL*19]. Small-multiple radar charts are used to bin and visualize multivariate time series, with anomalies represented in red within the radar charts. The system's effectiveness is evaluated using data from a high-performance computing monitoring use case. Another example within the domain of high-performance computing has been investigated by Nyguen et al. with CloudTraceViz to visualize multivariate time series [ND19]. The authors distinguish between bivariate data visualized by SPLOMs and multivariate data visualized by parallel coordinates and implement a scagnotstics-based anomaly detection method [WAG05]. CloudDet by Xu et al. is an interactive visual analytics system to detect, inspect and diagnose anomalies in multivariate time series data [XWY*20]. The system has been developed for a cloud computing use case and orchestrates multiple contextual views and rich interaction designs, which support the ranking and exploration of time series anomalies.

The work by Suschnigg et al. on visual exploration of anomalies in time series data contains a design study with experts from the automotive domain [SMK*21]. On the basis of the results, they propose a system allowing domain-expert users to self-label normal data for an anomaly detection task. They use a flexible glyph design to visualize and investigate the results of several anomaly detection methods. Liu et al. and their MTV system have proposed another system to visually detect, investigate and annotate anomalies in multivariate time series data [LAZV22]. In addition to efficient and intelligent anomaly identification, the work also focuses on in situ annotation and a workflow for collaboration.

Ko et al. proposed a comprehensive visual analytics system for analysing high-dimensional multivariate network data that integrates linked views such as heatmaps, calendar-based views, clock views, line plots and map-based views [KAW*14]. The calendar-based view in their system, which visualizes flight delay patterns, is similar to the anomaly heatmap view in MANDALA.

## 2.3. Visualization techniques

This section reviews related work on visualization techniques that inspired our proposed system. For example, CommAid by Fischer et al. [FSS*21a] is a system to analyze communication networks. The work drew our attention through its interactive multilevel matrix-based visualization and semantic zooming capabilities. Another example of semantic zooming in matrix-based visualizations is Timematrix by Yi et al. [EDG*08], which has been implemented using temporal data from social networks. The work employs the scalability of matrix representations to visualize large-scale network data. Another example of semantic zooming for scalable matrix representations has been investigated by Behrisch et al. [BDF*14], which focuses on comparative analysis of sets of matrices.

A hexagonal binning plot is a visualization technique to visualize the distribution of bivariate data developed by Carr et al. [CLNL87]. Generally, they are visual representations of binned scatterplots whose design space has been explored by Heimerl et al. [HCSG18]. We utilize them in our research work because of their ease of interpretation, making them practical for exploratory data analysis. Recent work on honeycomb plots by Trautner et al. highlights confusers of the hexagonal binning plots and proposes a novel system to improve the technique [TSSB22].

To conclude this section, while the related work outlined in this paper has significantly influenced our approach, much of the existing research does not focus on detecting and exploring multivariate anomalies comprehensively. Instead, it often analyses dimensions of multivariate data individually. To address this gap, our research focuses on identifying, analysing and visualizing multidimensional and contextual anomalies within cyclic time series data. The methods we have developed to achieve this are detailed in the following sections.

## 3. Design Requirements

The design requirements of our proposed visual design lie in the experiences we have gained in the past years in several research projects carried out in close collaboration with several companies in the industry. This research focuses on helping users detect and explore anomalies in cyclic multivariate time series data, which is likely to be acquired in industrial contexts. The dimensionality of time series data originates from multiple sensors and other data sources found in industrial applications. Going from anomaly detection in univariate data to multivariate data introduces many additional problems requiring interactive and user-in-the-loop approaches. Based on our past experience with the industry in previous research and the literature, we define five requirements to justify the design choices for our proposed visual analytics design and automated data analysis.

**R1 user-defined normal behaviour**. The literature defines anomalies in time series as points or sequences of points that deviate from the normal behaviour of the data [SAL*21]. However, from a data perspective, normal or expected behaviour can be ambiguous, especially if the ground truth or labels are not available [SAL*21]. In that common case, supervised anomaly detection approaches do not apply. Instead, unsupervised approaches make the implicit assumption that normal instances are far more frequent anomalies in the data [CBK09a]. In industrial contexts, this assumption may not be applicable in cases where (1) abnormal behaviour slightly evolves (i.e. wearing), or (2) abnormal behaviour stays undetected, and therefore, normal data is not more frequent than anomalies in the dataset. Therefore, the analysis process strongly relies on the definition and interpretation through domain-specific knowledge of domain experts. Accordingly, we require an approach where the user selects/labels the normal behaviour as reference data. Domain experts are qualified to define and interpret normal behaviour due to their extensive experience and the knowledge they acquire on-site. However, the automated selection or recommendation of reference cycles shows promise in uncovering insights that might otherwise be missed. We plan to focus on this topic in future research.

**R2 detection and exploration of different types of anomalies**. According to the literature, multivariate time series data can contain different anomaly types. Blázquez-García et al. propose a taxonomy of anomaly types which are characterized by the number of dimensions involved (univariate or multivariate) and the temporal duration of anomalous data within the time series (point, subsequence or whole time series) [BGCML21]. We define detecting and exploring different types of time series anomalies as the second requirement.

**R3 exploration of anomaly contexts**. Another attribute of multidimensional time series anomalies on which we focus is anomaly contexts. Contextual attributes are used to determine the context of anomalies [CBK09a]. In industrial applications, exploring contexts of multivariate anomalies can be significantly relevant; for example, contextual anomalies can be related to a specific environment temperature in testing facilities, engine speed in combustion engines [SMK*21], or the operating altitude of an aircraft turbine. The third requirement of our visual design is the possibility of detecting and exploring contexts in multivariate time series anomalies.

**R4 rank and suggest candidate views**. Exploring multivariate time series anomalies can be especially complex. In particular, if the investigated dataset differs extensively from normal behaviour, the amount of interesting or deviating data can overwhelm users. Therefore, we rank and suggest candidate views for user inspection to facilitate the exploration loop.

**R5 visual scalability**. Scalability in visual analytics is an ongoing long-term challenge in visual analytics [REE*09, Cui19]. As far as possible, our design should be scalable in terms of available screen space and the number of dimensions in the multivariate time series dataset. Therefore, the fifth requirement targets the visual scalability of our visual design and highly interactive zoomability.

To conclude this section, we emphasize two limitations of the requirements. First, we explicitly do not focus on the collective anomaly type to reduce the complexity of our design. Those anomalies capture the temporal dynamics of a time series, i.e. by utilizing functional data analysis [SFS*22, DG18]. They are characterized by a set of data points that can be observed individually as normal but together represent an anomaly [SAL*21]. The second limitation that we emphasize is the scalability limitation. **R5** focuses on the scalability of visual representations and should be distinguished from more holistic approaches (e.g. progressive data analytics [HBS22]). In the next section, we discuss the underlying automated data analysis algorithm that satisfies the requirements of the proposed visual analytics system.

## 4. Methodology for Detecting Anomalies in Cyclic Data

This section describes the method we utilize to detect anomalies on a cycle level (multivariate anomaly time series) and within a cycle (bivariate anomaly time series). One cycle represents data of recurring processes in industrial applications. We first explain the preprocessing of the dataset and the application of a sliding windows approach. Then, we explain KDE, which is utilized to estimate the probability density function (PDF) for each time window. Finally, we demonstrate the method for scoring each time window within a given cycle and the process for aggregating these scores to compute an overall anomaly score for the cycle. To assess the algorithm's effectiveness, we evaluate the types of anomalies that our approach can detect using a synthetic dataset generated with the TimeEval benchmarking toolkit [WSP22].

### 4.1. Kernel-density estimation

KDE [Par62] is a well-established approach for estimating densities. We chose KDE for MANDALA due to its intuitive understandability and close relation to histograms and hexagonal binning plots, which we employ in the main visual. Besides, KDE is a non-parametric density estimator that does not require assumptions about the underlying distribution of the input data [Che17, CBK09b, Zha13]. Unlike deep learning methods, KDE does not require much training data and is efficient for low-dimensional data. Note that other density-based methods for anomaly detection can also work within our framework, including more advanced but computationally expensive approaches, such as the method proposed by Zhang et al. [ZKH04].

**Data preprocessing**. First, we standardize the entire dataset to get a mean of 0 and a standard deviation of 1 for the whole dataset. This is necessary because different variables in the dataset may have different measurement units and different value ranges. If the variables are not standardized, the ones with larger values and wider ranges can dominate the analysis, and the ones with smaller values can be overlooked, using a distance-based approach such as KDE.

**Definition of sliding windows**. By implementing a sliding window approach, the smoothness of the KDE models can be ensured over time, and the influence of individual samples in the reference cycles can be reduced. Partitioning cycles into windows ensures that sequences within the same cycle are being compared to each other rather than a more heterogeneous set of data. Based on our analysis of multiple industrial datasets, we found that a window size of 30 s and a stride distance of 10 s produce adequate outcomes. However, these parameters can be modified to serve the requirements of the specific use case.

**Train KDE with reference cycles**. Our approach requires the user to specify one or several training cycles, which serve as the reference to model the normal data behaviour. Note that we show later in Section 5.1 how the anomaly heatmap view aids the selection of reference cycles and how users can try and compare different selections.

User-selected cycles are concatenated and used to determine the PDF of each sliding window. Since the lengths of cycles are identical, we create one independent KDE model per sliding window. For brevity, we present KDE with univariate random variables. Let $\{x_1, x_2, \ldots, x_N\}$ be $n$ independent and identically distributed samples, the PDF is $f$, and the kernel density is estimated as follows:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{1}$$

where $K(\cdot)$ denotes the kernel function (e.g. Gaussian), and $h$ is the bandwidth that represents a smoothing parameter of the function. With KDE, we construct PDFs that model the typical behaviour of a reference cycle in a particular sliding window. Given a non-reference cycle, we can measure its likelihood under KDE at a certain time step. The returned likelihood will denote how 'similar' the given cycle is to typical reference cycles at that particular time point.

**Multivariate anomaly score time series**. To assign an anomaly score for each time step from a given non-reference cycle, we compute the likelihood of the sample at that time step under the KDE model obtained from the corresponding window in the reference cycles. We invert this score to obtain the *anomaly score* in the [0,1] range. A higher anomaly score indicates a higher likelihood that the given cycle does not follow the estimated distribution $\hat{f}_h(.)$ at that time step. This time-wise scoring helps the user understand when a certain cycle deviates from the typical behaviour observed in reference cycles. Visual inspection and assessment of deviations are the main issues we address with MANDALA.

**Cycle-wise anomaly score**. A cycle's anomaly score is computed by averaging the multivariate anomaly score time series. Cycle-wise scoring makes it possible to visualize and compare different cycle anomaly scores.

**Bivariate anomaly time series**. Above, we explain how we calculate the multivariate anomaly score involving all dimensions. The method is applied to each pair of dimensions individually to calculate the bivariate anomaly score time series.

### 4.2. Anomaly type evaluation

As a preliminary step, we evaluate our algorithm's effectiveness by assessing the types of anomalies it can accurately identify. To achieve this, we generate a synthetic dataset using the GutenTAG anomaly and time series generator, a component of the TimeEval benchmarking toolkit [WSP22]. In a semi-supervised setting, we create 10 time series with the GutenTAG generator, each containing a normal time series for training our model and a time series with an injected anomaly of a specific type. This systematic approach enables us to evaluate our algorithm's performance across both individual and contextual dimensions within each time series.

The results, detailed in Table 1, outline the types of anomalies that have been successfully detected, indicating whether they were identified in a univariate or multivariate setting. Additionally, some anomalies could only be conditionally detected. For instance, the extremum anomaly type in a univariate setting can only be detected if it represents a global extremum. Note that the results are not definitive and are highly dependent on the specific time series and anomaly configuration. For comprehensive results and visualizations of the experiments conducted, refer to the supplementary material.

## 5. MANDALA Visualization and Interaction Design

In this section, we introduce and describe our proposed visualization, interaction design and analysis approach based on the design requirements and automated data analysis. Key interactions are detailed in the text and labelled as **I1-7**, with an overview provided in Table 2.

**Table 1:** *Evaluation of anomaly types shows that the KDE-based approach can detect anomalies either by analyzing dimensions individually (univariate) or in the context of other dimensions (multivariate).*

| Anomaly type | Univariate detection | Multivariate detection |
|---|---|---|
| Amplitude | Cond. | Yes |
| Extremum | No | Yes |
| Frequency | No | Yes |
| Mean | Yes | Yes |
| Pattern | No | Yes |
| Pattern-shift | No | Yes |
| Platform | Cond. | Yes |
| Trend | Cond. | Yes |
| Variance | Cond. | Yes |
| Mode-correlation | No | No |

*Note*: Some anomaly types can be detected under specific conditions (cond.). Further details are available in the supplementary material.

**Table 2:** *Summary of interactions of the proposed design.*

| Label | Type | Purpose |
|---|---|---|
| **I1** | Selecting | Select reference data for modelling the KDE |
| **I2** | Filtering | Filter heatmap view by anomaly score |
| **I3** | Selecting | Select cycle to explore in the SPLOM with selected reference data |
| **I4** | Semantic Zooming | Semantic zooming to explore SPLOM in greater detail |
| **I5** | Brushing and linking | Select data in lineplot views to filter points in SPLOM to investigate data within a temporal scope |
| **I6** | Annotating | Preliminary work on interactive labelling of anomalies |
| **I7** | Filtering | Hovering over suggestions highlight the corresponding data in the SPLOM |
| **I8** | Encoding | Change scope of color-encoding of SPLOM points |
| **I9** | Visualize | Switch to forward-looking upper triangle design of SPLOM |

### 5.1. Anomaly heatmap view

As the starting point for the analysis, **R1** will be met by selecting the reference cycles through the anomaly heatmap view Figure 1(a). It offers a drop-down list that allows users to select one or multiple cycles as reference data (**I1**). After selection, a KDE model is trained with the reference data using the method explained in Section 4. The model is then applied to all other cycles, resulting in an anomaly score between 0 and 1 for all cycles. These values are further mapped to the anomaly heat map visualization, which shows all cycles in temporal order (from left top to right bottom), whereas their colour-encoding visualizes their anomaly score between 0 (no anomaly—white) and 1 (maximum anomaly score—red) through a gradient colour map. A from-to-ruler can filter the heat map to visualize cycles within an anomaly score range to facilitate the exploration of cycles for further investigation (**I2**). Anomaly candidate cycles in the heat map can be clicked to be investigated in the SPLOM in the centre view (**I3**).

### 5.2. KDE—Analysis matrix view

The main view in the centre enables the exploration of anomaly candidate cycles detected by the KDE-based approach. Practically, these cycles show high anomaly scores compared to the reference data. Considering multivariate data, we use an SPLOM as the main visual to show all pairwise data distributions of the reference and a candidate cycle for comparison Figure 1(b). The goal of users is to assess and possibly confirm the existence of anomalies, the dimensions and the data intervals in which they occur.

Points in scatterplots show the candidate cycle's data and visualize an anomaly score between 0 and 1 for each point using the same colour-encoded gradient colour map used in the cycle anomaly heatmap view. Each scatterplot is placed in the superposition of a hexagonal binning plot, visualizing the reference data for quick comparison of a cycle's data with the user-selected reference data. In each cell on the diagonal, two histograms show the distribution of the according dimension (1) of the reference data in green and (2) of the candidate cycles data in blue.

Generally, our design requires the comparison of two two-dimensional datasets with each other. On the contrary, the main focus is to detect data points of a candidate cycle that do not lie within the distribution of the reference dataset. The design decision of using hexagonal binning plots to visualize reference data for each scatterplot within the matrix lies in the following considerations.

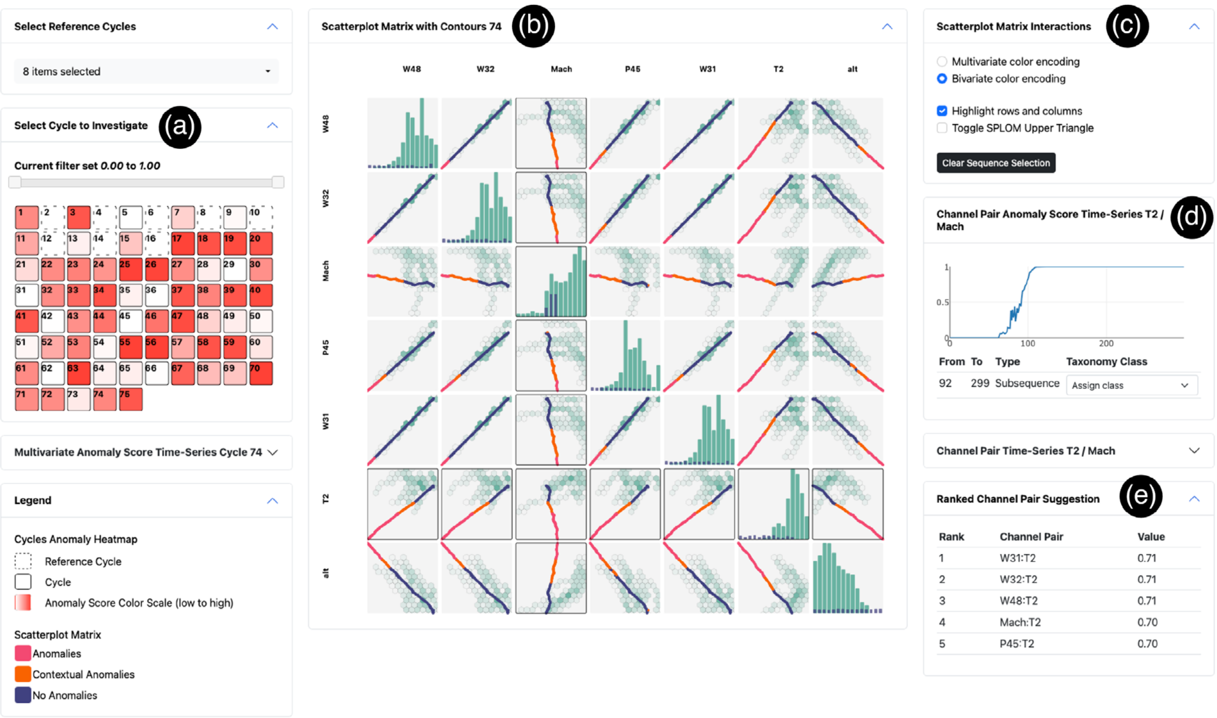
#### 5.2.1. *Design considerations*

One design alternative is using scatterplots to visualize reference data (Figure 2a). Visualizing two scatterplots in superposition would be an alternative if the number of data points in the scatterplots were not variable and relatively small. However, users can select as many reference cycles as required; therefore, a more scalable visualization is needed. Another alternative we discussed is contour plots, which aim to visualize bivariate distributions. They work well to detect data points in a superpositioned scatterplot that are not within the visualized distribution (Figure 2b). However, finding the right parameter set for this visualization technique was difficult, especially in defining the number of contours. We also considered Splatterplots [MG13] as an alternative due to its high scalability and intuitive design (Figure 2c). Like the contour plot, finding a set of generic parameters (bandwidth and threshold) for various reference data and several zoom levels was difficult. Therefore, we use hexagonal binning plots due to their simple design and advantages using a semantic zooming approach for the SPLOM (Figure 3). When zoomed out, the distributions are easy to perceive through the coloured areas of the hexagons, whereas when zoomed in, the size of hexbins can be decreased to show more details of the distribution on demand.

#### 5.2.2. *Semantic zooming features*

We added semantic zooming features to improve a simple geometric zooming approach and to add scalability to the SPLOM (**I4**). Generally, they provide more details when zooming in, whereas visualizing a fair amount of details on a lower zooming level (see Figure 3).

**Hexagon size**. The size of hexagons in the hexagonal binning plot changes depending on the zoom level. A low zoom visualizes



○ Multivariate color encoding
● Bivariate color encoding

☑ Highlight rows and columns
☐ Toggle SPLOM Upper Triangle

| From | To | Type | Taxonomy Class |
|---|---|---|---|
| 92 | 299 | Subsequence | Assign class |

| Rank | Channel Pair | Value |
|---|---|---|
| 1 | W31:T2 | 0.71 |
| 2 | W32:T2 | 0.71 |
| 3 | W48:T2 | 0.71 |
| 4 | Mach:T2 | 0.70 |
| 5 | P45:T2 | 0.70 |

**Figure 1:** *The user interface of MANDALA: (a) The tool requires users to select reference data. After the selection, the red colour-encoding highlights anomaly candidate cycles that differ from reference data. (b) The scatterplot matrix features semantic zooming and visualizes bivariate data of investigated anomaly candidate cycles as scatterplots and the distribution of bivariates of reference data as hexagonal binning plots. The diagonal shows the univariate distribution for each dimension of the reference data and the anomaly candidate cycle as two histograms in superposition. (c) The colour of scatterplot points can be changed to multivariate or bivariate colour-encoding, changing the scope of the analysis. (d) The temporal scope of anomalies can be explored through line plot views. These views can filter the scatterplot matrix by brushing and linking. (e) Bivariate anomalies are ranked by their anomaly score in the ranked channel pair suggestion view.*

fewer hexagons, whereas a high zoom level allows the visualization of more details and, therefore, more hexagons.

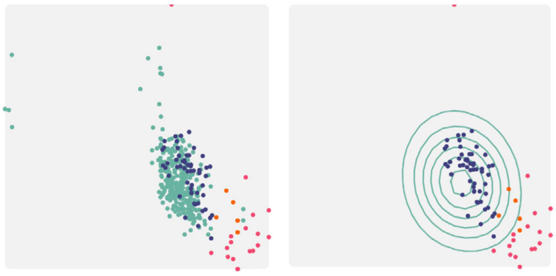
**Axes**. To save screen space, the SPLOM visualizes global axes for each dimension on the top and left. No axes or scales are shown for the possibly small scatterplots on the lowest zoom level. When zooming in, the axes become visible, whereas more and more ticks for a refined scale are displayed the closer the user zooms in to a cell.

**Scatterplot points**. Points in scatterplots can become too small, especially if the SPLOM consists of many dimensions. Therefore, we increase the size of the scatterplot points at the lowest zoom level and decrease the size when zooming in.

### 5.2.3. *Scatterplot colour-encoding*

MANDALA requires detecting and exploring different anomaly types (**R2, R3**). To facilitate the exploration of anomalies, contextual anomalies and non-anomalous data, we use the results of the KDE-based automatic anomaly detection and colour-encode scatterplot data points in three ways: (1) Anomalous data points are visualized in a red colour-encoding, which is a complement to the green colour-encoding of the reference data visualized through hexagonal binning plots, (2) contextual anomalies are visualized in orange and (3) non-anomalous data points are visualized in blue colour (see Figure 4). Nevertheless, those colour choices are not definitive and can be programmatically changed in the prototype if required. Furthermore, MANDALA distinguishes two different colour-encoding strategies for points in the SPLOM: (1) the multivariate colour-encoding and (2) the bivariate colour-encoding (Figure 1c, **I8**). Generally, multivariate anomaly detection techniques detect anomalies that involve all dimensions simultaneously, whereas detected anomalies lie in this multidimensional space. Using the multivariate colour-encoding strategy, all data points throughout the SPLOM are linked together using the same colour-encoding. This enables the exploration of 2D scatterplots and the anomaly score at a broader scope involving all dimensions. On the other hand, the bivariate colour-encoding strategy uses an anomaly score for each dimension pair individually, which omits the linking among scatterplots through colour-encoding. This implies that a KDE model needs to be trained individually for each bivariate through the reference data and applied for each point in each scatterplot.

To demonstrate the exploration of contextual anomalies, an example is given in Figure 4 visualizing histograms and scatterplots with hexagonal binning plots of two dimensions, *Mach* and *W32*. Anomalous data points in red are data points that have lower values



**(a) scatterplot** **(b) contourplot**

**(c) splatterplot** **(d) hexbinplot**

**Figure 2:** *Design alternatives, which we considered for comparing reference data and anomaly candidate data. The reference data is visualized with a green colour-encoding and anomaly candidate data is blue (close to reference) or orange and red (anomaly). Our goal is to visually indicate the different characteristics: Reference data visualizes many cycles. Hence, an aggregate view like dense point plots or hexbin plots is appropriate. Anomaly candidate data is more sparse, hence shown as points. The designs should allow for the contrast of both distributions for comparison.*

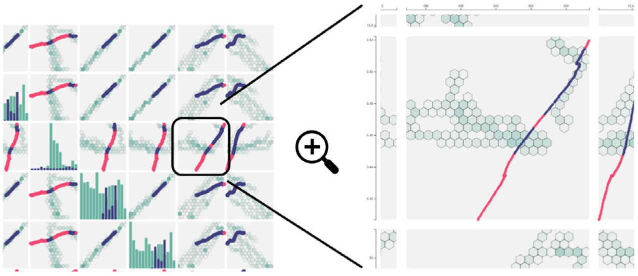


**Figure 3:** *Semantic zooming of SPLOM cells shows details on demand. The hexagonal binning plot is refined through smaller hexagon sizes, axes visualize a refined scale and the size of scatterplot points is smaller compared to a geometric zooming approach.*

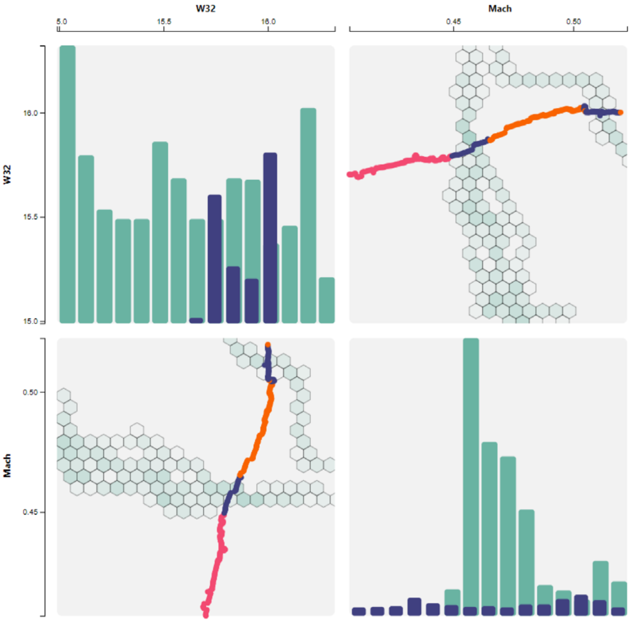


**Figure 4:** *Example of the SPLOM visualizing histogram and scatterplots of two dimensions. It shows normal scatterplot data in blue, anomalous data in red and contextual anomalies in orange. Additional examples of anomalies can be found in the supplementary material.*

than in the reference data regarding the *Mach* dimension. On the other hand, histograms show that the contextual anomaly data points in orange lie within the distribution of both dimensions. This means data points do not appear anomalous when looking at both dimensions individually. Nevertheless, scatterplots reveal that data points are anomalous in each other's contexts. Note that MANDALA supports the assessment of anomalies on three scopes: univariate (using histograms on the diagonal), bivariate (in the SPLOM cells) and multivariate (by multivariate colour-coding of points in all SPLOM cells).

The time dimension of the time series data is not visible when using a scatterplot representation. Nevertheless, the time dimension can be relevant in time series anomaly detection, for example, to understand an anomaly's temporal scope, duration, or correlation pattern. We added the multivariate anomaly time series view to the user interface (see Figure 1d). The view visualizes the time series of the anomaly score of a cycle, which has been calculated by the KDE anomaly detector (Section 4). It allows visually inspecting an anomaly's temporal occurrence and progress within a cycle. Furthermore, through this view, users can distinguish among 3 time series anomaly types (**R2**): point anomalies, subsequence anomalies and whole time series anomalies ([BGCML21]), which are calculated by the system and are listed below the time series view (Figure 6). Complementary to the multivariate and bivariate colour-encoding strategies, we added a second anomaly score time series line plot to the view. It visualizes the bivariate anomaly score time series to investigate the anomalous behaviour of a dimension pair over time. The juxtaposition to the multivariate anomaly score time series allows for a quick comparison between those two line plots, that is, to detect correlations. A positive correlation indicates a contribution of the channel pair to the multivariate anomaly score.

Furthermore, both line plots are interactive and implement a linking and brushing technique ([Kei02]). Anomalous sequences of interest can be selected in line plots to filter data points in the SPLOM (**I5**). This helps the user to quickly understand the relationships between anomalies and their temporal scope.

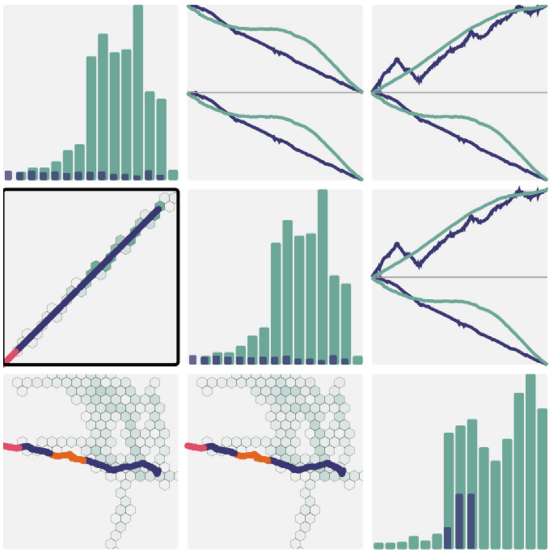


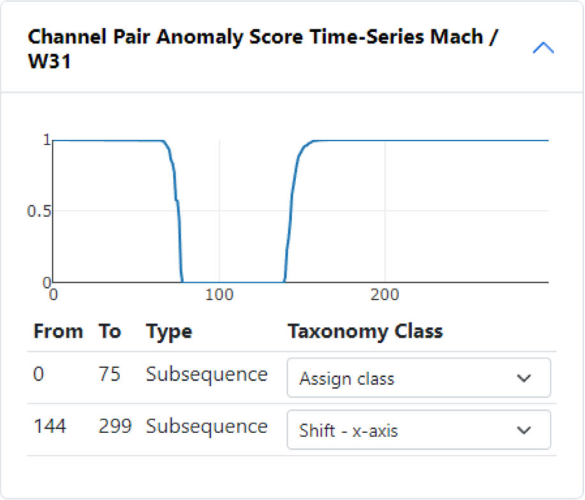
**Figure 5:** *Forward-looking design visualizing time series of single dimensions and the mean time series of reference data in the upper triangle of the SPLOM.*

### 5.3. Upper triangle of the scatterplot

Related work shows that the upper triangle of a scatterplot matrix can display additional information, such as correlation coefficients, to enhance the interpretation of variable relationships and to avoid redundancy. However, maintaining symmetry between the upper and lower triangles ensures visual consistency, making the matrix more intuitive for users. This is particularly advantageous in interactive settings, as it allows users to explore pairwise relationships without the cognitive burden of adapting to different visualizations. However, we created a forward-looking design where the upper triangle optionally displays the underlying time series data to utilize the space and avoid redundancy. In this design, each cell is split into two halves: the upper half shows the time series of the current cycle for the dimension represented on the y-axis, while the lower half displays the time series for the dimension corresponding to the x-axis of the cell in blue colour-encoded line plots. Additionally, we calculated the average time series of all reference cycles and overlaid them using green color-encoding for quick comparison. While this approach may not be ideal for detailed analysis, it gives users an overview of where and how the time series of individual dimensions deviate from the selected reference data. An example of this forward-looking design is shown in Figure 5.

### 5.4. Interactive labelling

Bivariate anomalies are calculated and listed in the bivariate anomaly time series view (Figure 6). As preliminary work on interactive labelling, we allow users to assign anomalies (**I6**) to four classes through the user interface (shift x-axis, shift y-axis, shift x-axis, shift y-axis, in between cycles, see Section 6.4). The system uses the classification system of Suschnigg et al. [SMK*21] to classify anomalies automatically. Labels are stored in a knowledge base and could serve (1) as training data for the classifier to classify future occurrences of similar patterns and (2) as the basis for collaborative analysis.



**Channel Pair Anomaly Score Time-Series Mach / W31**

| From | To | Type | Taxonomy Class |
|---|---|---|---|
| 0 | 75 | Subsequence | Assign class |
| 144 | 299 | Subsequence | Shift - x-axis |

**Figure 6:** *Interactive labelling: Anomalies are detected and assigned a type based on their duration. Additionally, users can label anomalies through the drop-down menu.*

### 5.5. Channel pair suggestions view

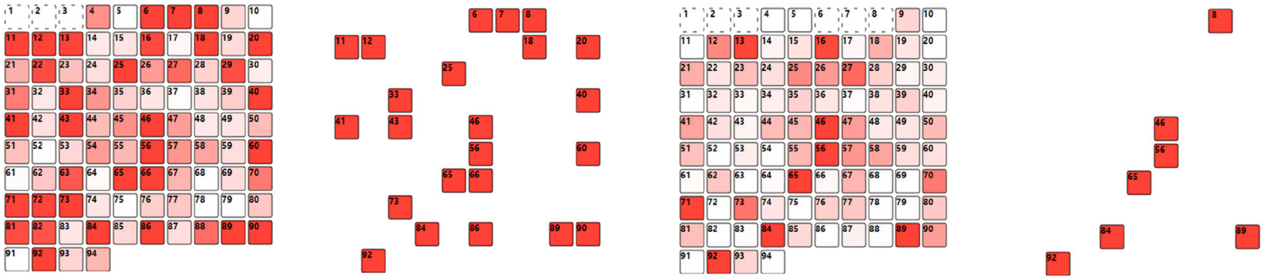
The SPLOM visualization of multivariate data becomes more overwhelming to the user as more dimensions are shown simultaneously. We implemented the channel pair suggestion view to offer starting points to explore the SPLOM (**R4**). Based on their bivariate anomaly score, it ranks and suggests candidate cells within the scatterplot in tabular form (Figure 1e). This view is connected to the anomaly time series view and is updated after a linking and brushing interaction. It can be effectively used by hovering over a suggestion to highlight the cell's row and column in the SPLOM (**I7**).

### 5.6. Implementation

MANDALA is a web application built on a Django web server using the Python programming language. The backend uses the Pandas and Scikit-learn libraries to calculate anomaly scores using KDE. The client communicates with the server via an API handled through the AngularJS front-end library. Most visual components and interactions are written in JavaScript through the D3.js framework. Line plots are rendered with the Plotly.js library. In addition, we use Bootstrap for the appearance of the HTML controls and the grid system.

## 6. Use Case Scenario

MANDALA has been designed for multivariate time series datasets, which display cyclic behaviour. We tackled such datasets in various research projects, e.g. automotive testing and production [SMF*20, GFS*19], manufacturing [SZB*20], or aluminium production [JSN*21]. This use case describes how MANDALA can assist in analyzing field data from electric vehicles (EVs) collected



**(a) Initial reference selection without filter** **(b) Initial reference selection with filter** **(c) Additional reference selection without filter** **(d) Additional reference selection with filter**

**Figure 7:** *Reference cycle selection: This figure shows how reference selection changes the anomaly heatmap view during the use case scenario. The initial selection shows many cycles with an anomaly score close to one (a and b), whereas extending the reference selection adds diversity to the reference data and, therefore, lowers the anomaly score throughout the whole dataset (c and d).*

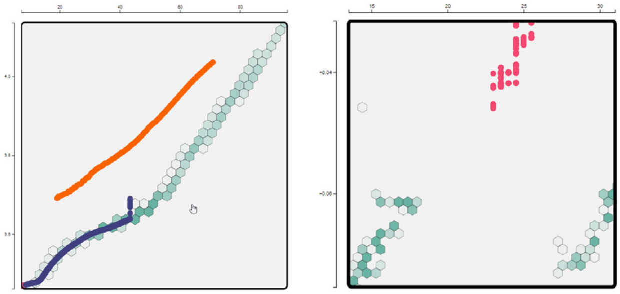
from real-world applications. The primary objective of our analysis is to identify intriguing patterns that can provide insights into the technology of EVs, especially in understanding battery thermal ageing phenomena. The use case scenario has been conducted with data and a data scientist from a global company specializing in developing and engineering powertrain systems. The company plays an important role in helping automotive manufacturers develop and optimize their vehicles for better performance, efficiency and environmental impact.

### 6.1. Electric vehicle dataset

This dataset comprises real-world field data collected from a large fleet of EVs (>1,000), featuring various sensor measurements, including battery temperature, battery voltage and state of charge, recorded at 30-s intervals. In particular, all trips have been recorded over a 2-year time span, extend beyond 4 h and represented journeys with charging events that are less prone to noise due to various driving scenarios and environmental conditions. For this use case scenario, we analyze data from 94 trips of a single EV, fulfilling the criteria. The dataset provides information to explore the behaviour and performance of EVs over extended periods of operation.

### 6.2. Reference cycle selection

First, we discuss the selection of the reference cycle and pick the first two cycles as references. From a domain perspective, selecting data early in the product lifetime and detecting deviations later on in the data is a common approach but not necessarily the only way to analyze industrial data. Discussions with the domain expert highlight that the problem of anomaly detection is difficult because it is not specified for many cases since the normal state of a machine is often unknown. After selection, many cycles appeared to be anomalous, and we selected the third cycle for analysis. The exploration of the third cycle showed that the detected anomalies are irrelevant from a domain perspective and, therefore, can be added to the selection of reference cycles. This process has been repeated with cycles 6, 7 and 8, whereas every time the reference cycle view is updated, fewer anomalies are shown because the diversity of non-anomalous data increases (see Figure 7). At first sight, the domain expert was sceptical because of the amount of anomalous data points with only two cycles selected as reference, but afterwards appreciated the progressive approach of analysis and exploration and the possibility of adding domain expertise at each iteration. As follows, we discuss two cases of anomalies that have been investigated.



**(a) Ambiguous anomaly** **(b) Strong anomaly**

**Figure 8:** *Two examples of anomalies explored with MANDALA in the use case scenario: (a) MANDALA has detected an ambiguous anomaly, which suggests a slight shift in the time series's y-axis dimension within a subsequence. (b) The strong anomaly suggests that the data lie within two references on the x-axis dimension with an upshift of the y-axis dimension.*

#### 6.2.1. *Exploration of anomalies*

One feature of MANDALA is the exploration of ambiguous multivariate anomalies through the small multiple view by comparing selected reference data with candidate anomalies across all dimensions. These anomalies are difficult to interpret by automated data analysis and, therefore, require SMEs. In Figure 8, two examples show the relationship between the most critical measuring parameters of the battery, such as state-of-charge, voltage and C-rate. In contrast to ambiguous anomalies that are depicted in panel (a), MANDALA can detect and highlight strong anomalies (see panel (b)) from the corresponding charging cycles. From a data perspective, those anomalies are clearly anomalous but still require domain experts to confirm or disprove these findings. For that purpose, users

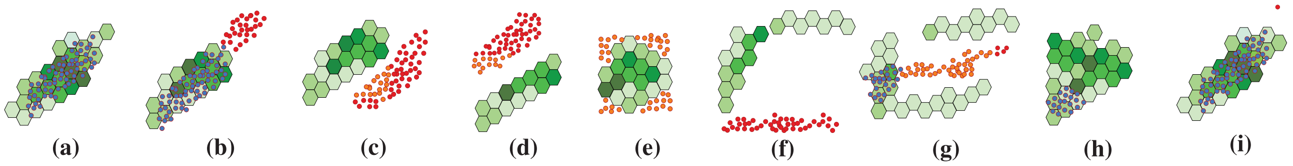


**Figure 9:** *Taxonomy: Exploring differences between reference data visualized as hexagonal binning plots and inspected data visualized as scatterplots show different patterns. In this taxonomy, we list and interpret the patterns that occurred during our experiments on different datasets. The figure uses the same colour-encoding as explained in Section 5.2.3, distinguishing between normal data, anomalies and contextual anomalies.*

can explore the temporal scope of anomalies through the time series views and the dimensions involved in the scatterplot view. Such in-depth investigation is critical for timely identifying joint deviations from their nominal range that might signify the upcoming thermal failure of EV batteries [LLAJ21].

### 6.3. Domain expert remarks

The use case scenario has been conducted with a SME who has extensive experience in data science and mechanical engineering yet has limited exposure to visual analytics. After the collaborative data pre-processing and exploration in MANDALA, follow-up discussions explore specific challenges analysts face, providing insights into potential improvements. Generally, the tool was determined to be well designed by the SME, demonstrating proficiency in handling complex time series data. The SME noted the intuitive interface, which facilitates seamless navigation, and highlighted the added value of heatmap filtering, which guides users through the exploration process. The histograms in the SPLOM was recognized for its effectiveness in rapidly displaying dimensional differences from the selected reference data. Additionally, scatterplots were identified as a valuable component, allowing a more detailed examination of specific data points and improving the understanding of anomalies by exploring bivariate subsets of the data. The SME concluded that the thoughtful integration of these features positions the tool as a valuable resource for deriving insights from the explored dataset. However, the domain expert prefers automatic reference selection and states that 'this feature would reduce cognitive load, enhance efficiency and allow a more targeted exploration of the time series data'.

### 6.4. Taxonomy hexagonal binning plot and scatterplot

During our work on MANDALA, we observed various hexagonal binning plots and scatterplots in superposition, showing unique shapes and patterns. We created a taxonomy as shown in Figure 9 to classify those patterns. The taxonomy illustrates reference data distributions and their position to scatterplots. Note that the distribution of reference data and scatterplots depicted in the taxonomy are simplified illustrations and can become more complex and/or mixed with actual application data. Additionally, the taxonomy highlights examples of contextual anomalies using orange colour-encoded points. As follows, we give short descriptions and interpretations of the patterns:

(a) **Full match**: Data fully lies within the reference data and does not visualize anomalous data.
(b) **Partial anomalous**: Data lies partially outside of the reference data.
(c) **Shift x-axis**: Values of the x-axis dimension are shifted to higher or lower values.
(d) **Shift y-axis**: Values of the y-axis dimension are shifted to higher or lower values.
(e) **Contextual**: Univariate data of both dimensions lie within the reference but do not lie within the bivariate distribution.
(f) **Different cycle**: Data distribution differs from the reference data and is probably created by a different underlying process.
(g) **In between cycles**: Data distribution lies between values of two reference cycles. Anomalies appear as contextual anomalies.
(h) **Diverse references**: Data distribution from many reference cycles covers a larger range in both dimensions than single cycles.
(i) **Point anomaly**: Only single points appear anomalous in the plot.

This taxonomy is based on our observation of many cycles and data subsets and is proposed as a reference for analysts on what to look for when starting an exploration. While we have preliminarily implemented and investigated the automatic recognition of taxonomy patterns through interactive labelling and time series views, the patterns should be useful for ranking and suggesting candidate views to the user and filtering large view spaces. They might also be the basis for computing features analogous to the Scagnostics approach [WAG05]. These, in turn, could be useful for ranking and grouping data and providing user guidance in the exploration process. We leave this to future work.

## 7. User Study

We conducted a user study to investigate how readily the user can identify and explore anomalies in multivariate times series using MANDALA. To do that, we defined 3 qualitative research questions (**RQ**)

**RQ1** How effective is MANDALA in assisting users to identify and explore anomalies in multivariate time series data?

**RQ2** How effectively does MANDALA support users in detecting anomalies within a temporal range in multivariate time series data?

**RQ3** Does the visualization provided by MANDALA enable users to identify the dimensions that contribute to the anomaly score in multivariate time series data?

The user study results should also help us identify usability gaps and assess the user experience when working with MANDALA. This section describes the methods utilized during the study and their qualitative and quantitative results.

### 7.1. Aircraft turbine dataset

The synthetic dataset used for the user study represents sensor data from an aircraft engine generated by a simulation that runs the aircraft to failure for diagnostic and prognostic purposes using real flight condition data as input [CKGF21]. Given the run-to-failure setup, this dataset is suitable for our research on anomaly detection. Specifically, we selected a simulation of 75 flight cycles and focused on the first 5 min of the climb phase after the liftoff. We align each flight cycle with the altitude dimension to facilitate better comparability of cycles. We select seven dimensions that do not require extensive domain knowledge in flight mechanics, such as the *Mach* number and *Altitude*. In addition, we select correlating and non-correlating dimensions to obtain a variety of patterns through visualization.

### 7.2. Procedure

The study included 15 participants, mainly post-graduate, 2 females and 13 males, between 18 and 37 years old. All participants had prior experience with visual and data analytics tools, and most were from the field of Computer Science. Furthermore, the collective experience of the participants included a wide range of industry-related projects. These projects ranged from predictive maintenance to condition monitoring. This means that the participants were involved in activities that included using analytical techniques to predict and monitor the condition of machines and systems in various industries. Their hands-on involvement in such projects contributed to their understanding and proficiency in the use of visual and data analytics tools. However, none of the participants had used MANDALA before. This ensured that their interactions and responses with the MANDALA tool during the study were unbiased and uninfluenced by familiarity with the tool.

The study began by showing participants a 9 min instruction video, including a video demonstration of the user interface features of MANDALA. After the demonstration, participants were given a subset of the dataset that is used in the study and had the opportunity to explore and use the tool on their own for 5 min. The example dataset is a smaller variant of the dataset used for the tasks, consisting of only four dimensions and twenty cycles. After 5 min had expired, the study officially began. During the study, participants were given four tasks to complete: Task 1 (**T1**), Task 2 (**T2**) and Task 3 (**T3**), comprising two subtasks. These tasks were carefully structured to increase in difficulty and complexity incrementally. This sequence served primarily to enable a step-by-step learning process for users who were not yet familiar with the tool. Furthermore, drawing on our previous experience working with professionals in various industries, we have deep insight into how professionals identify anomalies. This knowledge has guided us in the careful design of these tasks.

**T1** was designed to address **RQ1** and required the detection of cycles with two specific anomaly score thresholds using the anomaly heatmap view. **T2** aimed to answer **RQ2** and asked to investigate the temporal scope of the anomalies and required participants to detect the anomalous subsequence within two cycles. **T3** was geared toward **RQ3**, whereas its goal was to detect univariate dimensions that deviate from reference data within a specific cycle by investigating the SPLOM. Additionally, the task aimed to detect the dimension that exhibits values higher than those in the reference data. Task 4 (**T4**) also aimed to answer **RQ3** and asked participants to detect bivariates within a cycle that differs from the reference data. Participants were given 3 min to complete the first two tasks **T1** and **T2**, and 5 min to complete the tasks **T3** and **T4**. Once a task was completed or the allotted time was consumed, participants had to submit their responses digitally. After each task, subjective feedback was collected through a post-task questionnaire consisting of a 7-point Likert NASA-TLX scale covering six dimensions of workload (mental demand, physical demand, temporal demand, effort, frustration and perceived performance) [HS88]. In addition, users had the opportunity to give feedback on each task. At the end of the study, participants were asked to complete a positive SUS [B*96, SL09] using a 7-point Likert scale [JKCP15] to measure the overall usability of the proposed tool. In the context of our user study, we deliberately chose not to employ a baseline method for comparative analysis. Since this was an exploratory study, our goal was to examine the tool's functionality, user engagement, and overall user experience in depth. Including a baseline could have limited this qualitative investigation, as the focus could have changed from understanding the inherent qualities of the tool to simply comparing it to an external benchmark.

### 7.3. Pilot study

Before conducting the user study, we conducted a pilot study with two participants. Both were familiar with MANDALA, but were not directly involved in its development. The main goal of the pilot study was to refine the user study procedure, ensure its practicability, and eliminate potential problems. The results of the pilot study led to slight adaptations to the prototype, such as changing the appearance of the histograms in the scatterplots' diagonal cells. In addition, a significant modification on **T4** has been performed. Initially, an additional task was intended to assess how MANDALA can help detect contextual anomalies. However, during the pilot study, we found that explaining contextual anomalies in the instructional video was too difficult for participants to comprehend. As a result, the task was substituted with the second subtask in **T2**.

### 7.4. Quantitative results

Thirteen participants successfully completed both **T1a** and **T2a**. Similarly, 12 participants successfully completed **T1b** and **T4**. However, while fourteen participants were able to complete **T2a**, only six participants were able to complete **T4** successfully. The average time for **T1** was 1 min and 24 s, for **T2** 2 min and 16 s, for **T3** 3 min and 48 s and for **T4** 4 min and 3 s. The workload was estimated using the results of the NASA-TLX questionnaire. The results displayed in Figure 10 have revealed a low overall workload in all tasks (mean below 50): (**T1**) $\mu = 25$, $\sigma = 7.6$, (**T2**) $\mu = 42.7$, $\sigma = 16.3$, (**T3**) $\mu = 36.6$, $\sigma = 14.4$ and (**T4**) $\mu = 33.5$, $\sigma = 13.4$. In

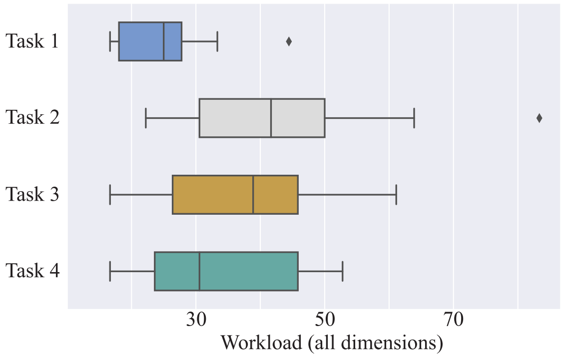


**Figure 10:** *Overall workload per task calculated based on NASA-TLX results: all dimensions have a low overall workload on all tasks (mean below 50).*

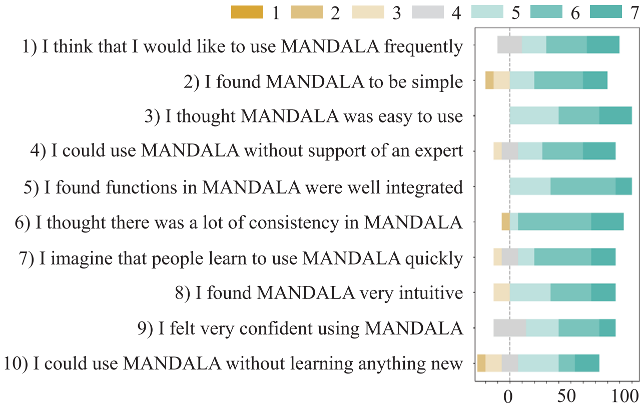


**Figure 11:** *The participants' responses to the system usability scale (SUS) questionnaire: MANDALA obtaining a B score for overall usability.*

summary, **T2** caused a higher workload than the other tasks, while **T4** was the most challenging task in terms of completion rate.

We averaged all the questions and participants to determine the average usability score based on the SUS questionnaire and obtained a mean raw score of 75.6 ($\sigma = 11.2$). Our design falls in the 70–79 percentile range in the curved grading scale interpretation of SUS scores, thus obtaining a B score [SL09]. Note that we used a 7-point Likert scale instead of a 5-point one ($1 =$ strongly disagree, $7 =$ strongly agree). Thus, user responses were multiplied by 1.66 instead of 2.5 to obtain overall SUS scores in a range between 0 and 100. These scores are subdivided into the Usable subscale (questions 1, 2, 3, 5, 6, 7, 8 and 9) and the Learnable subscale (questions 4 and 10) following the approach presented in [LS09]. After the multipliers are adjusted on the 7-point Likert scale (2.08 and 8.33), MANDALA scored a B for Usable ($\mu = 77.7$, $\sigma = 11.1$) and Learnable ($\mu = 70, 8$, $\sigma = 16.8$). The questions and results of the SUS questionnaire are visualized in Figure 11 by the diverging stacked bar graph. It represents the Likert scale of survey responses. The Likert bar chart shows the responses collected from the Likert scale survey, showing the percentage of responses for each of the seven options. The size of each bar corresponds to the proportion of respondents for that specific option.

Regarding **RQ1**, MANDALA demonstrated high effectiveness in assisting users in identifying and exploring anomalies in multivariate time series data, as indicated by the high competition rate and lower workload for **T1**. However, the higher workload for **T2** suggests that some users may struggle to identify anomalies within a temporal range in a multidimensional space. Furthermore, the effectiveness of MANDALA in allowing users to identify the dimensions that contribute to the anomaly score in multivariate time series data remains unclear, as evidenced by the higher workload for **T3** and by the lower competition rate for **T4**. Therefore, more research is needed to address the research questions **RQ2** and **RQ3**.

Overall, looking at the usability score of MANDALA (B), we can conclude that it is an effective and useful approach for detecting and exploring anomalies in multivariate time series data. Although a significant number of participants were successful in completing **T1**, **T2** and **T3**, the higher workload for **T2** and the lower competition rate for **T4** suggest some areas for improvement in the design of MANDALA. Participants provided feedback and suggestions for improvement, which we discuss in the next sections to identify specific areas that require further attention.

### 7.5. Qualitative results: User feedback

After completing the NASA-TLX questionnaire for each task, the participants were asked to give feedback, which we summarize as follows.

Feedback on **T1** suggests that the filtering feature was effective. Participants found the filter to be intuitive, smooth and helpful in completing the task. The use of a slider filter was especially appreciated by one participant who had initially gone through the cycles one by one. However, one participant felt that it would be beneficial to know more about the metrics used to determine anomaly scoring. Providing additional information regarding the automated data analysis could improve the overall understanding and appreciation of the work. The user feedback on **T2** suggests that there are several areas where improvements could be made in the anomaly time series view. Participants commented that it would be helpful to have the option to textually enter time ranges, making the task easier to complete. They also reported that a larger line plot view could be beneficial. Addressing these issues would help improve the research tool's overall usability and user experience. The participant feedback on **T3** highlights both positive aspects of solving **T3** and some areas where improvements could be made. Overall, participants found the tool to be helpful and effective in solving **T3**. However, several participants mentioned that the opacity of the histogram bars could occasionally make it difficult to distinguish between the test data and reference data. Participants suggested that making the bars more transparent or showing them as separate bars would help to address this issue. One participant also noted that the video instructions were unclear and that better instruction could be provided to clarify the intended task, e.g. by providing examples. Regarding **T4**, several participants reported that they found the MANDALA visualization helpful, which suggests that the tool effectively presents data in an understandable and interpretable way. However, some participants also commented that improvements concerning the user study could be made. E.g. one participant was unsure about how to classify a particular data point, highlighting the need for clearer guidance or instructions. The following section will discuss limitations and potential future work identified during the development of MANDALA and through the results obtained from the user study and the use case scenario.

## 8. Limitations and Future Work

While the system was designed with scalability in mind, the lack of emphasis on this aspect resulted in certain limitations. This suggests a design implication that future systems should prioritize scalability as a core design principle to ensure robust performance, especially when dealing with large or complex datasets. Additionally, the system's current limitation in detecting and visualizing only a single context of contextual anomalies indicates the need for more advanced visualization techniques that can handle multidimensional contexts. The challenge of handling high frequency time series data reveals another design implication: the necessity of developing adaptive visualization and automated data analysis techniques that can dynamically adjust to different data frequencies. This adaptability is crucial for making the system more versatile and effective across varying types of time series data.

Overall, we achieved good results, especially considering the time of 8 min participants had to be introduced to the system and the complexity of the multivariate time series anomaly detection problem. Generally, users understood the system, as it is easy to use, and mostly felt confident using the tool without much training.

The results of our studies demonstrate the potential for visual exploration of anomalies in multivariate time series through our system. However, we identified several directions for future research. Furthermore, the reliance on specifying normal behaviour through reference cycles highlights the need for more automated approaches. The SPLOM was selected as the primary visual component due to its scalability and clear visual structure, whereas a semantic zooming technique was also implemented. Focusing on those two methods in future research can significantly improve scalability for datasets with higher dimensions.

A limitation of this work is its dependence on cyclic multivariate time-series data. This limitation can potentially be addressed through data preprocessing techniques that segment the data into equal time intervals (such as hours or days) to account for seasonality. However, the effectiveness of this approach is contingent upon the characteristics of the dataset and the specific use case. For non-time-series data, utilizing a scatterplot matrix with hexagonal binning plots may be applicable, but this approach lacks interactive capabilities, particularly with line plot views visualizing the temporal context of data.

In our current approach, we emphasize the expertise of domain experts in defining and interpreting normal behaviour, as their contextual knowledge ensures the relevance of selected reference cycles. This method aligns with the real-world complexities of many domains where human understanding is indispensable. However, we acknowledge that automated techniques for selecting or suggesting reference cycles can guide users to uncover patterns or potential insights. Future work could explore unsupervised techniques, such as clustering algorithms, to group similar cycles and suggest candidate references for identifying subsets of data that best represent typical behaviour.

The preliminary work conducted in this research on the automatic detection of taxonomy-based patterns Figure 9 to guide users in exploration needs further investigation. For example, by drawing attention to interesting classes and their relevant dimensions by highlighting remaining cells in the SPLOM [SVLB10], or by investigating other industrial datasets for taxonomy-based error classes. To provide a more comprehensive understanding of the effectiveness of the techniques used in MANDALA, future research could systematically compare design alternatives (Figure 2) and conduct a more comprehensive user study. To date, MANDALA has been comprehensively applied to the aircraft turbine dataset from the user study and the EV dataset featured in the use case scenario. To further assess the generalizability of our findings, we are conducting experiments with additional industrial datasets and collaborating with experts who have in-depth knowledge of the application's specific requirements. Extending the evaluation to include other use cases and industrial domains may provide valuable insights and inform further improvements to MANDALA.

## 9. Conclusion

We introduced and implemented MANDALA, a visual interactive approach for detecting anomalies in multivariate time series data, mainly in industrial settings. The tool uses a learning approach using KDE for anomaly calculation and proposes SPLOMs, hexagonal binning plots and line plots for visualization and exploration. It provides effective techniques to compare data along all dimensions of the dataset and along the timeline. The exploration allows us to validate, explain and improve anomaly detection provided by automatic KDE-based anomaly detection.

We demonstrated the effectiveness of MANDALA and presented patterns that emerged during our work using the taxonomy. A user study involving 15 participants confirmed its usability. The results suggest possible future directions, such as reference cycle recommendation and user guidance, advanced visualization techniques for exploring contextual anomalies with more than two contexts, and improved scalability through matrix representations and semantic zooming.

## Acknowledgements

This work has been supported by the FFG, Contract No. 881844: 'Pro$^2$Future is funded within the Austrian COMET Program Competence Centers for Excellent Technologies under the auspices of the Austrian Federal Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology, the Austrian Federal Ministry for Digital and Economic Affairs and of the Provinces of Upper Austria and Styria. COMET is managed by the Austrian Research Promotion Agency FFG.' We thank Julian Rakuschek for his insights during the writing process and the participants from Pro$^2$Future and the Institute of Visual Computing for their contributions to the evaluation. We also thank the reviewers of the previous and current submissions for their constructive feedback, which helped improve this study.

**Conflict of Interest Statement**

The authors declare no conflict of interest.

**References**

[BGCML21] Blazquez-Garcia A., Conde A., Mori U., Lozano J. A.: A review on outlier/anomaly detection in time series data. *ACM Computing Surveys 54*, 3 (2021). https://doi.org/10.1145/3444690.

[BDF*14] Behrisch M., Davey J., Fischer F., Thonnard O., Schreck T., Keim D., Kohlhammer J.: Visual analysis of sets of heterogeneous matrices using projection-based distance functions and semantic zoom. *Computer Graphics Forum 33*, 3 (2014), 411–420. https://doi.org/10.1111/cgf.12397.

[B*96] Brooke J.: SUS-A quick and dirty usability scale. *Usability Evaluation in Industry 189*, 194 (1996), 4–7.

[CBK09a] Chandola V., Banerjee A., Kumar V.: Anomaly detection. *ACM Computing Surveys 41* (2009), 1–58. https://doi.org/10.1145/1541880.1541882.

[CBK09b] Chandola V., Banerjee A., Kumar V.: Anomaly detection: A survey. *ACM Computing Surveys (CSUR) 41*, 3 (2009), 1–58.

[Che17] Chen Y.-C.: A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology 1*, 1 (2017), 161–187. https://doi.org/10.1080/24709360.2017.1396742.

[CKGF21] Chao M. A., Kulkarni C., Goebel K., Fink O.: Aircraft engine run-to-failure dataset under real flight conditions for prognostics and diagnostics. *MDPI Data 6* (2021), 1–14. https://doi.org/10.3390/data6010005.

[CLNL87] Carr D. B., Littlefield R. J., Nicholson W. L., Littlefield J. S.: Scatterplot matrix techniques for large n. *Journal of the American Statistical Association 82*, 398 (1987), 424–436. https://doi.org/10.1080/01621459.1987.10478445.

[CMF20] Cook A. A., Misirli G., Fan Z.: Anomaly detection for IOT time-series data: A survey. *IEEE Internet of Things 7* (2020), 6481–6494. https://doi.org/10.1109/JIOT.2019.2958185.

[Cui19] Cui W.: Visual analytics: A comprehensive overview. *IEEE Access 7* (2019), 81555–81573. https://doi.org/10.1109/ACCESS.2019.2923736.

[DG18] Dai W., Genton M. G.: Multivariate functional data visualization and outlier detection. *Journal of Computational and Graphical Statistics 27*, 4 (2018), 923–934. https://doi.org/10.1080/10618600.2018.1473781.

[DMSO10] Das S., Matthews B. L., Srivastava A. N., Oza N. C.: Multiple kernel learning for heterogeneous anomaly detection: Algorithm and aviation safety case study. In *Proceedings of the SIGKDD International Conference on Knowledge Discovery and Data Mining* (2010), ACM, pp. 47–56. https://doi.org/10.1145/1835804.1835813.

[EDG*08] Elmqvist N., Do T.-N., Goodell H., Henry N., Fekete J.-D.: ZAME: Interactive large-scale graph visualization. In *2008 IEEE Pacific Visualization Symposium* (2008), IEEE, pp. 215–222. https://doi.org/10.1109/PACIFICVIS.2008.4475479.

[FFV*12] Fischer F., Fuchs J., Vervier P.-A., Mansmann F., Thonnard O.: VisTracer: A visual analytics tool to investigate routing anomalies in traceroutes. In *Proceedings of the International Symposium on Visualization for Cyber Security* (2012), VizSec '12, ACM, pp. 80–87. https://doi.org/10.1145/2379690.2379701.

[FSS*21a] Fischer M. T., Seebacher D., Sevastjanova R., Keim D. A., El-Assady M.: CommAID: Visual analytics for communication analysis through interactive dynamics modeling. *Computer Graphics Forum 40*, 3 (2021), 25–36. https://doi.org/10.1111/cgf.14286.

[FSS*21b] Fujiwara T., Shilpika, Sakamoto N., Nonaka J., Yamamoto K., Ma K.-L.: A visual analytics framework for reviewing multivariate time-series data with dimensionality reduction. *IEEE Transactions on Visualization and Computer Graphics 27*, 02 (2021), 1601–1611. https://doi.org/10.1109/TVCG.2020.3028889.

[GD21] Gonzalez J., Dang T.: OutViz: Visualizing the outliers of multivariate time series. In *Proceedings of the 12th International Conference on Advances in Information Technology* (2021), IAIT '21, ACM. https://doi.org/10.1145/3468784.3471606.

[GFS*19] Gutschi C., Furian N., Suschnigg J., Neubacher D., Voessner S.: Log-based predictive maintenance in discrete parts manufacturing. In *Procedia CIRP - Conference on Intelligent Computation in Manufacturing Engineering* (2019), vol. *79*, Elsevier, pp. 528–533. https://doi.org/10.1016/j.procir.2019.02.098.

[GGAH14] Gupta M., Gao J., Aggarwal C. C., Han J.: Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering 26*, 9 (2014), 2250–2267. https://doi.org/10.1109/TKDE.2013.184.

[HBS22] Hogräfer M., Burkhardt J., Schulz H.-J.: A pipeline for tailored sampling for progressive visual analytics. In *EuroVis Workshop on Visual Analytics (EuroVA)* (2022), Bernard J. and Angelini M. (Eds.). The Eurographics Association. https://doi.org/10.2312/eurova.20221079.

[HC14] Hayes M. A., Capretz M. A.: Contextual anomaly detection in big sensor data. In *International Congress on Big Data* (2014), IEEE, pp. 64–71. https://doi.org/10.1109/BigData.Congress.2014.19.

[HCSG18] Heimerl F., Chang C.-C., Sarikaya A., Gleicher M.: Visual designs for binned aggregation of multi-class scatterplots, 2018. https://doi.org/10.48550/ARXIV.1810.02445.

[HS88] HART S. G., STAVELAND L. E.: Development of NASA-TLX (task load index): Results of empirical and theoretical research. *Advances in Psychology 52* (1988), 139–183.

[JKCP15] JOSHI A., KALE S., CHANDEL S., PAL D. K.: Likert scale: Explored and explained. *British Journal of Applied Science & Technology 7*, 4 (2015), 396–403. https://doi.org/10.9734/BJAST/2015/14975.

[JSN*21] JEKIC N., SCHREYER M., NEUBERT S., MUTLU B., SCHRECK T.: Visual data analysis of production quality data for aluminum casting. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (2021), ScholarSpace, pp. 1487–1495.

[KAW*14] KO S., AFZAL S., WALTON S., YANG Y., CHAE J., MALIK A., JANG Y., CHEN M., EBERT D.: Analyzing high-dimensional multivariate network links with integrated anomaly detection, highlighting and exploration. In *Conference on Visual Analytics Science and Technology (VAST)* (2014), IEEE, pp. 83–92. https://doi.org/10.1109/VAST.2014.7042484.

[Kei02] KEIM D.: Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics 8*, 1 (2002), 1–8. https://doi.org/10.1109/2945.981847.

[LAZV22] LIU D., ALNEGHEIMISH S., ZYTEK A., VEERAMACHANENI K.: MTV: Visual analytics for detecting, investigating, and annotating anomalies in multivariate time series. In *Proceedings ACM Human-Computer Interactions* (2022), vol. *6*, ACM. https://doi.org/10.1145/3512950.

[LLAJ21] LI X., LI J., ABDOLLAHI A., JONES T.: Data-driven thermal anomaly detection for batteries using unsupervised shape clustering. In *2021 IEEE 30th International Symposium on Industrial Electronics (ISIE)* (2021), IEEE, pp. 1–6. https://doi.org/10.1109/ISIE45552.2021.9576348.

[LLP07] LATECKI L. J., LAZAREVIC A., POKRAJAC D.: Outlier detection with kernel density functions. In *Machine Learning and Data Mining in Pattern Recognition* (2007), Lecture Notes in Computer Science, 4571, Springer, pp. 61–75. https://doi.org/10.1007/978-3-540-73499-4_6.

[LS09] LEWIS J. R., SAURO J.: The factor structure of the system usability scale. In *Human Centered Design* (2009), Lecture Notes in Computer Science, 5619, Springer, pp. 94–103. https://doi.org/10.1007/978-3-642-02806-9_12.

[MG13] MAYORGA A., GLEICHER M.: Splatterplots: Overcoming overdraw in scatter plots. *IEEE Transactions on Visualization and Computer Graphics 19*, 9 (2013), 1526–1538. https://doi.org/10.1109/TVCG.2013.65.

[MHK*19] MATHISEN A., HORAK T., KLOKMOSE C. N., GRØNBÆK K., ELMQVIST N.: Insideinsights: Integrating data-driven reporting in collaborative visual analytics. *Computer Graphics Forum 38*, 3 (2019), 649–661. https://doi.org/10.1111/cgf.13717.

[ND19] NGUYEN V. T., DANG T.: CloudTraceViz: A visualization tool for tracing dynamic usage of cloud computing resources. In *IEEE/ACM Industry/University Joint International Workshop on Data-center Automation, Analytics, and Control (DAAC)* (2019), pp. 1–6. https://doi.org/10.1109/DAAC49578.2019.00006.

[Par62] PARZEN E.: On estimation of a probability density function and mode. *The annals of mathematical statistics 33*, 3 (1962), 1065–1076. https://doi.org/10.1214/aoms/1177704472.

[PNL*19] PHAM V., NGUYEN N., LI J., HASS J., CHEN Y., DANG T.: MTSAD: Multivariate time series abnormality detection and visualization. In *International Conference on Big Data* (2019), IEEE, pp. 3267–3276. https://doi.org/10.1109/BigData47090.2019.9006559.

[REE*09] ROBERTSON G., EBERT D., EICK S., KEIM D., JOY K.: Scale and complexity in visual analytics. *Information Visualization 8* (2009), 247–253. https://doi.org/10.1057/ivs.2009.23.

[SAL*21] SHAUKAT K., ALAM T. M., LUO S., SHABBIR S., HAMEED I. A., LI J., ABBAS S. K., JAVED U.: A review of time-series anomaly detection techniques: A step to future perspectives. In *Proceedings of the Future of Information and Communication Conference* (2021), Springer, pp. 865–877. https://doi.org/10.1007/978-3-030-73100-7_60.

[SDS17] STOJANOVIC N., DINIC M., STOJANOVIC L.: A data-driven approach for multivariate contextualized anomaly detection: Industry use case. In *International Conference on Big Data (Big Data)* (2017), IEEE, pp. 1560–1569. https://doi.org/10.1109/BigData.2017.8258090.

[SFS*22] SHILPIKA, FUJIWARA T., SAKAMOTO N., NONAKA J., MA K.-L.: A visual analytics approach for hardware system monitoring with streaming functional data analysis. *IEEE Transactions on Visualization and Computer Graphics 28*, 6 (2022), 2338–2349. https://doi.org/10.1109/TVCG.2022.3165348.

[SL09] SAURO J., LEWIS J. R.: Correlations among prototypical usability metrics: Evidence for the construct of usability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2009), CHI '09, ACM, pp. 1609–1618. https://doi.org/10.1145/1518701.1518947.

[SLT*22] SHI Y., LIU Y., TONG H., HE J., YAN G., CAO N.: Visual analytics of anomalous user behaviors: A survey. *IEEE Transactions on Big Data 8*, 2 (2022), 377–396. https://doi.org/10.1109/TBDATA.2020.2964169.

[SMF*20] SUSCHNIGG J., MUTLU B., FUCHS A. K., SABOL V., THALMANN S., SCHRECK T.: Exploration of anomalies in cyclic multivariate industrial time series data for condition monitoring. In *BigVis 2020-3rd International Workshop on Big Data Visual Exploration and Analytics* (2020), CEUR-WS.org, p. 8.

[SMK*21] SUSCHNIGG J., MUTLU B., KOUTROULIS G., SABOL V., THALMANN S., SCHRECK T.: Visual exploration of anomalies in cyclic time series data with matrix and glyph representations. *Big Data Research 26* (2021), 100251. https://doi.org/10.1016/j.bdr.2021.100251.

[SMS23] SCHRECK T., MUTLU B., STREIT M.: *Visual Data Science for Industrial Applications*. Springer, Berlin (2023), pp. 447–471. https://doi.org/10.1007/978-3-662-65004-2_18.

[SPP*06] SUBRAMANIAM S., PALPANAS T., PAPADOPOULOS D., KALOGERAKI V., GUNOPULOS D.: Online outlier detection in sensor data using non-parametric models. In *Proceedings of the International Conference on Very Large Data Bases* (2006), VLDB Endowment, pp. 187–198.

[SVLB10] SCHRECK T., VON LANDESBERGER T., BREMM S.: Techniques for precision-based visual analysis of projected data. *Information Visualization 9*, 3 (2010), 181–193. https://doi.org/10.1057/ivs.2010.2.

[SWP22] SCHMIDL S., WENIG P., PAPENBROCK T.: Anomaly detection in time series: A comprehensive evaluation. *Proceedings of the VLDB Endowment 15*, 9 (2022), 1779–1797. https://doi.org/10.14778/3538598.3538602.

[SZB*20] SUSCHNIGG J., ZIESSLER F., BRILLINGER M., VUKOVIC M., MANGLER J., SCHRECK T.: Industrial production process improvement by a process engine visual analytics dashboard. In *Proceedings of the 53rd Hawaii International Conference on System Sciences* (2020), pp. 1320–1329. https://doi.org/10.24251/HICSS.2020.163.

[THS*19] THALMANN S., HEIMO G., SUSCHNIGG J., GASHI M., ENNSBRUNNER H., FUCHS A. K., SCHRECK T., MUTLU B., MANGLER J., KAPPL G., HUEMER C., LINDSTAEDT S.: Cognitive decision support for industrial product life cycles: A position paper. In *International Conference on Advanced Cognitive Technologies and Applications* (2019), vol. *11*, IARIA, pp. 3–9.

[TSSB22] TRAUTNER T., SBARDELLATI M., STOPPEL S., BRUCKNER S.: Honeycomb Plots: Visual Enhancements for Hexagonal Maps. In *Vision, Modeling, and Visualization* (2022), The Eurographics Association. https://doi.org/10.2312/vmv.20221205.

[WAG05] WILKINSON L., ANAND A., GROSSMAN R.: Graph-theoretic scagnostics. In *Symposium on Information Visualization* (2005), IEEE, pp. 157–164. https://doi.org/10.1109/INFVIS.2005.1532142.

[Weg18] WEGLARCZYK S.: Kernel density estimation and its application. *ITM Web of Conferences 23* (2018), 00037. https://doi.org/10.1051/itmconf/20182300037.

[WSP22] WENIG P., SCHMIDL S., PAPENBROCK T.: TimeEval: A benchmarking toolkit for time series anomaly detection algorithms. *Proceedings of the VLDB Endowment*, *15*, 12 (2022), 3678–3681. https://doi.org/10.14778/3554821.3554873.

[WZC*18] WU W., ZHENG Y., CHEN K., WANG X., CAO N.: A visual analytics approach for equipment condition monitoring in smart factories of process industry. In *2018 IEEE Pacific Visualization Symposium (PacificVis)* (2018), pp. 140–149. https://doi.org/10.1109/PacificVis.2018.00026.

[XWY*20] XU K., WANG Y., YANG L., WANG Y., QIAO B., QIN S., XU Y., ZHANG H., QU H.: CloudDet: Interactive visual analysis of anomalous performances in cloud computing systems. *IEEE Transactions on Visualization & Computer Graphics 26*, 01 (2020), 1107–1117. https://doi.org/10.1109/TVCG.2019.2934613.

[Zha13] ZHANG J.: Advancements of outlier detection: A survey. *EAI Endorsed Transactions on Scalable Information Systems 1*, 1 (2013), 1–26. https://doi.org/10.4108/trans.sis.2013.01-03.e2.

[ZKH04] ZHANG X., KING M. L., HYNDMAN R. J.: Bandwidth Selection for Multivariate Kernel Density Estimation Using MCMC. Monash University. 2004. https://doi.org/10.4225/03/59389cae32a30.

[ZLL*19] ZHOU F., LIN X., LIU C., ZHAO Y., XU P., REN L., XUE T., REN L.: A survey of visualization for smart manufacturing. *Journal of Visualization 22* (2019), 419–435. https://doi.org/10.1007/s12650-018-0530-2.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supporting Information

Supplemental Video 1