# MANDALA

## Multivariate ANomaly Detection& expLorAtion

Pro²Future

**Josef Suschnigg[1], Belgin Mutlu[1], Roman Kern[2], Tobias Schreck[3]**

Pro2Future GmbH[1], TUG-ISDS (Institute of Interactive Systems and Data Science)[2], TUG-CGV (Institute of Computer Graphics and Knowledge Visualization)[3]

[1] Inffeldgasse 25F, 8010 Graz

[2] Sandgasse 36/III, 8010 Graz

[3] Inffeldgasse 16c, 8010 Graz

## MOTIVATION & GOALS

- The increase in large and complex data sets across various industries has led to a growing need for data analytics tools that provide practical insights to facilitate **decision-making**
- MANDALA is a comprehensive platform for **seamless identification**, **analysis**, and **comprehensive understanding** of **anomalies** in **complex multivariate time series data**
- The platform subsequently **facilitates** the exploration and **comparative assessment of anomaly candidates**, **their related dimensions, and temporal aspects**
- Through its **user-friendly interface** and a **range of features**, MANDALA offers a toolset to effortlessly tackle the complexity of large and high dimensional sensor data.

### Project FactBox

| | |
|---|---|
| **Project Name** | SERAM |
| **Project ID** | StratP II 3.4.1 |
| **Duration** | 21 Months |

**Area 3**
Cognitive Decision Making

**Project Lead**
Dr. Belgin Mutlu

## APPROACH

- Semi-supervised learning method using kernel density estimation (KDE) to detect anomalies
- Scatterplot matrices (SPLOMs), hexagonal binning plots and line plots as visualization and exploration aids
- These techniques allow versatile comparisons between individual dimensions, pairs of dimensions and the whole dataset, while also facilitating temporal analysis of anomalies.

## CONTRIBUTION

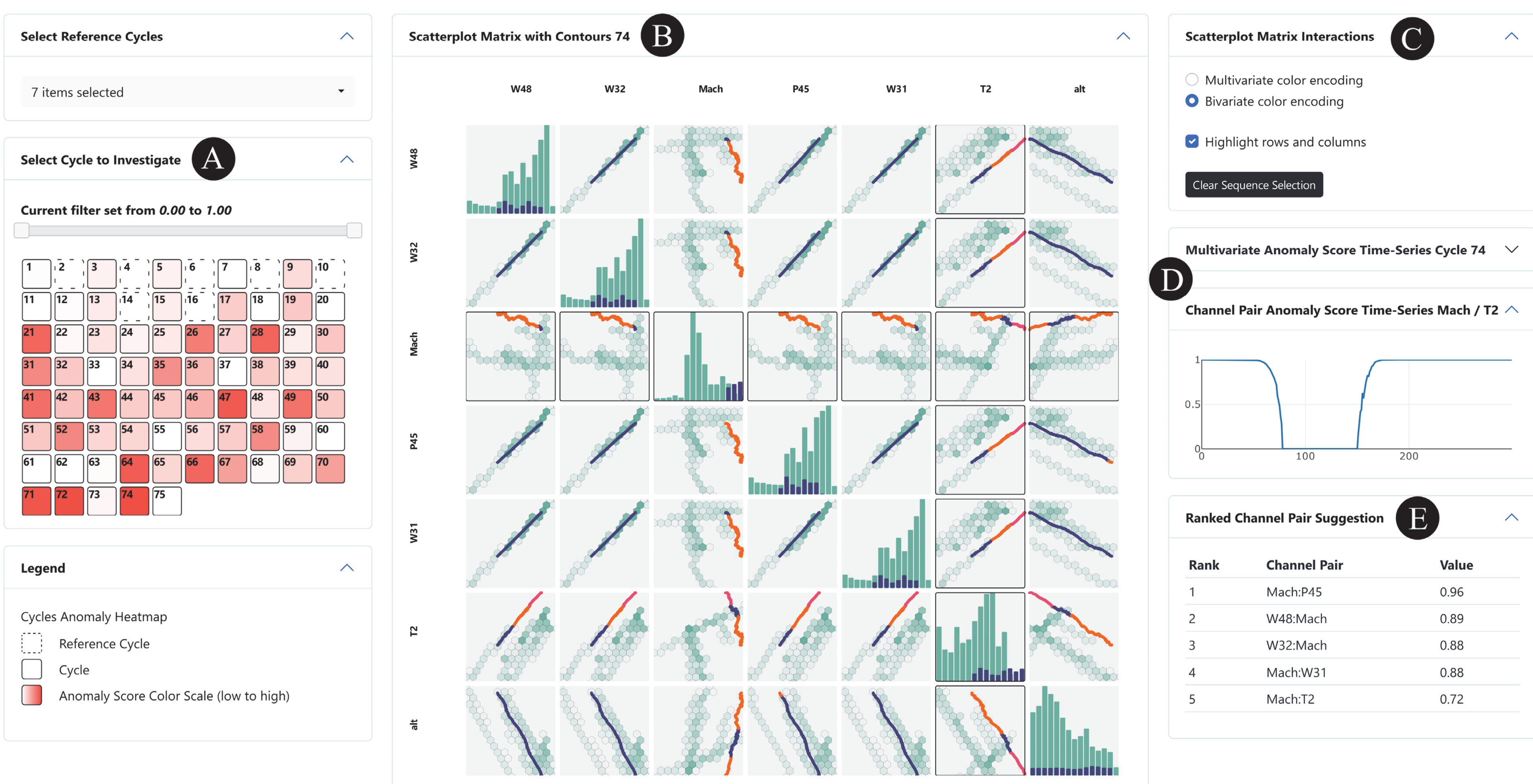
**Scientific contribution**

- User guidance in identifying, assessing, and exploring anomalies in multivariate time series data
- Novel methods for detecting anomalies within a temporal range in a multidimensional space
- Helping users in pinpointing contributing parameters to anomaly scores in multivariate time series data

**Economic contribution**

- Minimize downtime, reduce operational disruptions, and optimize resource allocation
- Ensure product quality, reducing defects and minimizing costly recalls
- Timely maintenance, prolonging equipment life and preventing expensive breakdowns

## FRAMEWORK




(A) The semi-supervised approach involves users in selecting reference data. Anomaly candidate cycles, distinct from the reference data, are highlighted using a red color scheme.

(B) The scatterplot matrix offers semantic zooming and presents bivariate data of the analyzed anomaly candidate cycles through scatterplots. The distribution of bivariate data in the reference set is depicted as hexagonal binning plots. Along the diagonal, separate histograms compare the univariate distribution of each reference data dimension with the anomaly candidate cycle.

(C) The color of scatterplot points can be customized for multivariate or bivariate color-encoding, allowing flexibility in the analysis scope.

(D) The temporal scope of anomalies can be investigated via line plot views, which also interactively impact the scatterplot matrix through linking and brushing functionalities.

(E) Bivariate anomalies are ranked by their anomaly score, and this ranking is presented in the ranked channel pair suggestion view.

**Contact:** DI Josef Suschnigg, Pro2Future GmbH, josef.suschnigg@pro2future.at

**Acknowledgement**: This work was supported by Pro²Future (FFG, 881844).

JOHANNES KEPLER UNIVERSITÄT LINZ · TU Graz · AVL · Fronius · UAR Upper Austrian Research GmbH · Bundesministerium Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie · Bundesministerium Digitalisierung und Wirtschaftsstandort · FFG Forschung wirkt. · ober österreich · Das Land Steiermark Wissenschaft und Forschung · SFG Neues denken. Neues fördern.