CEFET/RJ

DATA CENTRIC AI APPROACHES FOR TIME SERIES EVENT DETECTION

Eduardo Ogasawara
eogasawara@ieee.org
https://eic.cefet-rj.br/~eogasawara

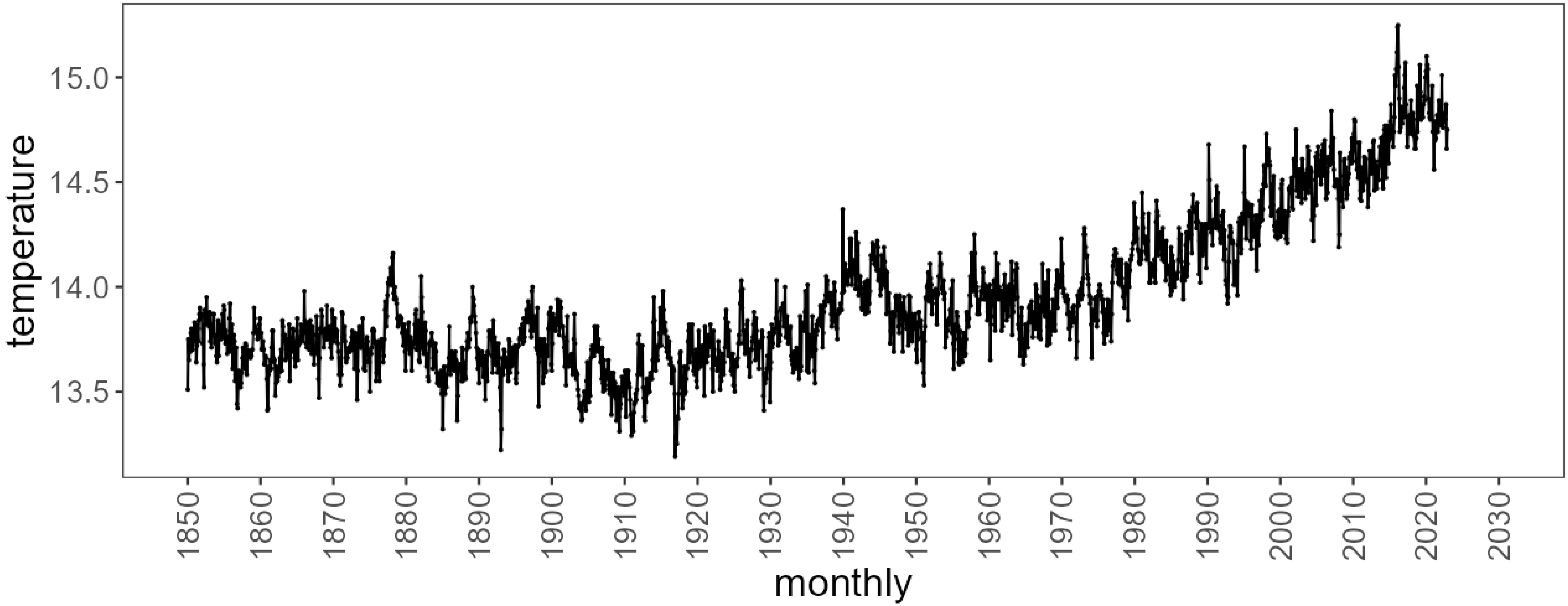
Road map

- Overview of Time Series Event Detection
- Data-Centric AI Initiatives
- Challenges

ChatGPT. (2024). Illustration of a data-centric AI for time series event detection

Time Series Events

- Time series events are commonly instants or intervals in the time series where observations change in a manner that is considered important for analysis or decision-making processes
  - The interpretation of an event can vary significantly across different domains
  - They can be categorized into main types: anomalies, change points, and motifs



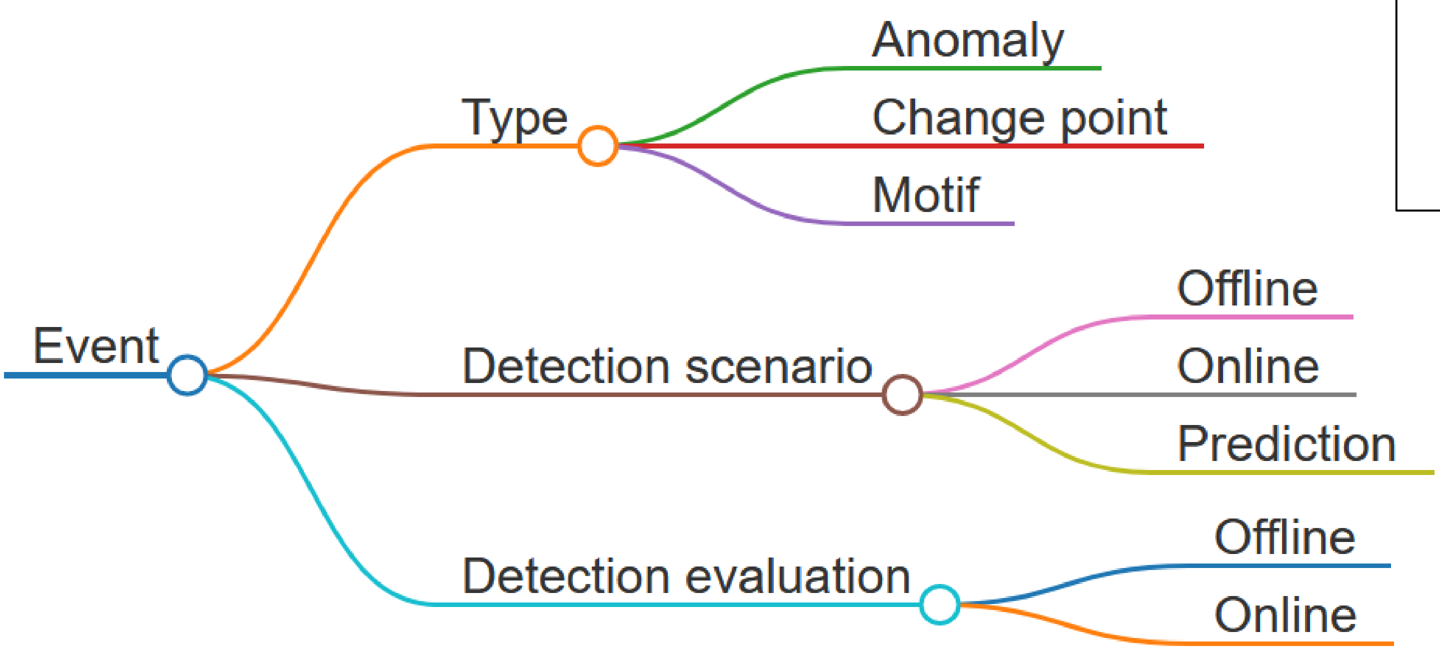
[1] V. Guralnik and J. Srivastava, “Event Detection from Time Series Data,” in *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, in KDD ’99. New York, NY, USA: ACM, 1999, pp. 33–42. doi: 10.1145/312129.312190.

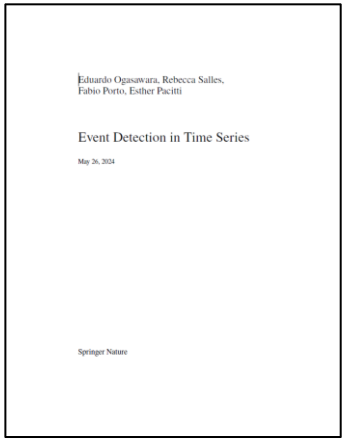
Event Detection

- Process of identifying events
- Important for monitoring and surveillance
  - Industry, seismic, oil exploration, epidemiology, climate
- There are many studies, but
  - Focused on specific types of events
  - Lacking a holistic view of the problem

ChatGPT. (2024). Illustration of a time series event detection

Taxonomy

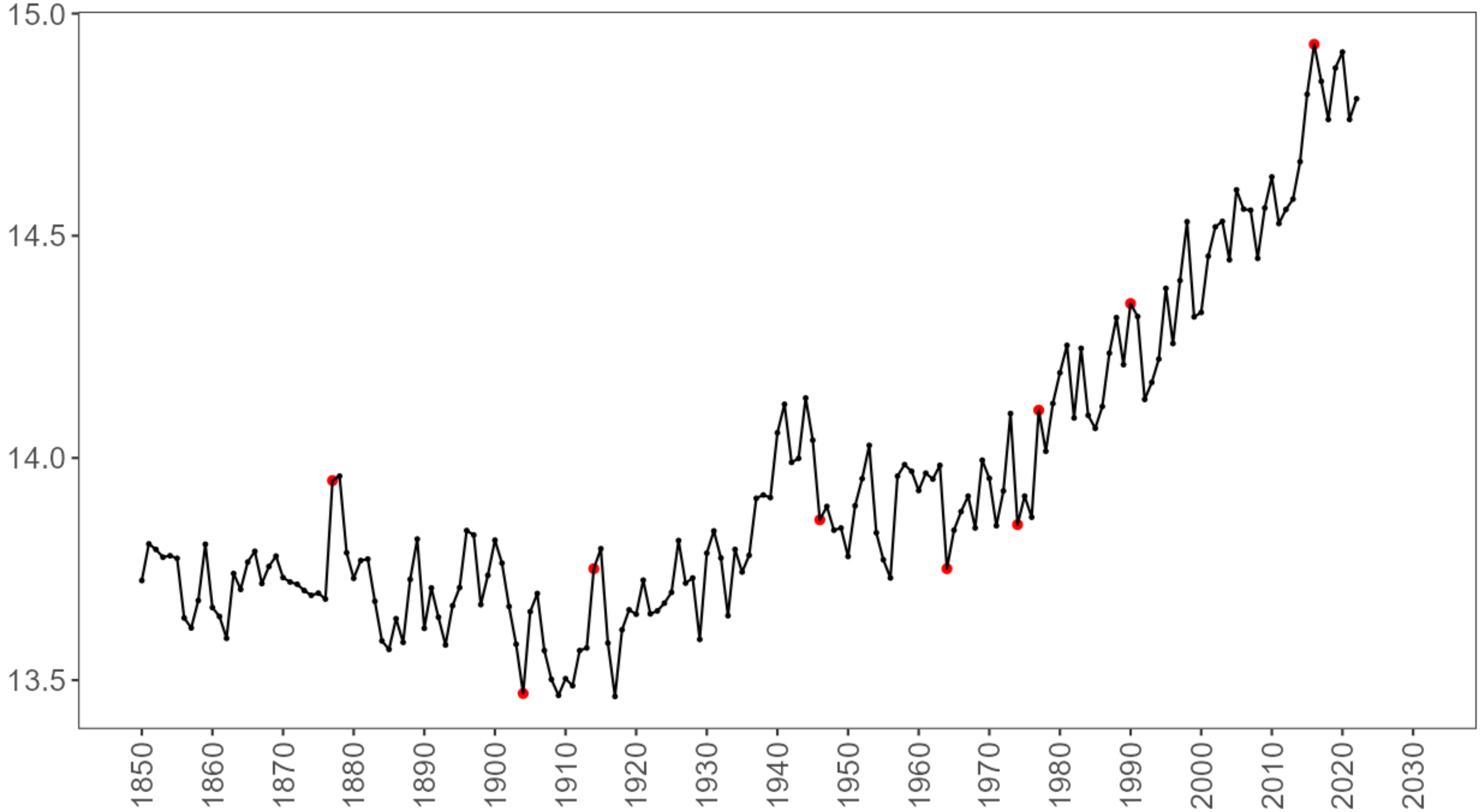




[1] E. Ogasawara, R. Salles, F. Porto, and E. Pacitti, *Time Series Event Detection*. Springer, (to appear).

Anomalies

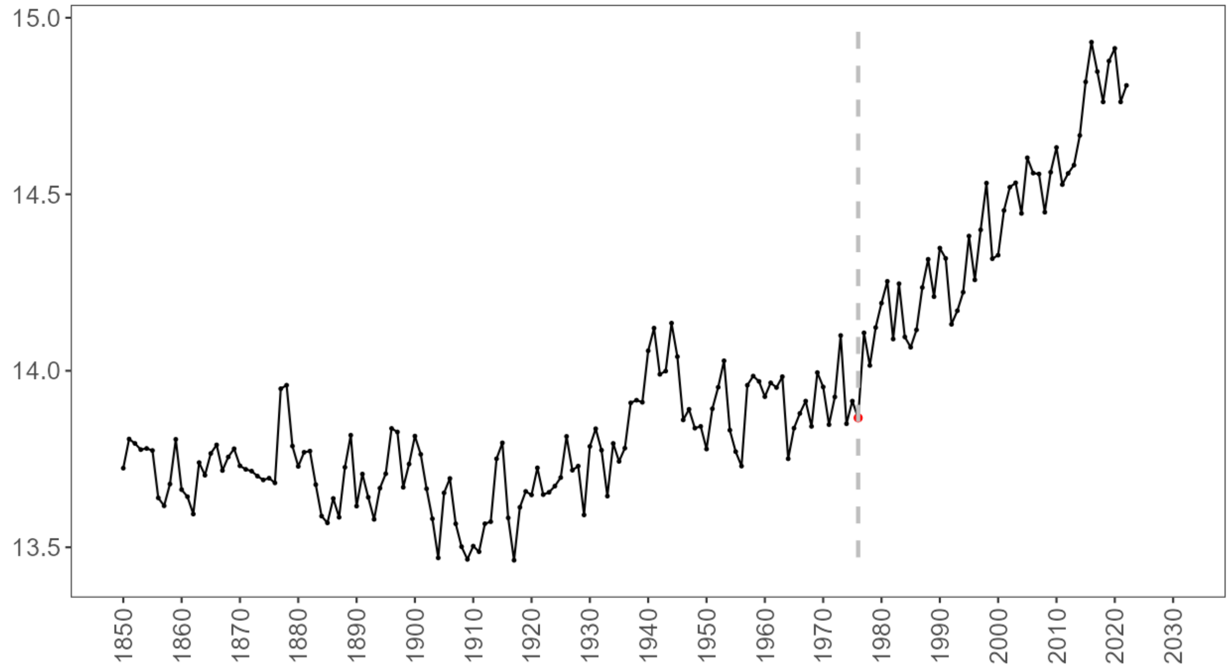
- Anomalies are observations that do not conform to the typical ones at the time series [1]



[1] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM Computing Surveys, vol. 41, no. 3. 2009. doi: 10.1145/1541880.1541882.

Change points

- Change points are time intervals where there is a significant change in the statistical properties in a time series [1]
  - This can include changes in mean, variance, correlation, distribution
- They represent a transition between different states in a process that generates the time series [2]

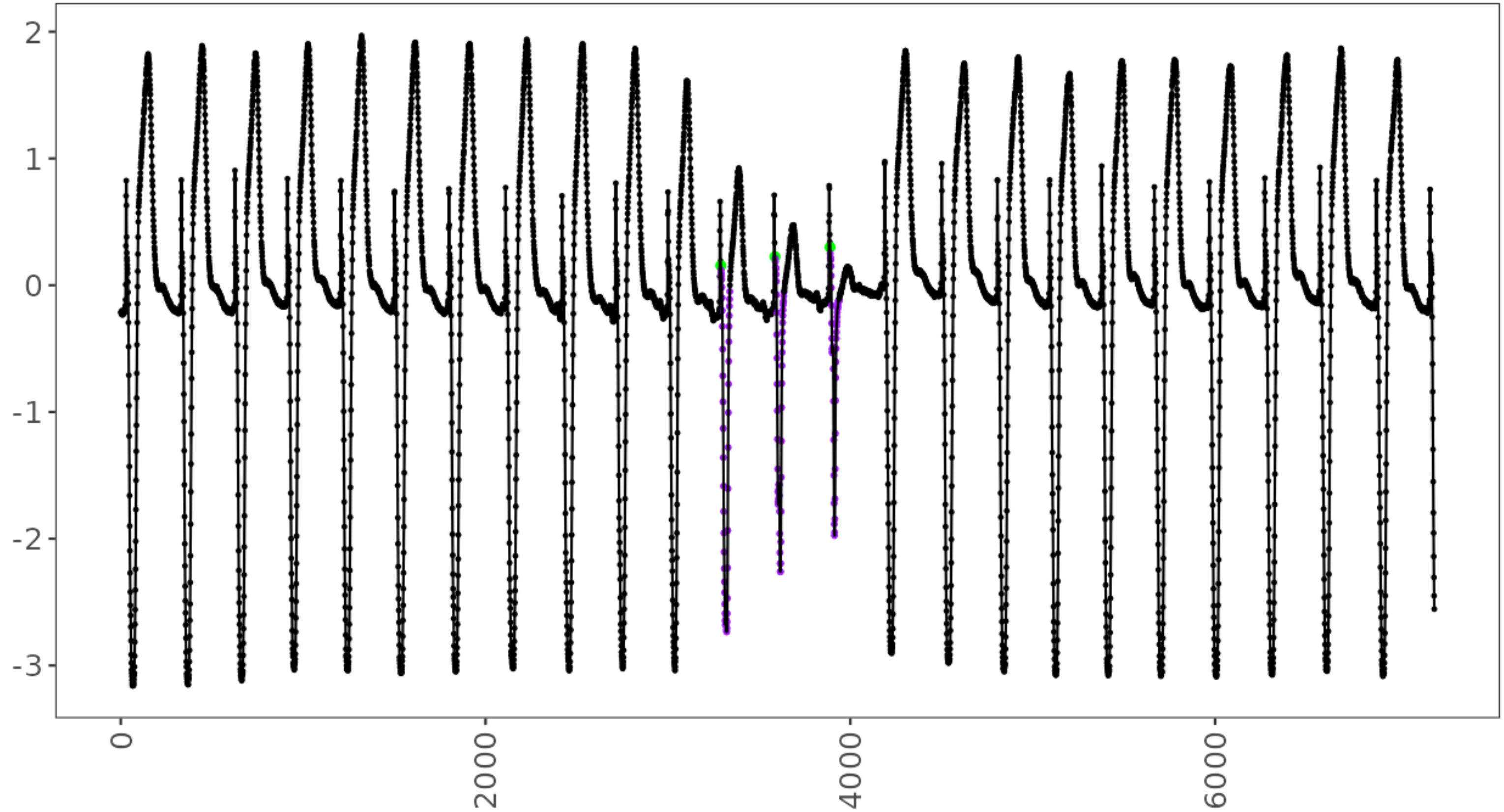


[1] T. Górecki, L. Horváth, and P. Kokoszka, "Change point detection in heteroscedastic time series," Econometrics and Statistics, vol. 7. pp. 63–88, 2018. doi: 10.1016/j.ecosta.2017.07.005.

[2] C. Truong, L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods," Signal Processing, vol. 167. 2020. doi: 10.1016/j.sigpro.2019.107299.

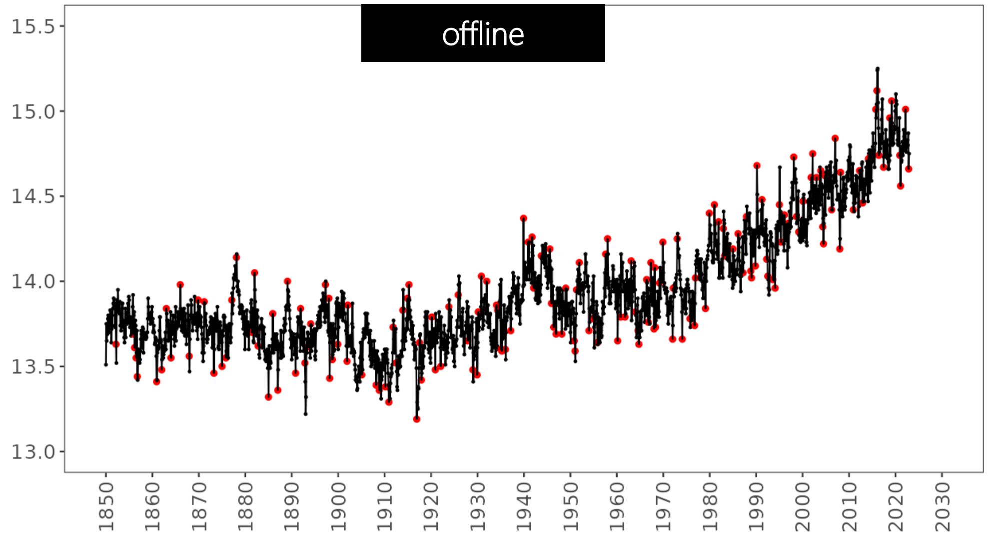
Motifs

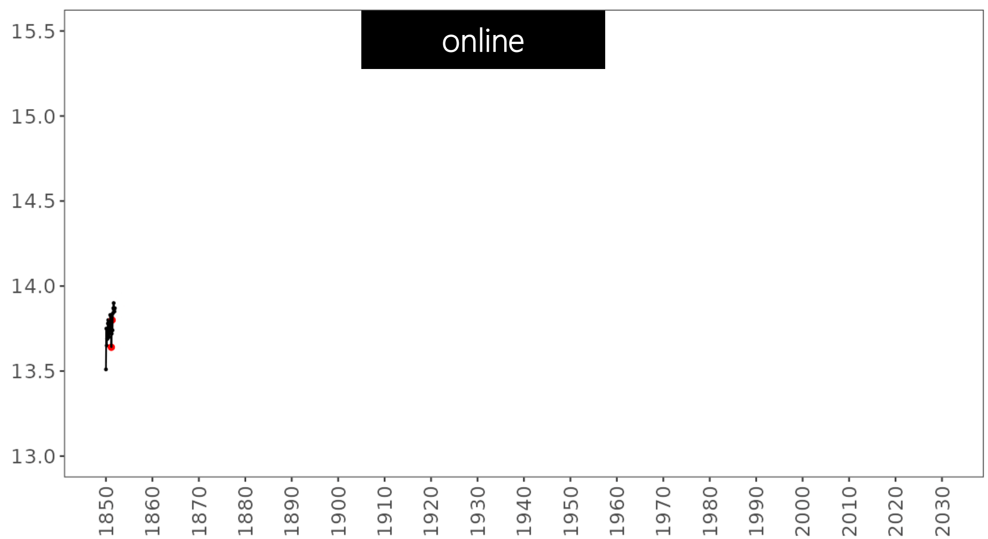
- Time series motifs are sequences of significantly similar observations within a time series
  - It is an approximately repeated subsequence within a longer time series [1]



[1] A. Mueen, "Time series motif discovery: Dimensions and applications," Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 4, no. 2. pp. 152–159, 2014. doi: 10.1002/widm.1119.

Offline versus online detection



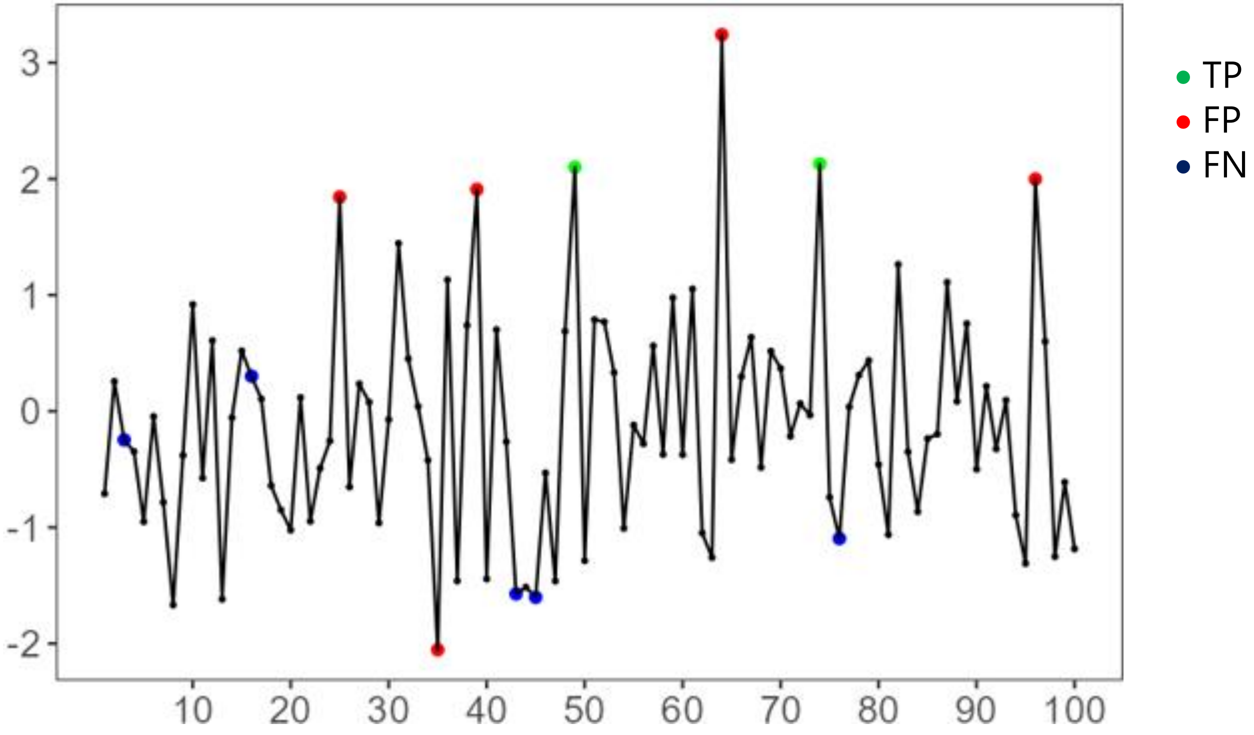


prediction

[1] E. Ogasawara, R. Salles, F. Porto, and E. Pacitti, *Time Series Event Detection*. Springer, (to appear).

Basic metrics for event detection

- $\text{precision} = \frac{TP}{TP+FP}$
- $\text{recall} = \frac{TP}{TP+FN}$
- $F_1 = \frac{2 \cdot \text{precision} \cdot \text{recall}}{\text{precision}+\text{recall}}$



Data-Centric AI Initiatives

Data Centric AI

- Data-centric AI is an approach that emphasizes data preparation
  - Data Quality: accurate, complete, and representative data
  - Data Transformation: normalization, encoding categorical variables, and transforming features to improve model performance
  - Feature Engineering: new features to better capture the underlying patterns
  - Data labeling: Maintaining consistent data labels
  - Bias mitigation: Identifying and addressing biases in the data
  - Data augmentation: Using techniques to increase dataset size artificially

It might be a buzzword for data preprocessing

ChatGPT. (2024). Illustration of a data-centric AI

Adaptive normalization

- Integrated normalization for sliding windows
- Compute a moving average for each sliding window
- Differentiate in each sliding window observation relative to its moving average
- Remove windows with outliers
- Scale each window between 0 and 1 with respect to the maximum and minimum differences of all windows


Adaptive Normalization: A Novel Data Normalization Approach for Non-Stationary Time Series

Eduardo Ogasawara, Leonardo C. Martinez, Daniel de Oliveira, Geraldo Zimbrão, Gisele L. Pappa and Marta Mattoso

PERFORMANCE OF ALGORITHMS TO FORECAST THE MONTHLY AVERAGE EXCHANGE RATE OF U.S. DOLLAR TO BRAZILIAN REAL TIME SERIES

Algorithm	RMSE 1-step	RMSE 12-step
AR	0.082	0.545
NN-MM	0.177	1.173
NN-DS	0.094	1.444
NN-ZS	0.126	0.814
NN-SW	0.088	0.451
NN-AN	0.062	0.345

[1] E. Ogasawara, L. C. Martinez, D. De Oliveira, G. Zimbrão, G. L. Pappa, and M. Mattoso, "Adaptive Normalization: A novel data normalization approach for non-stationary time series," Proceedings of the International Joint Conference on Neural Networks. 2010. doi: 10.1109/IJCNN.2010.5596746.

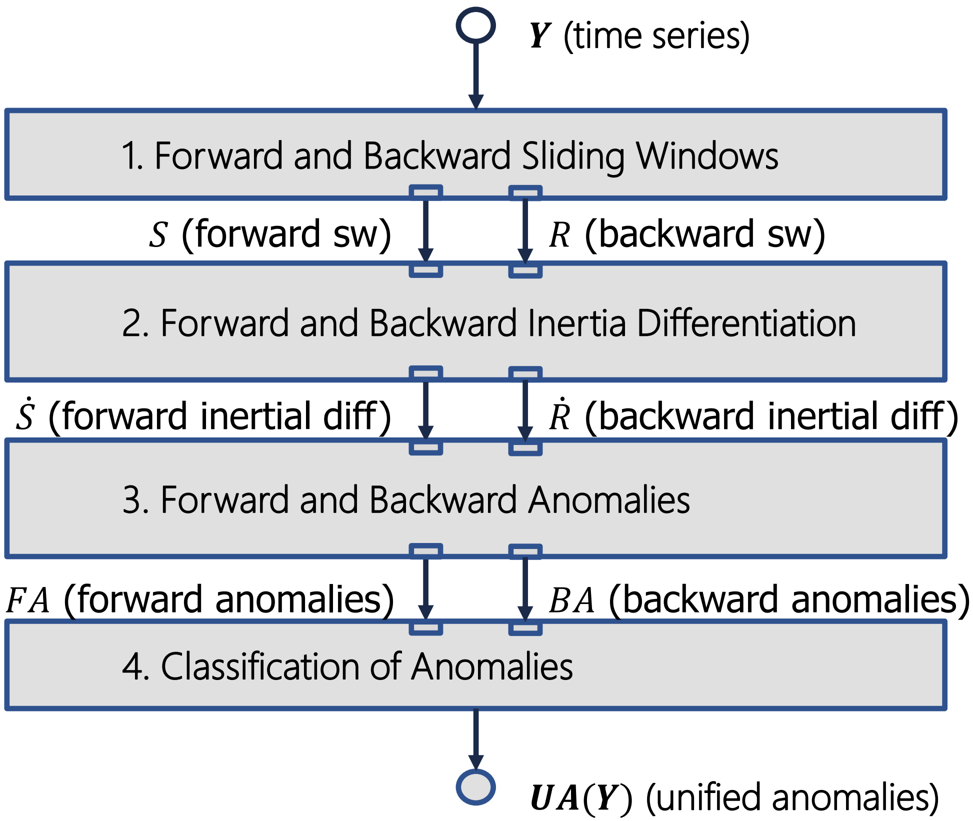
AN Properties

- Provides inertia during time series analysis
  - Higher moving average, higher inertia
- It usually provides good step-ahead predictions using machine learning
- It enables outlier removal (could be used for anomaly detection)
- Limitations
  - Should establish the moving average

[1] E. Ogasawara, L. Murta, G. Zimbrão, and M. Mattoso, “Neural networks cartridges for data mining on time series,” Proceedings of the International Joint Conference on Neural Networks. pp. 2302–2309, 2009. doi: 10.1109/IJCNN.2009.5178615.

FBIAD: Forward-Backward Inertia Anomaly Detection

- Use AN ideas for anomaly detection



Forward and Backward Inertial Anomaly Detector: A Novel Time Series Event Detection Method

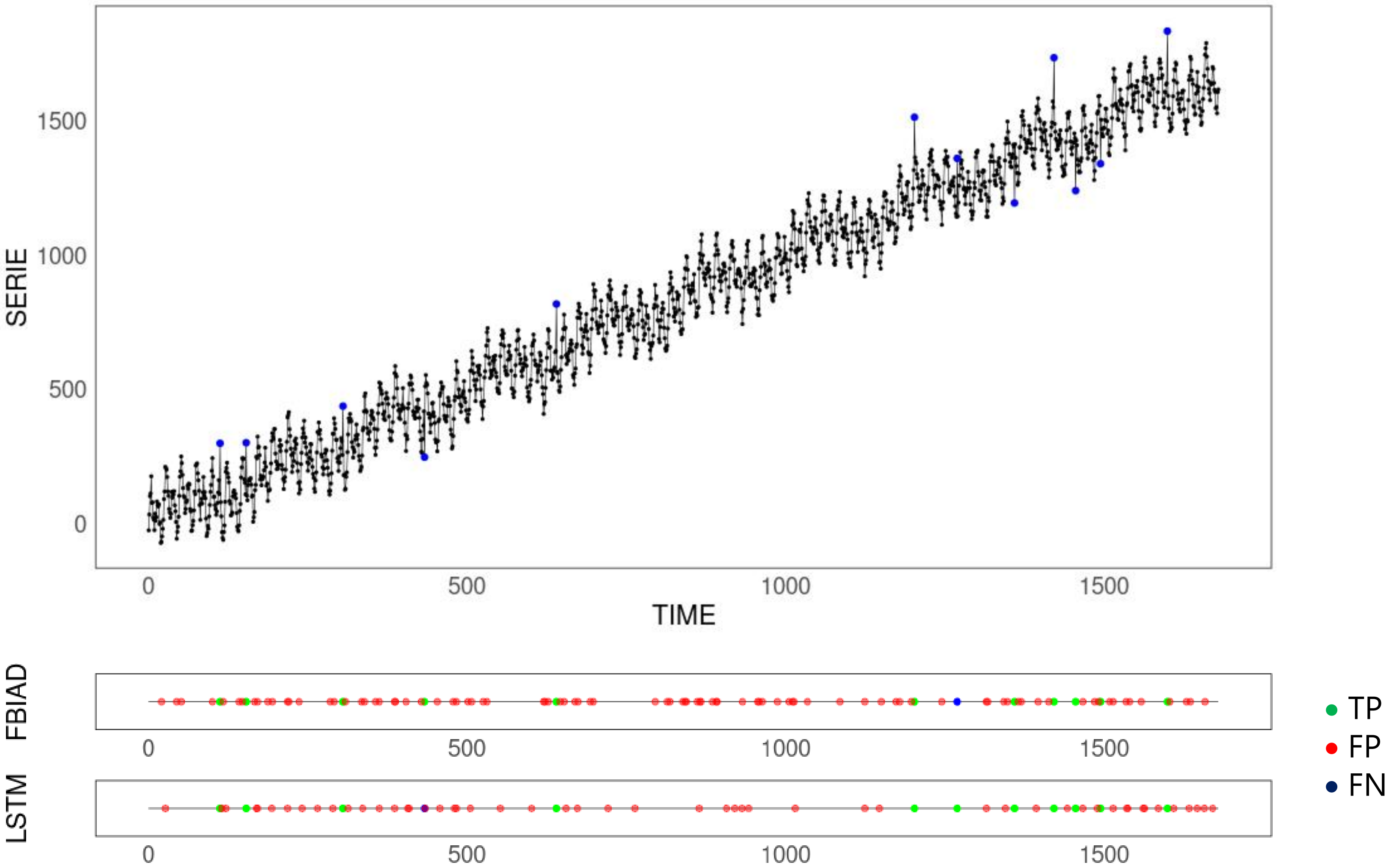
[1] J. Lima, R. Salles, F. Porto, R. Coutinho, P. Alpis, L. Escobar, E. Pacitti, and E. Ogasawara, "Forward and Backward Inertial Anomaly Detector: A Novel Time Series Event Detection Method," Proceedings of the International Joint Conference on Neural Networks, vol. 2022-July. 2022. doi: 10.1109/IJCNN55064.2022.9892088.

Comparison

- Dataset studied (Yahoo, Numenta and Gecco)

Method	Precision	Recall	F1	Accuracy
FBIAD	**0.066**	0.528	**0.085**	0.731
ARIMA	0.045	**0.556**	0.067	**0.746**
LSTM	0.041	0.534	0.062	0.735
ELM	0.041	0.517	0.063	0.726
Conv1D	0.036	0.519	0.055	0.724
SVM	0.030	0.542	0.049	0.732

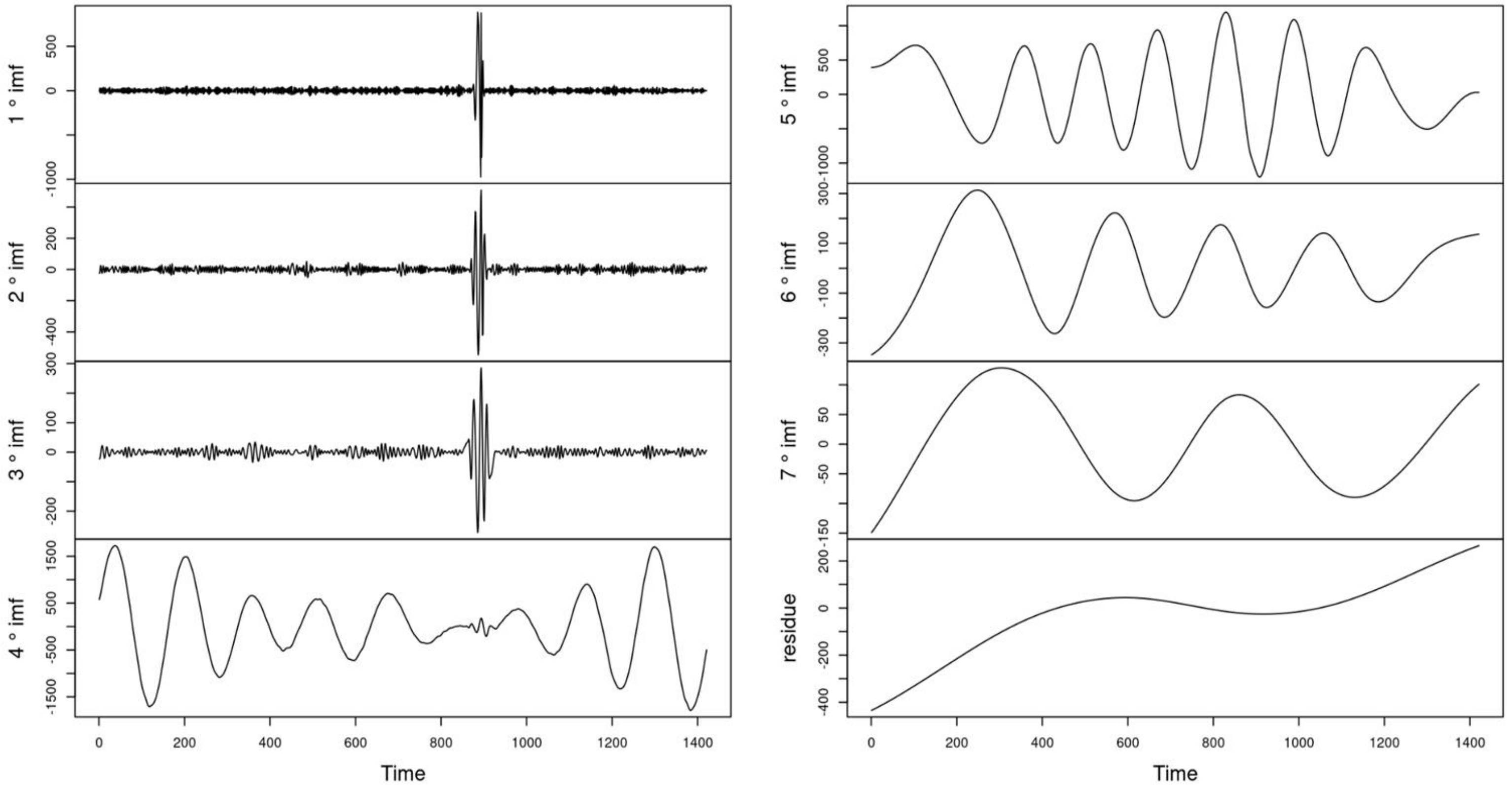
Inspecting Performance Comparison



[1] J. Lima, R. Salles, F. Porto, R. Coutinho, P. Alpis, L. Escobar, E. Pacitti, and E. Ogasawara, "Forward and Backward Inertial Anomaly Detector: A Novel Time Series Event Detection Method," Proceedings of the International Joint Conference on Neural Networks, vol. 2022-July. 2022. doi: 10.1109/IJCNN55064.2022.9892088.

Addressing moving average limitation using EMD

- Empirical Mode Decomposition (EMD) is a technique for decomposing non-linear and non-stationary series into a series of functions called Intrinsic Mode Functions (IMFs)

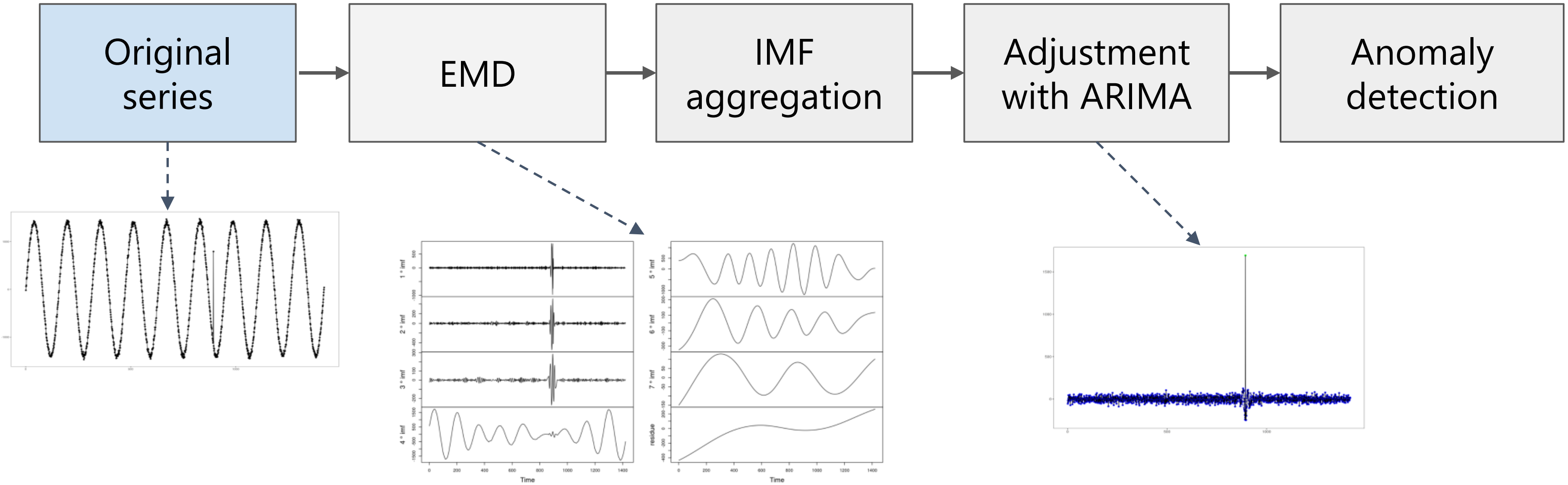


[1] Y. Lei, J. Lin, Z. He, and M. J. Zuo, "A review on empirical mode decomposition in fault diagnosis of rotating machinery," Mechanical Systems and Signal Processing, vol. 35, no. 1–2. pp. 108–126, 2013. doi: 10.1016/j.ymssp.2012.09.015.

REMD

- REMD: A hybrid method consisting of four steps
  - EMD decomposition
  - IMF aggregation
  - ARIMA adjustment
  - Anomaly detection: analysis of distribution error


REMD: A Novel Hybrid Anomaly Detection Method Based on EMD and ARIMA

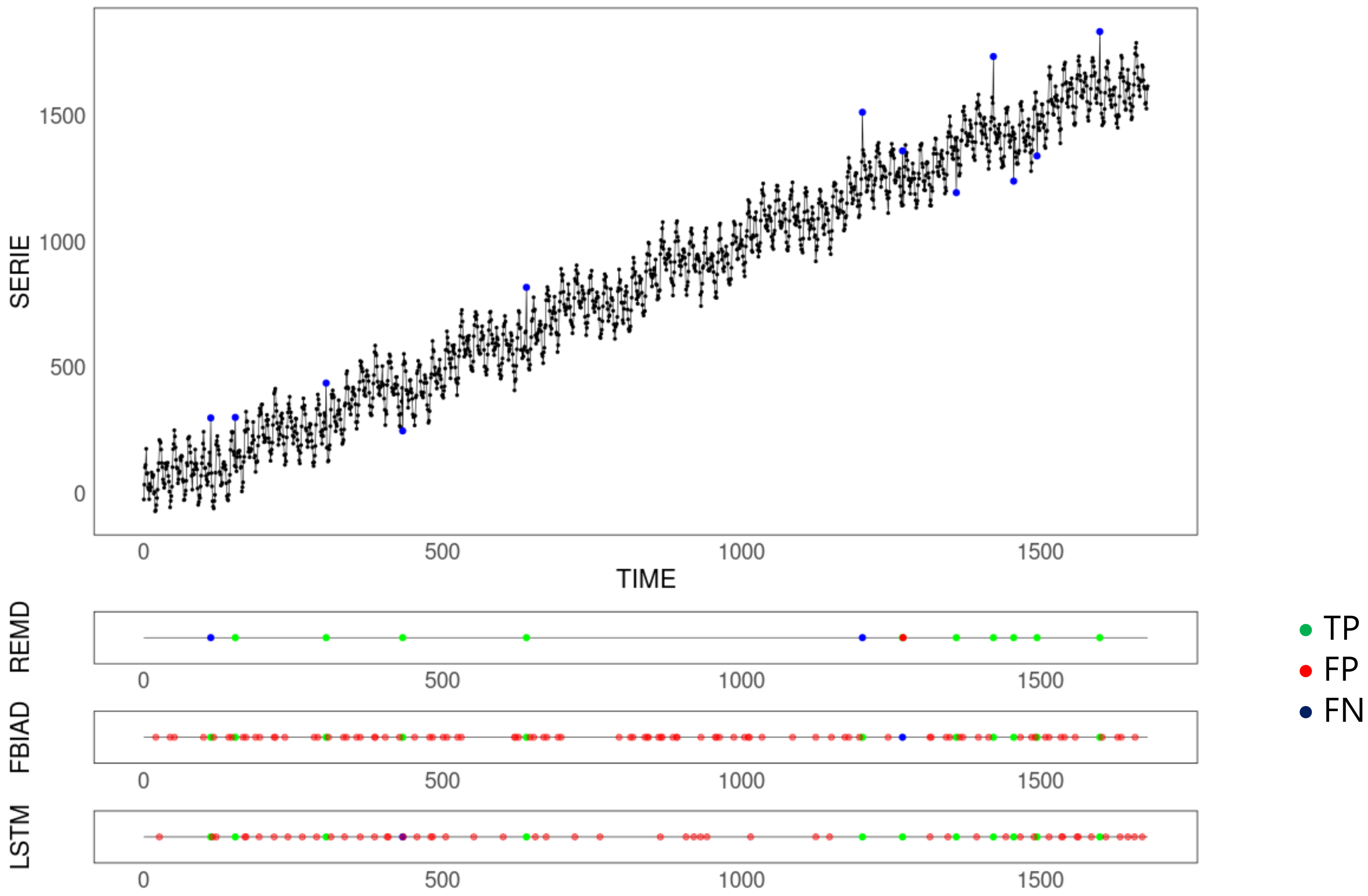


Comparison

- Datasets: Yahoo, Numenta, and Gecco
- REMD presents a much better performance than the second-placed method
  - EMD-based method, when we use F1 as the main selection criterion

Method	Precision	Recall	F1
REMD	**0.684**	0.386	**0.448**
EMD	0.243	0.408	0.207
FBIAD	0.066	0.528	0.085
ARIMA	0.045	**0.556**	0.067
LSTM	0.041	0.534	0.062
ELM	0.041	0.517	0.063
Conv1D	0.036	0.519	0.055
SVM	0.030	0.542	0.049

Inspecting performance comparison

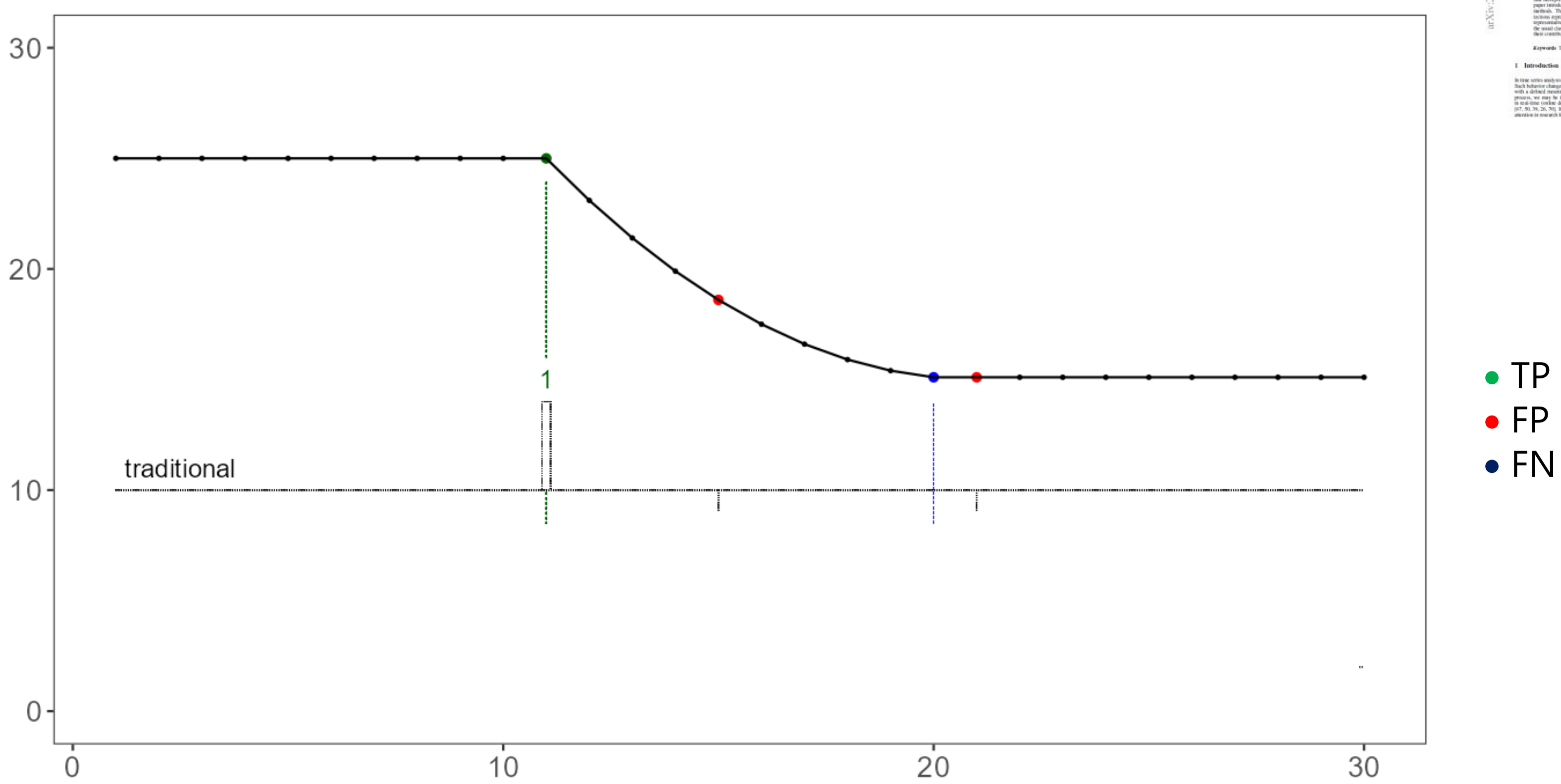


[1] J. Souza, E. Paixão, F. Fraga, L. Baroni, R. F. S. Alves, K. Belloze, J. Santos, E. Bezerra, F. Porto, and E. Ogasawara, "REMD: A Novel Hybrid Anomaly Detection Method Based on EMD and ARIMA," *Proceedings of the International Joint Conference on Neural Networks*, vol. 2024-July. pp. 1–8, 2024.

Time should count while evaluating events

- Traditional scoring methods, such as precision and recall, are not sufficient to assess the performance of event detection
  - They do not incorporate time and do not reward close detections.
  - True positives are rewarded
  - All other outcomes are equally penalized


SoftED: Metrics for Soft Evaluation of Time Series Event Detection

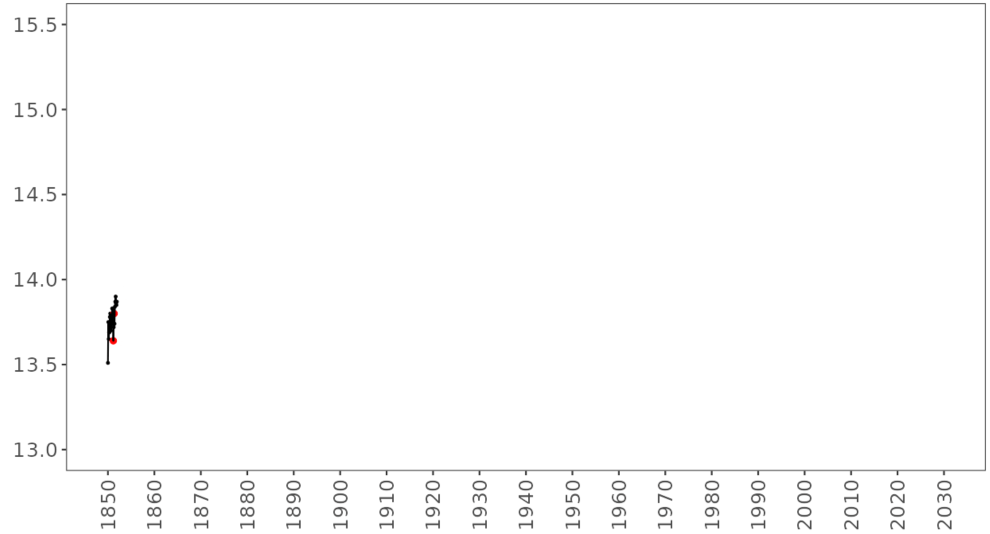


[1] R. Salles, J. Lima, R. Coutinho, E. Pacitti, F. Masseglia, R. Akbarinia, C. Chen, J. Garibaldi, F. Porto, and E. Ogasawara, "SoftED: Metrics for Soft Evaluation of Time Series Event Detection." arXiv, Apr. 01, 2023. doi: 10.48550/arXiv.2304.00439.

Still need improvements when it comes to streaming and online events

- The schizophrenic behavior of detectors in online detection
- Detection Probability: $DP(x_i) = \frac{df(x_i)}{bf(x_i)}$
  - df: detection frequency
  - bf: batch frequency
- Detection Lag: $Lag_i^s = fdb_i - sb_i$
  - fdb (first detection batch)
  - sb (start batch)

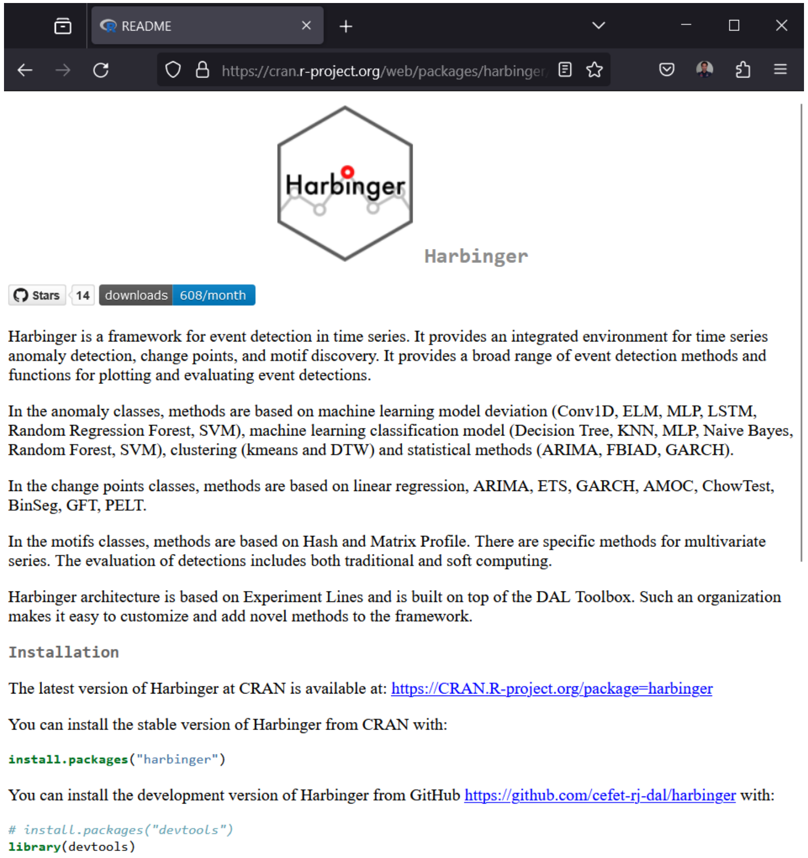

Online Event Detection in Streaming Time Series:
Novel Metrics and Practical Insights



[1] J. Lima, L. G. Tavares, E. Pacitti, J. E. Ferreira, I. Santos, I. G. Siqueira, D. Carvalho, F. Porto, R. Coutinho, and E. Ogasawara, “Online Event Detection in Streaming Time Series: Novel Metrics and Practical Insights,” Proceedings of the International Joint Conference on Neural Networks, vol. 2024-July. pp. 1–8, 2024.

Harbinger: Framework for Time Series Event Detection

- Holistic view of the problem
  - Anomalies
  - Change points
  - Motif discovery
- Properties
  - Uniform Data Model
  - Rigid interface (algebraic)
  - Expansible
  - Based on experimental line
- Inspiration from Sci-Kit Learn
  - Fit()
  - Detection()
- More than 50 event detectors
- R Package available at CRAN



https://cran.r-project.org/web/packages/harbinger/index.html

CEFET/RJ Team

D.Sc. students

Ellen Paixão

Janio Lima

Lais Baroni

Lucas Giusti*

Paulo Elias*
(UFF)

M.Sc. students

Antônio Mello

Arthur Garcia

Cristiane Gea

Edson Sobrinho

Fabiana Santos*

Frank Faisca

Jéssica de Souza

Josélia Rabelo

Luiz Oliveira

Michel Reis*

Ricardo Buçard*

Rodrigo Machado

Thiago Marques

Challenges for Collaboration

- Multivariate Change-point Detection and Concept Drift
- Volatility Anomalies and Change Points
- Model Management Under Concept Drift
- Detection Metrics for Interval Events
- Data Augmentation for Time Series

[1] E. Ogasawara, L. Murta, G. Zimbrão, and M. Mattoso, "Neural networks cartridges for data mining on time series," Proceedings of the International Joint Conference on Neural Networks. pp. 2302–2309, 2009. doi: 10.1109/IJCNN.2009.5178615.

DATA CENTRIC AI APPROACHES FOR TIME SERIES EVENT DETECTION

CEFET/RJ

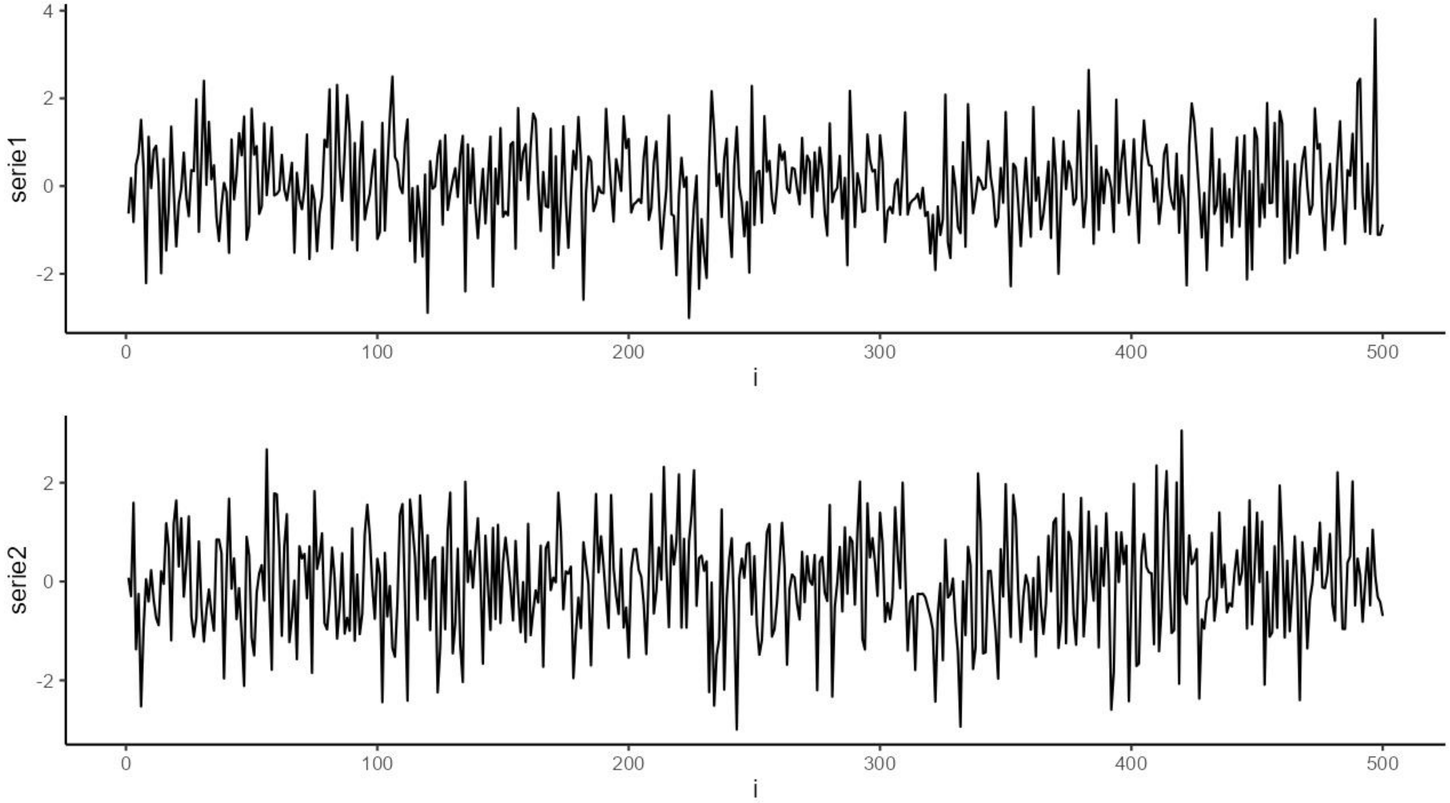
Eduardo Ogasawara
eogasawara@ieee.org
https://eic.cefet-rj.br/~eogasawara

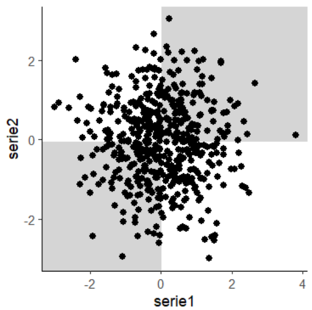
Challenges: Multivariate Change-point Detection and Concept Drift

Change-point & Concept Drift

- Virtual Concept Drift (data distribution)
- Real Concept Drift (model change)
- Virtual Concept Drift
  - Univariate time series (detection in isolation)
  - Multivariate time series

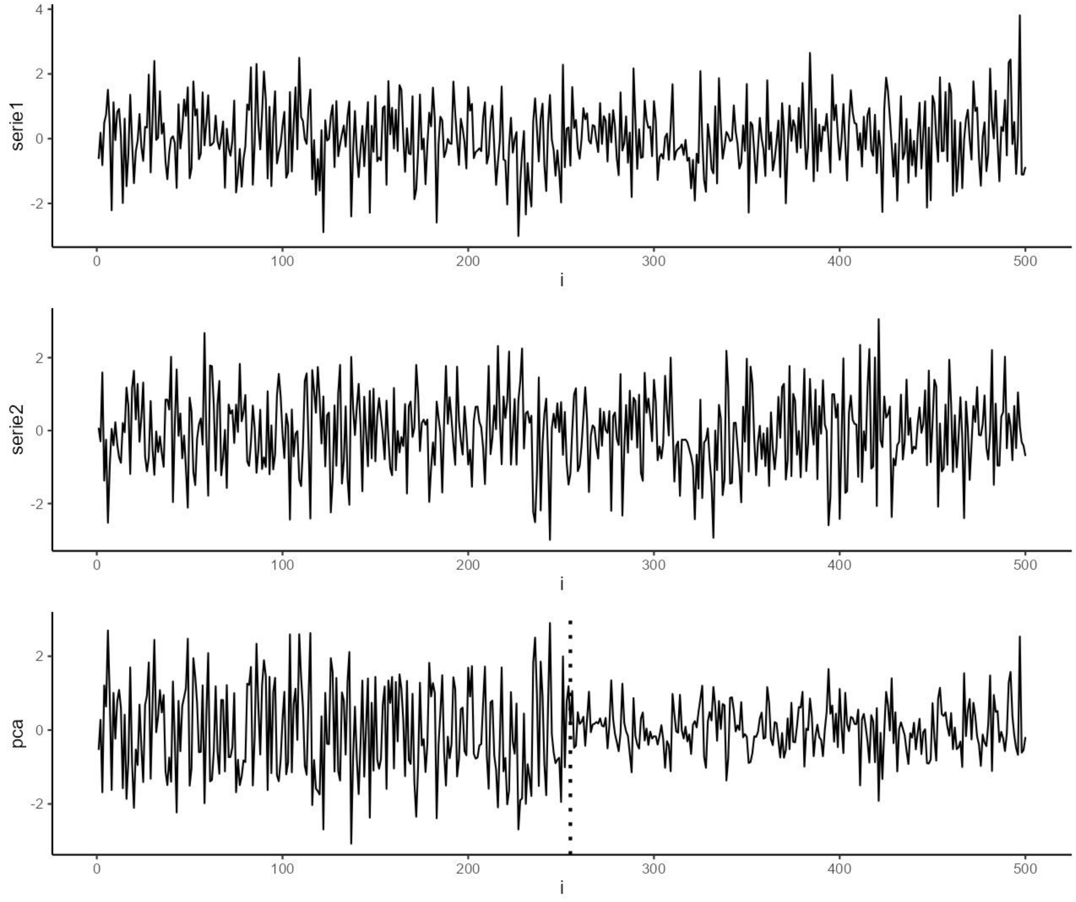
Multivariate Change-point & Concept Drift



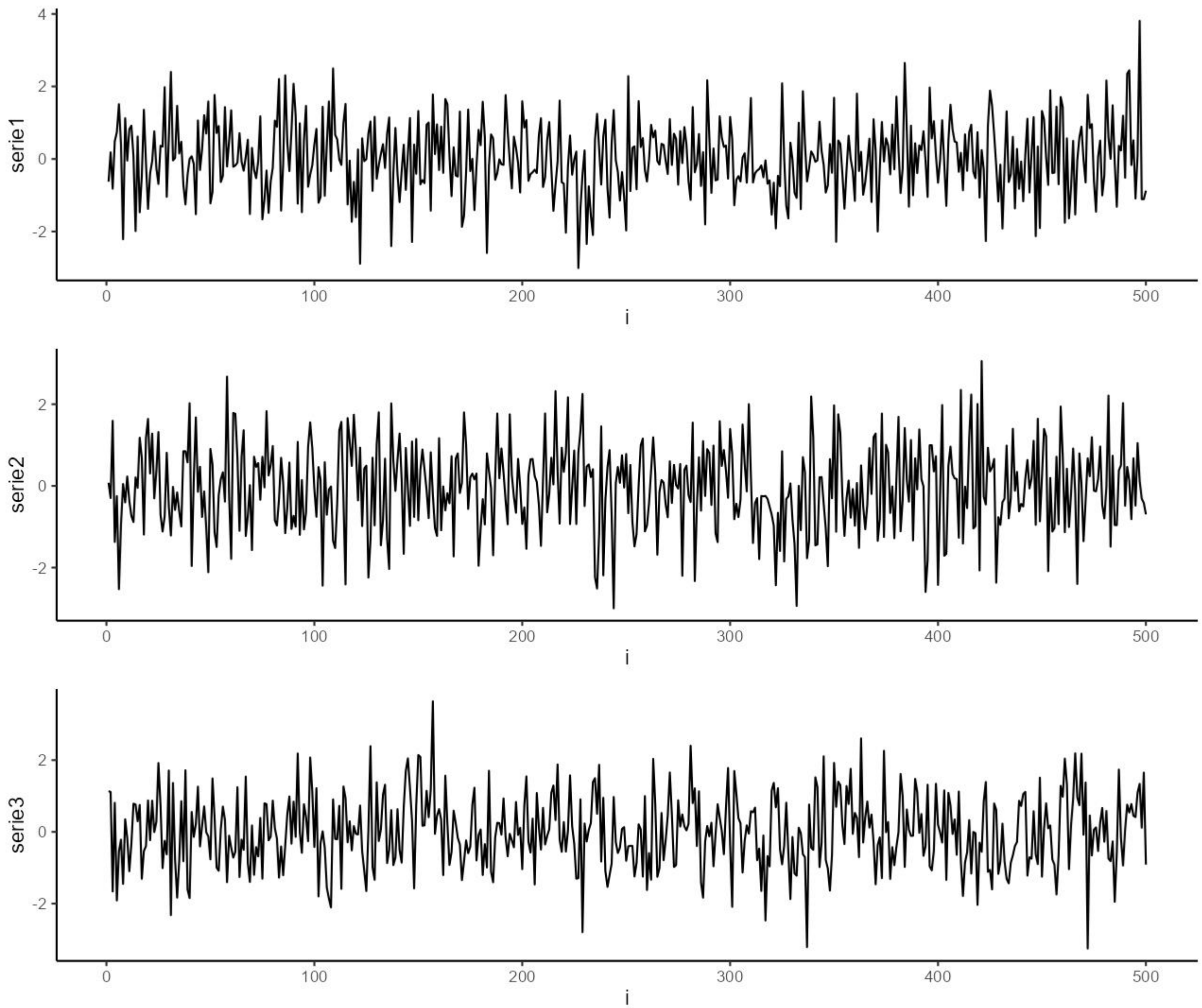


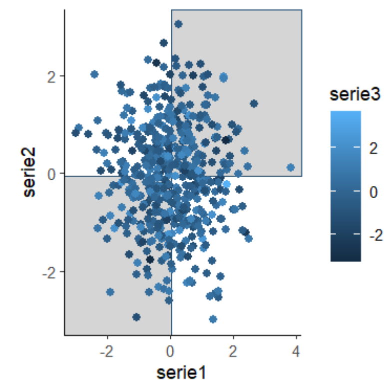
PCA as dimensionality reduction

- Compute 1st principal component
- Analyze transformed time series

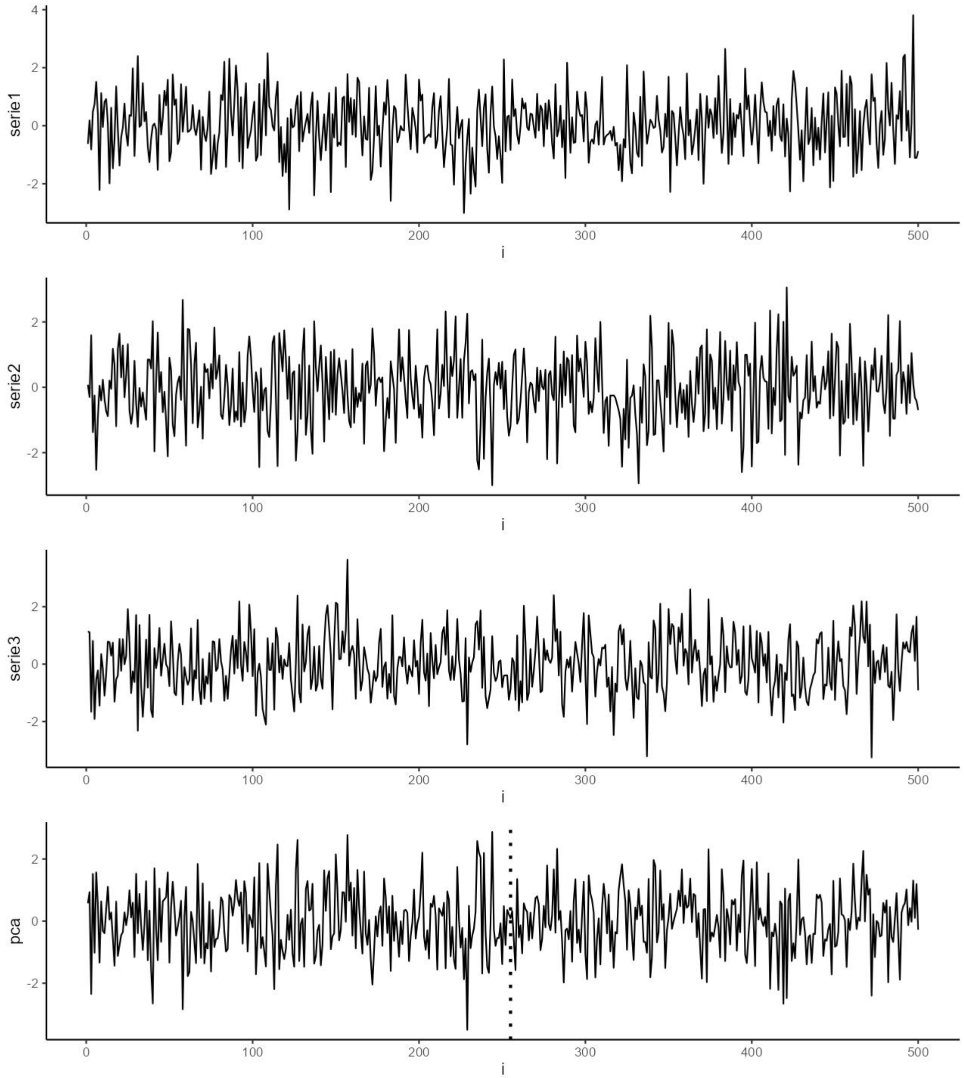


Consider a 3D time series



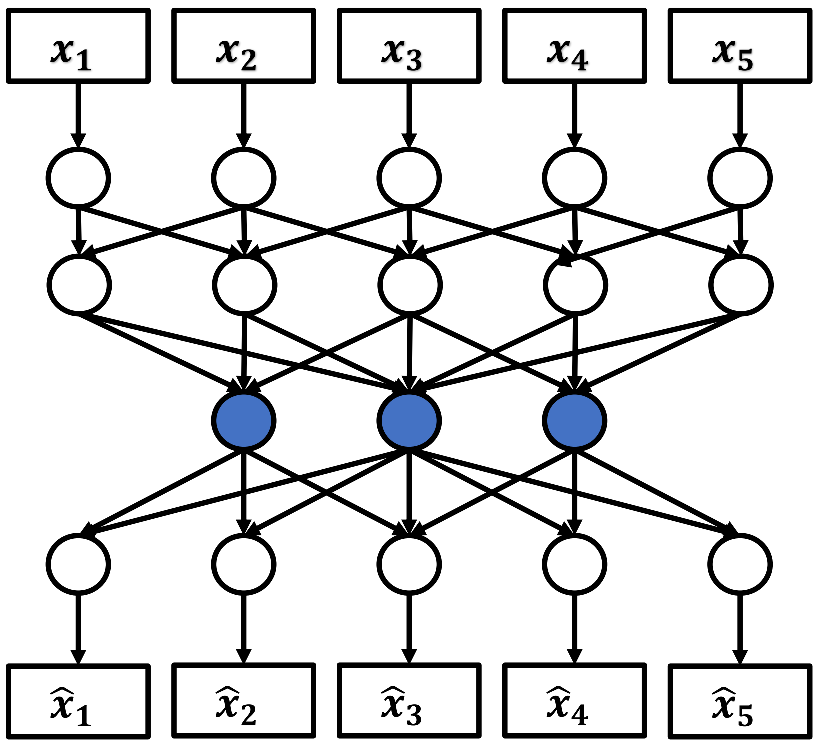


PCA does not capture change points

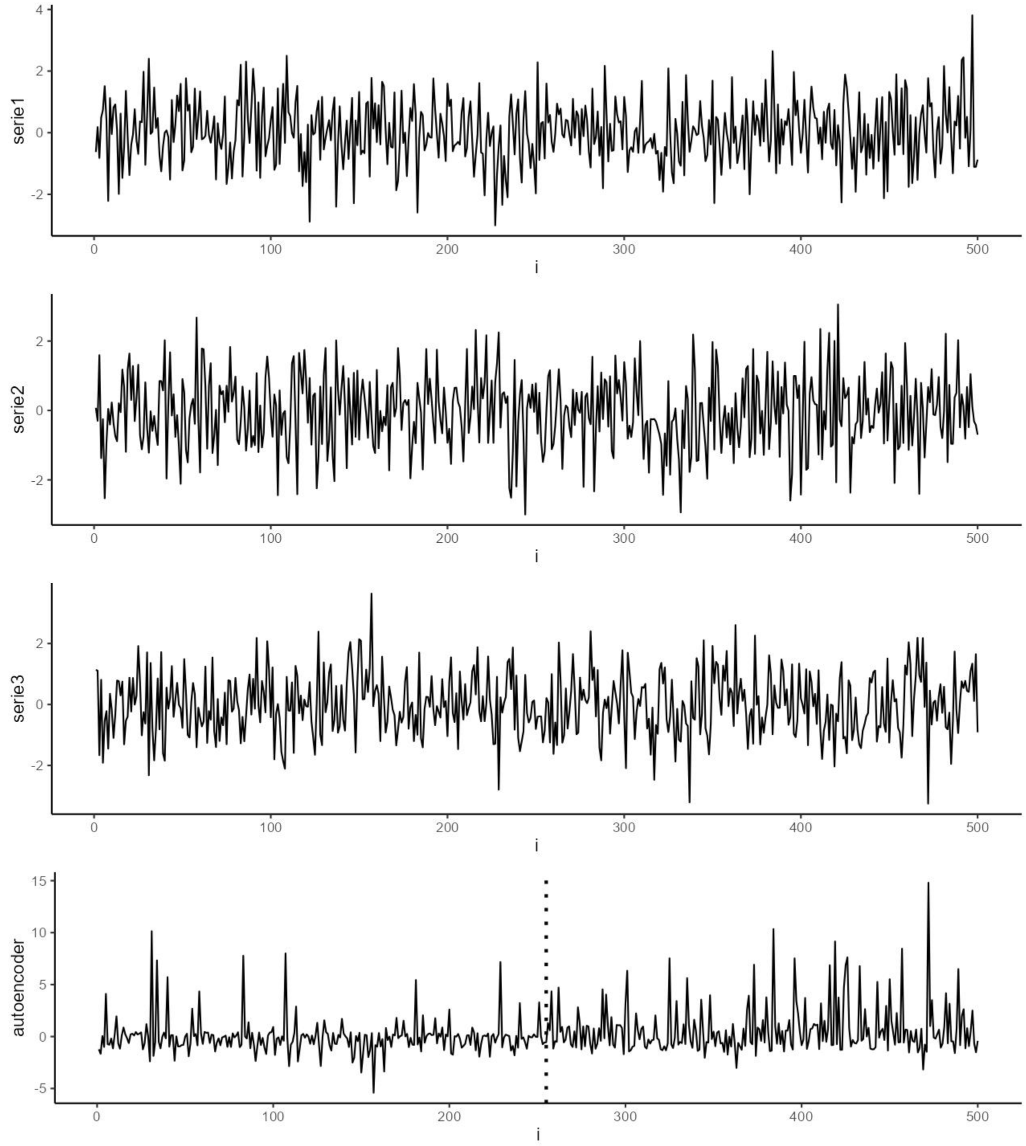


Using autoenconders

- Consider a multivariate time series X = $\{x_1, x_2, x_3, x_4, x_5\}$
- Train multivariate to produce same output
- Extract encoded layer
- Detect change points in encoded time series



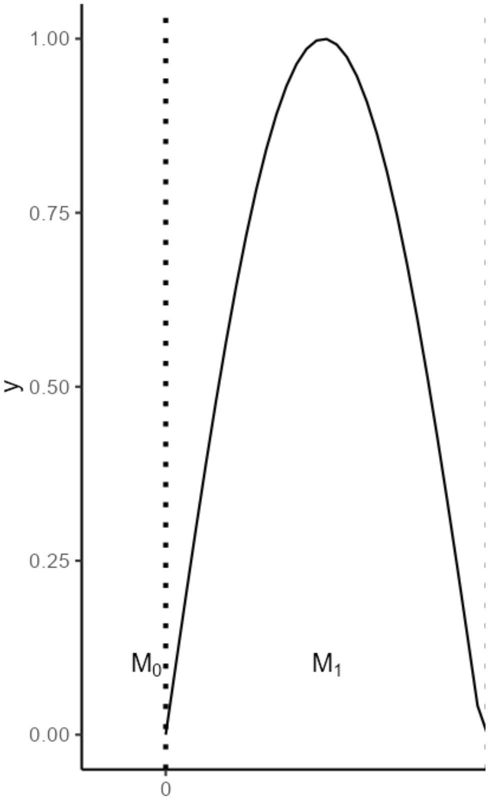
Detection with autoencoder



Challenges: Model management under concept drift

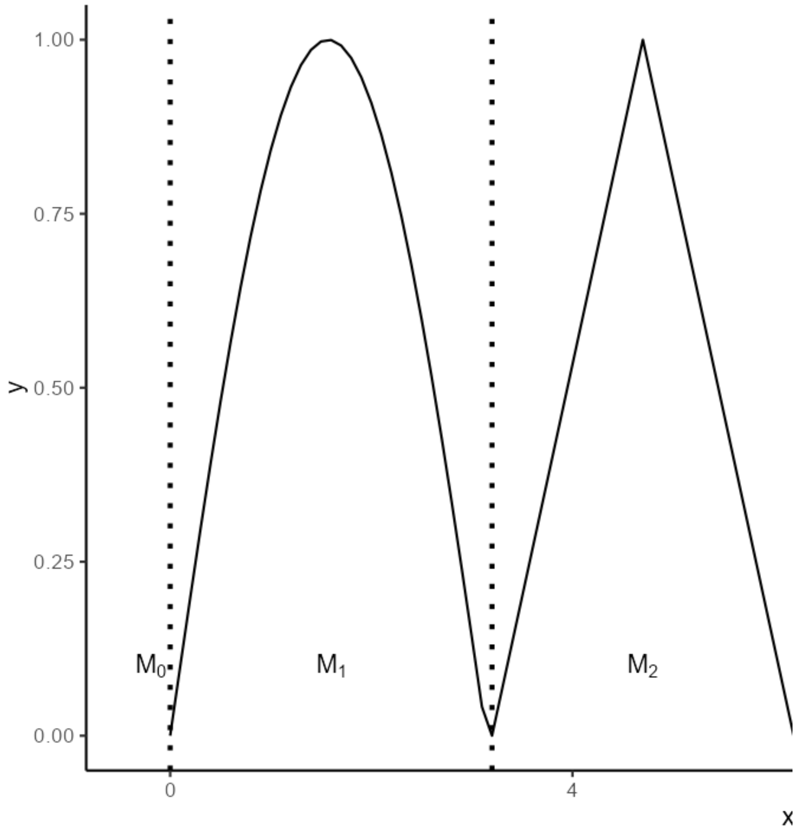
Model management under concept drift

- Assume we have model M_0
- Novel concept appear and new model is built M_1



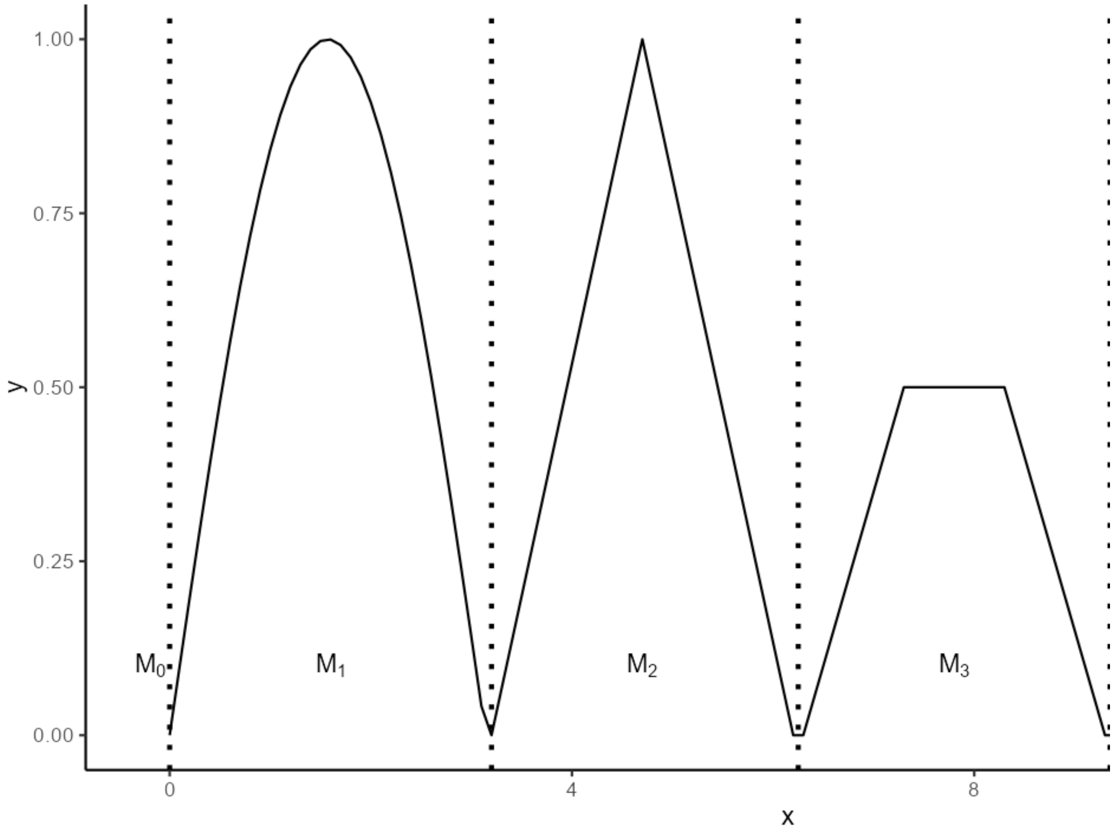
Model management under concept drift

- New batch introduces new concept
- New model is built M_2



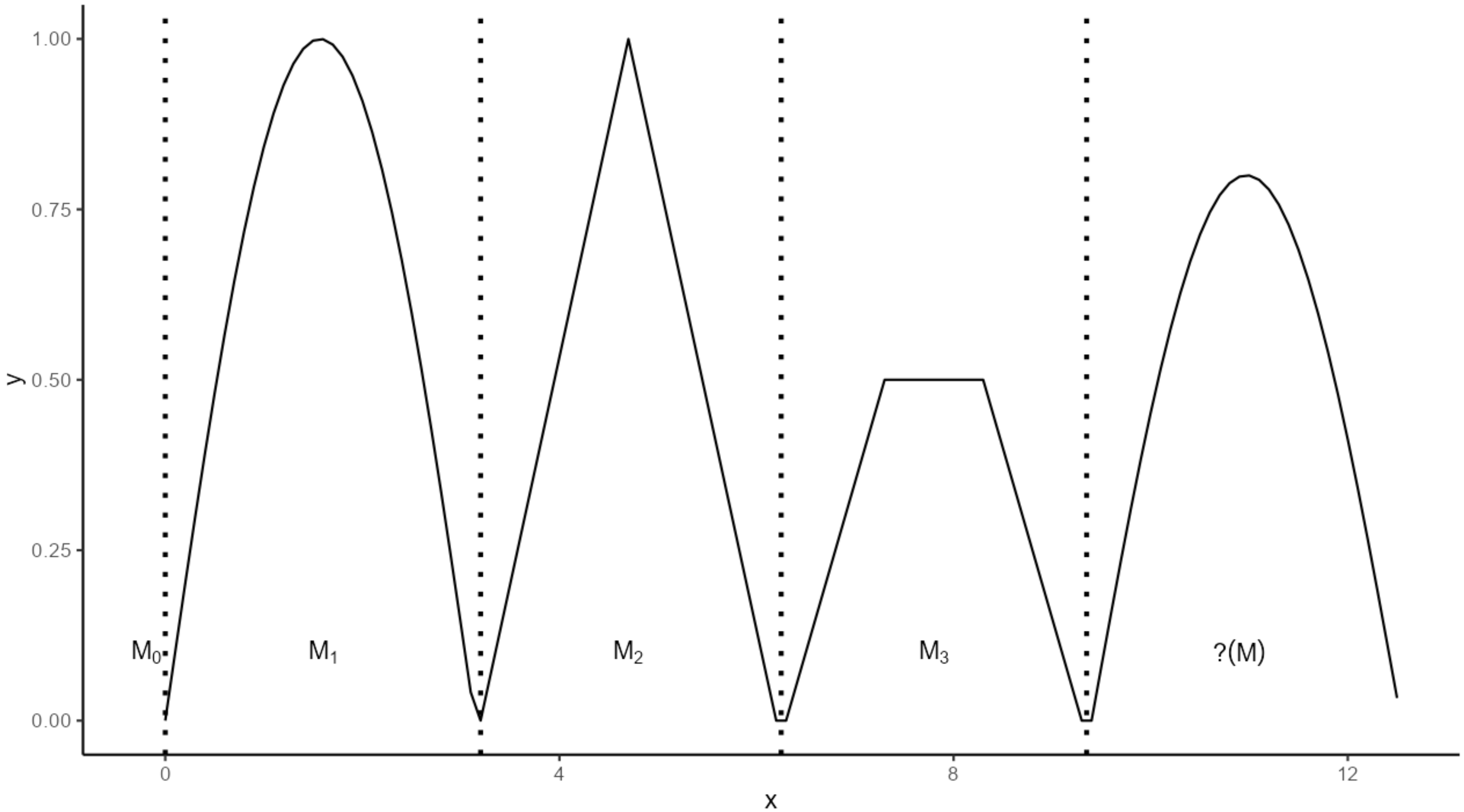
Model management under concept drift

- New batch introduces new concept
- New model is built M_3



Model management under concept drift

- New batch introduces new concept
- Should we build new model (M_4) or can we reuse M_1?
- Could we forget M_0?

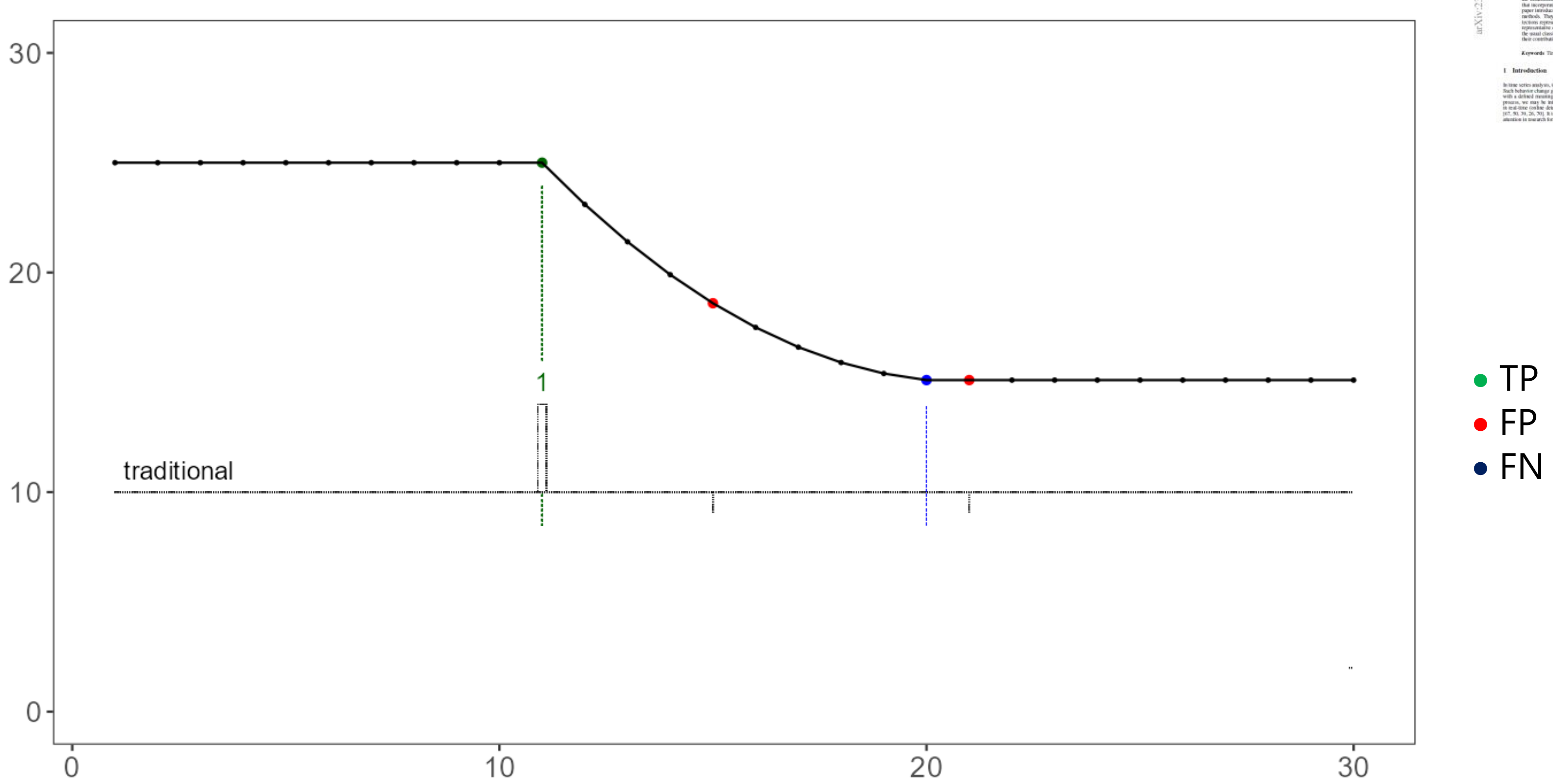


Challenges: Detection Metrics

Detection Metrics for Interval Events

- Traditional scoring methods, such as precision and recall, are not sufficient to assess the performance of event detection
  - They do not incorporate time and do not reward close detections.
  - True positives are rewarded
  - All other outcomes are equally penalized


SoftED: Metrics for Soft Evaluation of Time Series Event Detection



[1] R. Salles, J. Lima, R. Coutinho, E. Pacitti, F. Masseglia, R. Akbarinia, C. Chen, J. Garibaldi, F. Porto, and E. Ogasawara, “SoftED: Metrics for Soft Evaluation of Time Series Event Detection.” arXiv, Apr. 01, 2023. doi: 10.48550/arXiv.2304.00439.

Challenges: Anomalies and Change Points related to Volatility

Detection of anomalies and change points related to volatility