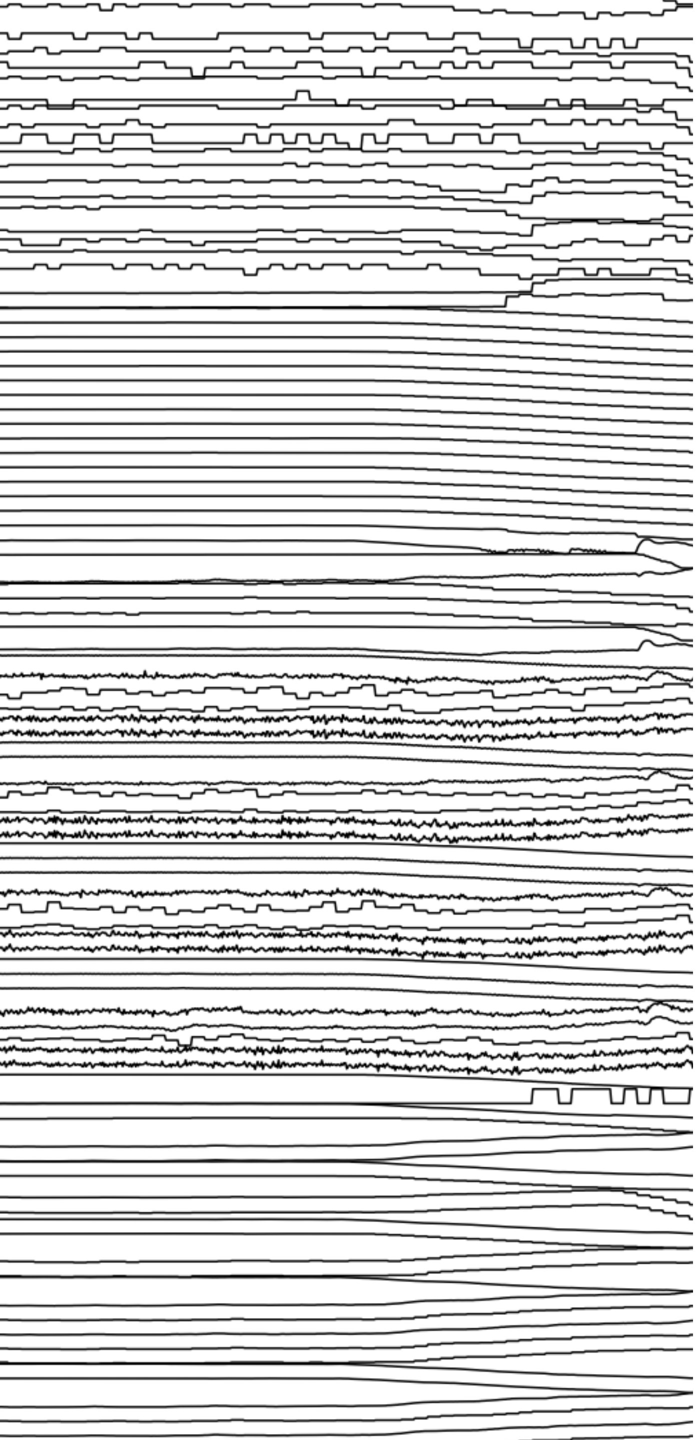


An Interactive Dive into Time Series Anomaly Detection

Paul Boniol
Inria, ENS, PSL University
Paul.boniol@inria.fr

John Paparrizos
Ohio State University
paparrizos.1@osu.edu

Themis Palpanas
Université Paris Cité; IUF
themis@mi.parisdescartes.fr

Inria | ENS École Normale Supérieure | Université Paris Cité | The Ohio State University

Introduction: *Time series are Everywhere*

Energy Production

Edf.fr: tinyurl.com/yc7x5xje

Astrophysics

Virgo: https://www.virgo-gw.eu/

Medicine

tinyurl.com/39dx2us4

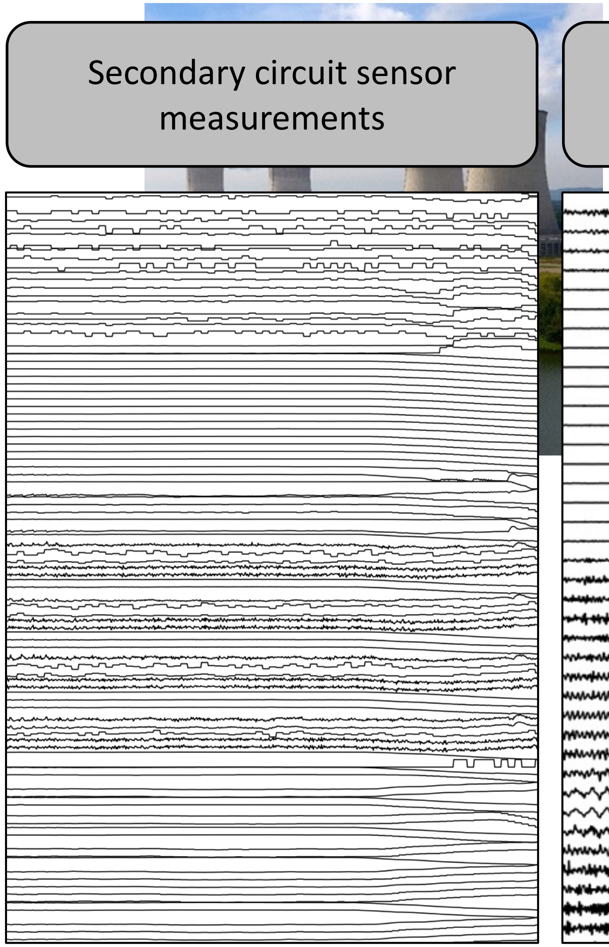
Volcanology

tinyurl.com/ybcttmfz

Introduction: *Time series are Everywhere*

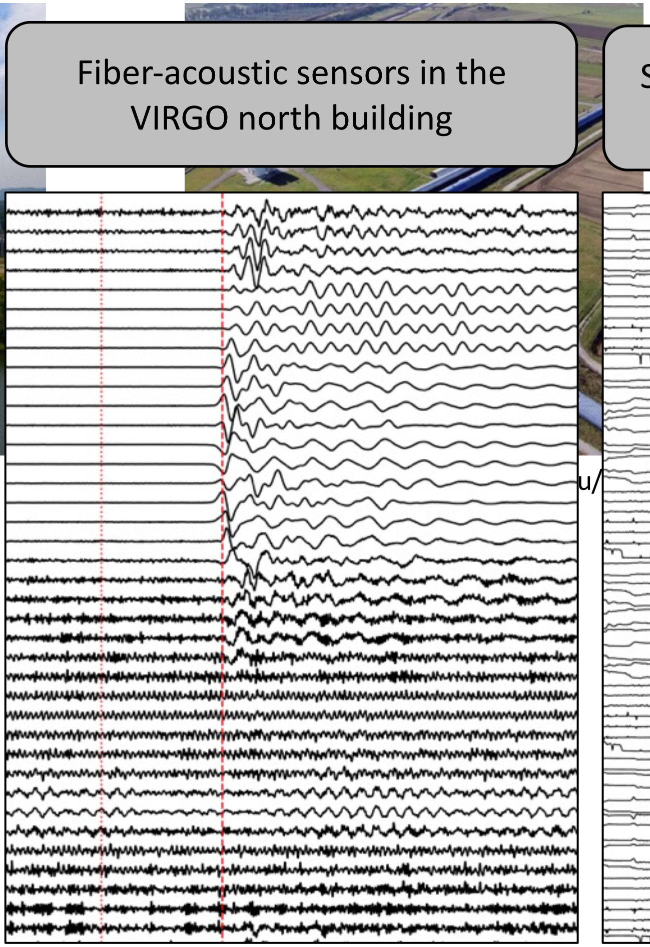
Energy Production

Secondary circuit sensor measurements



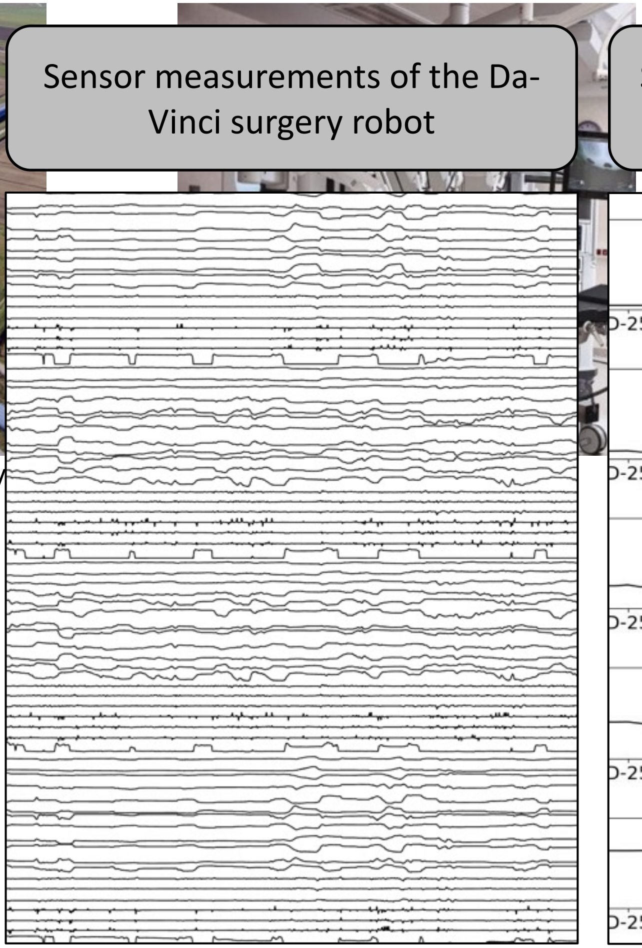
Astrophysics

Fiber-acoustic sensors in the VIRGO north building



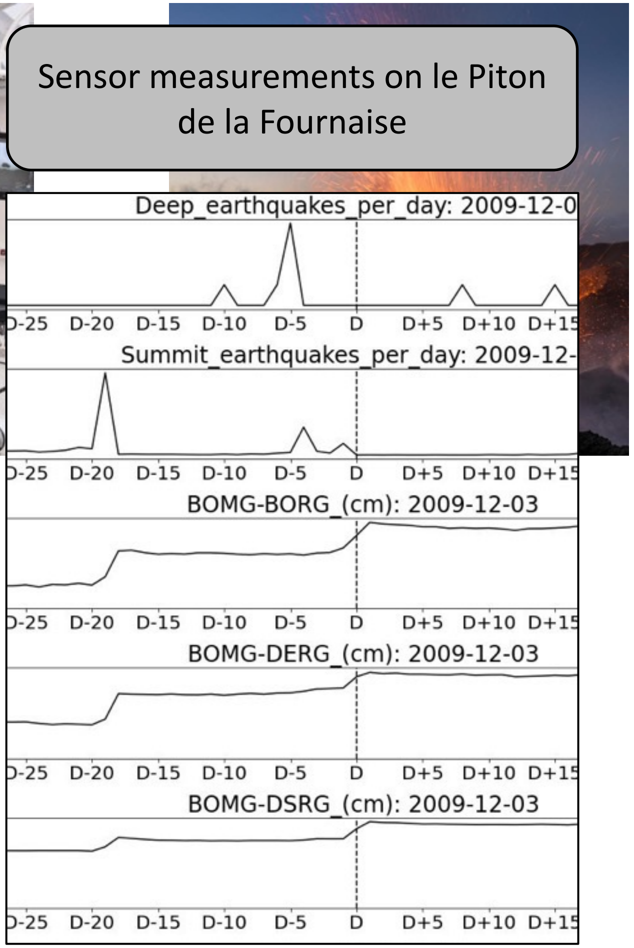
Medicine

Sensor measurements of the Da-Vinci surgery robot



Volcanology

Sensor measurements on le Piton de la Fournaise

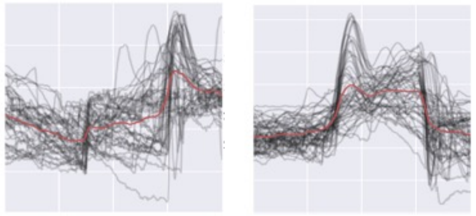


Introduction: *with Important Challenges*

Energy Production

Secondary circuit sensor measurements

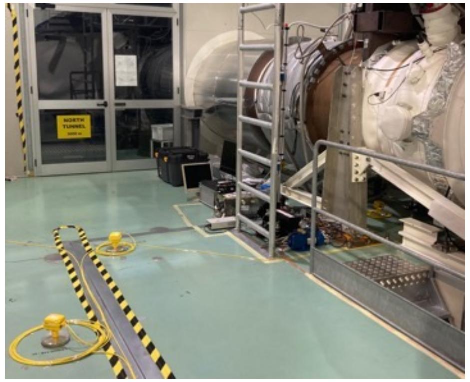
Identification of precursors of feed-water pumps vibrations



Astrophysics

Fiber-acoustic sensors in the VIRGO north building

Noise detection in VIRGO interferometer north building



Medicine

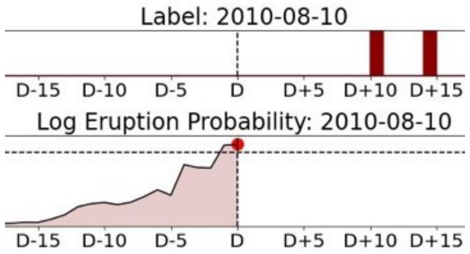
Sensor measurements of the Da-Vinci surgery robot

Unusual surgeons gestures detection

Volcanology

Sensor measurements on le Piton de la Fournaise

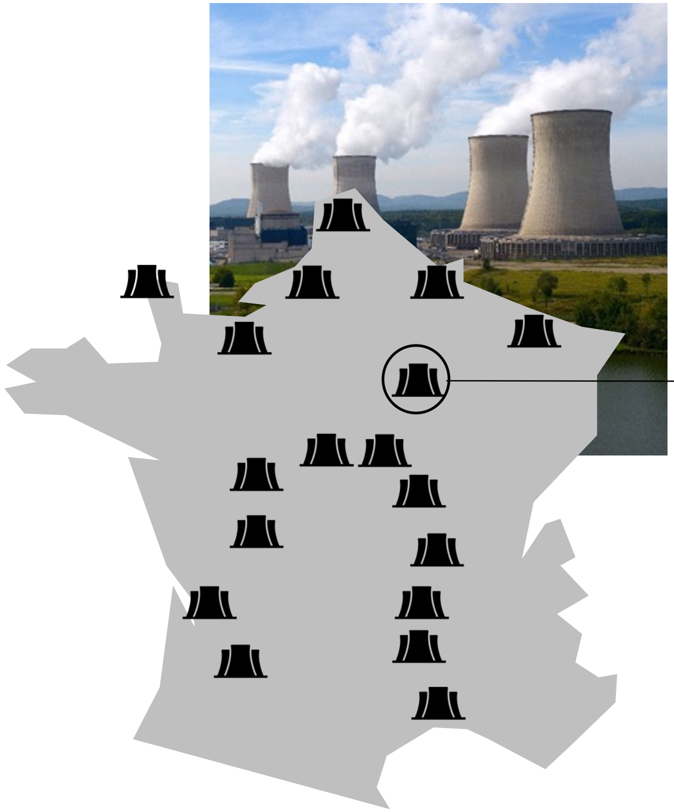
Detection of abnormal events on the volcano for predicting eruptions



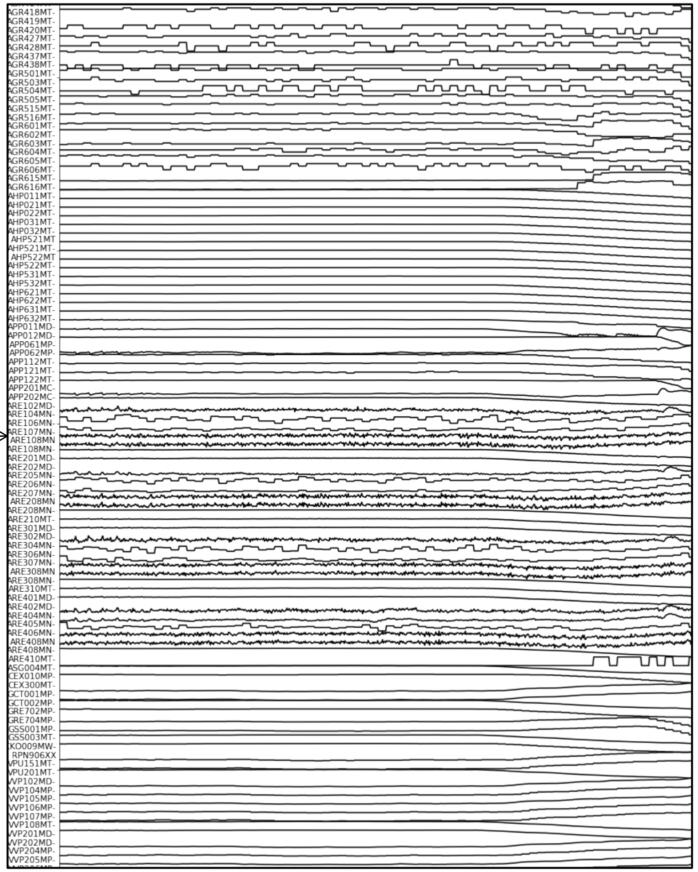
Introduction: *with Important Challenges*

Large-scale time series database

Energy Production



Edf.fr: tinyurl.com/yc7x5xje



Example of Nuclear production

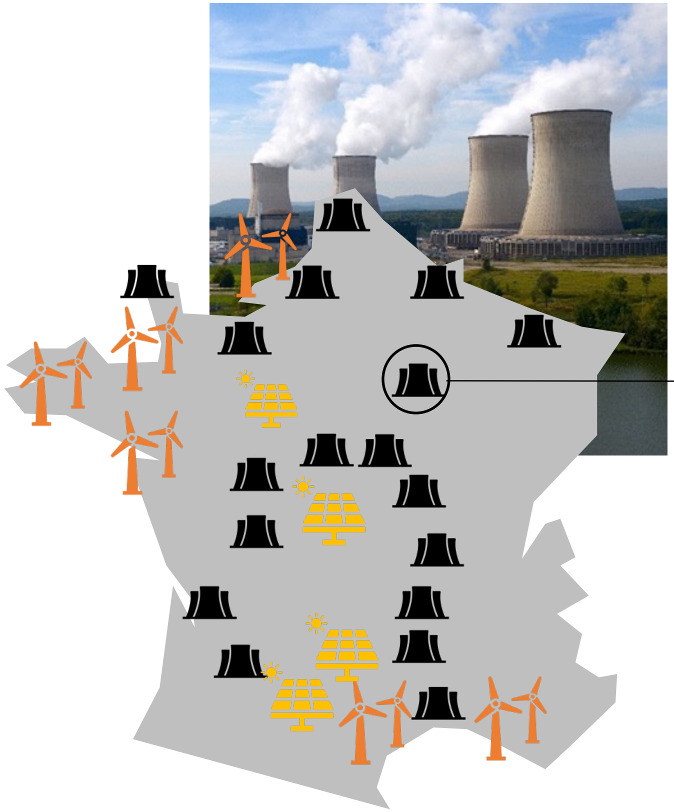
- 58 nuclear power plants across France
- 2000+ sensors per power plant
- 30 years of data collections

A total of 500 TeraBytes

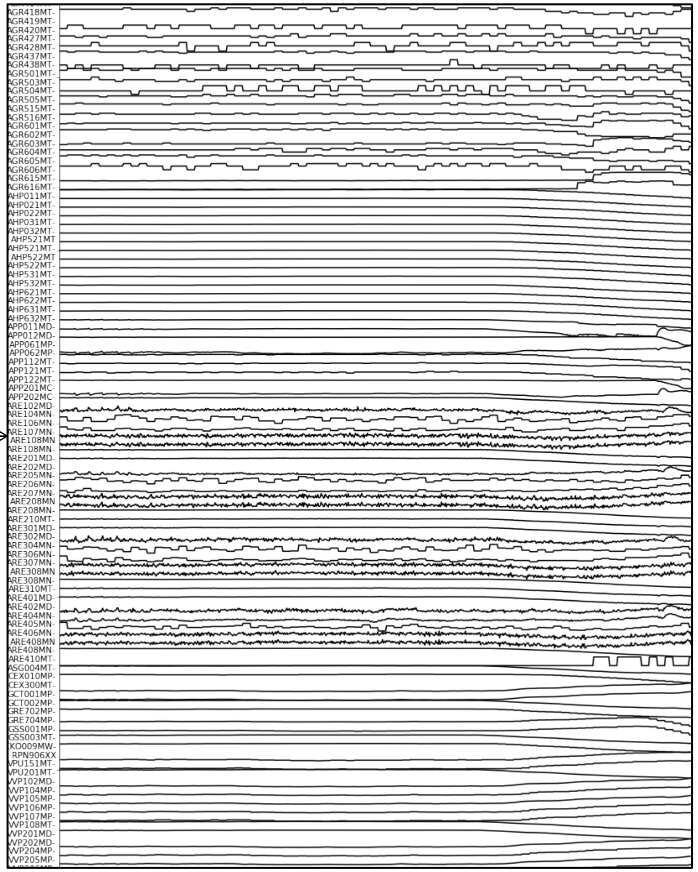
Introduction: *with Important Challenges*

Large-scale time series database

Energy Production



Edf.fr: tinyurl.com/yc7x5xje



Example of Nuclear production

- 58 nuclear power plants across France
- 2000+ sensors per power plant
- 30 years of data collections

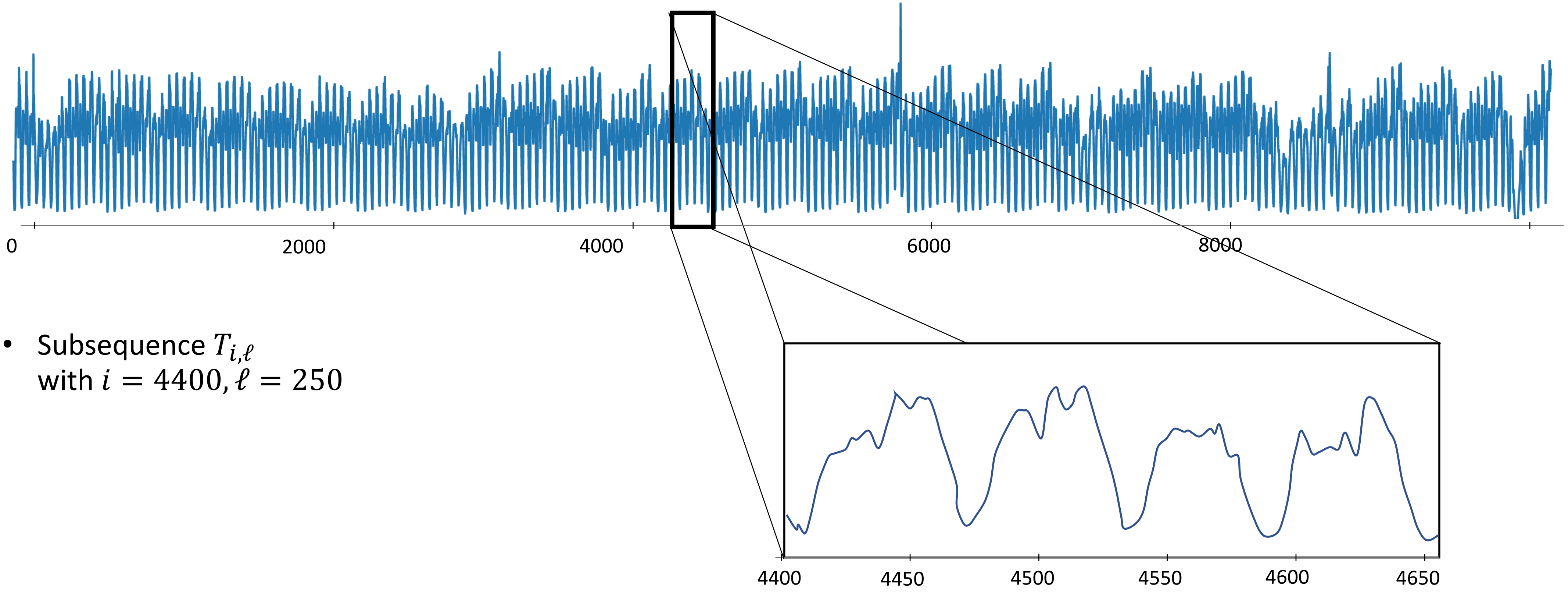
A total of 500 TeraBytes

Other source of production

- New sensors with higher acquisition rate

Introduction: *Anomaly Detection in Time Series*

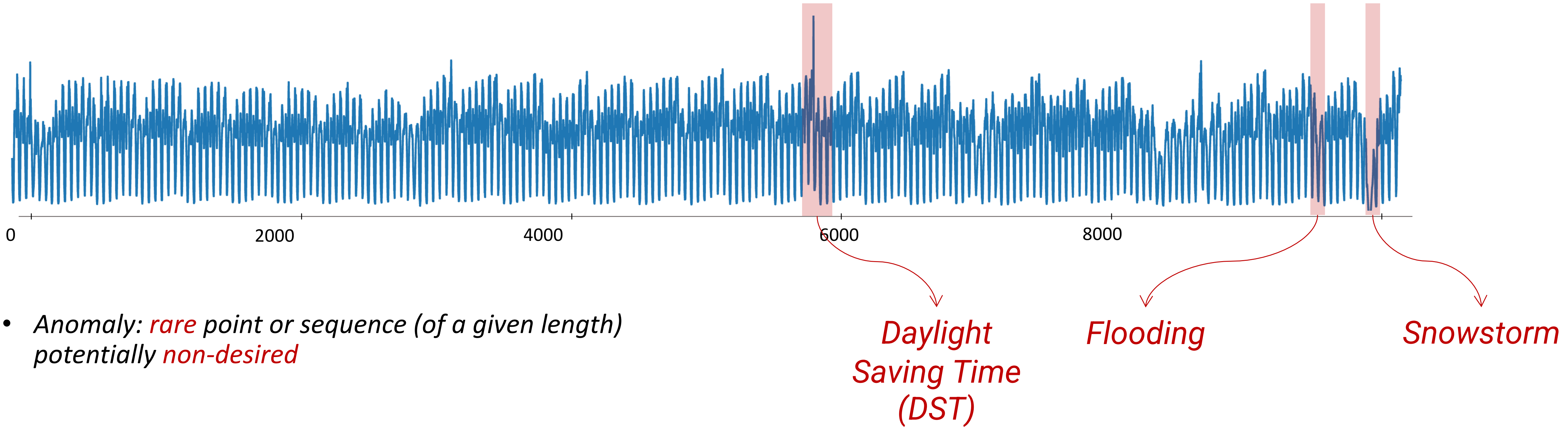
- Time series T *(example : number of taxi passengers in New York City)*



- Subsequence $T_{i,\ell}$
  with $i = 4400, \ell = 250$

Introduction: *Anomaly Detection in Time Series*

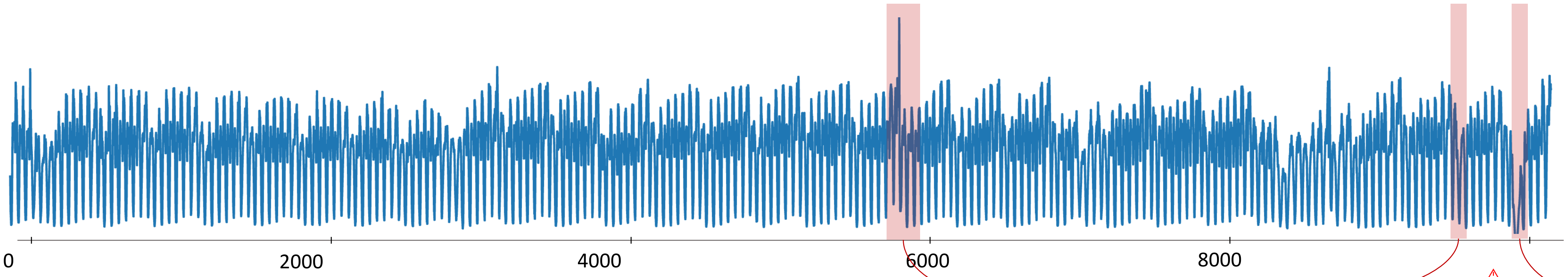
- Time series T *(example : number of taxi passengers in New York City)*



- *Anomaly: rare point or sequence (of a given length) potentially non-desired*

Introduction: *Anomaly Detection in Time Series*

- Time series T *(example : number of taxi passengers in New York City)*

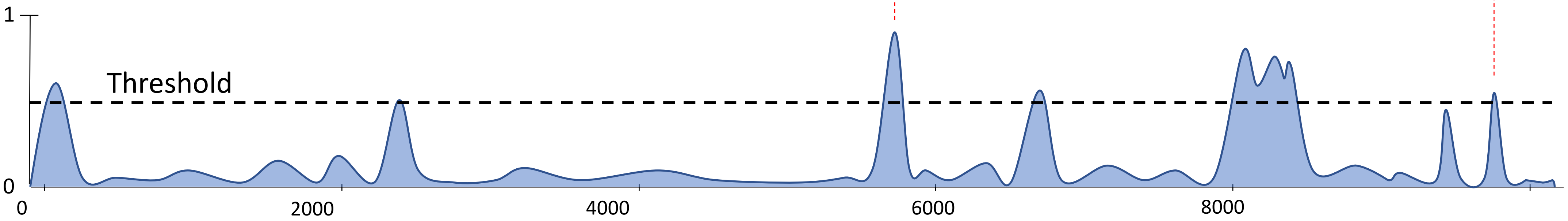


Daylight Saving Time (DST)

Flooding

Snowstorm

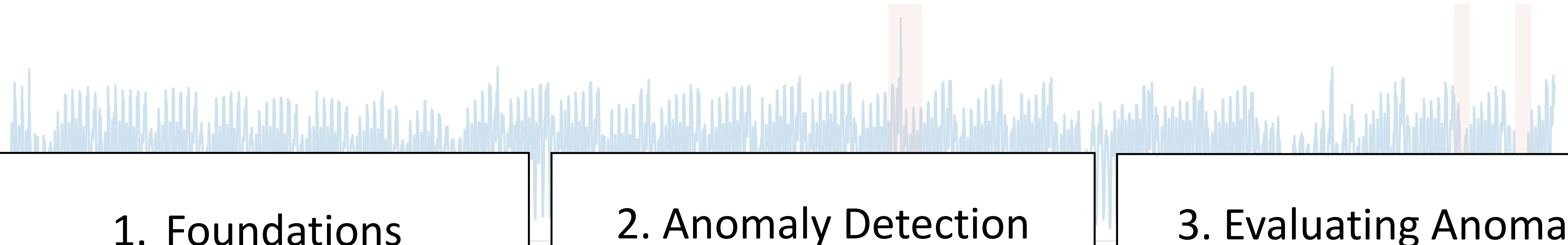
- *Anomaly: rare point or sequence (of a given length) potentially non-desired*



Anomaly score S_T

Introduction: *Outline*

- Time series *(example : number of taxi passengers in New York City)*



1. Foundations

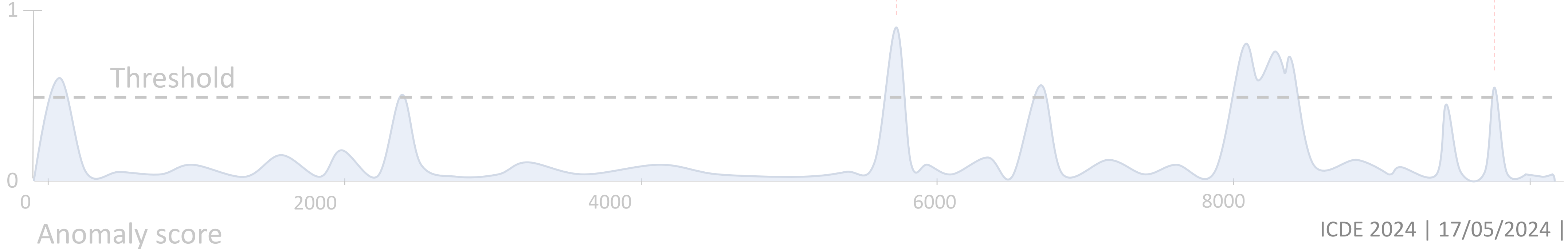
1.1. Type of Time Series
1.2. Type of Anomalies

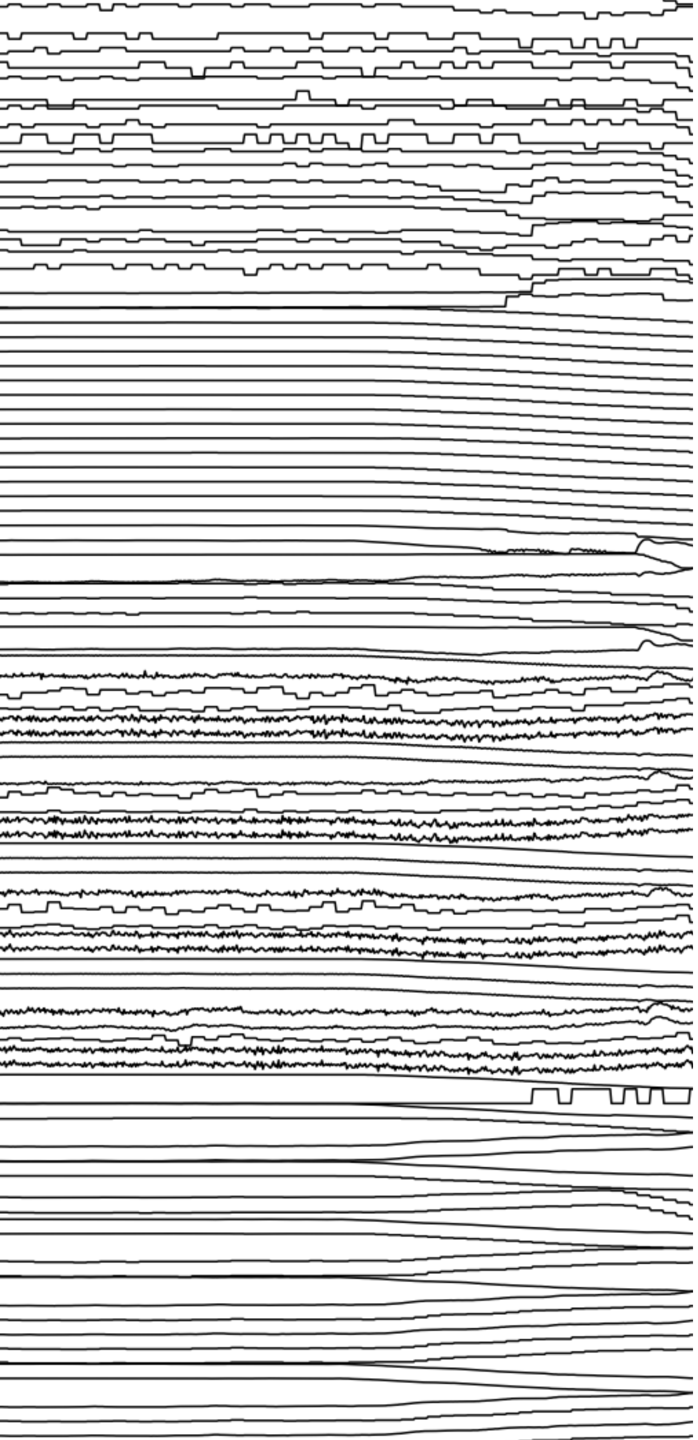
2. Anomaly Detection Methods

2.1. A Taxonomy of Methods
2.2. Existing Benchmarks

3. Evaluating Anomaly Detection

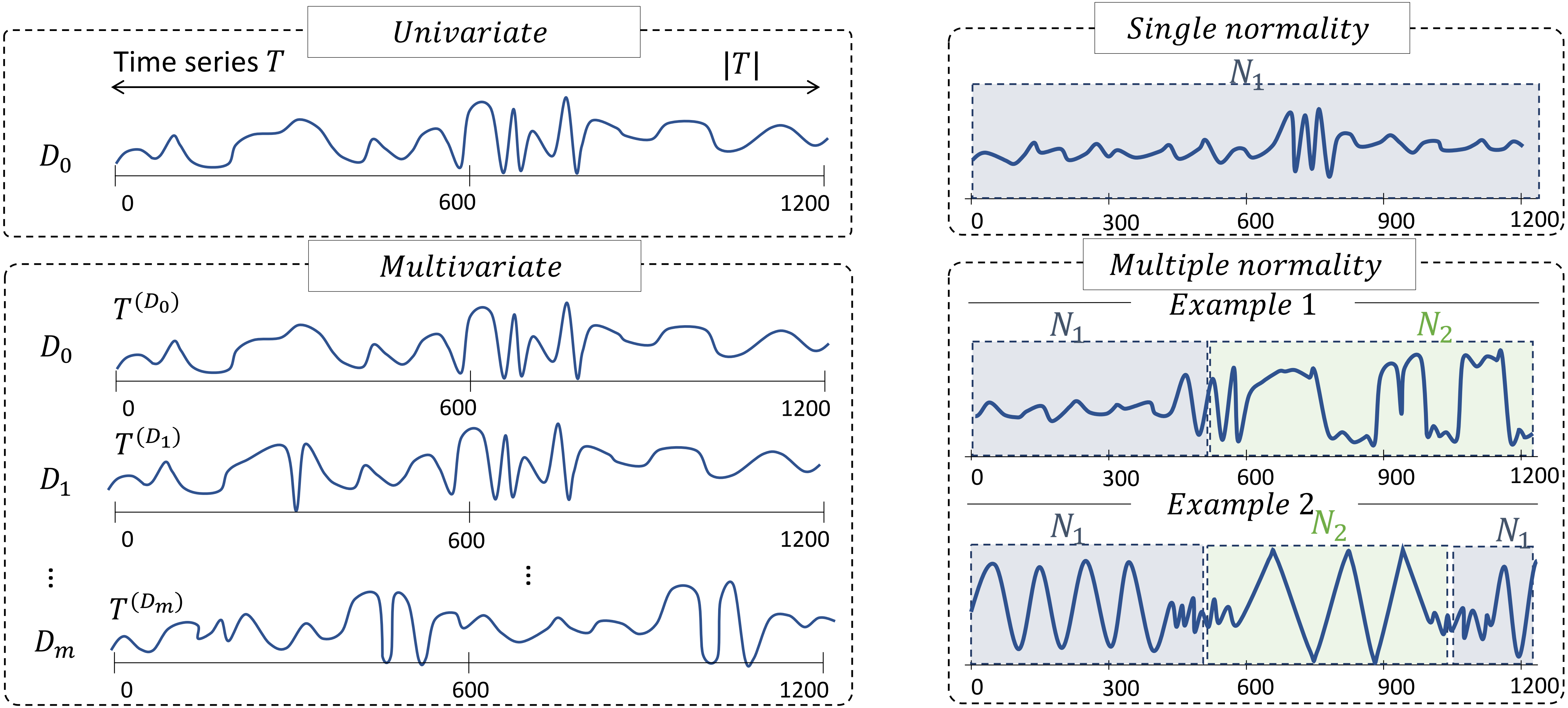
3.1. Threshold-based
3.2. Time series labeling issues



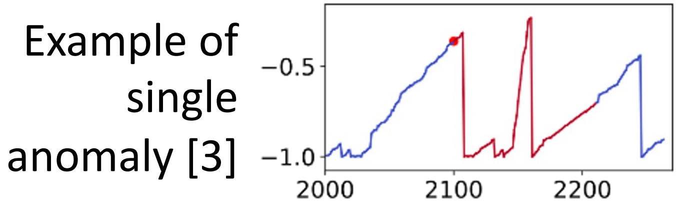


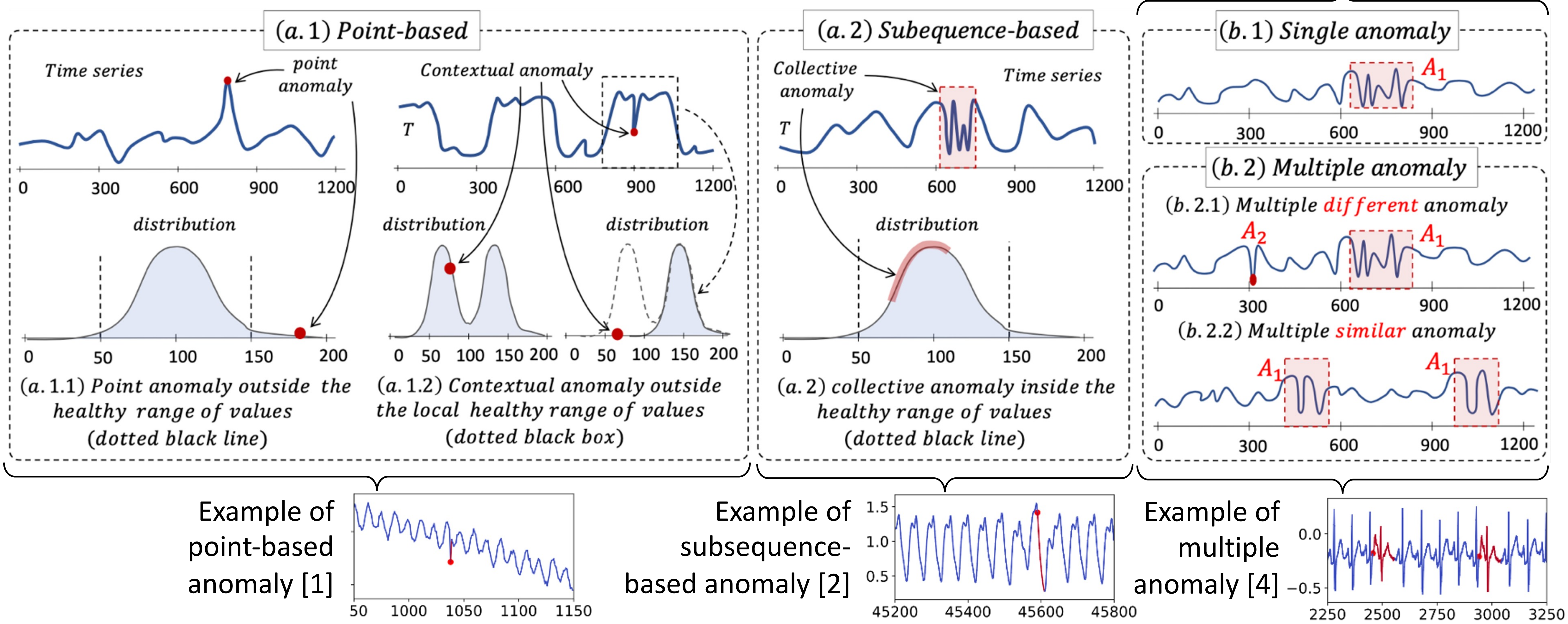
Foundations

Foundations: *Type of time series*

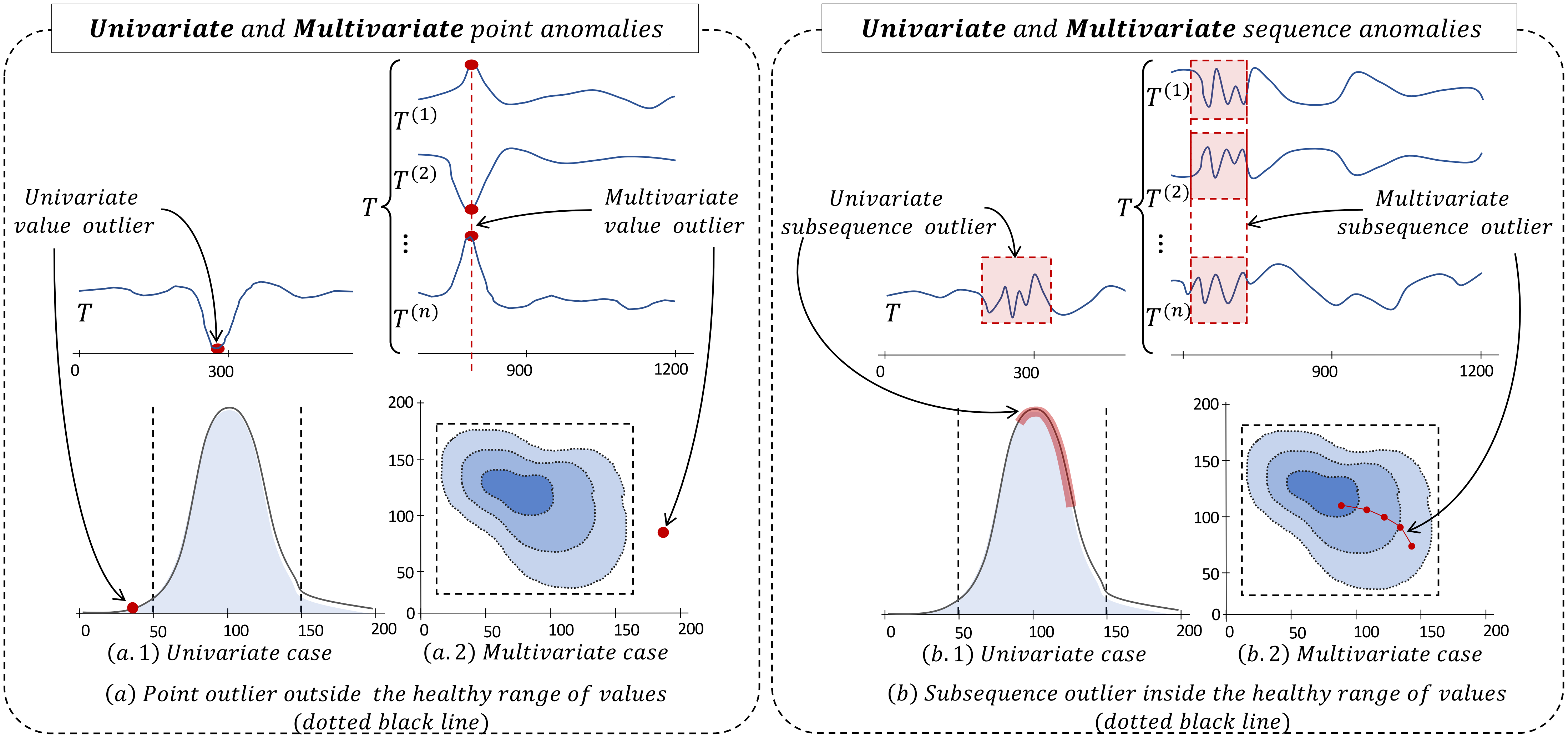


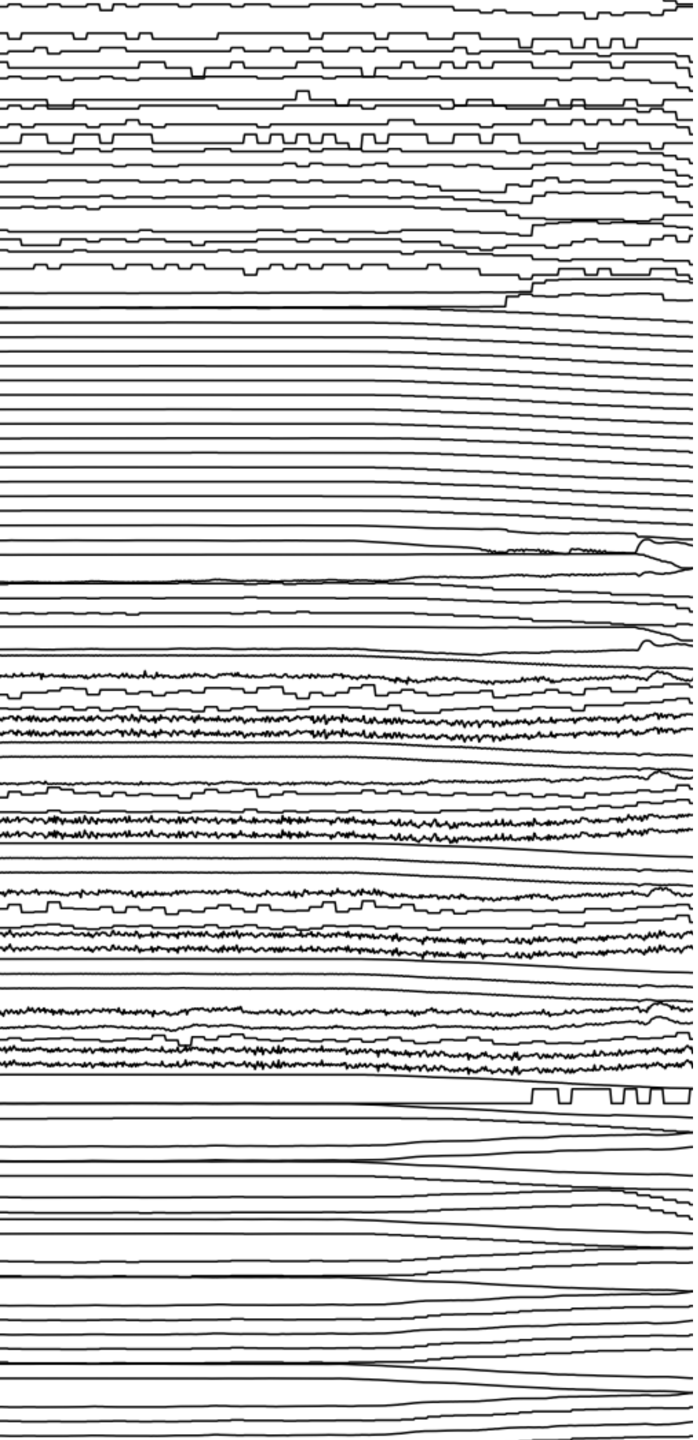
Foundations: *Type of anomalies*





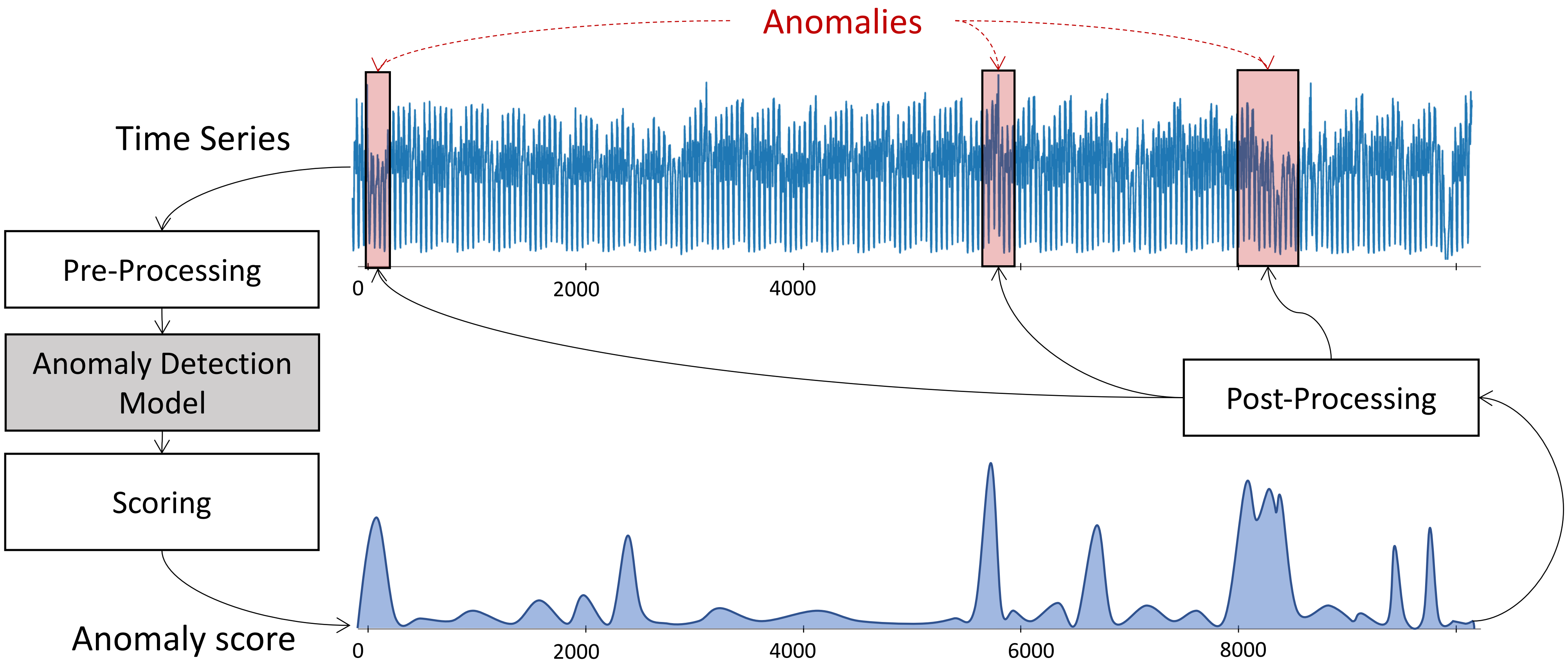
Foundations: *Type of anomalies*





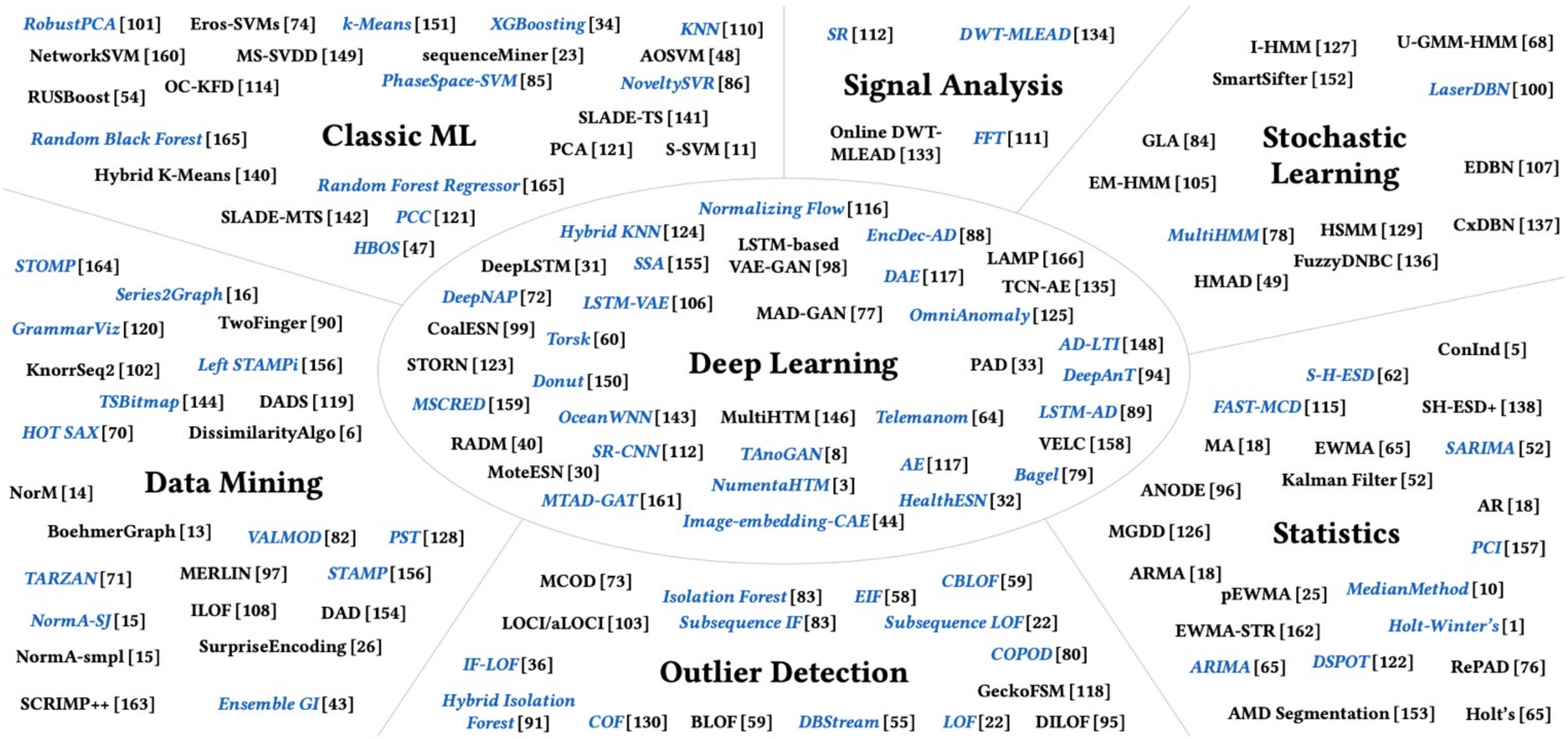
Anomaly Detection Methods

Anomaly Detection methods: *A taxonomy*



Anomaly Detection methods: *A taxonomy*

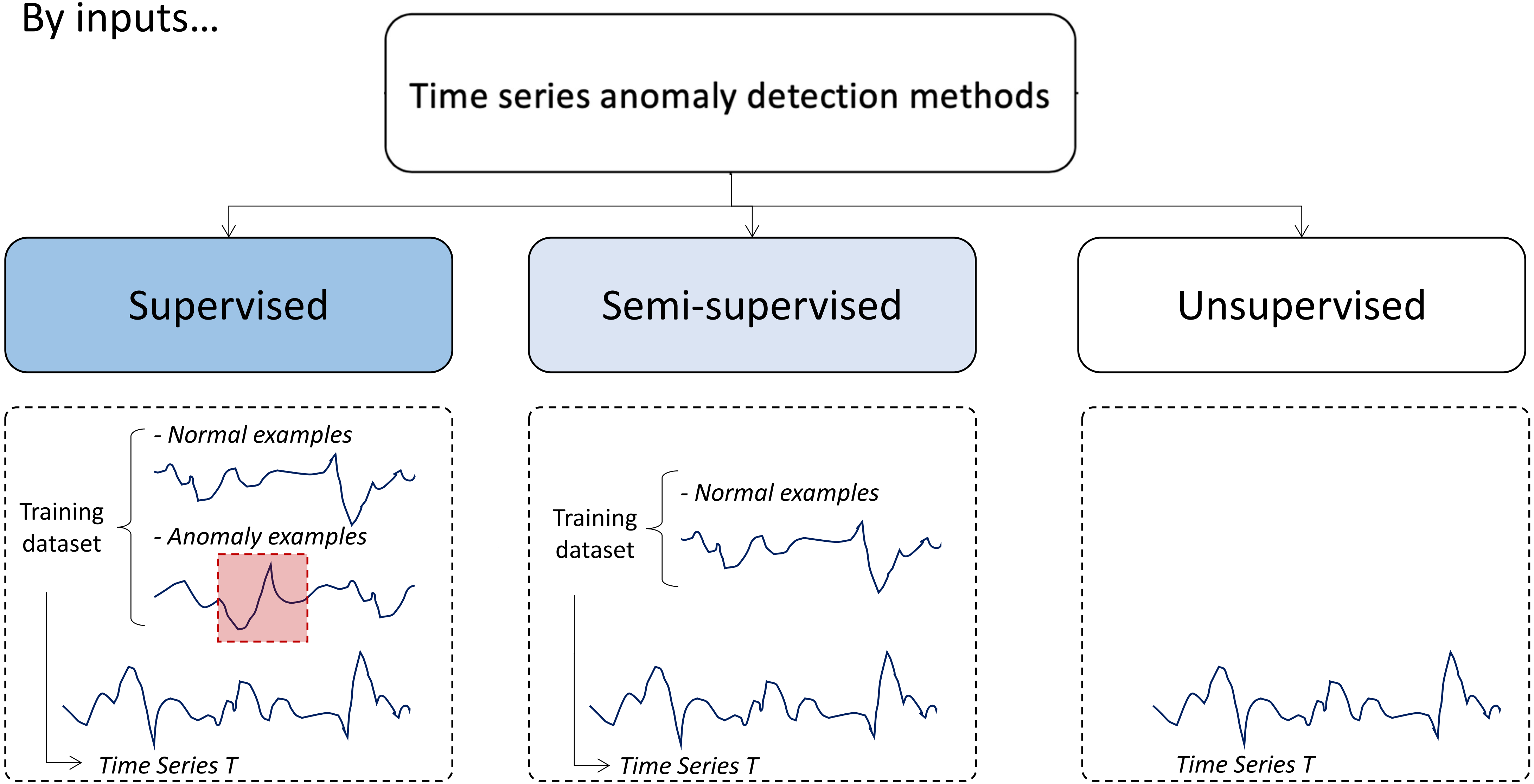
By domains [5] …



[5] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. 2022. Anomaly detection in time series: a comprehensive evaluation. Proc. VLDB Endow. 15, 9 (May 2022), 1779–1797.

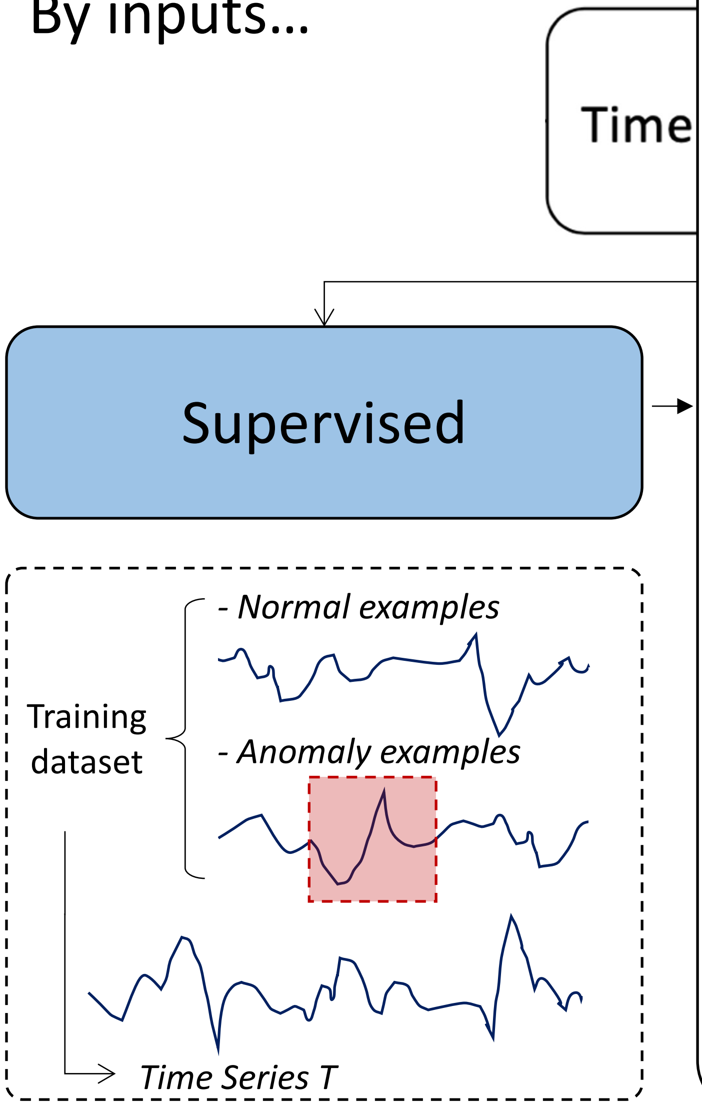
Anomaly Detection methods: *A taxonomy*

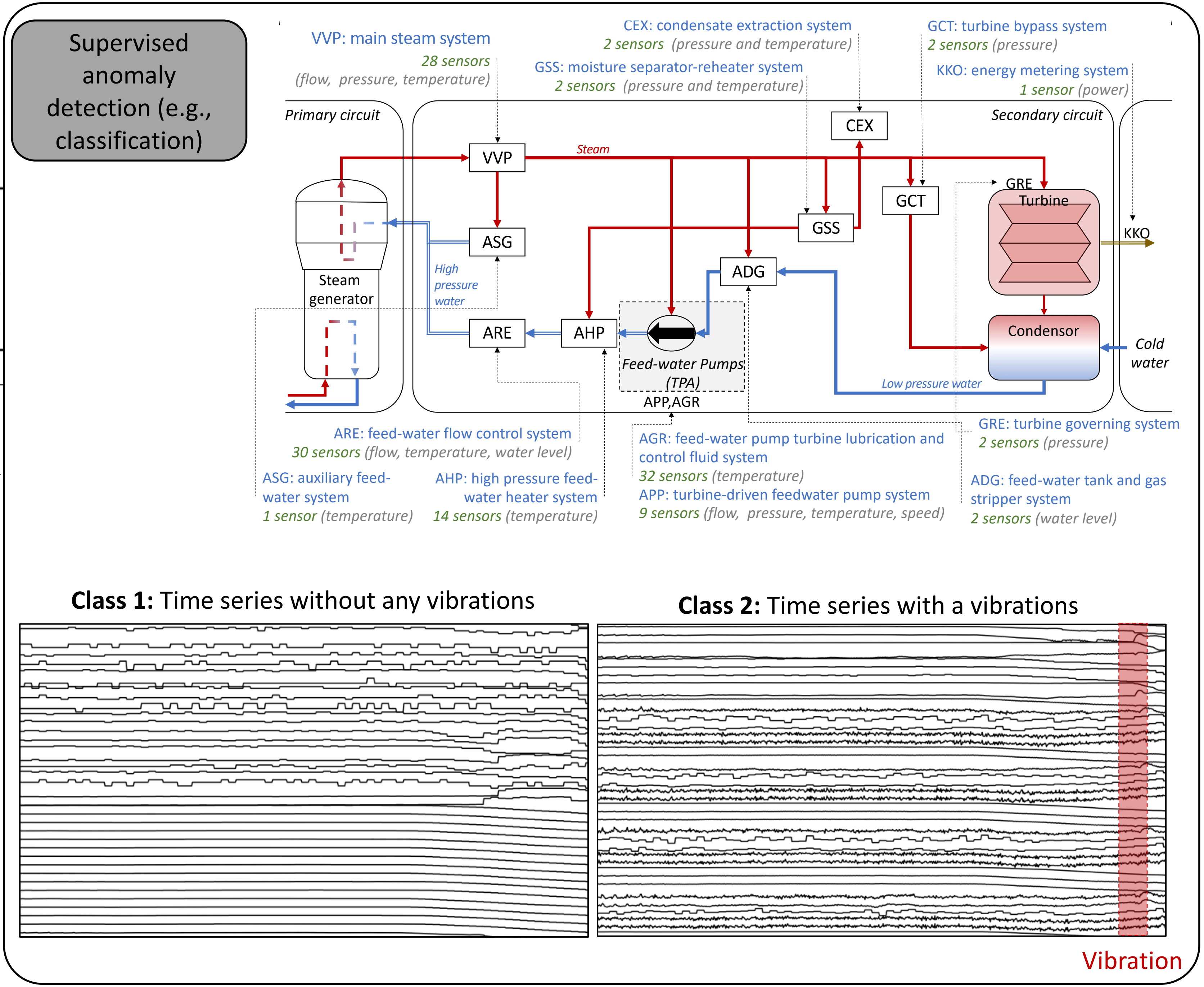
By inputs…



Anomaly Dete

By inputs…



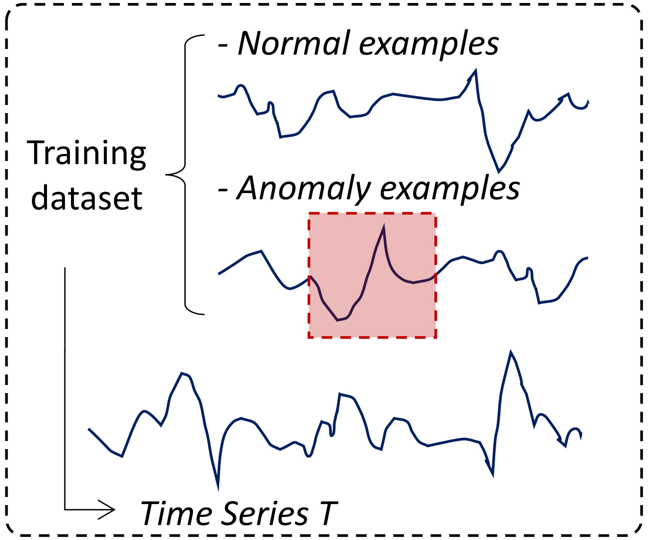


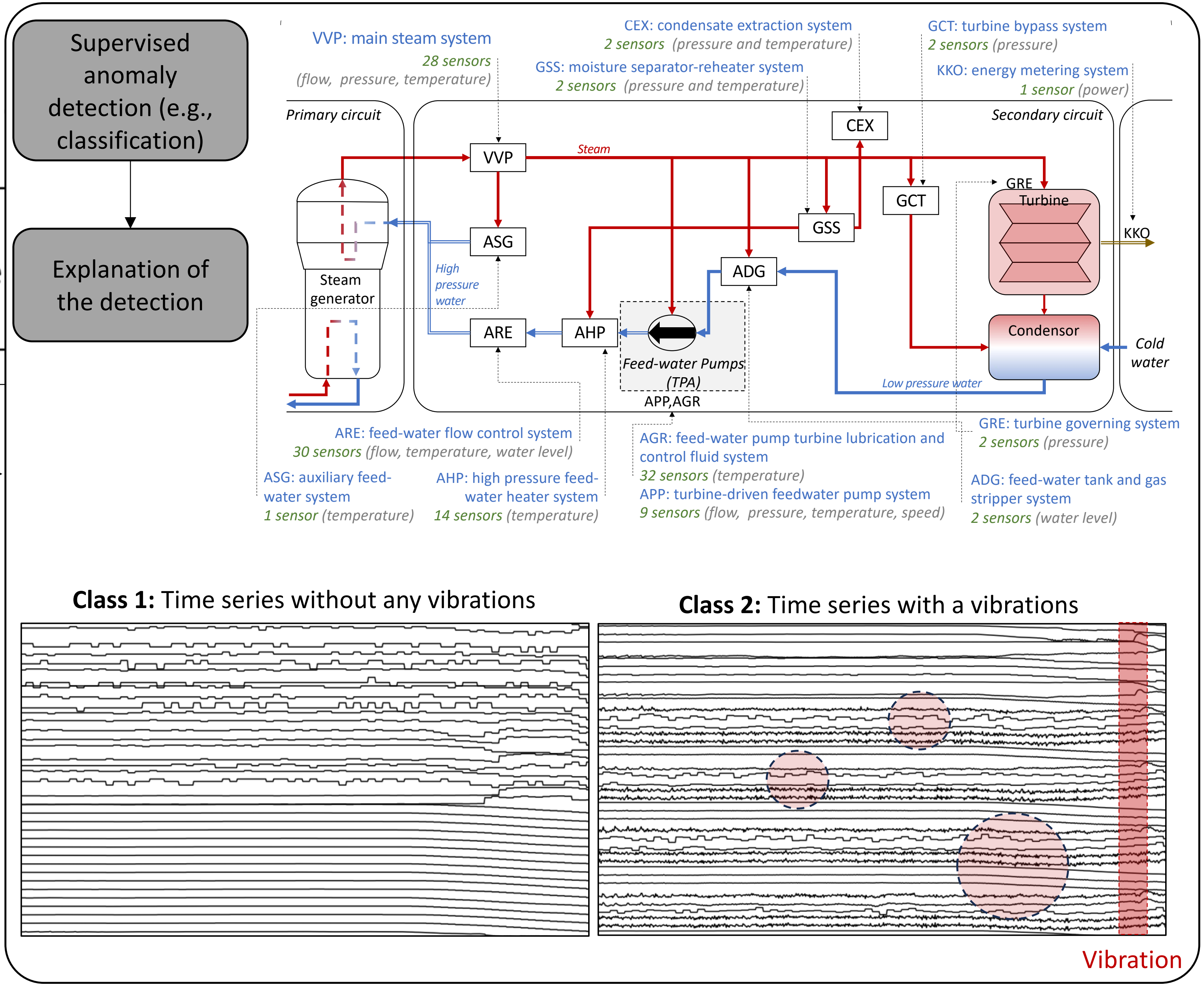
Anomaly Dete

By inputs…

Time

Supervised



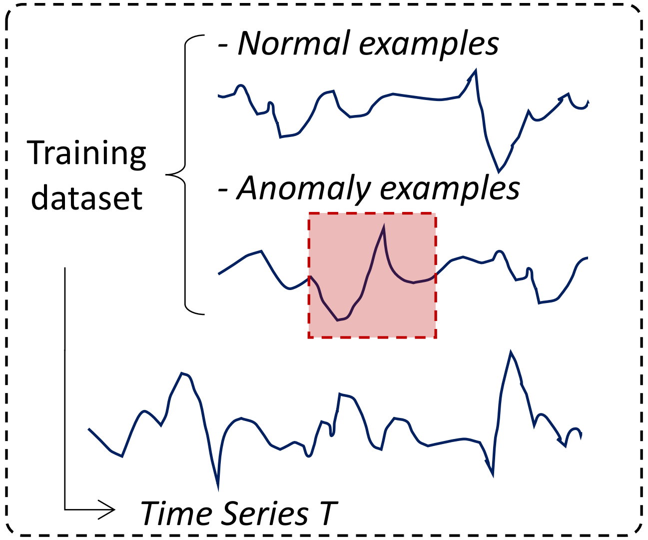


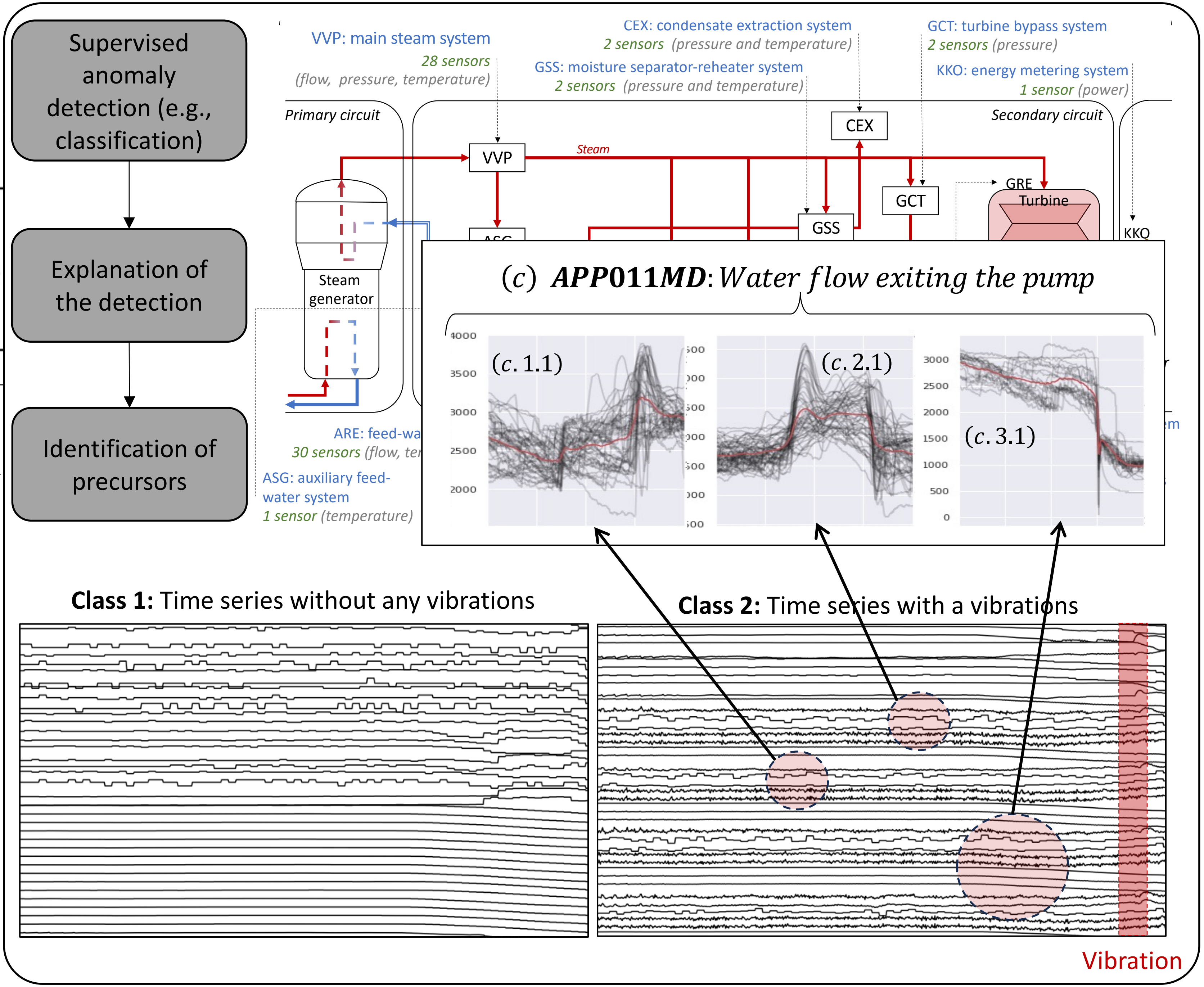
Anomaly Dete

By inputs…

Time

Supervised



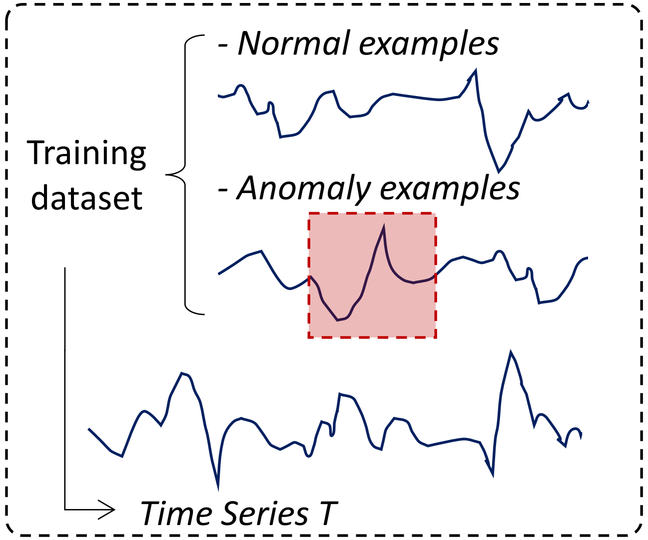


Anomaly Dete

By inputs…

Time

Supervised



More info :

On the use case

DCE journal 2023

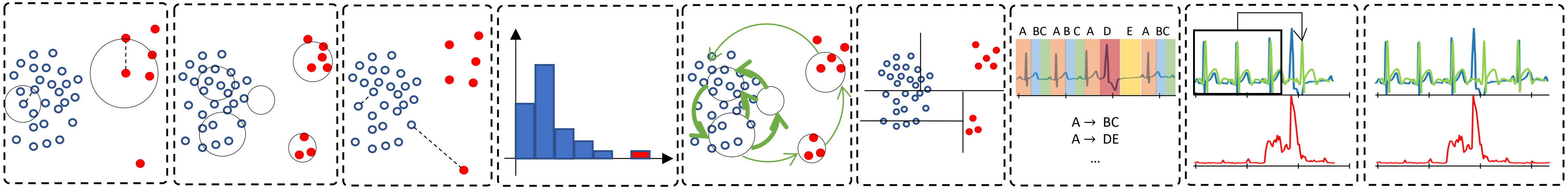
On the method

SIGMOD 2022

Anomaly Detection methods: *A taxonomy*

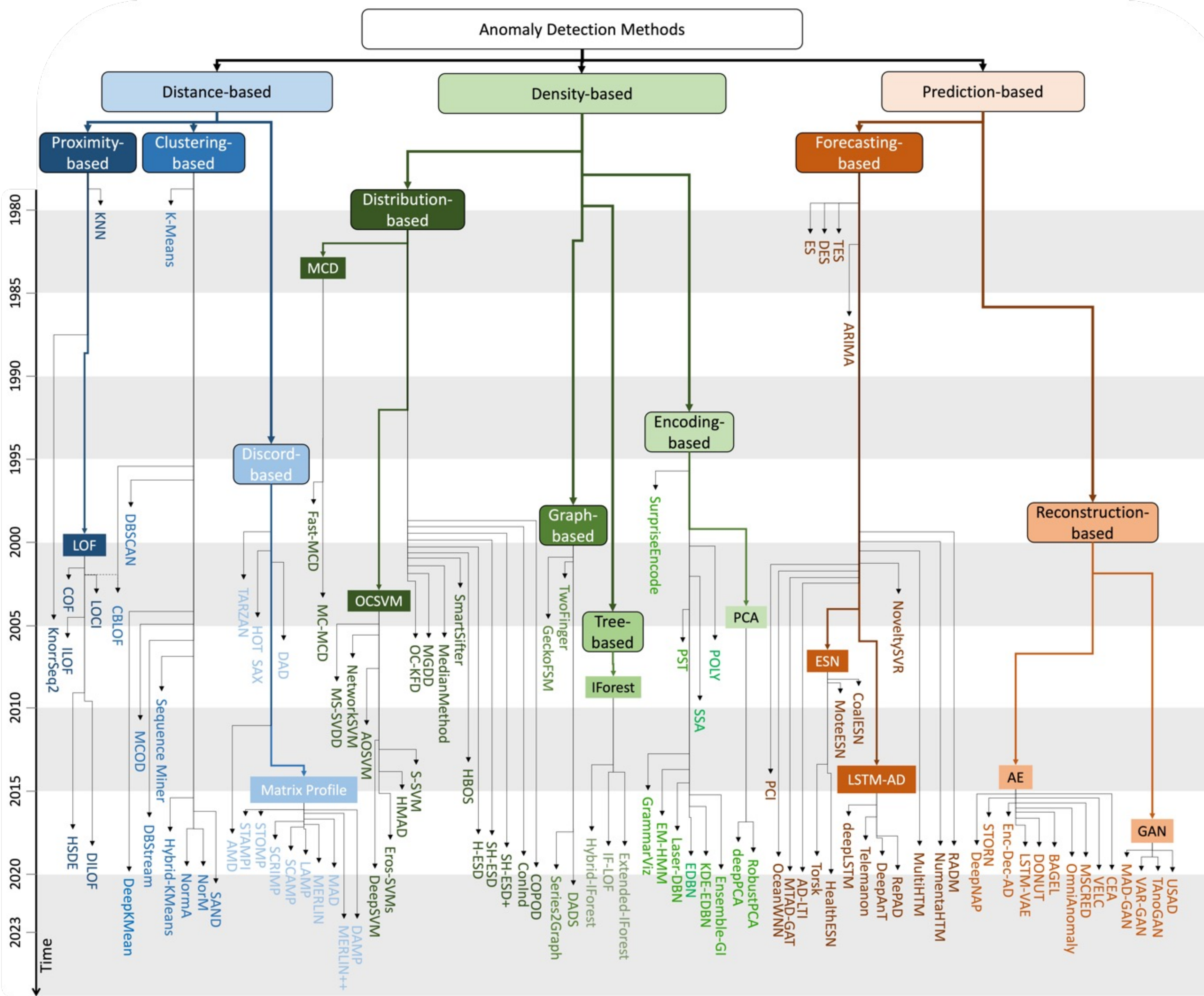
By methods…

- Time series anomaly detection methods
  - Distance-based
    - Proximity-based — *E.g.,* LOF
    - Clustering-based — *E.g.,* NormA, SAND
    - Discord-based — *E.g.,* MP, DAMP
  - Density-based
    - Distribution-based — *E.g.,* HOBS, OCSVM
    - Graph-based — *E.g.,* Series2Graph
    - Tree-based — *E.g.,* Isolation-Forest
    - Encoding-based — *E.g.,* GrammarViz
  - Prediction-based
    - Forecasting-based — *E.g.,* LSTM,CNN, POLY
    - Reconstruction-based — *E.g.,* PCA, AutoEncoder



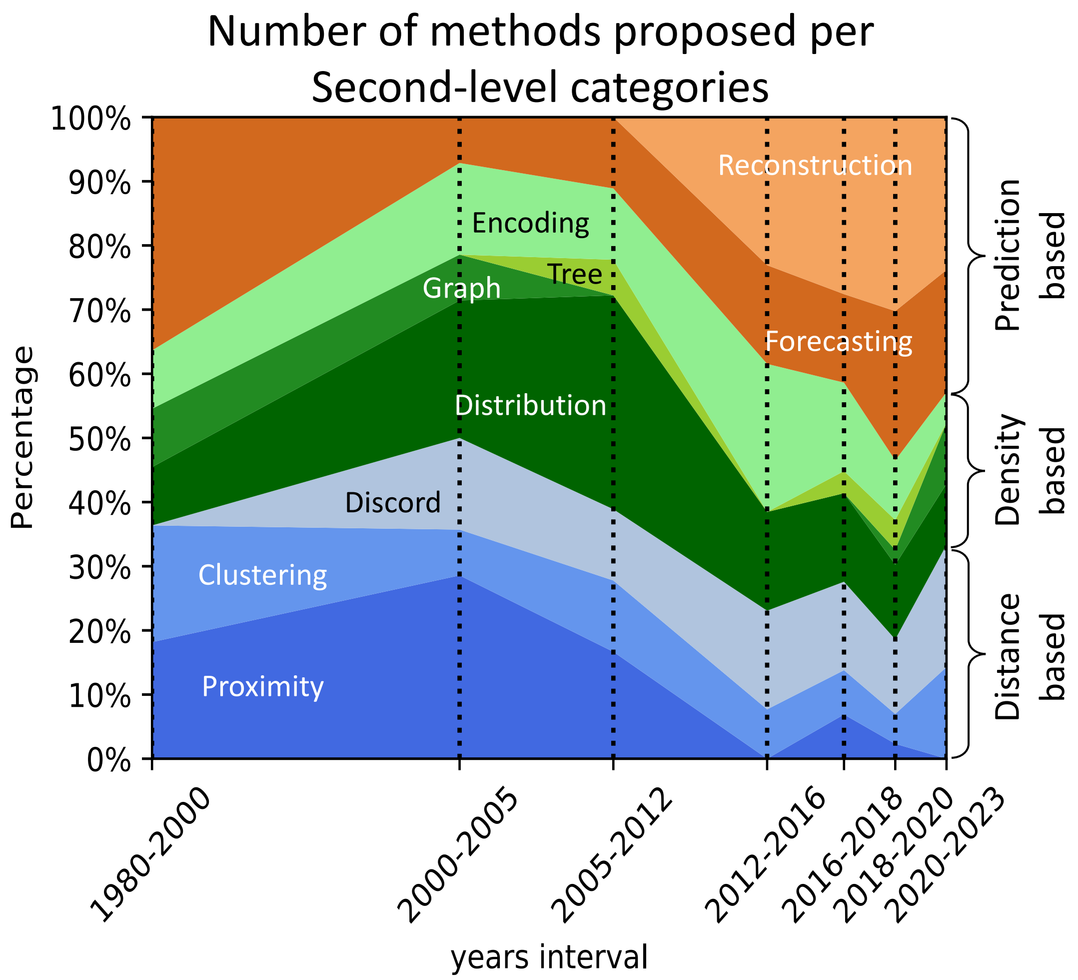
Anomaly Detection methods:
A taxonomy

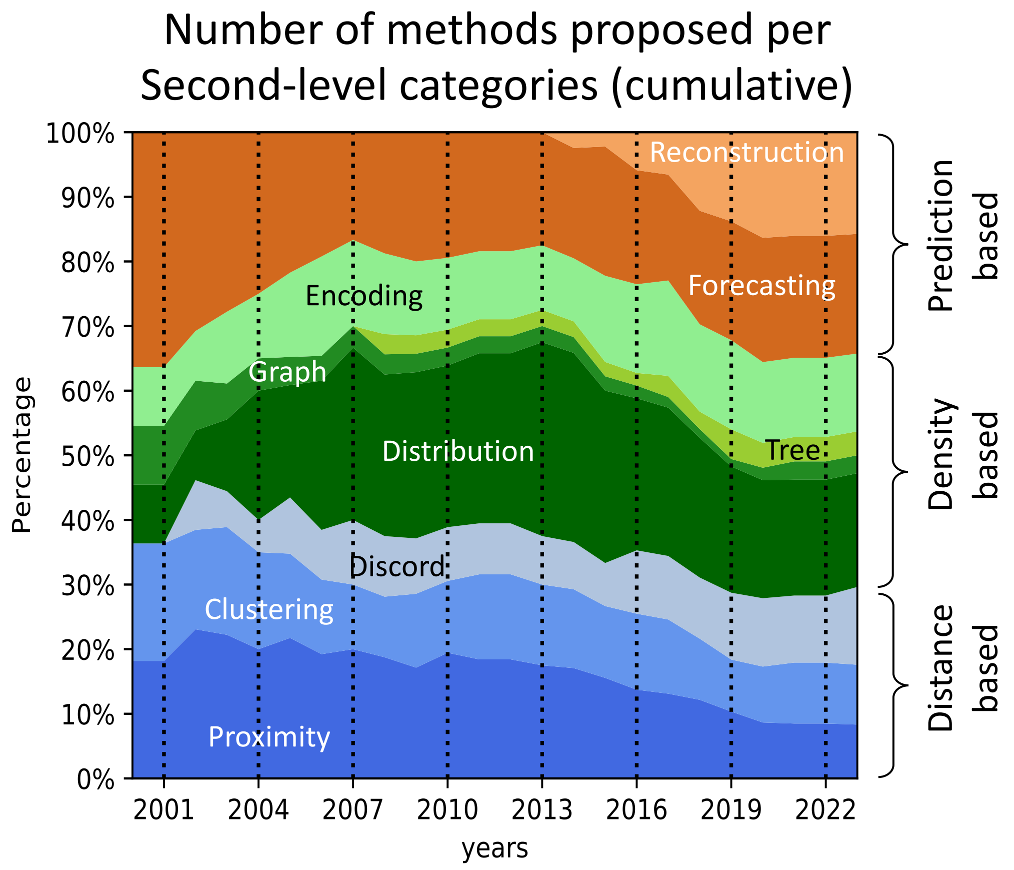
By time…



Anomaly Detection methods: *A taxonomy*

By time…

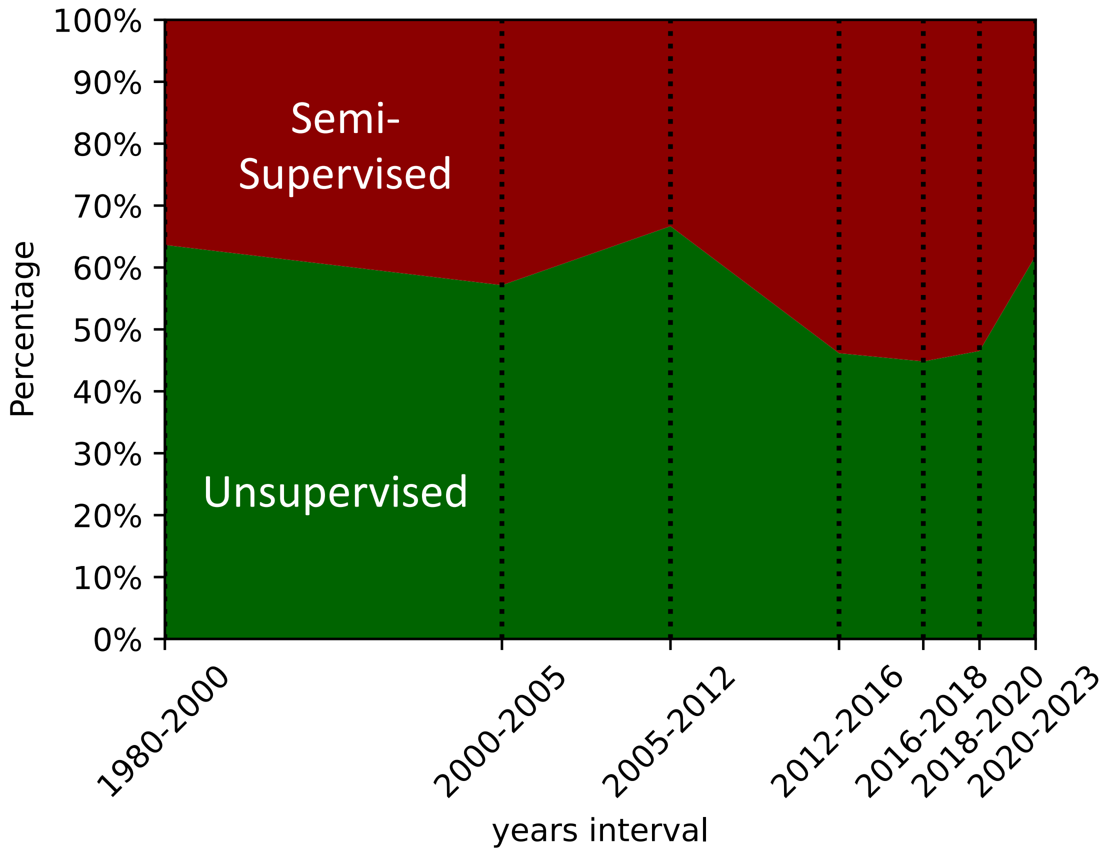




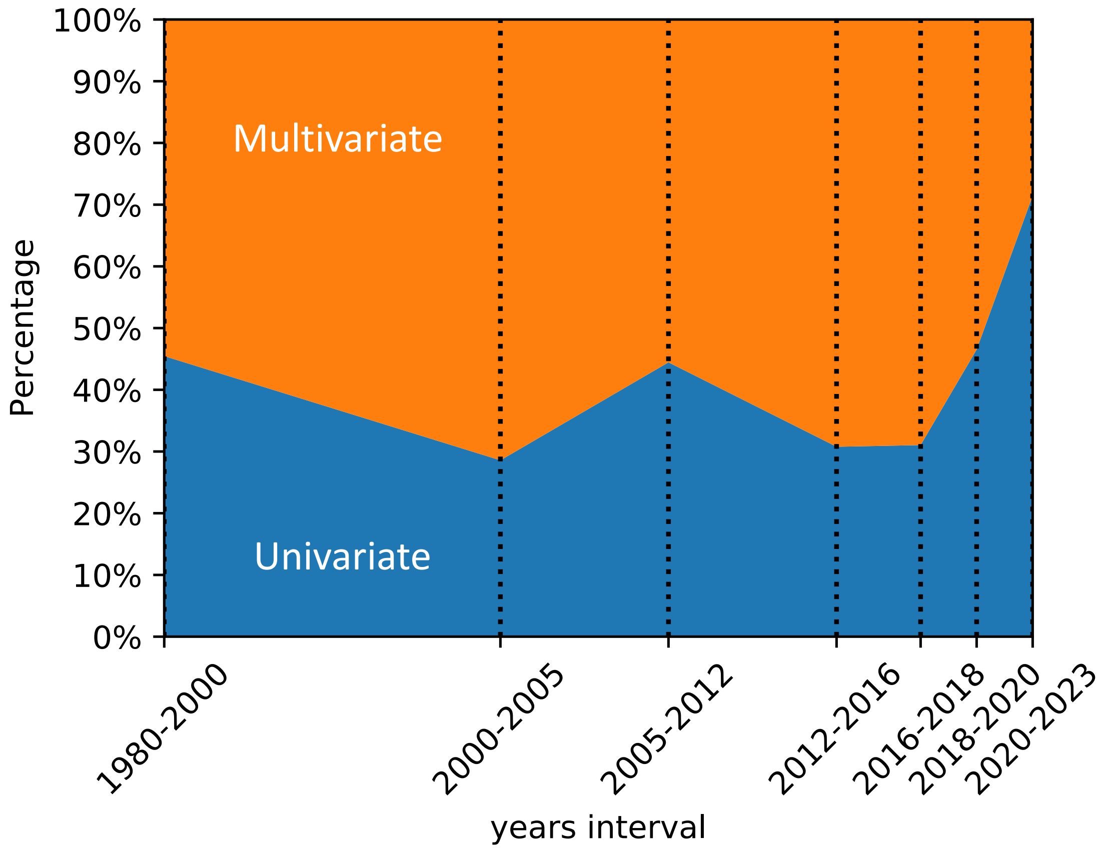
Anomaly Detection methods: *A taxonomy*

By time…

Number of methods proposed that are
Unsupervised or *Semi-Supervised*

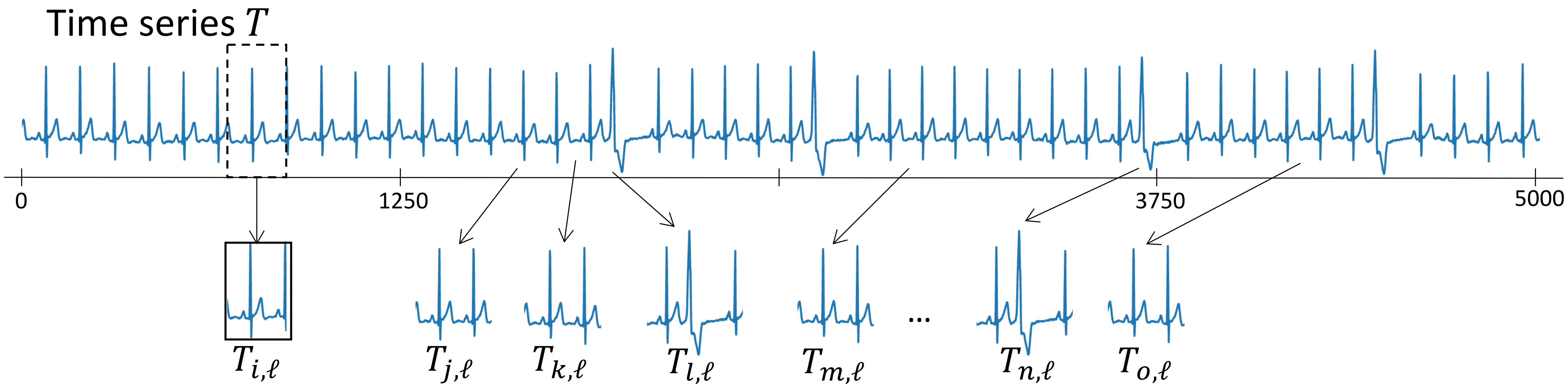


Number of methods proposed that can handle
Univariate or *Multivariate* time series



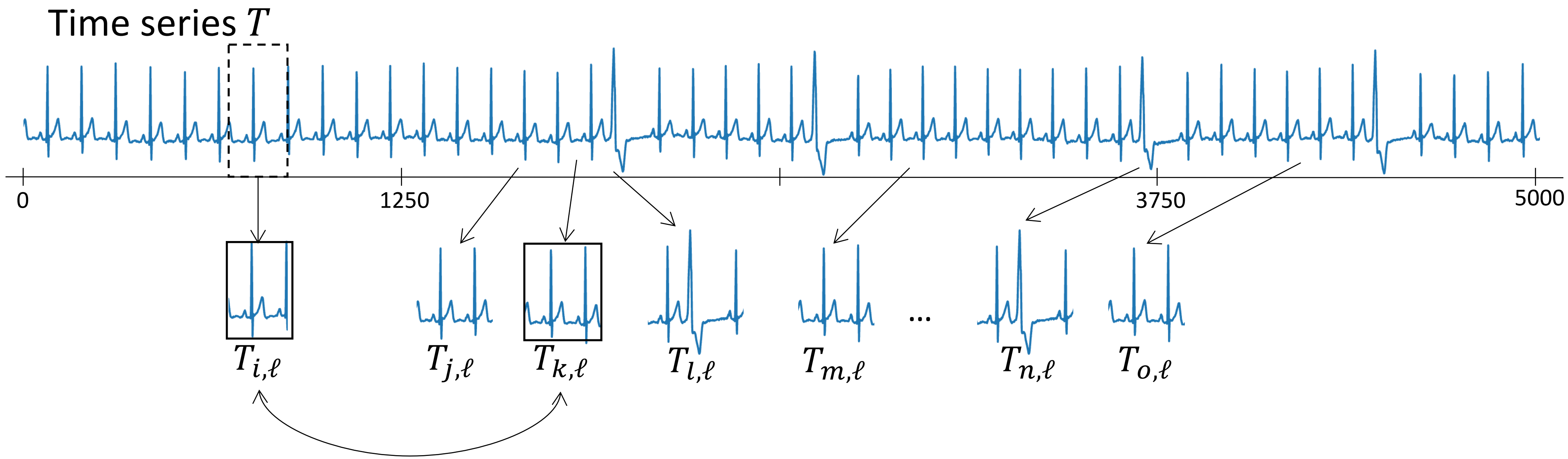
Anomaly Detection methods: *Distance-based*

Methods that use distance computation between subsequences (or group of subsequences) to detect anomalies.



Anomaly Detection methods: *Distance-based*

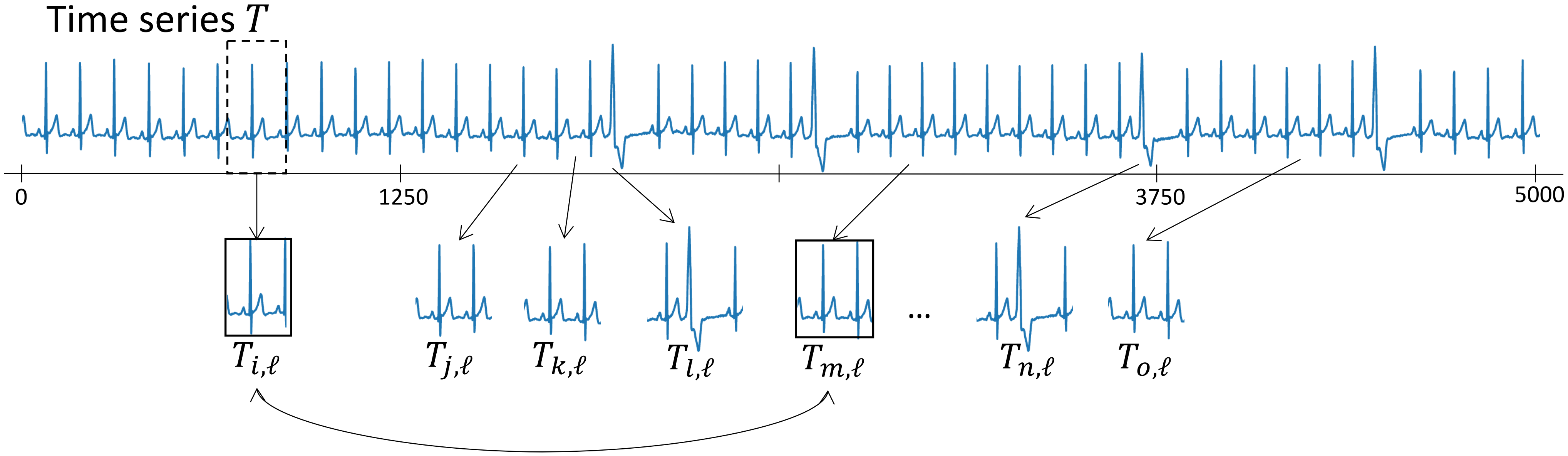
Methods that use distance computation between subsequences (or group of subsequences) to detect anomalies.



$S_{T,i} = d(T_{i,\ell}, T_{j,\ell})$

Anomaly Detection methods: *Distance-based*

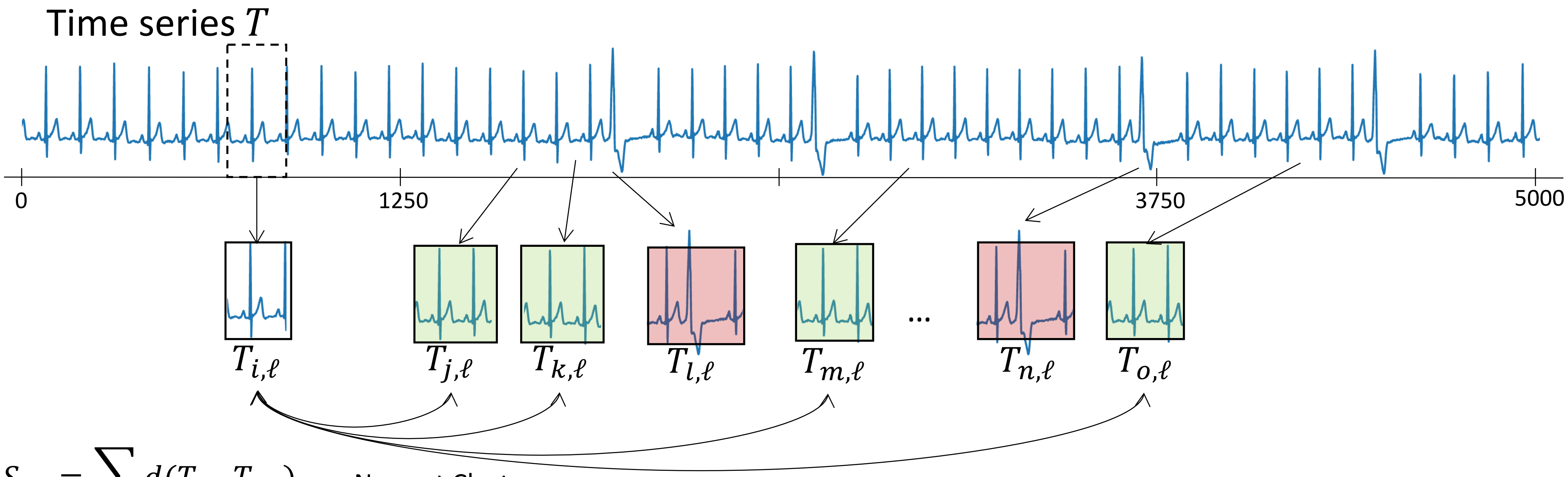
Methods that use distance computation between subsequences (or group of subsequences) to detect anomalies.



$$S_{T,i} = d(T_{i,\ell}, T_{m,\ell})$$

Anomaly Detection methods: *Distance-based*

Methods that use distance computation between subsequences (or group of subsequences) to detect anomalies.

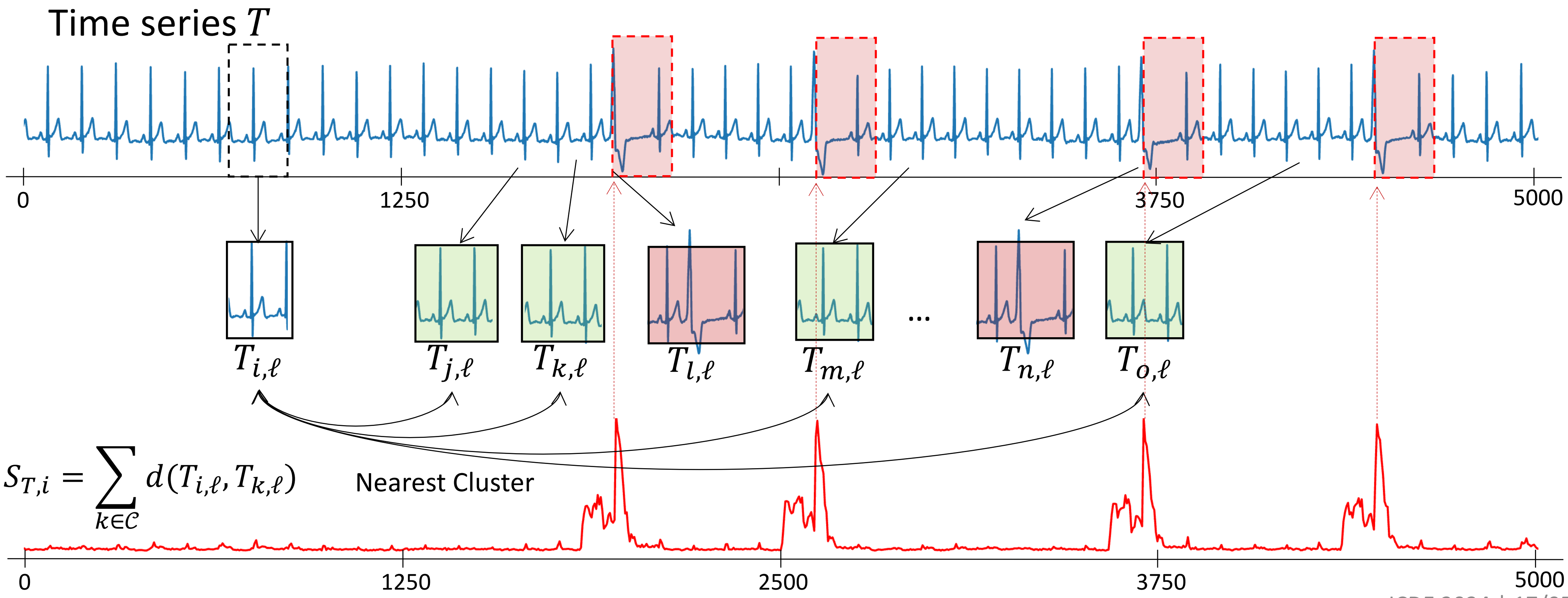


$$S_{T,i} = \sum_{k \in \mathcal{C}} d(T_{i,\ell}, T_{k,\ell})$$

Nearest Cluster

Anomaly Detection methods: *Distance-based*

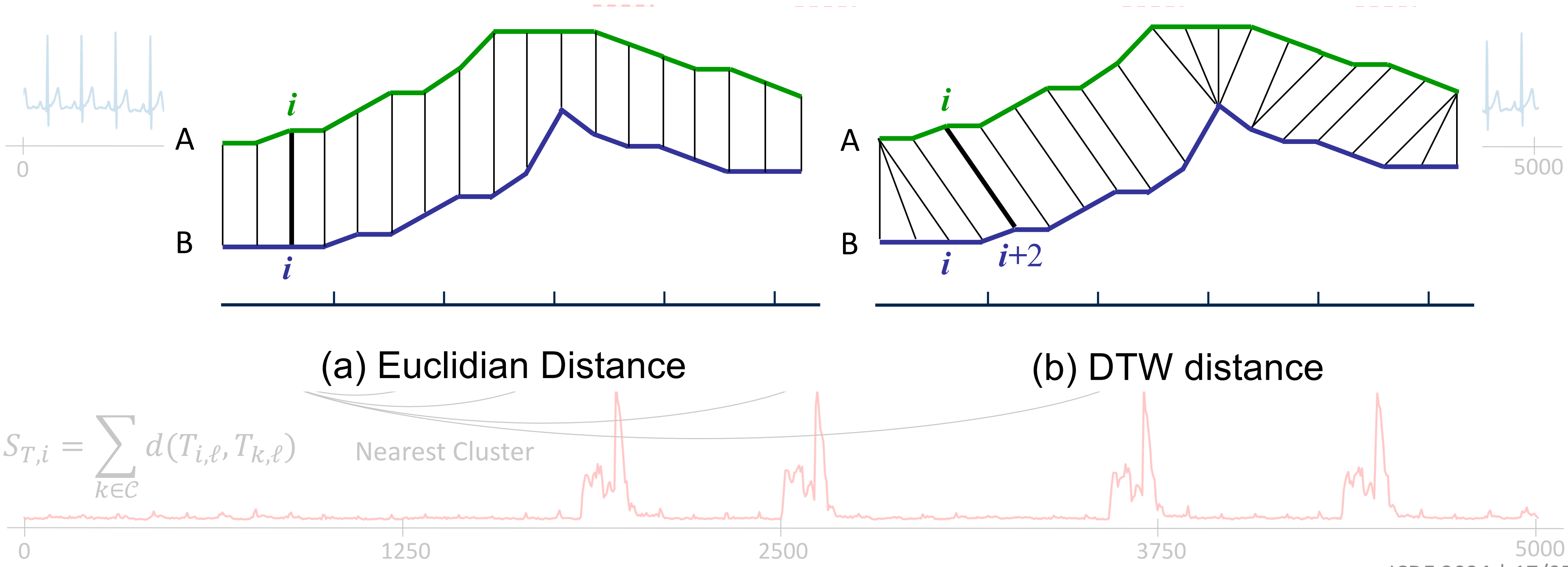
Methods that use distance computation between subsequences (or group of subsequences) to detect anomalies.



Anomaly Detection methods: *Distance-based*

Methods that use distance computation between subsequences (or group of subsequences) to detect anomalies.

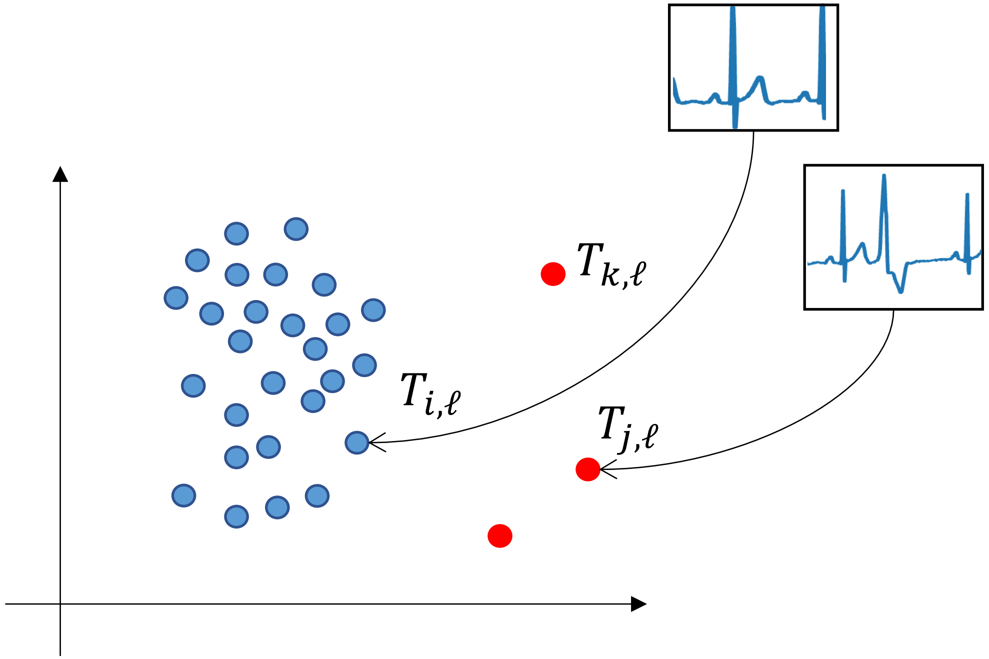
Example of distance computation



(a) Euclidian Distance

(b) DTW distance

Anomaly Detection methods: *an Example*



Matrix Profile [6] (MP)

Compute the distance to the nearest neighbor (using the MASS algorithm z-norm Euclidean distance computation) and use it as anomaly score

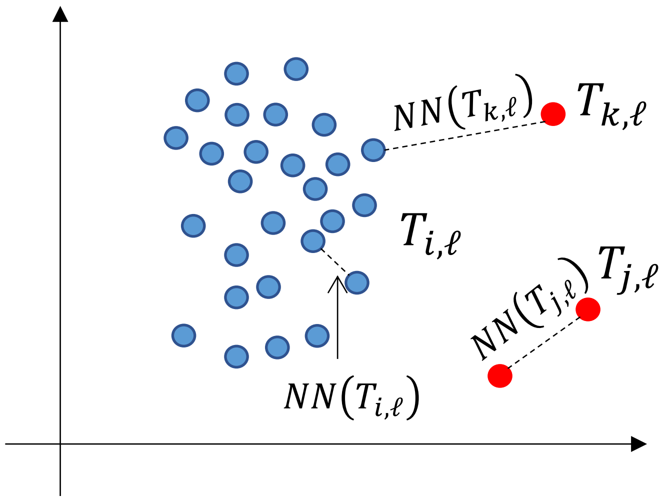
Unsupervised

Univariate

sequence

[6] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn J. Keogh. 2016. Matrix Prole I: All Pairs Similarity Joins for Time Series. In ICDM.

Anomaly Detection methods: *an Example*



The matrix Profile is computed as follows:

$$S_T = [NN(T_{0,\ell}), NN(T_{1,\ell}), \dots, NN(T_{|T|-\ell,\ell})]$$

Matrix Profile [6] (MP)

Compute the distance to the nearest neighbor (using the MASS algorithm z-norm Euclidean distance computation) and use it as anomaly score

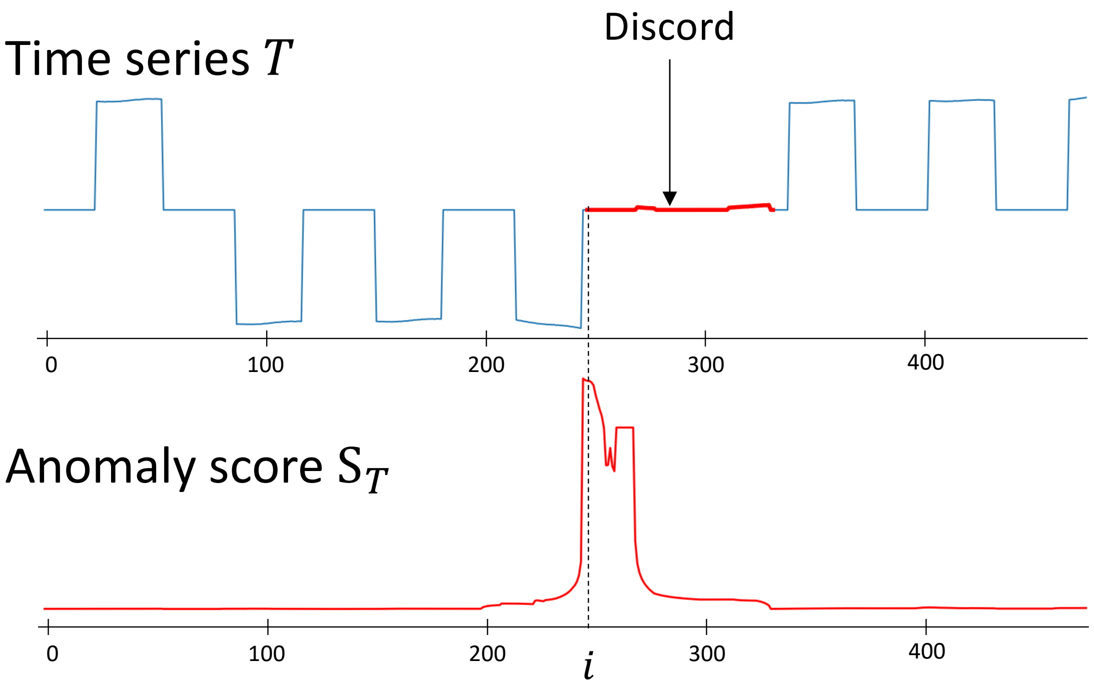
Unsupervised

Univariate

sequence

[6] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn J. Keogh. 2016. Matrix Prole I: All Pairs Similarity Joins for Time Series. In ICDM.

Anomaly Detection methods: *an Example*



Matrix Profile [6] (MP)

Compute the distance to the nearest neighbor (using the MASS algorithm z-norm Euclidean distance computation) and use it as anomaly score

Unsupervised

Univariate

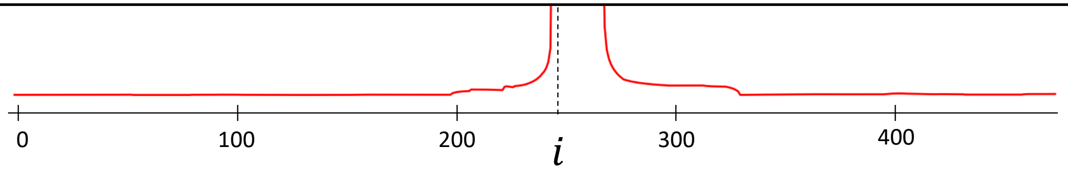
sequence

[6] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn J. Keogh. 2016. Matrix Prole I: All Pairs Similarity Joins for Time Series. In ICDM.

Anomaly Detection methods: *an Example*

Many different extensions…

- For streaming time series: **STAMPi** [6], **DAMP** [8]
- For similar recurrent anomalies: **left-STAMP** [6]
- Anytime or ordered: **STAMP** [6], **STOMP** [7]
- For multivariate time series: **mSTAMP** [9]



Matrix Profile [6] (MP)

Compute the distance to the nearest neighbor (using the MASS algorithm z-norm Euclidean distance computation) and use it as anomaly score

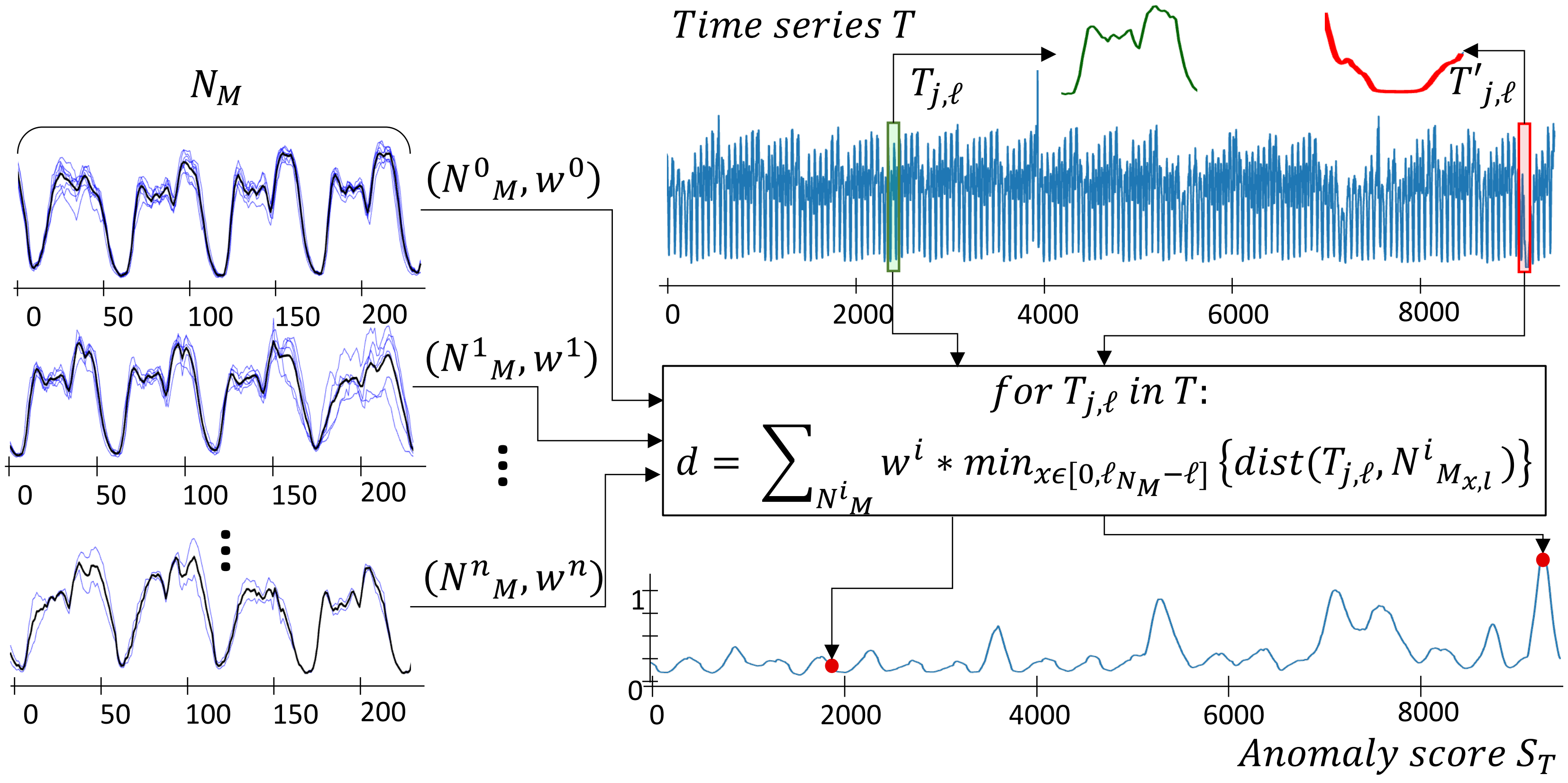
Unsupervised

Univariate

sequence

[6] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn J. Keogh. 2016. Matrix Prole I: All Pairs Similarity Joins for Time Series. In ICDM.

Anomaly Detection methods: *an Example*



NormA [10]

Distance-based approach that summarize the time series into a weighted set of subsequences and use the distance to them as anomaly score

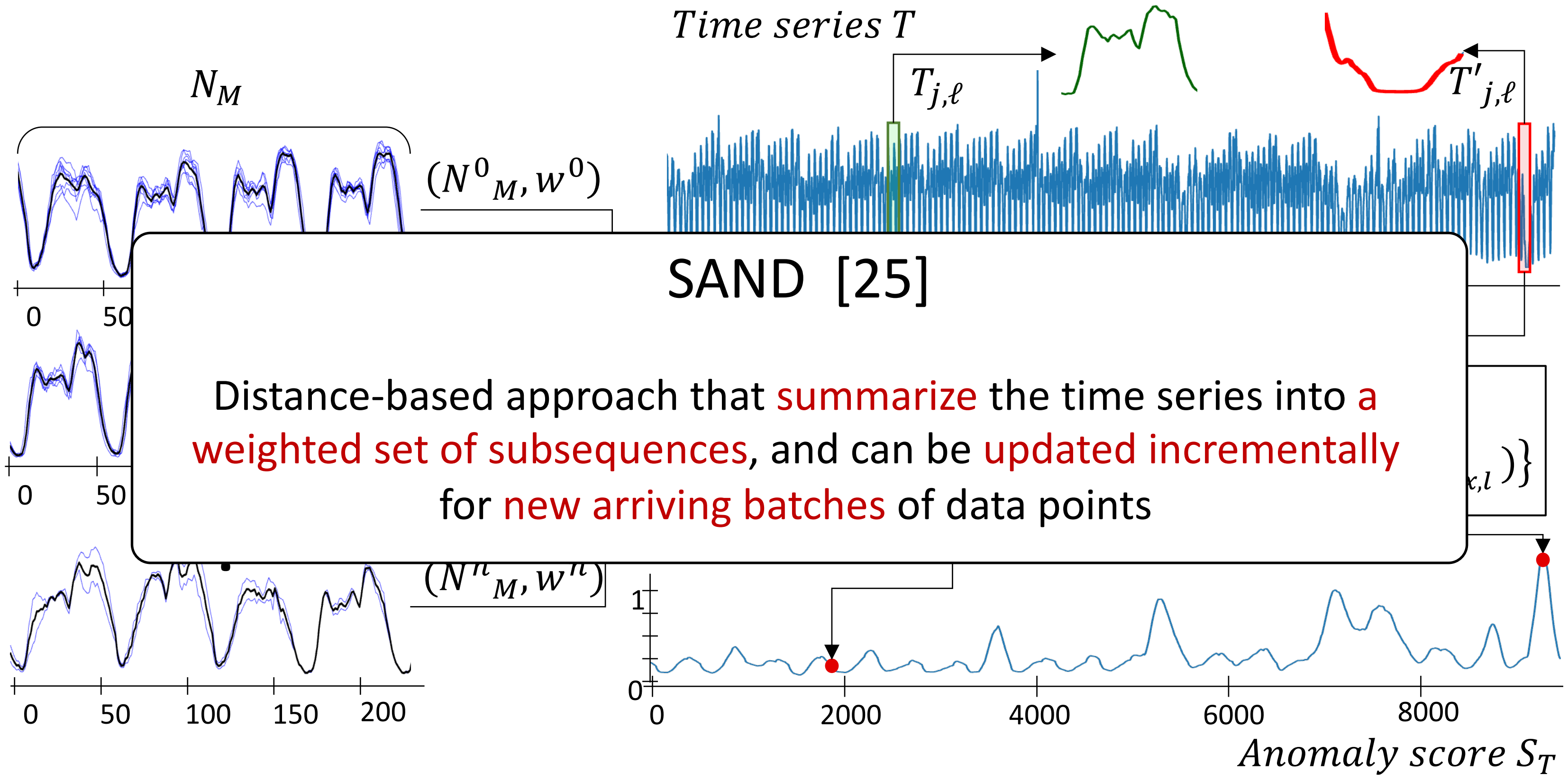
Unsupervised

Univariate

sequence

[10] Paul Boniol, Michele Linardi, Federico Roncallo, Themis Palpanas, Mohammed Meftah, and Emmanuel Remy. 2021. Unsupervised and scalable subsequence anomaly detection in large data series. The VLDB Journal 30, 6 (Nov 2021), 909–931.

Anomaly Detection methods: *an Example*



SAND [25]

Distance-based approach that summarize the time series into a weighted set of subsequences, and can be updated incrementally for new arriving batches of data points

NormA [10]

Distance-based approach that summarize the time series into a weighted set of subsequences and use the distance to them as anomaly score

Unsupervised

Univariate

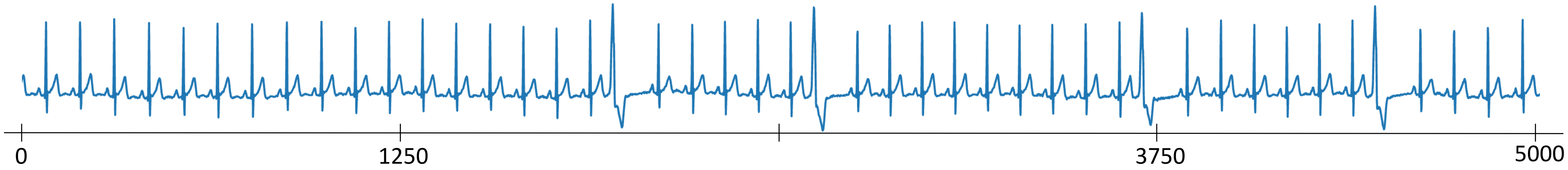
sequence

[25] Paul Boniol, John Paparrizos, Themis Palpanas, and Michael J. Franklin. 2021. SAND: streaming subsequence anomaly detection. Proc. VLDB Endow. 14, 10 (June 2021), 1717–1729.

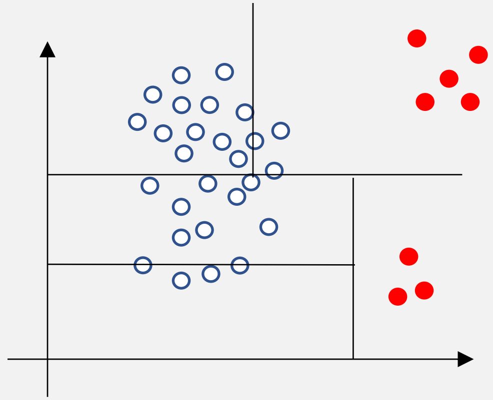
Anomaly Detection methods: *Density-based*

Methods that estimate the density of the space (points or subsequences) and identify as anomalies points (or sequences)that are in low-density subspace.

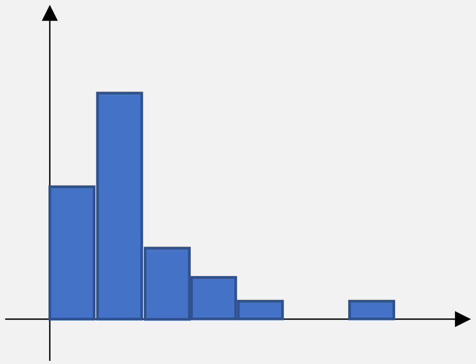
Time series T



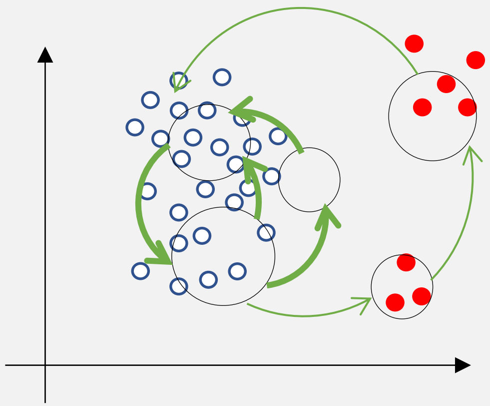
Tree-based approaches [11]



Distribution-based Approaches [12]



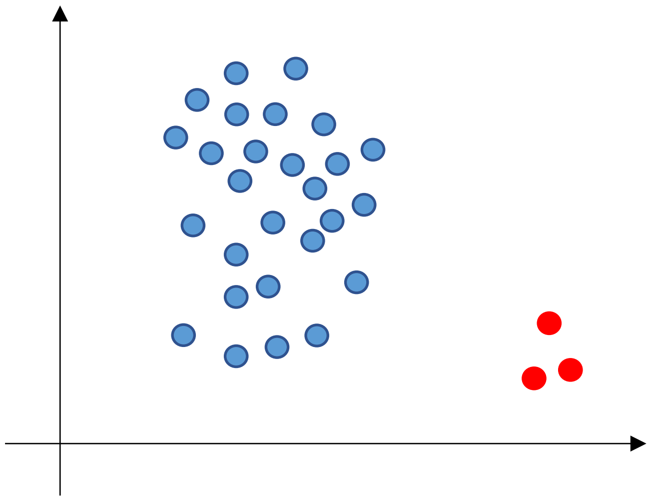
Graph-based approaches [13]



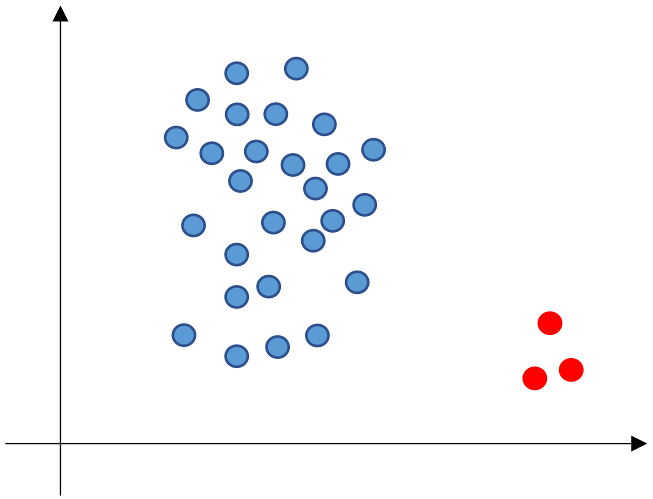
…

Anomaly Detection methods: *an Example*

0 splits



0 splits



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

Unsupervised

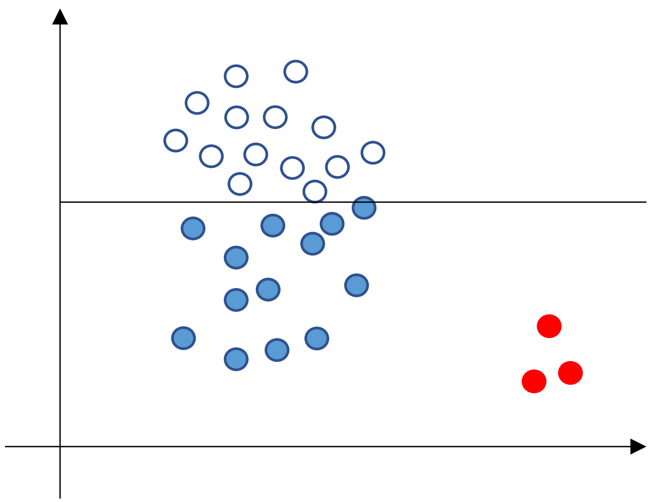
Univariate/Multivariate

Point/sequence

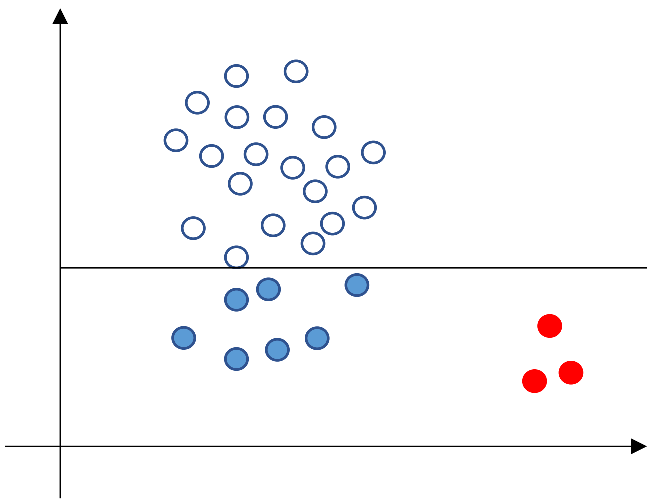
[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*

1 splits



1 splits



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

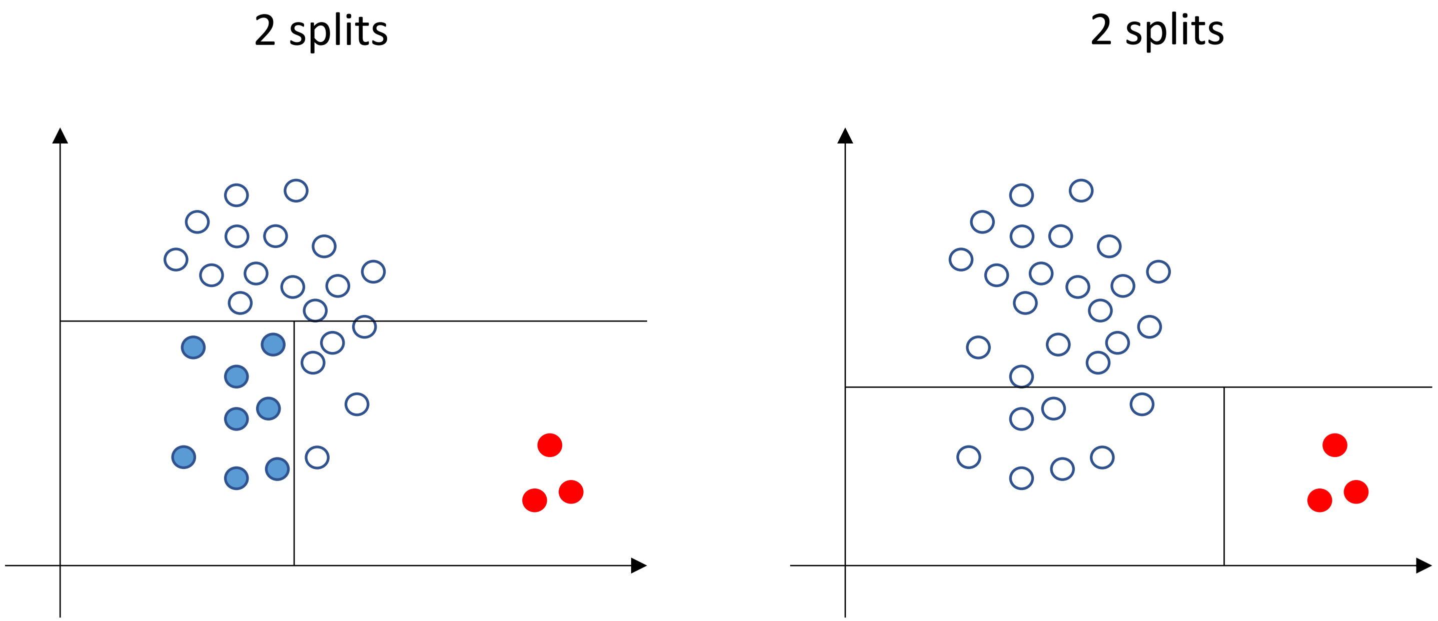
Unsupervised

Univariate/Multivariate

Point/sequence

[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

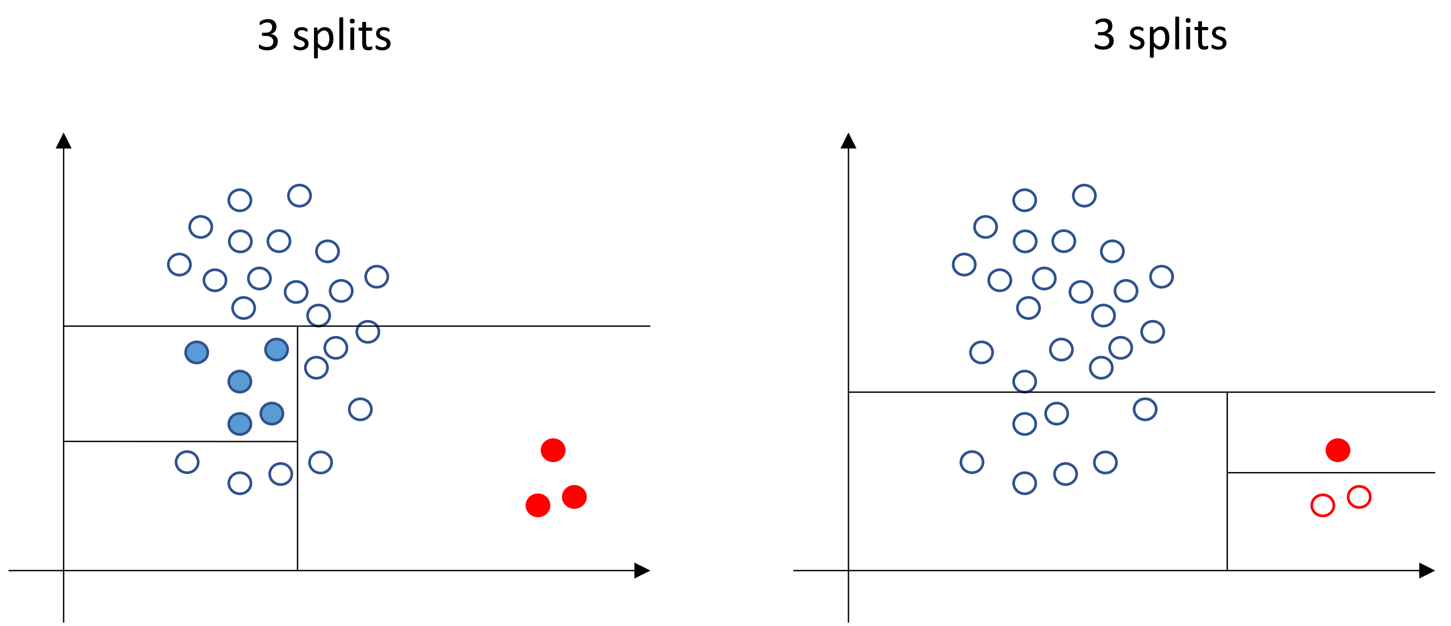
Unsupervised

Univariate/Multivariate

Point/sequence

[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

Unsupervised

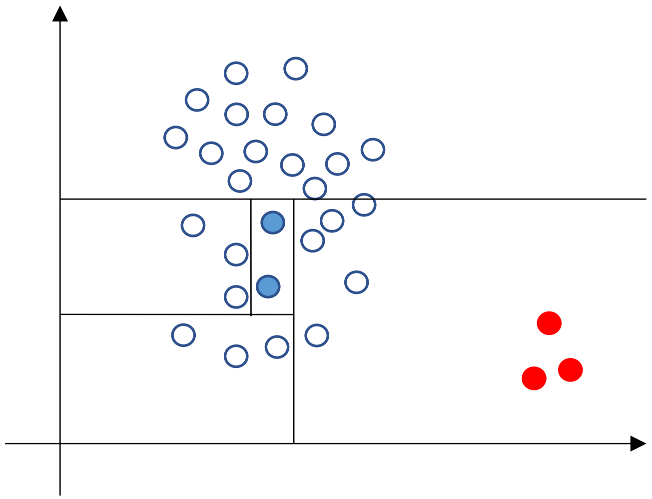
Univariate/Multivariate

Point/sequence

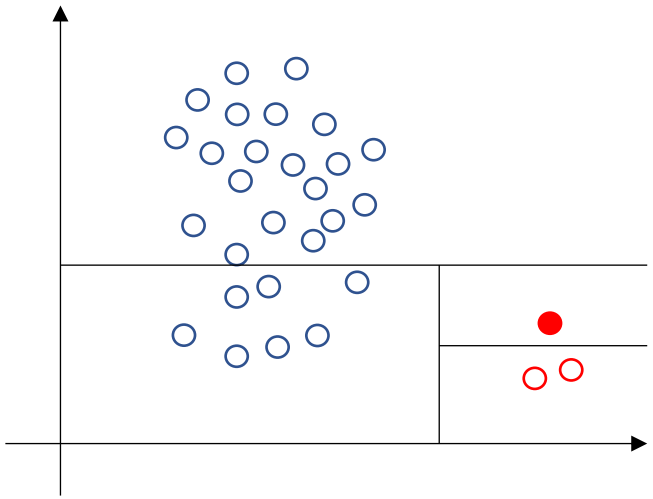
[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*

4 splits



3 splits



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

Unsupervised

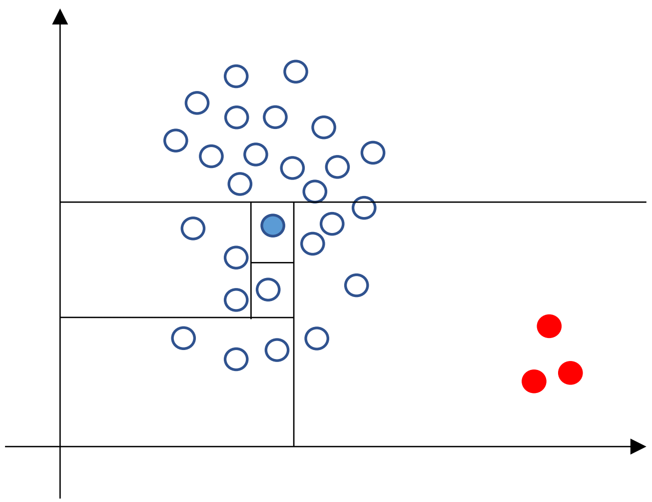
Univariate/Multivariate

Point/sequence

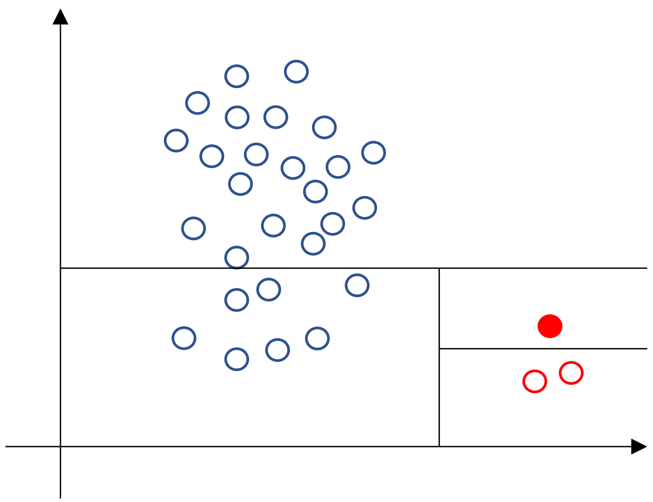
[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*

5 splits



3 splits



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

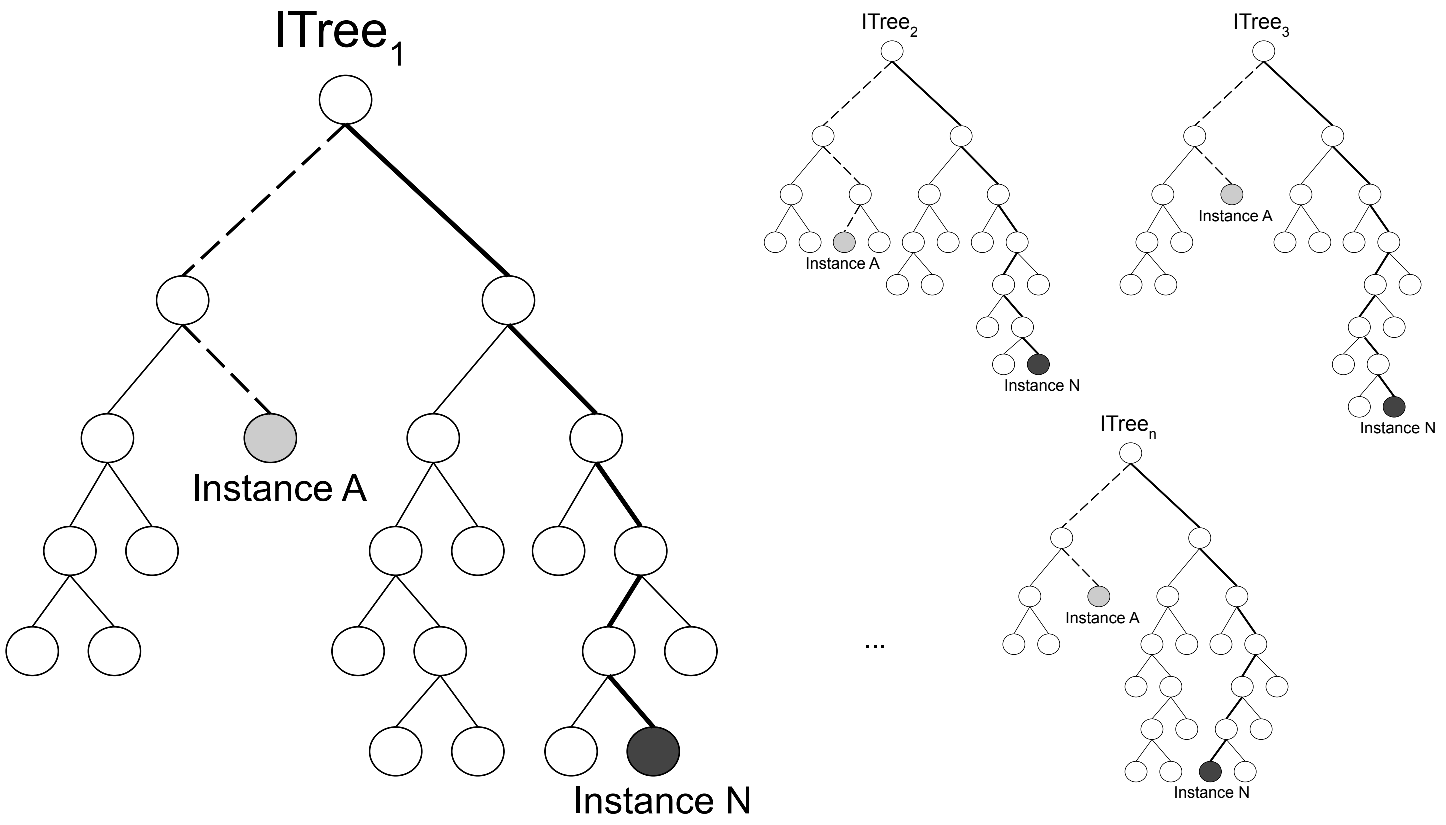
Unsupervised

Univariate/Multivariate

Point/sequence

[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*



Isolation Forest [11]

Density-based approach that split the space randomly and using the depth of the trees to identify anomalies

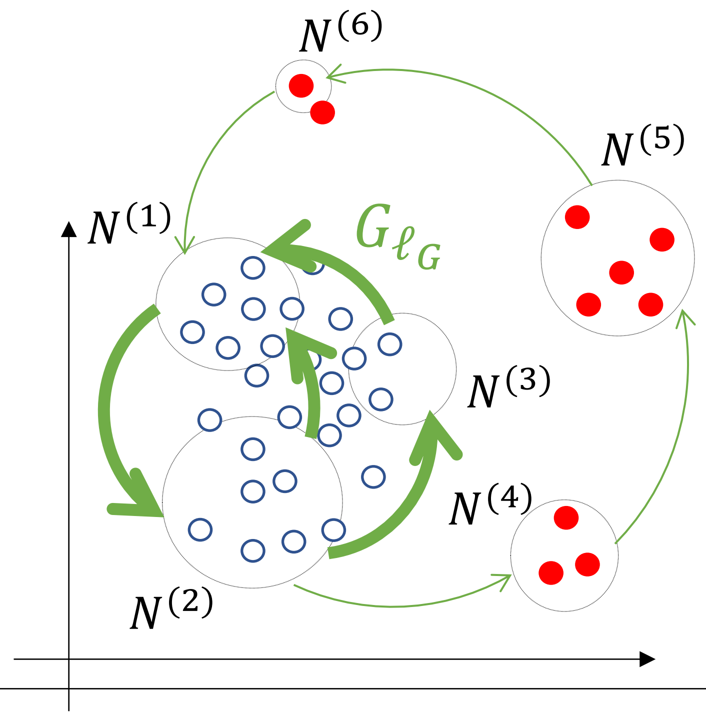
Unsupervised

Univariate/Multivariate

Point/sequence

[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, "Isolation Forest," 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

Anomaly Detection methods: *an Example*



Each **node** is an ensemble of similar subsequences.

Each **edge** is associated to a weight w that corresponds to the number of times a subsequence move from one node to another.

For a given subsequence $T_{i,\ell}$ and its corresponding path $P_{th} =< N^{(i)}, N^{(i+1)}, \dots, N^{(i+\ell)} >$, we define the normality score as follows:

$$Norm(P_{th}) = \sum_{j=i}^{i+\ell-1} \frac{w\left(N^{(j)}, N^{(j+1)}\right) \deg(N^{(j)} - 1)}{\ell}$$

Series2Graph [13]

Density-based approach that convert the time series into a graph and detect unusual trajectories

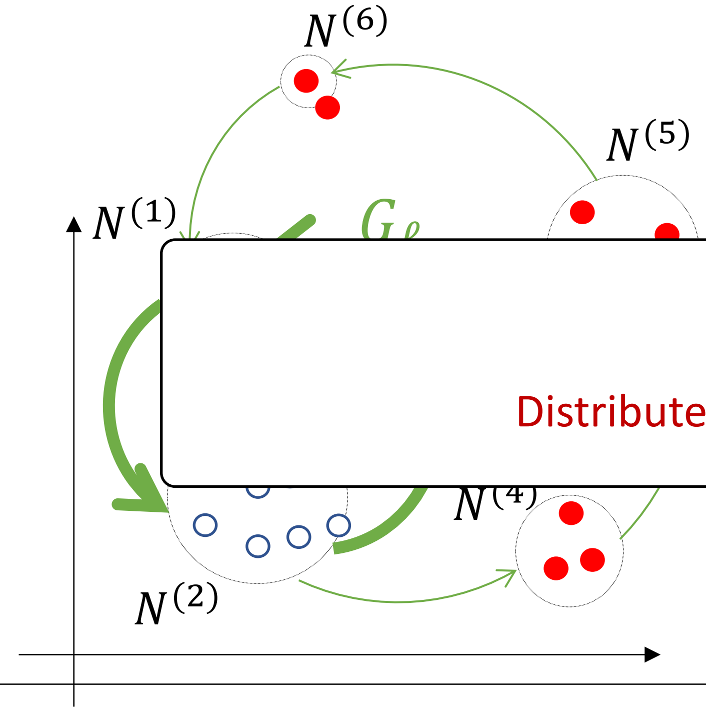
Unsupervised

Univariate

subsequence

[13] Paul Boniol and Themis Palpanas. 2020. Series2Graph: graph-based subsequence anomaly detection for time series. Proc. VLDB Endow. 13, 12 (August 2020), 1821–1834

Anomaly Detection methods: *an Example*



Each **node** is an ensemble of similar subsequences.

DADS [26]

Distributed version of Series2Graph

For a given subsequence $T_{i,\ell}$ and its corresponding path $P_{th} =< N^{(i)}, N^{(i+1)}, \ldots, N^{(i+\ell)} >$, we define the normality score as follows:

$$Norm(P_{th}) = \sum_{j=i}^{i+\ell-1} \frac{w\left(N^{(j)}, N^{(j+1)}\right) \deg(N^{(j)} - 1)}{\ell}$$

Series2Graph [13]

Density-based approach that convert the time series into a graph and detect unusual trajectories

Unsupervised

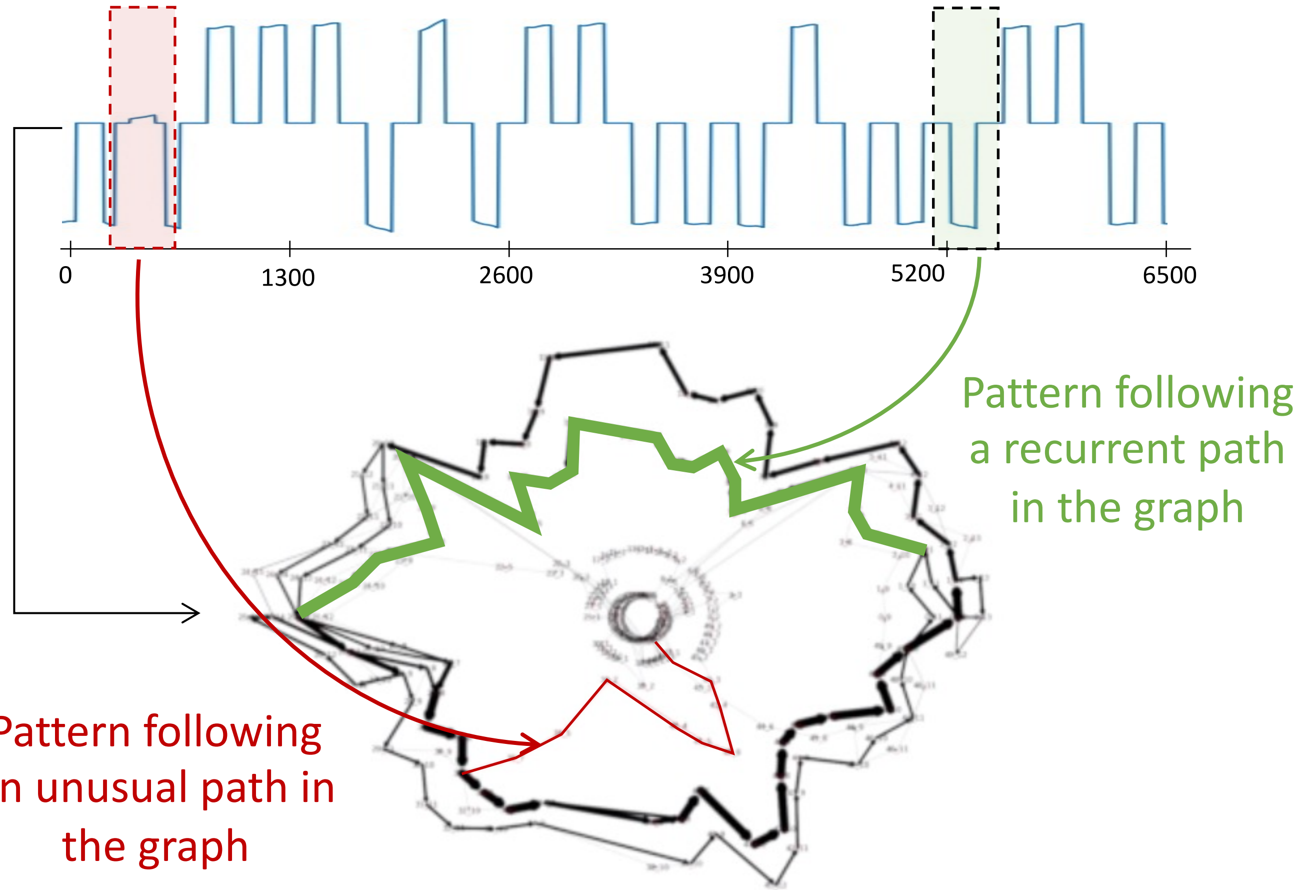
Univariate

subsequence

[26] Schneider, J., Wenig, P. & Papenbrock, T. Distributed detection of sequential anomalies in univariate time series. The VLDB Journal **30**, 579–602 (2021).

Anomaly Detection methods: *an Example*

Snippet of SED time series [14]



Series2Graph [13]

Density-based approach that convert the time series into a graph and detect unusual trajectories

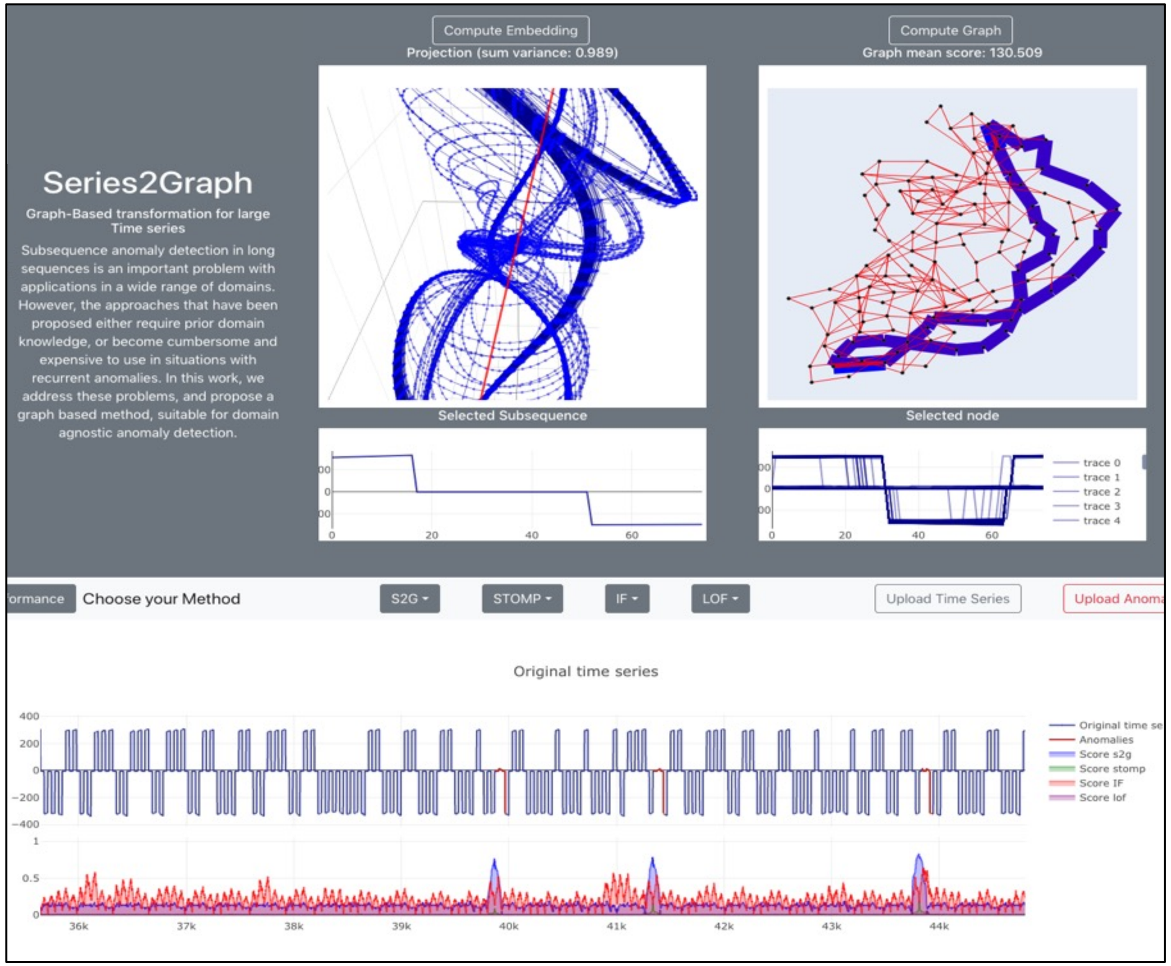
Unsupervised

Univariate

subsequence

[13] Paul Boniol and Themis Palpanas. 2020. Series2Graph: graph-based subsequence anomaly detection for time series. Proc. VLDB Endow. 13, 12 (August 2020), 1821–1834

Anomaly Detection methods: *an Example*



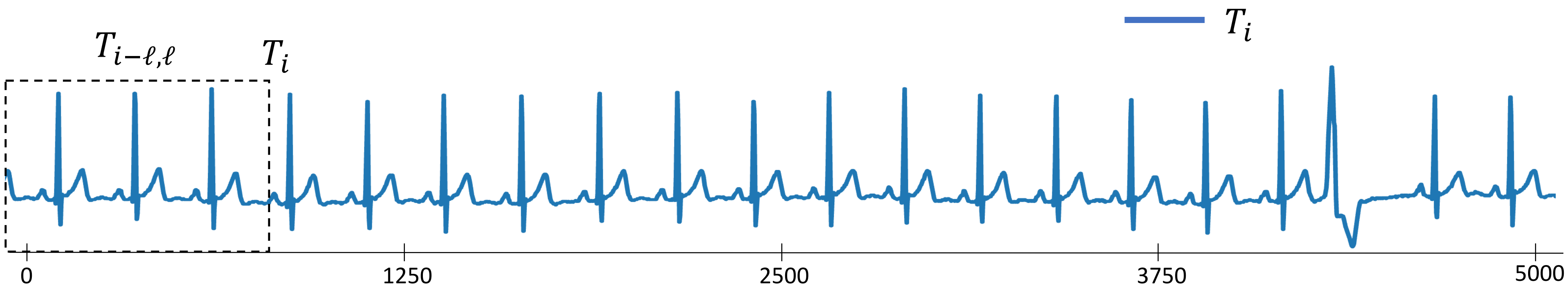
GraphAn [28]

An interactive tool to dive into the computation steps of Series2Graph :

[28] Paul Boniol, Themis Palpanas, Mohammed Meftah, and Emmanuel Remy. 2020. GraphAn: graph-based subsequence anomaly detection. Proc. VLDB Endow. 13, 12 (August 2020), 2941–2944.

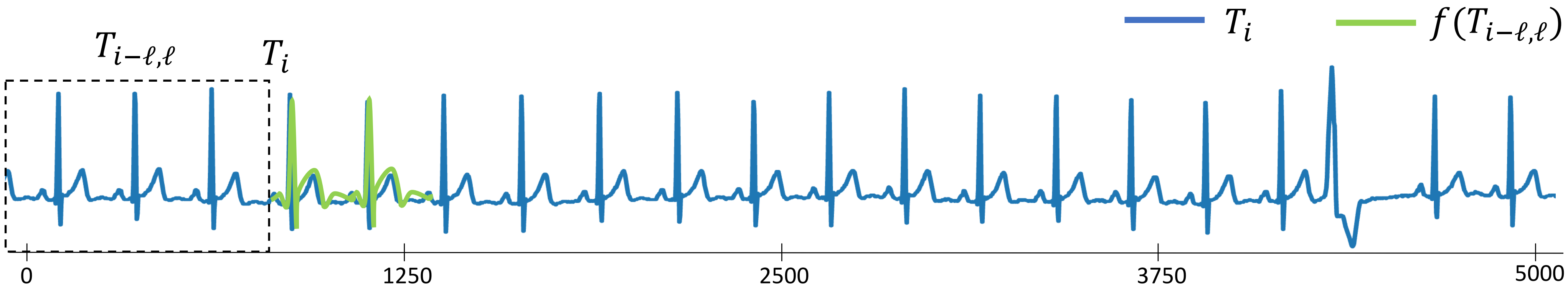
Anomaly Detection methods: *Forecasting-based*

Methods that aims to predict the next points based on the previous ones. The prediction error is used to detect if there is an anomaly or not.



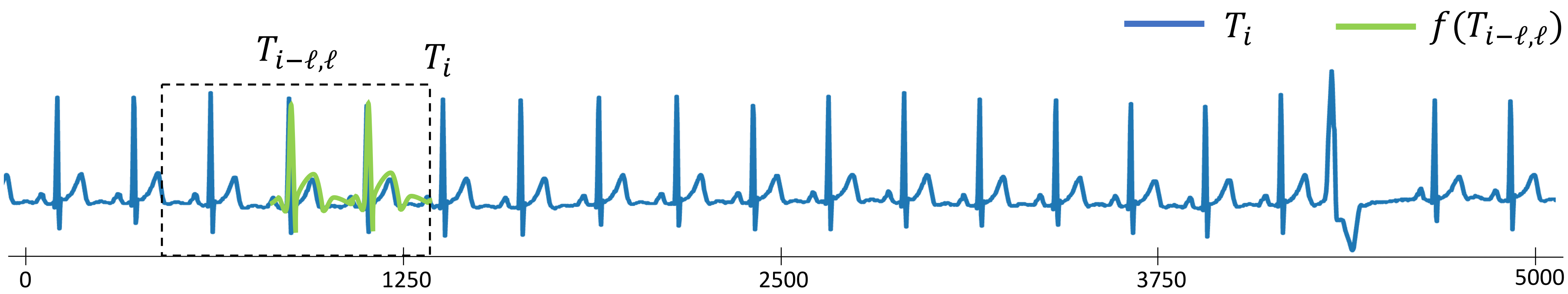
Anomaly Detection methods: *Forecasting-based*

Methods that aims to predict the next points based on the previous ones. The prediction error is used to detect if there is an anomaly or not.



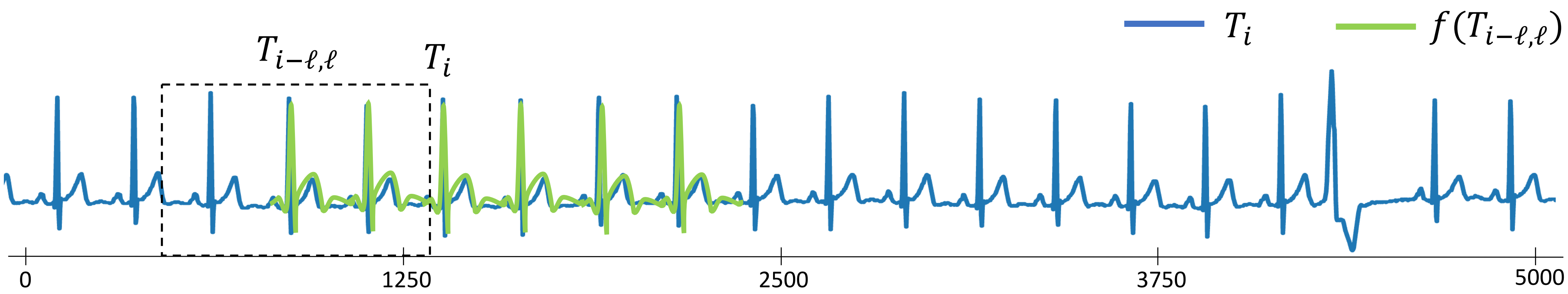
Anomaly Detection methods: *Forecasting-based*

Methods that aims to predict the next points based on the previous ones. The prediction error is used to detect if there is an anomaly or not.



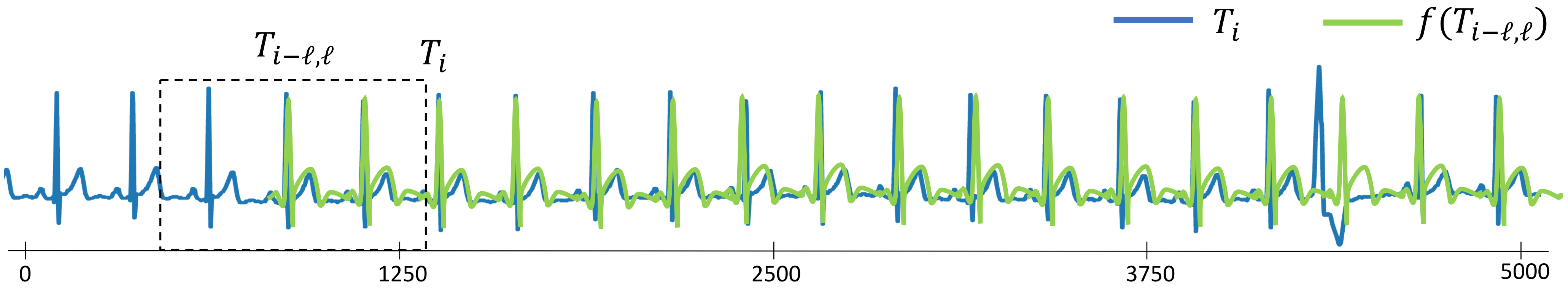
Anomaly Detection methods: *Forecasting-based*

Methods that aims to predict the next points based on the previous ones. The prediction error is used to detect if there is an anomaly or not.



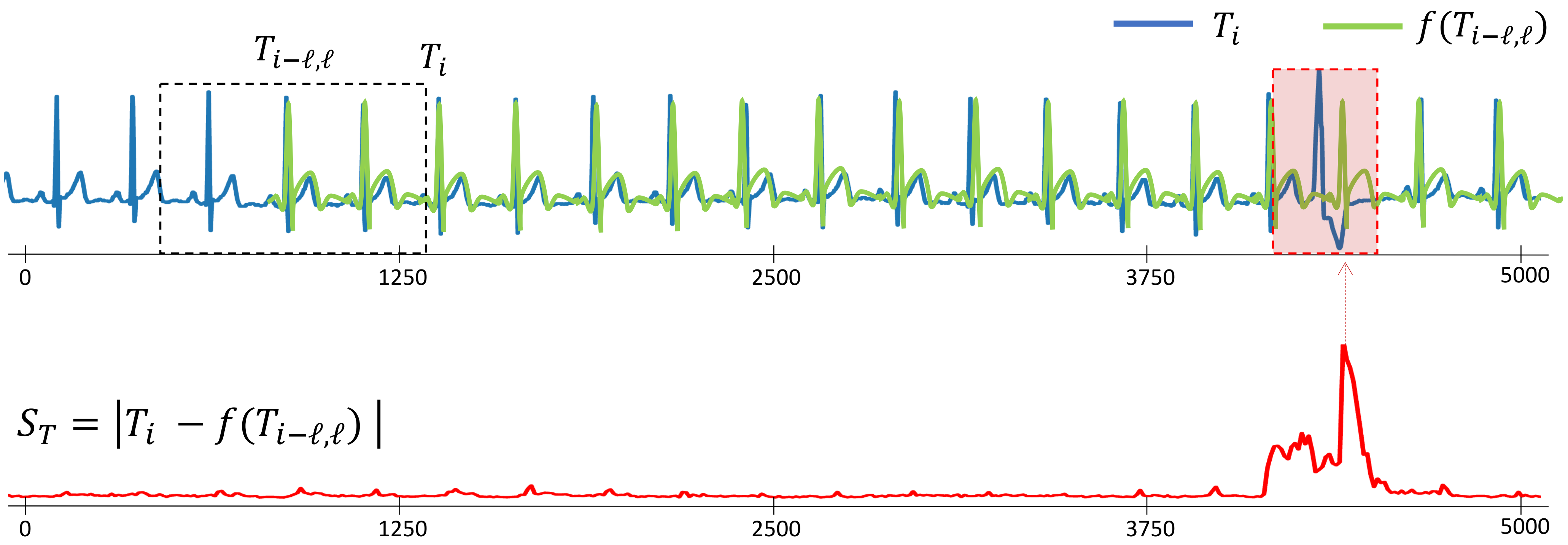
Anomaly Detection methods: *Forecasting-based*

Methods that aims to predict the next points based on the previous ones. The prediction error is used to detect if there is an anomaly or not.

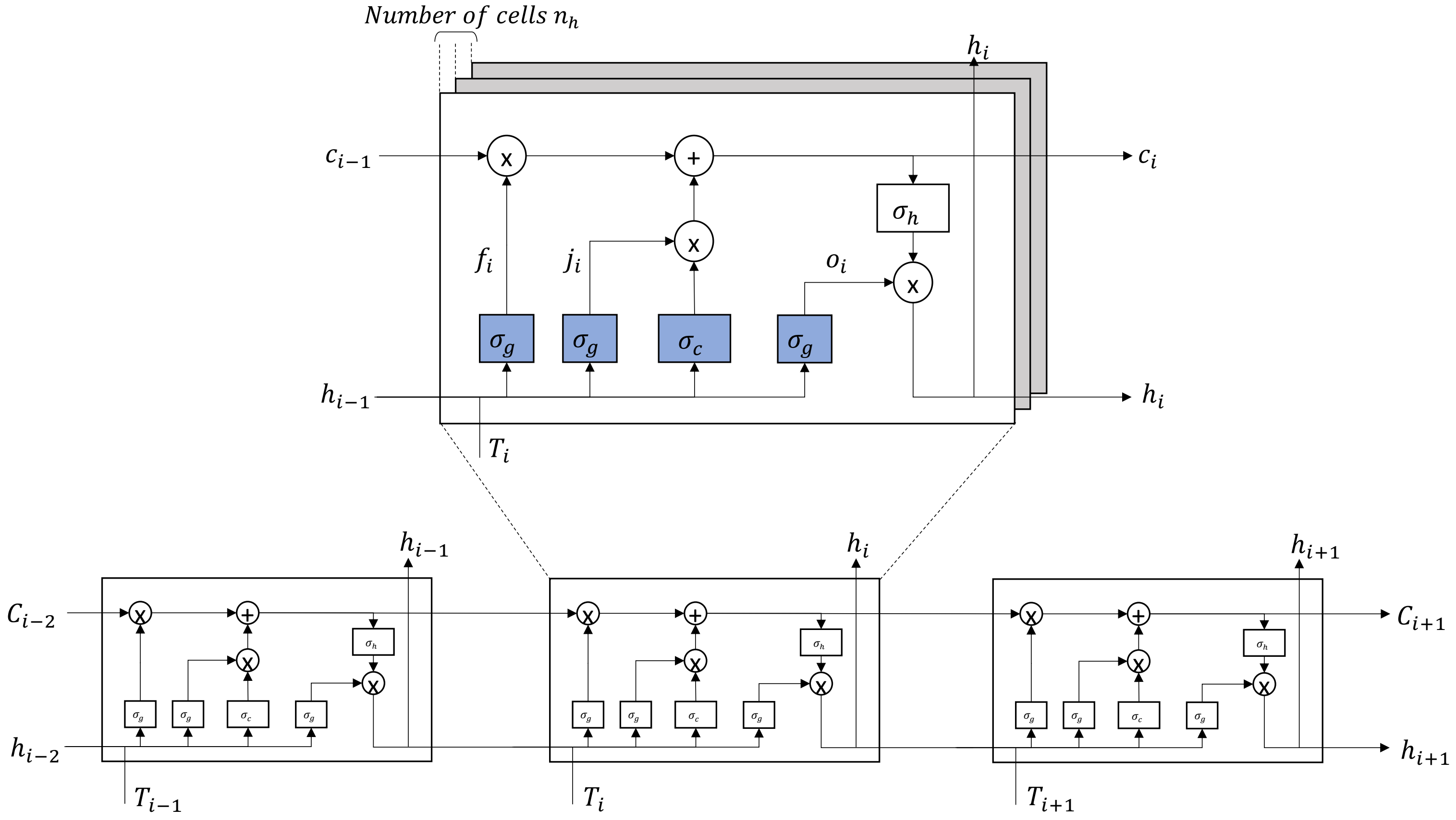


Anomaly Detection methods: *Forecasting-based*

Methods that aims to predict the next points based on the previous ones. The prediction error is used to detect if there is an anomaly or not.



Anomaly Detection methods: *an Example*



LSTM-AD [15]

Model that stack multiple LSTM cell and use the output to predict the next value

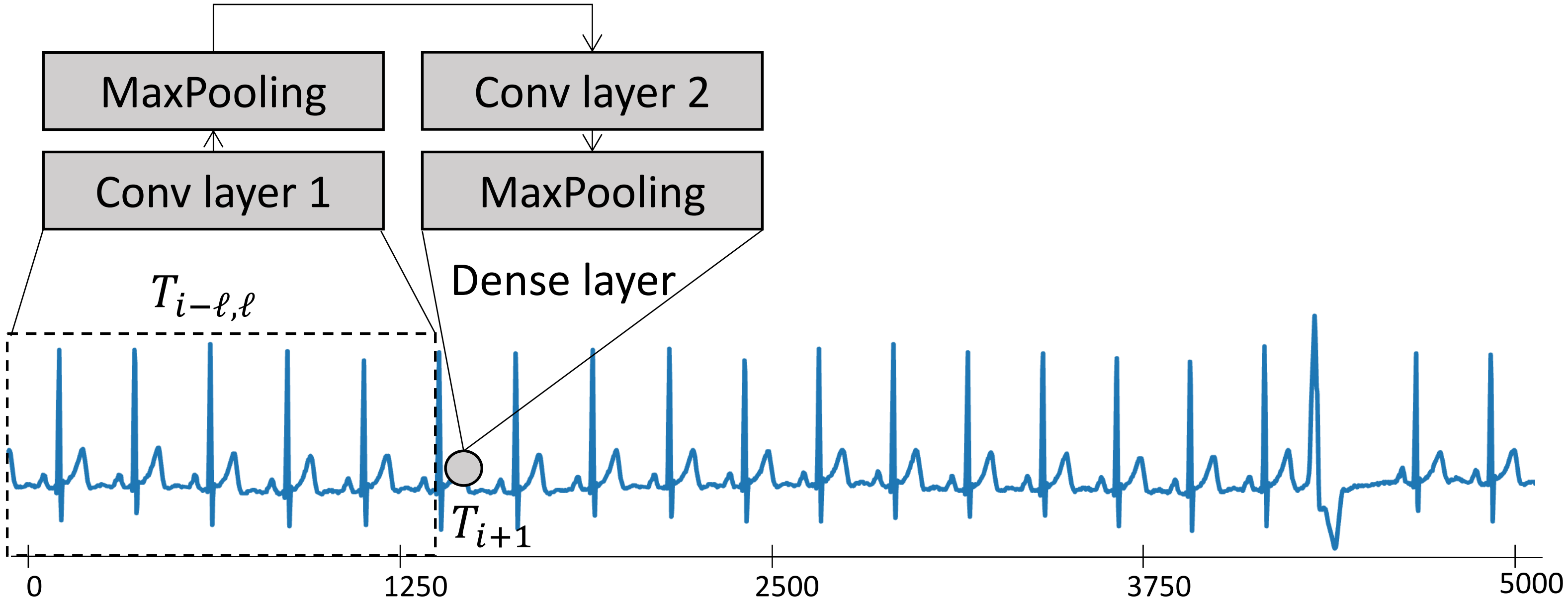
Semi-supervised

Univariate/Multivariate

Point/sequence

[15] Pankaj Malhotra, Lovekesh Vig, Gautam Shro, and Puneet Agarwal. 2015. Long Short Term Memory Networks for Anomaly Detection in Time Series. (2015).

Anomaly Detection methods: *an Example*



DeepAnT [16] (CNN)

Convolutional-based approach (2 convolutional layers) taking as input a sequence and aims to predict the next value.

Semi-supervised

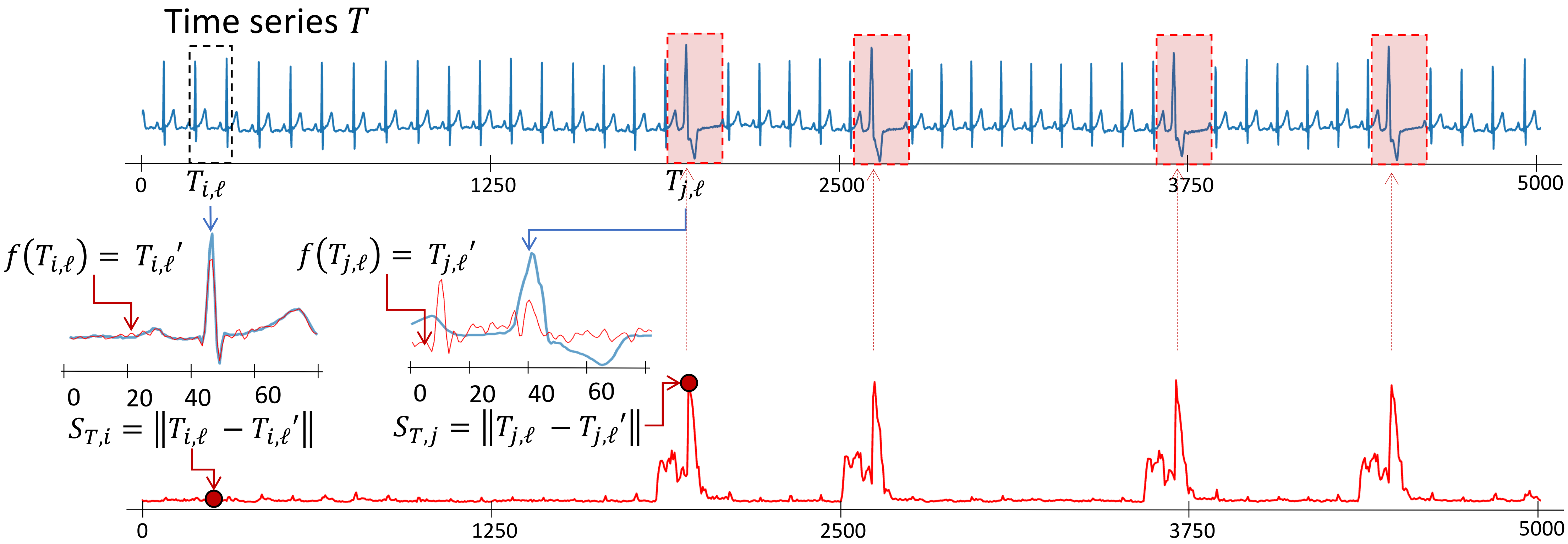
Univariate/Multivariate

Point/sequence

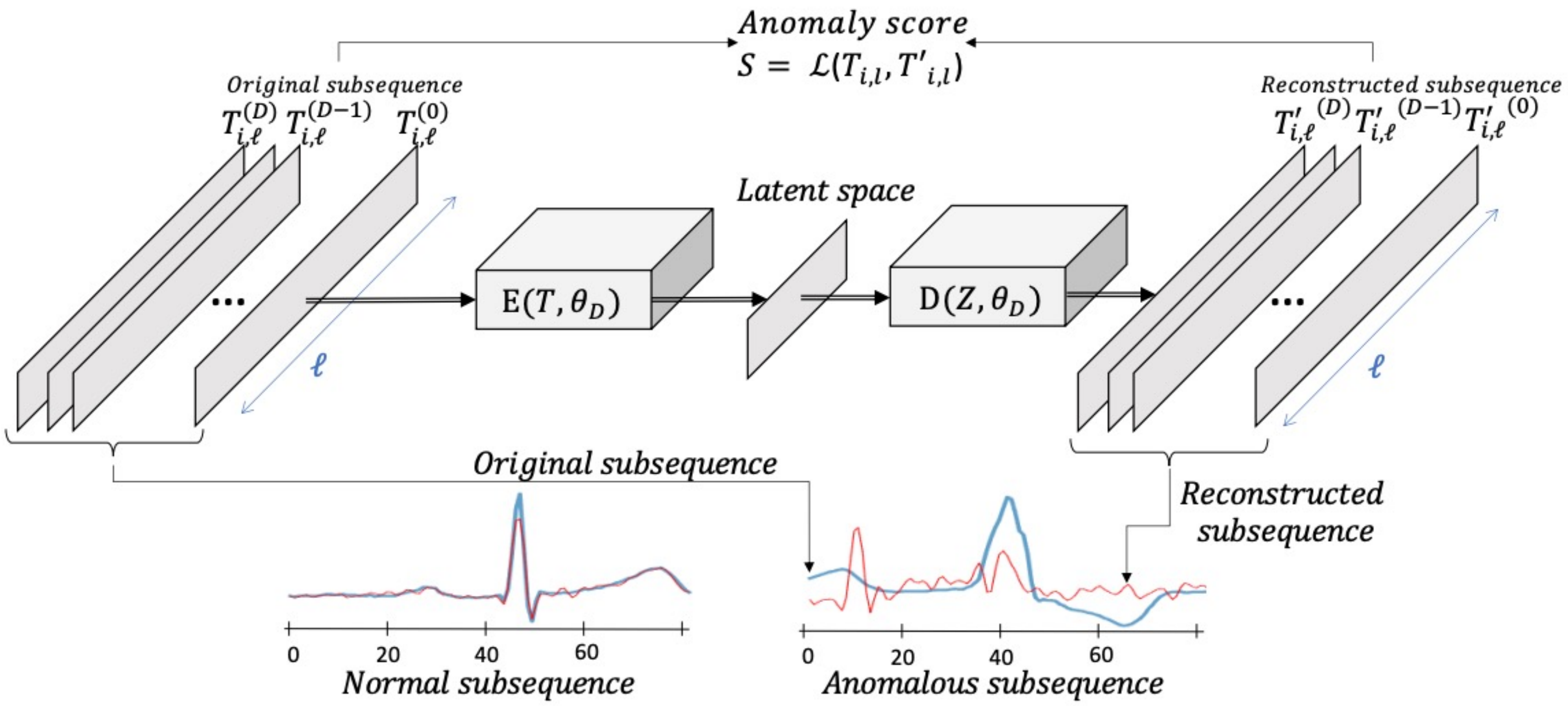
[16] M. Munir, S. A. Siddiqui, A. Dengel, and S. Ahmed. 2019. DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series. IEEE Access 7 (2019), 1991–2005.

Anomaly Detection methods: *Reconstruction-based*

Methods that aims to reconstruct the time series T and use the reconstruction error to detect if the time series is an anomaly or not.



Anomaly Detection methods: *an Example*



AutoEncoders [17] (AE)

Neural Network composed of an encoder (that reduce the dimensionality) and decoder that reconstruct the time series. The objective is to minimize the reconstruction error.

Semi-supervised

Univariate/Multivariate

Point/sequence

[17] Mayu Sakurada and Takehisa Yairi. 2014. Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction. In Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (Gold Coast, Australia QLD, Australia) (MLSDA'14).

Anomaly Detection methods: *Existing benchmark*

HEX/UCR [18]

Set of 250 time series with labels.

Details

- The labels have been manually checked and are reliable
- Each time series contains only 1 labeled anomaly

TimeEval [5]

Set of 976 time series with labels.

Details

- New synthetic benchmark GutenTag used to tune parameters
- Only Time series with low contamination rate (< 0.1)
- Time series with at least one methods above 0.8 AUC-ROC

TSB-UAD [19]

Set of 2000 time series with labels.

Details

- Collected as proposed in the literature (no filtering based on contamination, size or label quality)
- Artificial and synthetic data generation methods for reliable labels

Anomaly Detection methods: *Existing benchmark*

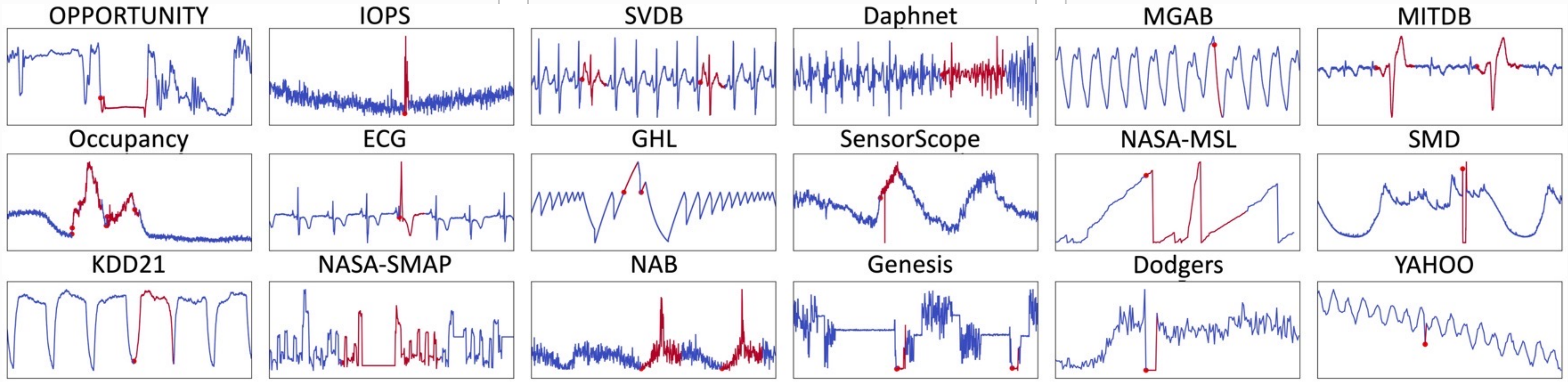
HEX/UCR [18]

Set of 250 time series with

only 1 labeled anomaly

TimeEval [5]

Real datasets collection



contamination rate (< 0.1)

- Time series with at least one methods above 0.8 AUC-ROC

TSB-UAD [19]

Set of 2000 time series with

contamination, size of label quality.

Anomaly Detection methods: *Existing benchmark*

HEX/UCR [18]

Set of 250

TimeEval [5]

of 976 time series with labels

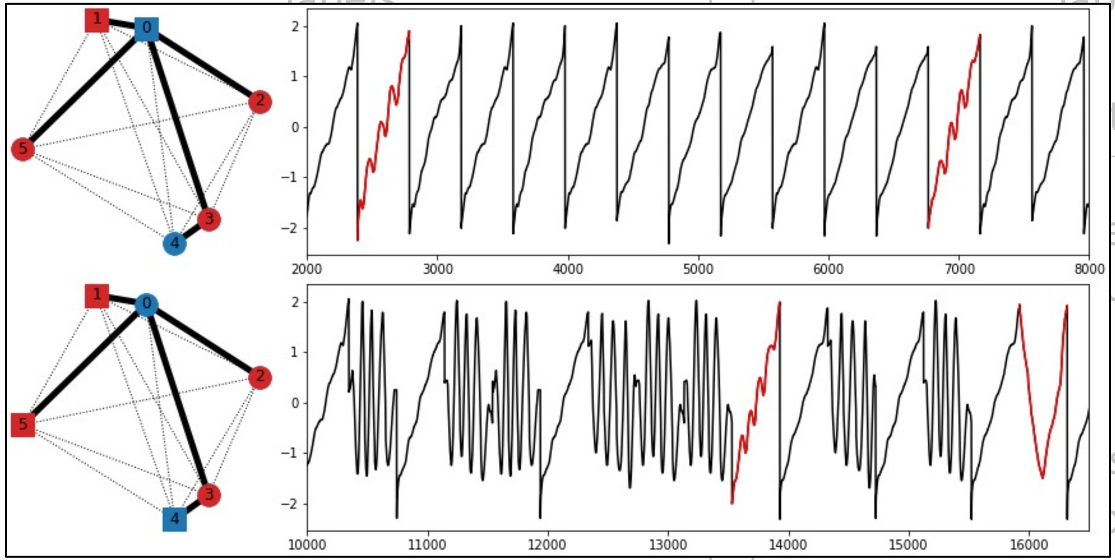
- Time series with at least one methods above 0.8 AUC-ROC

TSB-UAD [19]

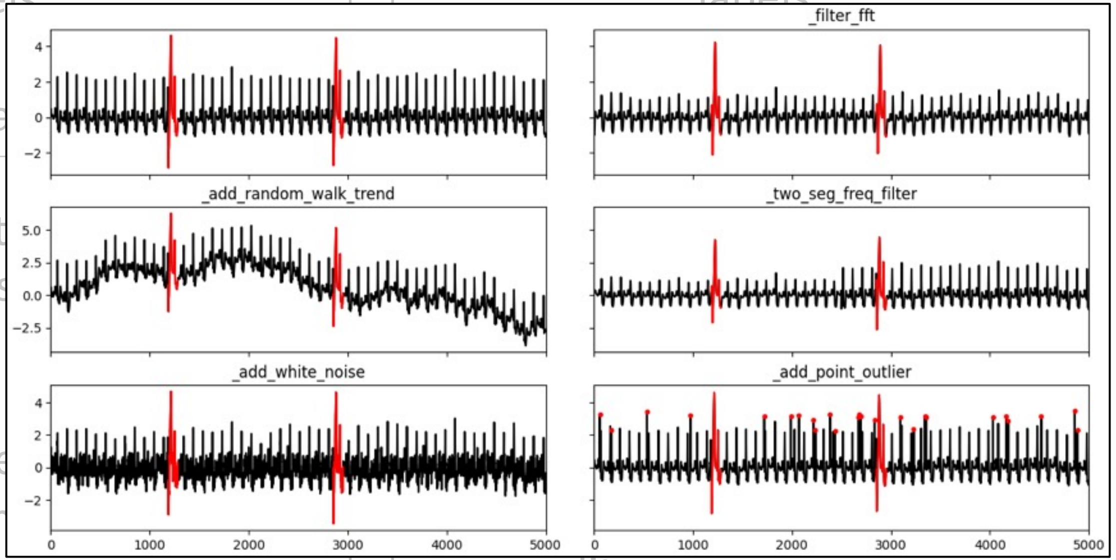
ies with

quality.

Artificial dataset generation



Synthetic dataset generation



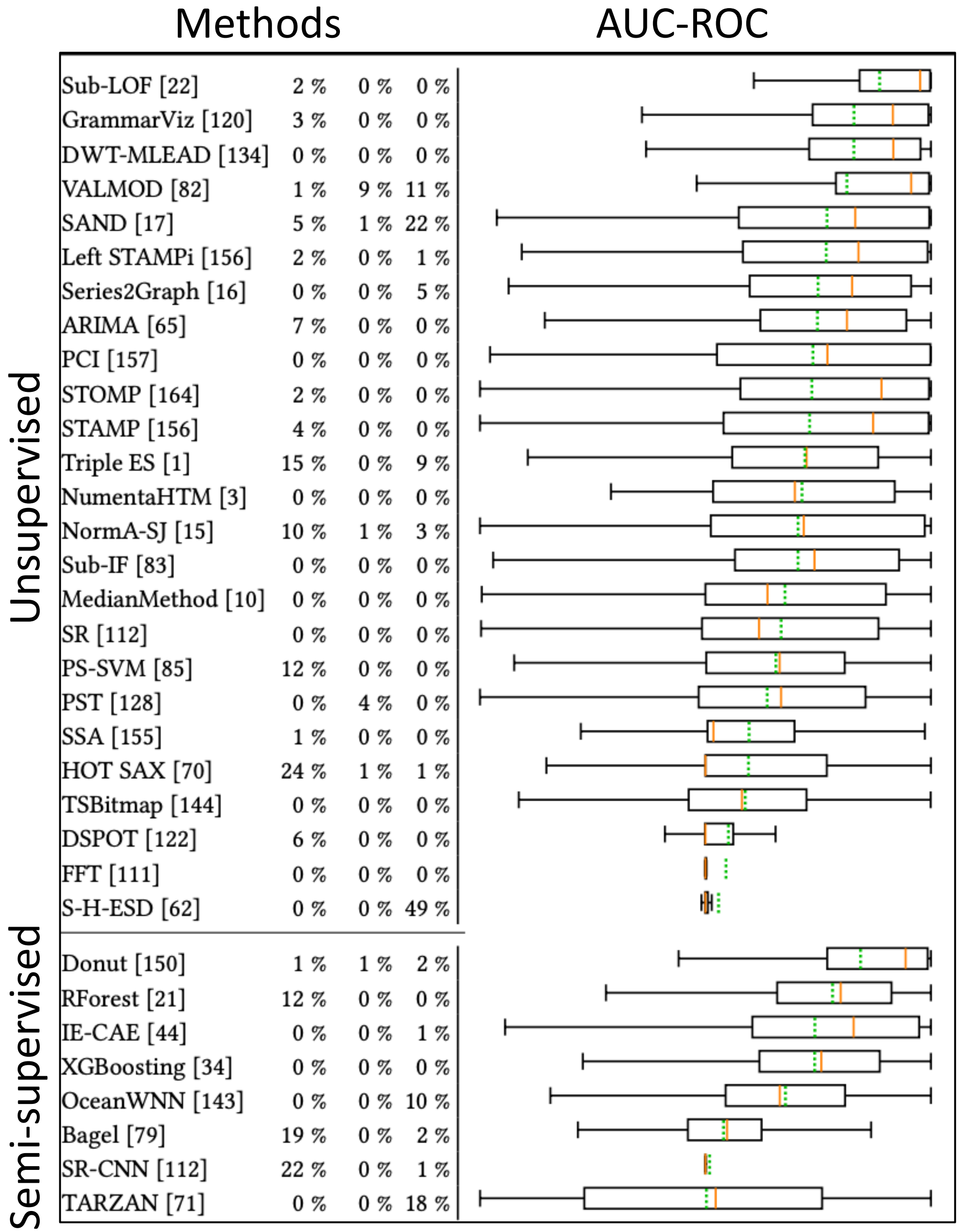
Anomaly Detection methods:

Experimental evaluation

Observations on TimeEval [5]:

- Distance-based and Density-based methods have a better accuracy (AUC-ROC) than forecasting and reconstruction-based approaches

- Semi-supervised methods are not outperforming Unsupervised approaches

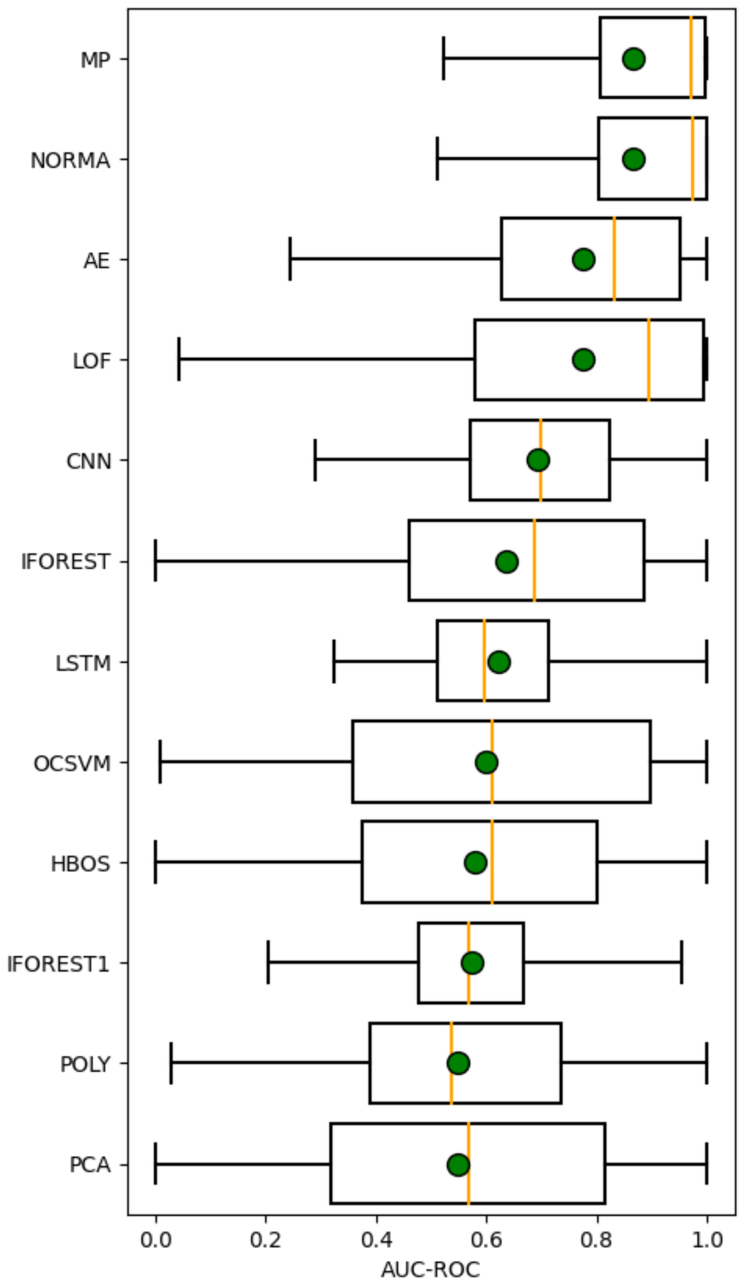


[5] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. 2022. Anomaly detection in time series: a comprehensive evaluation. Proc. VLDB Endow. 15, 9 (May 2022), 1779–1797.

Anomaly Detection methods:
Experimental evaluation

Observations on HEX/UCR [18]:

- Distance-based methods have a better accuracy (AUC-ROC) than forecasting and distribution-based approaches



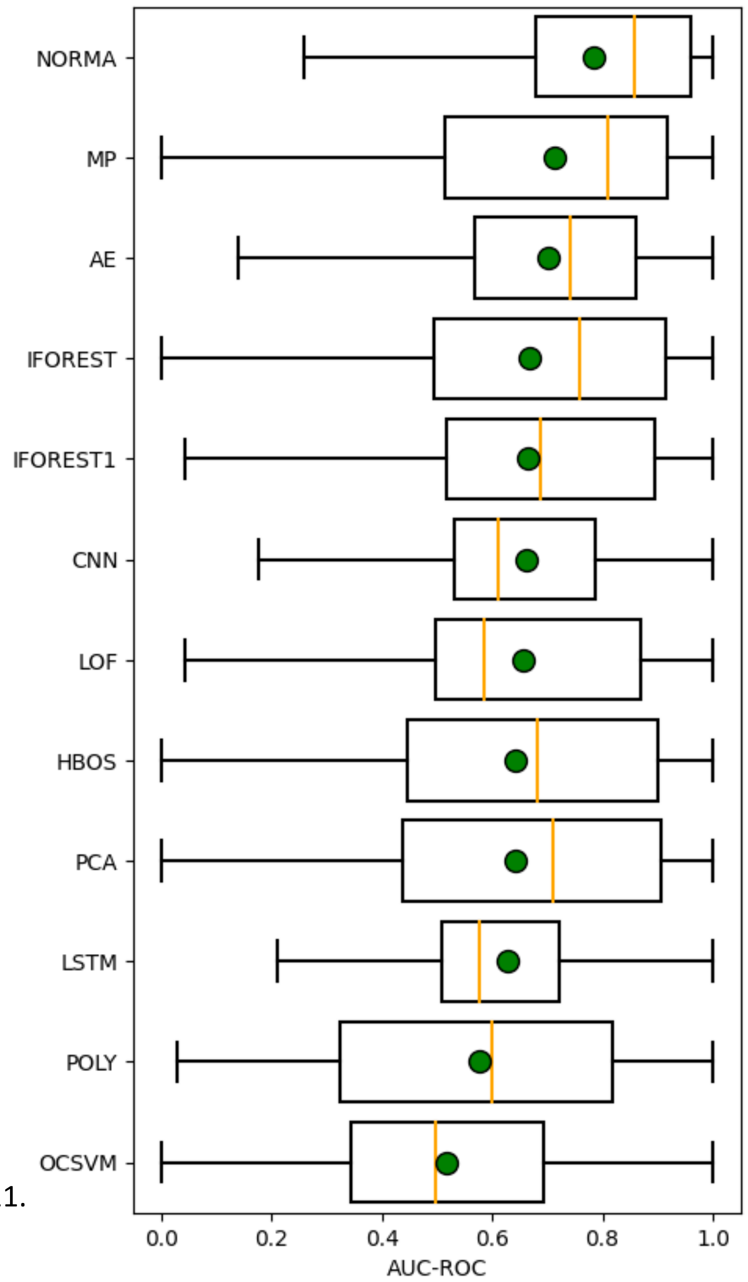
[18] R. Wu and E. Keogh, "Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress" in IEEE Transactions on Knowledge & Data Engineering, vol. 35, no. 03, pp. 2421-2429, 2023.

Anomaly Detection methods: *Experimental evaluation*

Observations on TSB-UAD [19]:

- Distance-based methods have a better accuracy (AUC-ROC) than forecasting-based methods.
- Isolation Forest (distribution-based and not proposed for time series) have also a strong accuracy

- AutoEncoder (AE) is also very accurate.



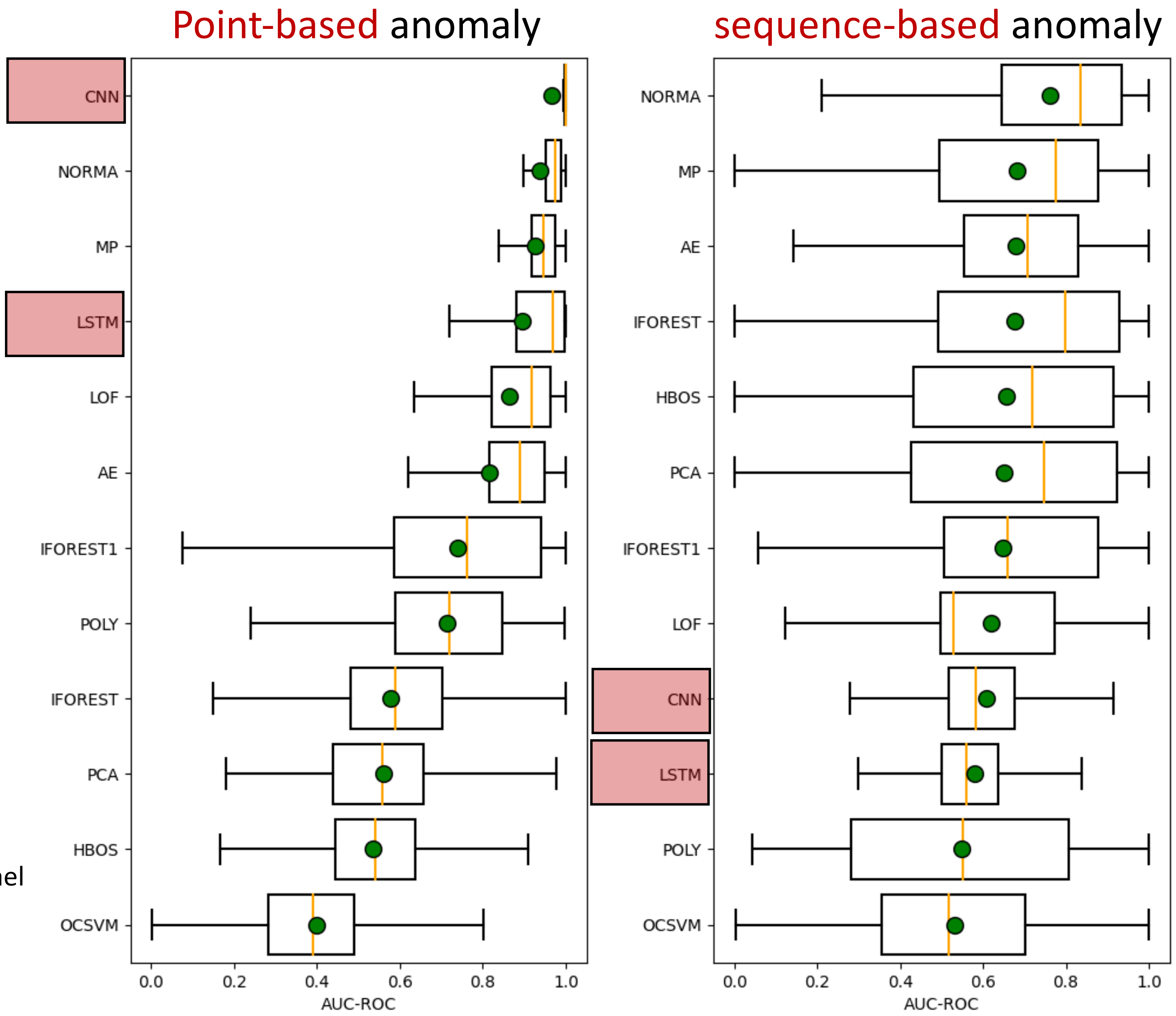
[19] John Paparrizos, Yuhao Kang, Paul Boniol, Ruey S. Tsay, Themis Palpanas, and Michael J. Franklin. 2022. TSB-UAD: an end-to-end benchmark suite for univariate time-series anomaly detection. Proc. VLDB Endow. 15, 8 (April 2022), 1697–1711.

Anomaly Detection methods:

Experimental evaluation

Observations on TSB-UAD [19]:

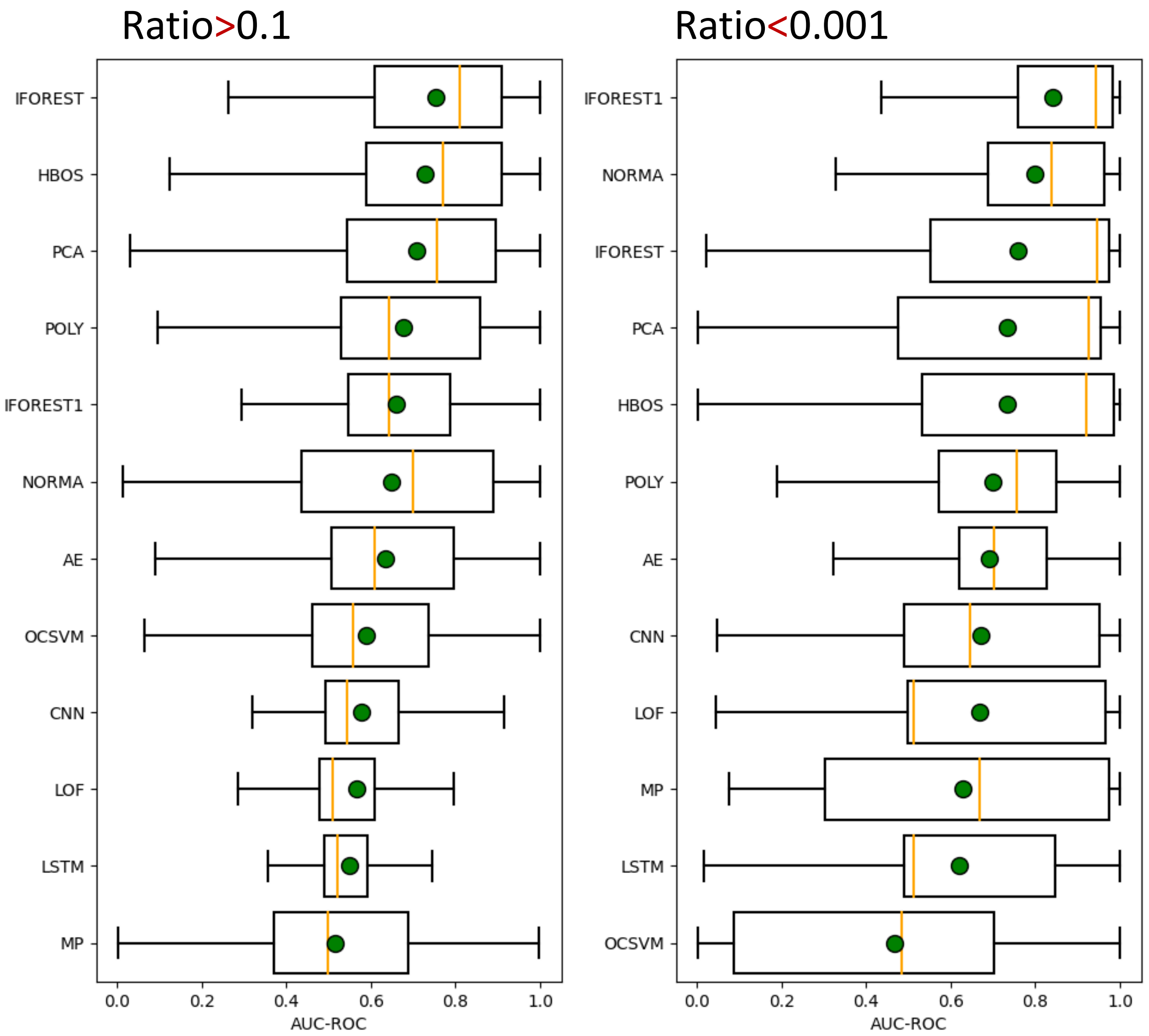
- Forecasting methods (LSTM and CNN) are very accurate for point anomalies
- But have poor performances on sequence-based anomalies.



[19] John Paparrizos, Yuhao Kang, Paul Boniol, Ruey S. Tsay, Themis Palpanas, and Michael J. Franklin. 2022. TSB-UAD: an end-to-end benchmark suite for univariate time-series anomaly detection. Proc. VLDB Endow. 15, 8 (April 2022), 1697–1711.

Anomaly Detection methods:

Experimental evaluation



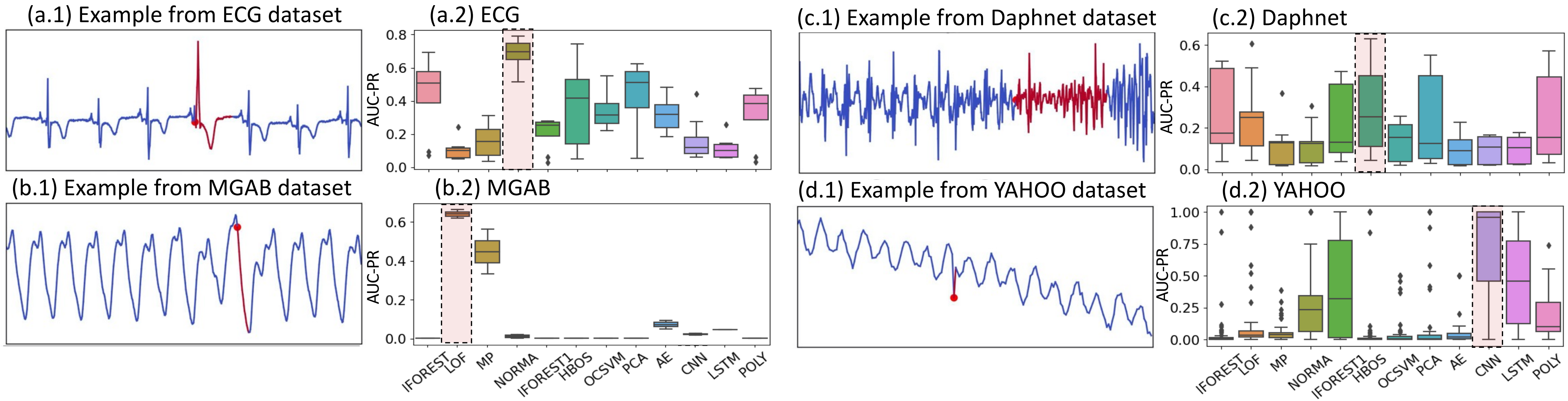
Observations on TSB-UAD [19]:

- The ratio of normal/abnormal points has a strong impact on the methods ranking.

[19] John Paparrizos, Yuhao Kang, Paul Boniol, Ruey S. Tsay, Themis Palpanas, and Michael J. Franklin. 2022. TSB-UAD: an end-to-end benchmark suite for univariate time-series anomaly detection. Proc. VLDB Endow. 15, 8 (April 2022), 1697–1711.

Anomaly Detection methods: *Experimental evaluation*

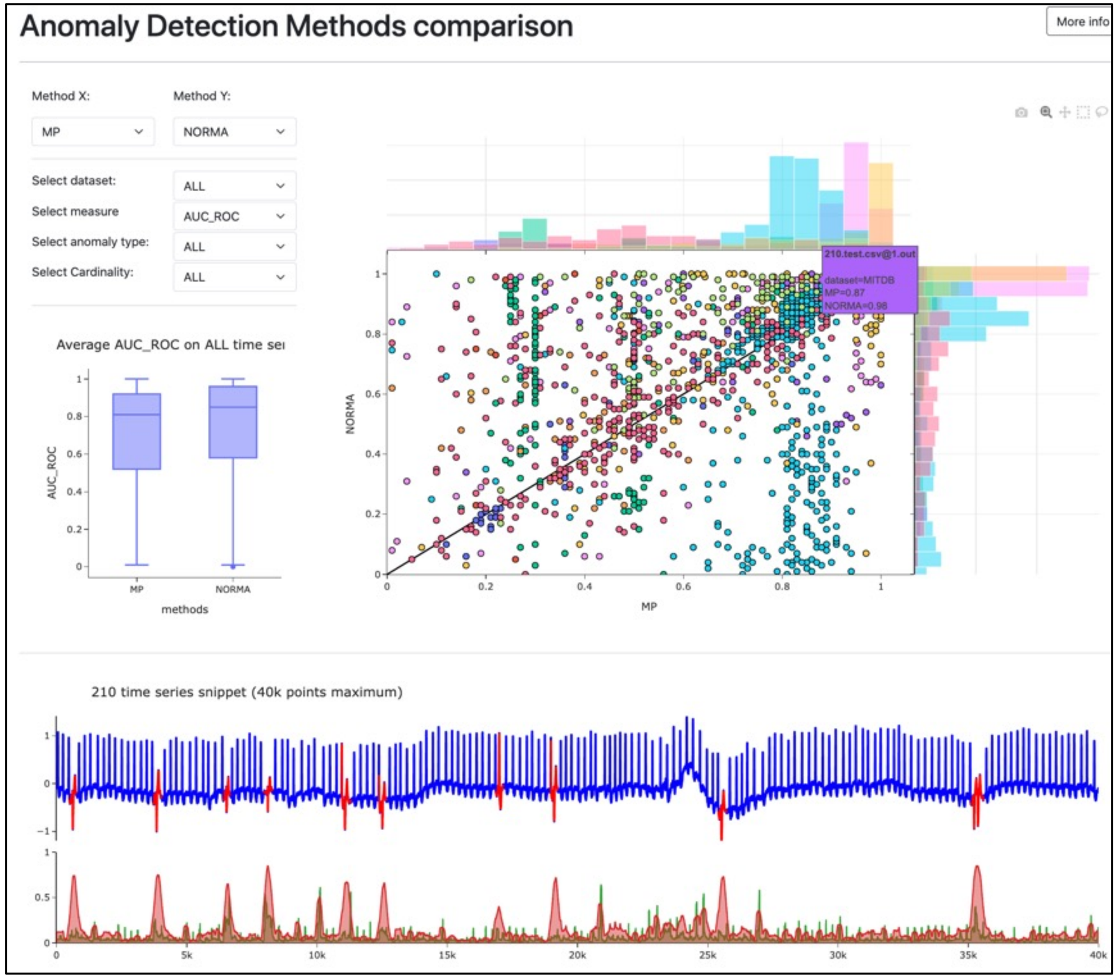
Observation from the results applied on specific datasets (TSB-UAD [19])



There is no overall winner.

[19] John Paparrizos, Yuhao Kang, Paul Boniol, Ruey S. Tsay, Themis Palpanas, and Michael J. Franklin. 2022. TSB-UAD: an end-to-end benchmark suite for univariate time-series anomaly detection. Proc. VLDB Endow. 15, 8 (April 2022), 1697–1711.

Anomaly Detection methods: *Experimental evaluation*



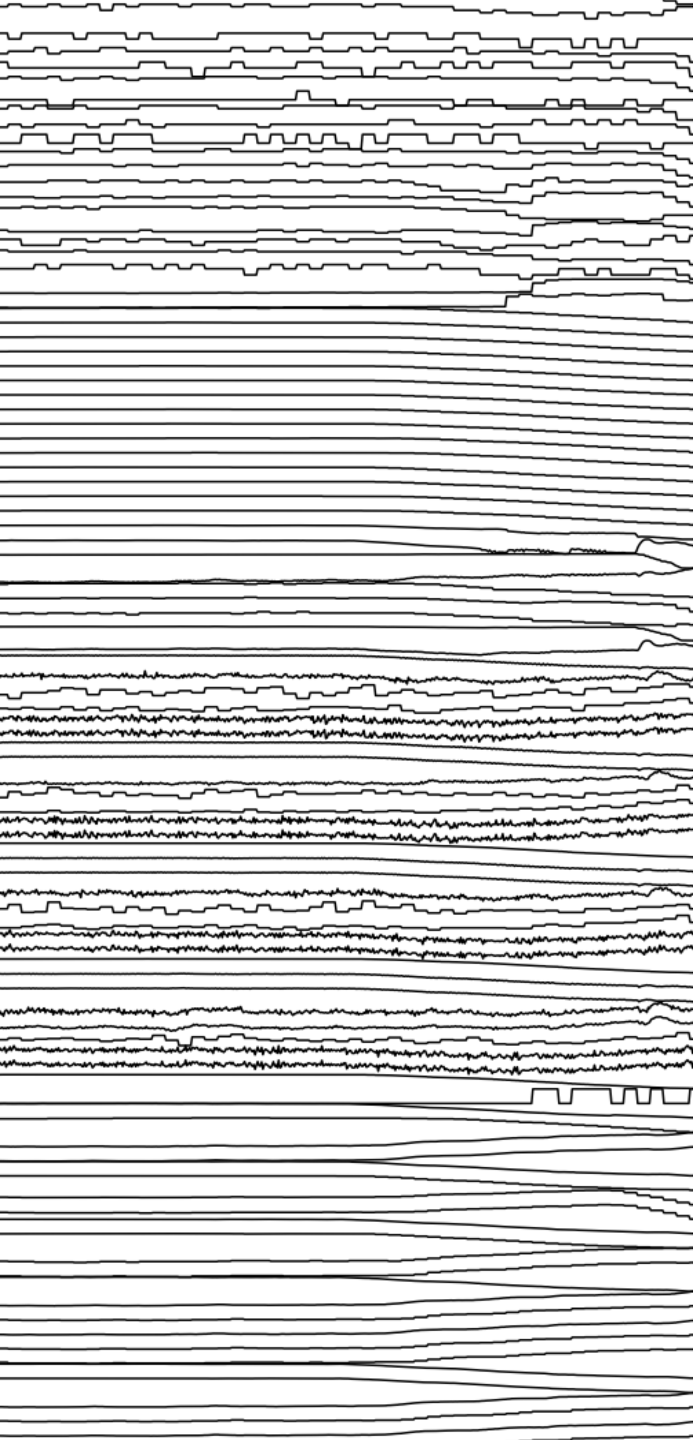
Theseus [27]

An interactive tool to compare anomaly detection methods

VLDB 2022

Github repo

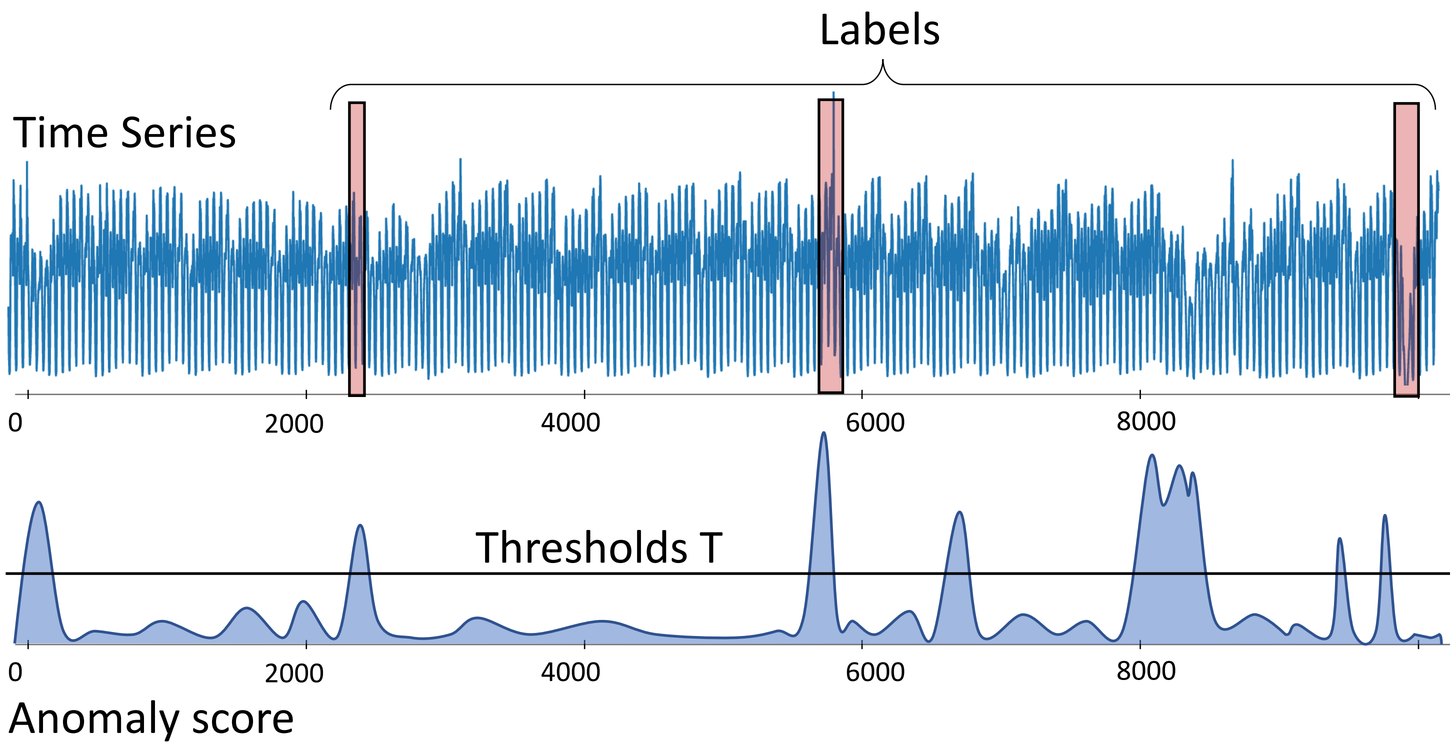
[27] Paul Boniol, John Paparrizos, Yuhao Kang, Themis Palpanas, Ruey S. Tsay, Aaron J. Elmore, and Michael J. Franklin. 2022. Theseus: navigating the labyrinth of time-series anomaly detection. Proc. VLDB Endow. 15, 12 (August 2022), 3702–3705.



Evaluation Measures

Evaluation measures: *Threshold-based*

Threshold-based Evaluation Measures:

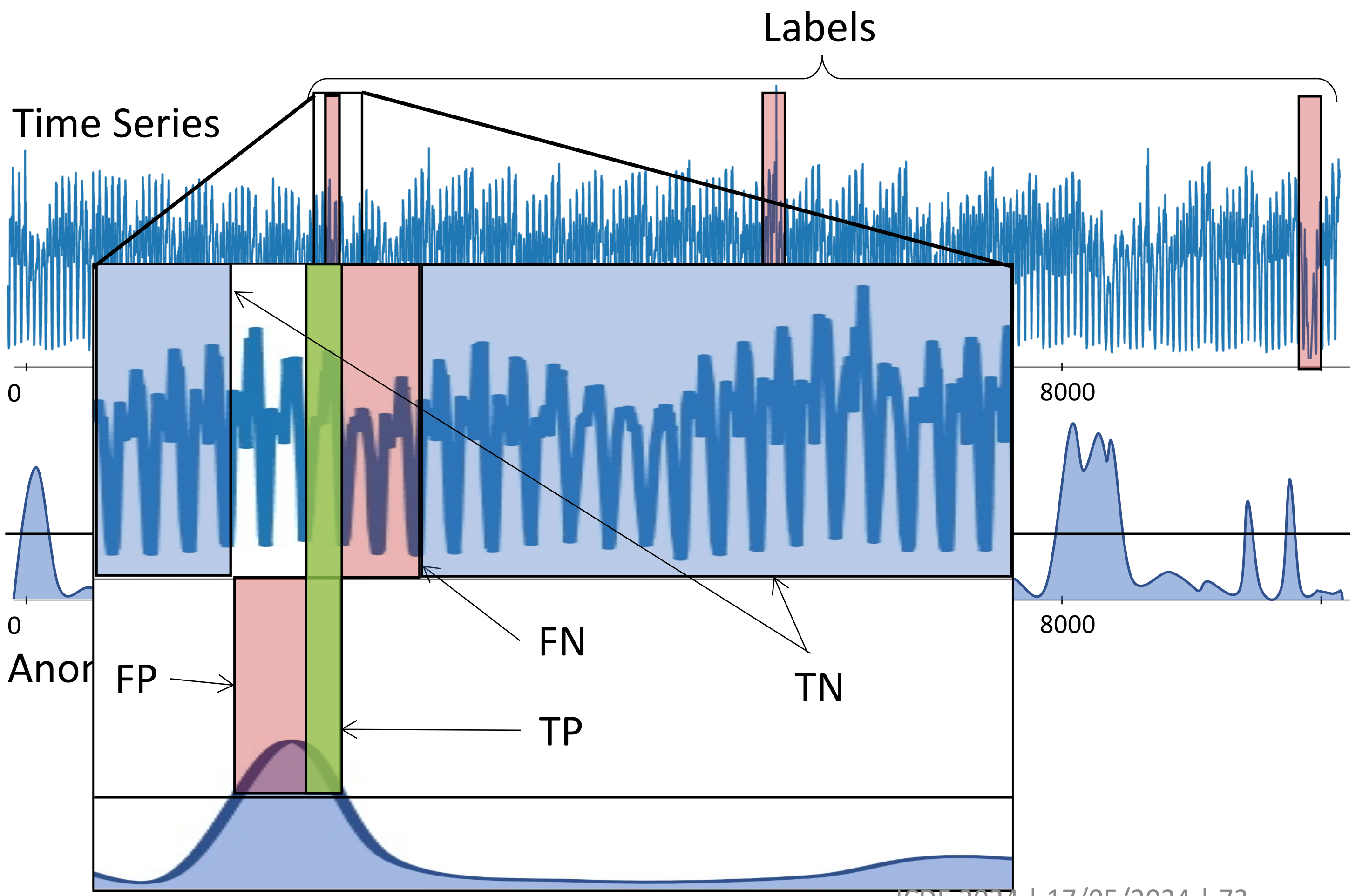


Evaluation measures: *Threshold-based*

Threshold-based Evaluation Measures:

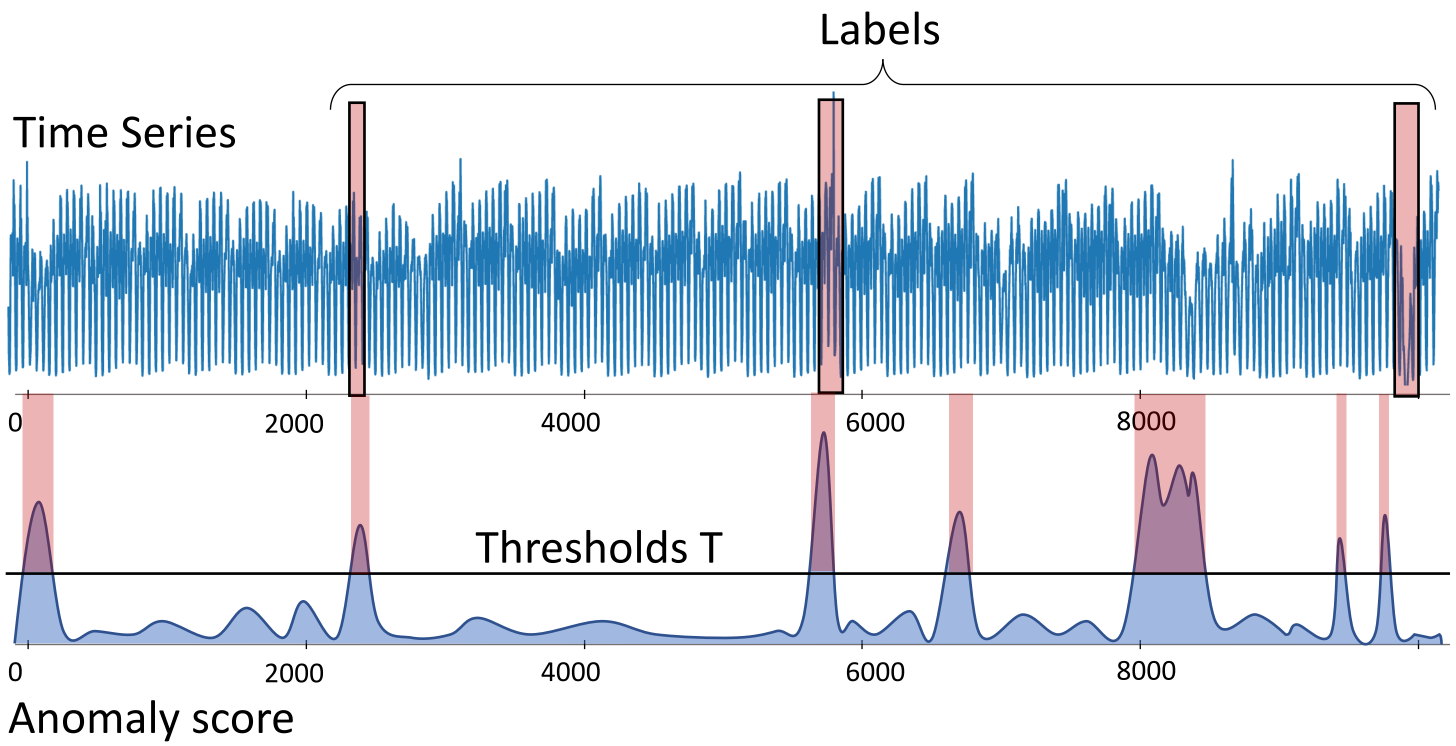
- Precision: $\frac{TP}{TP+FP}$
- Recall (true positive rate): $\frac{TP}{TP+FN}$
- False positive rate: $\frac{FP}{FP+TN}$

- F-score: $\frac{(1+\beta^2)*Precision}{\beta^2*Precision+Recall}$



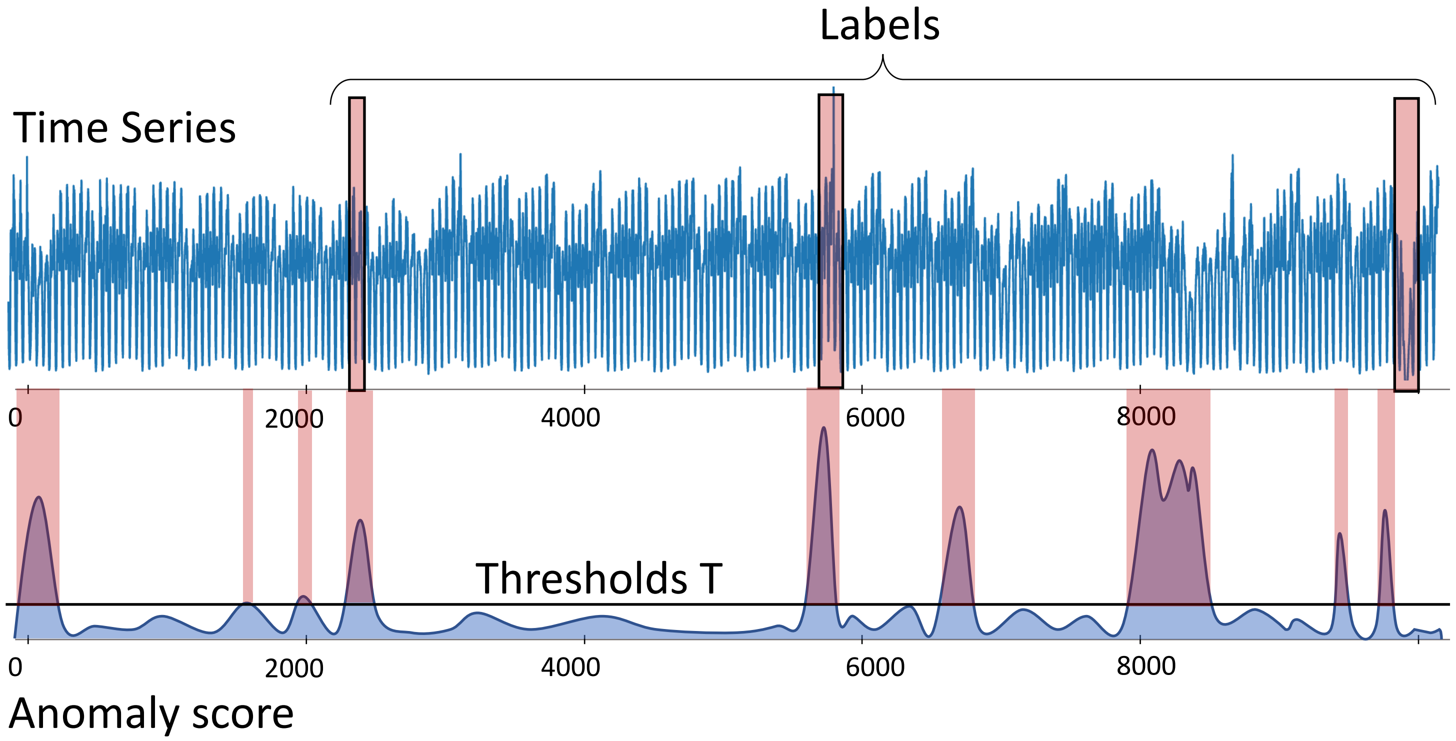
Evaluation measures: *AUC-based*

How do we set the threshold?



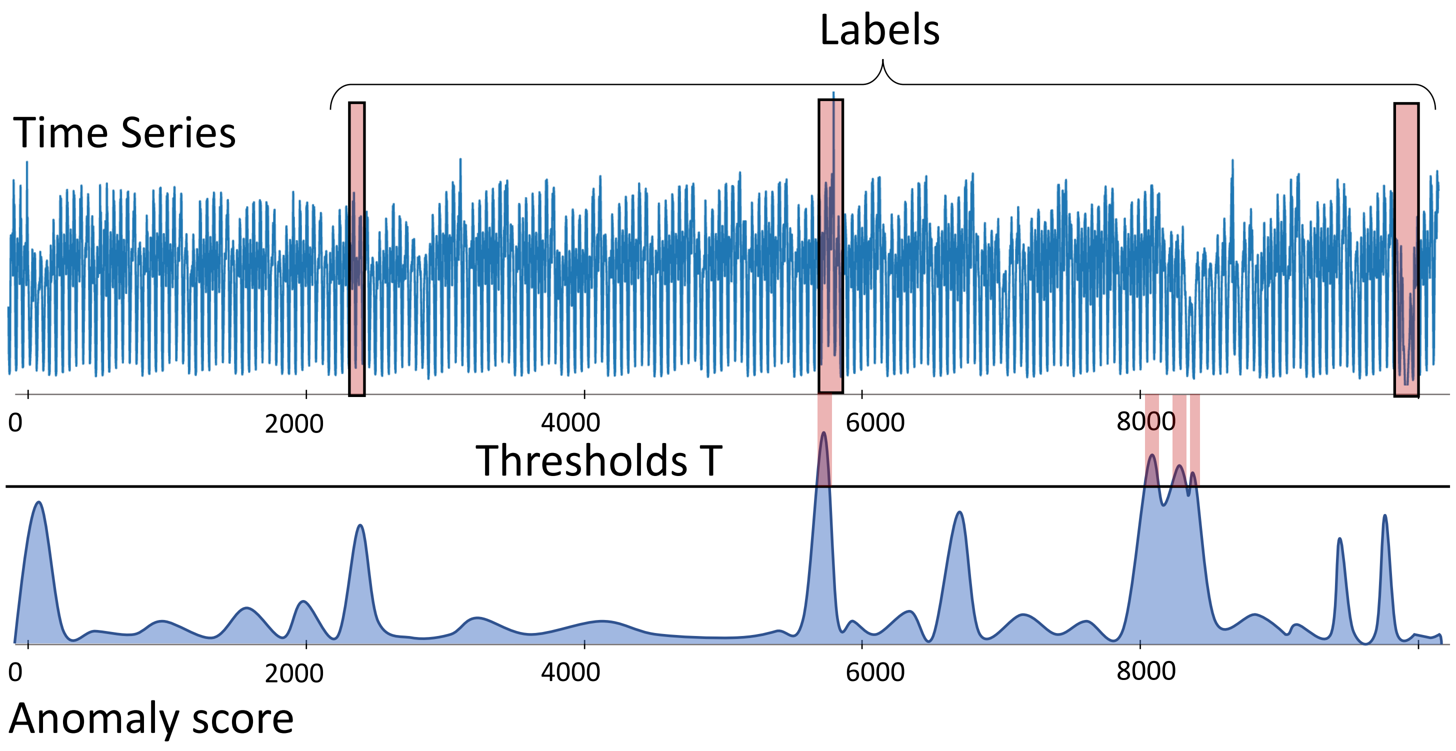
Evaluation measures: *AUC-based*

How do we set the threshold?



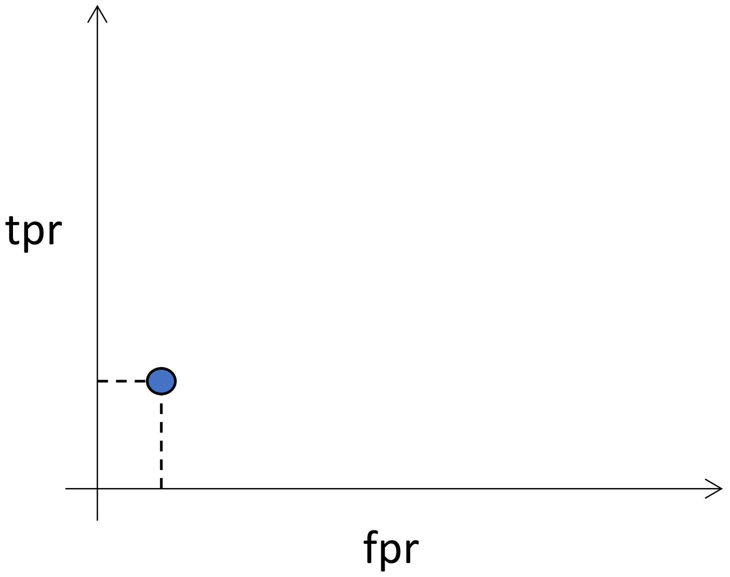
Evaluation measures: *AUC-based*

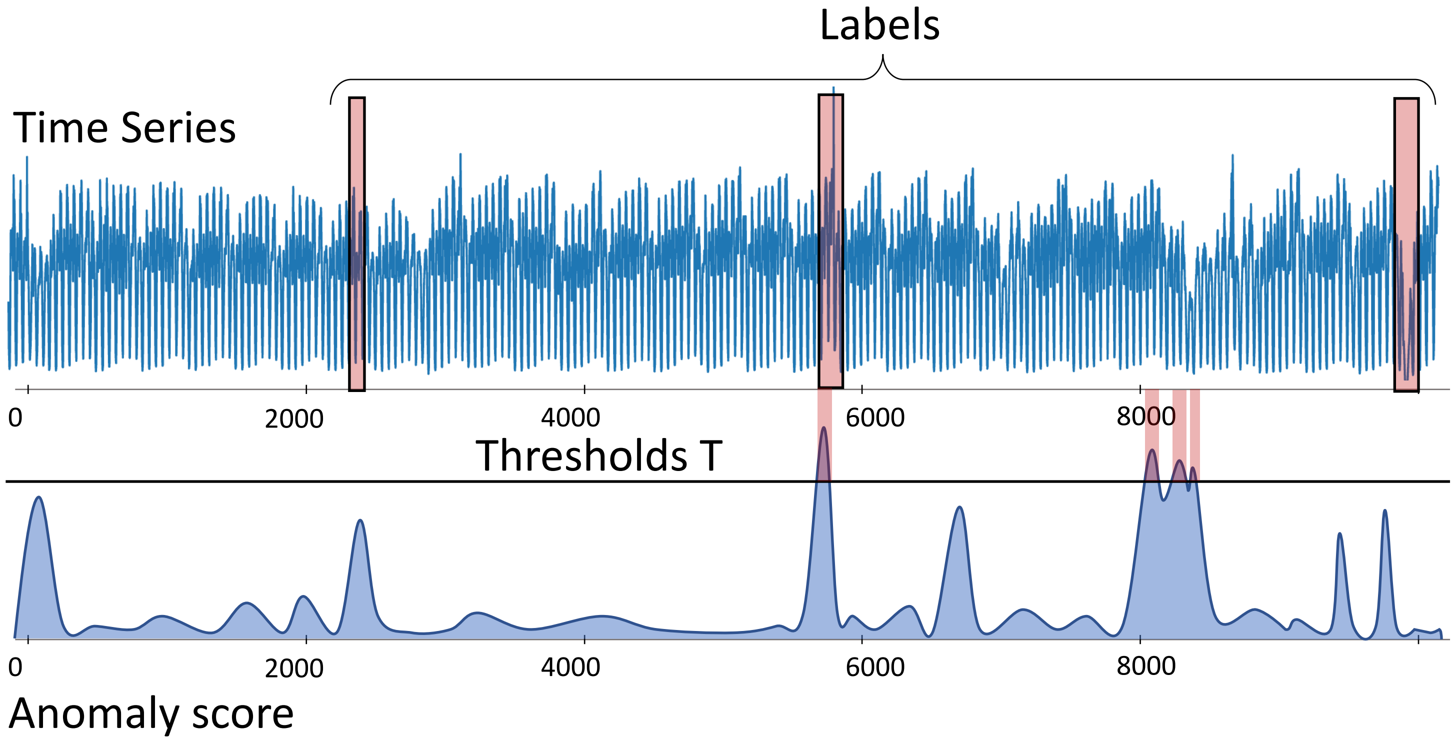
How do we set the threshold?



Evaluation measures: *AUC-based*

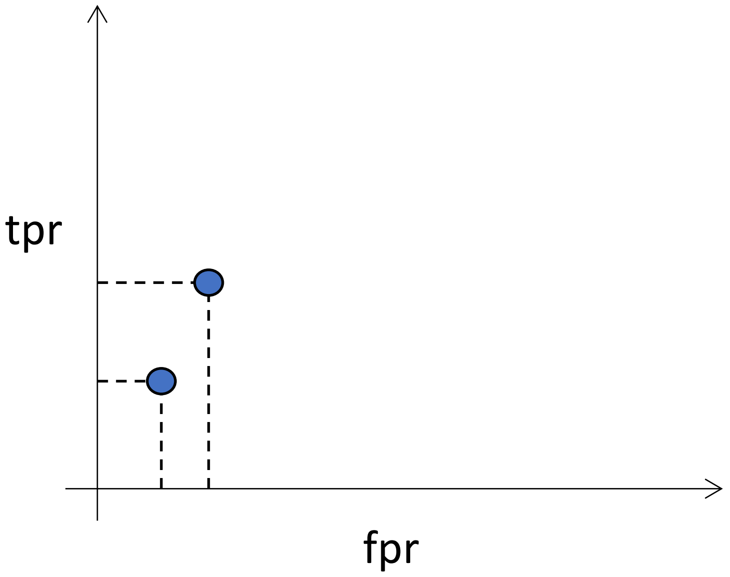
AUC-based Evaluation Measures:

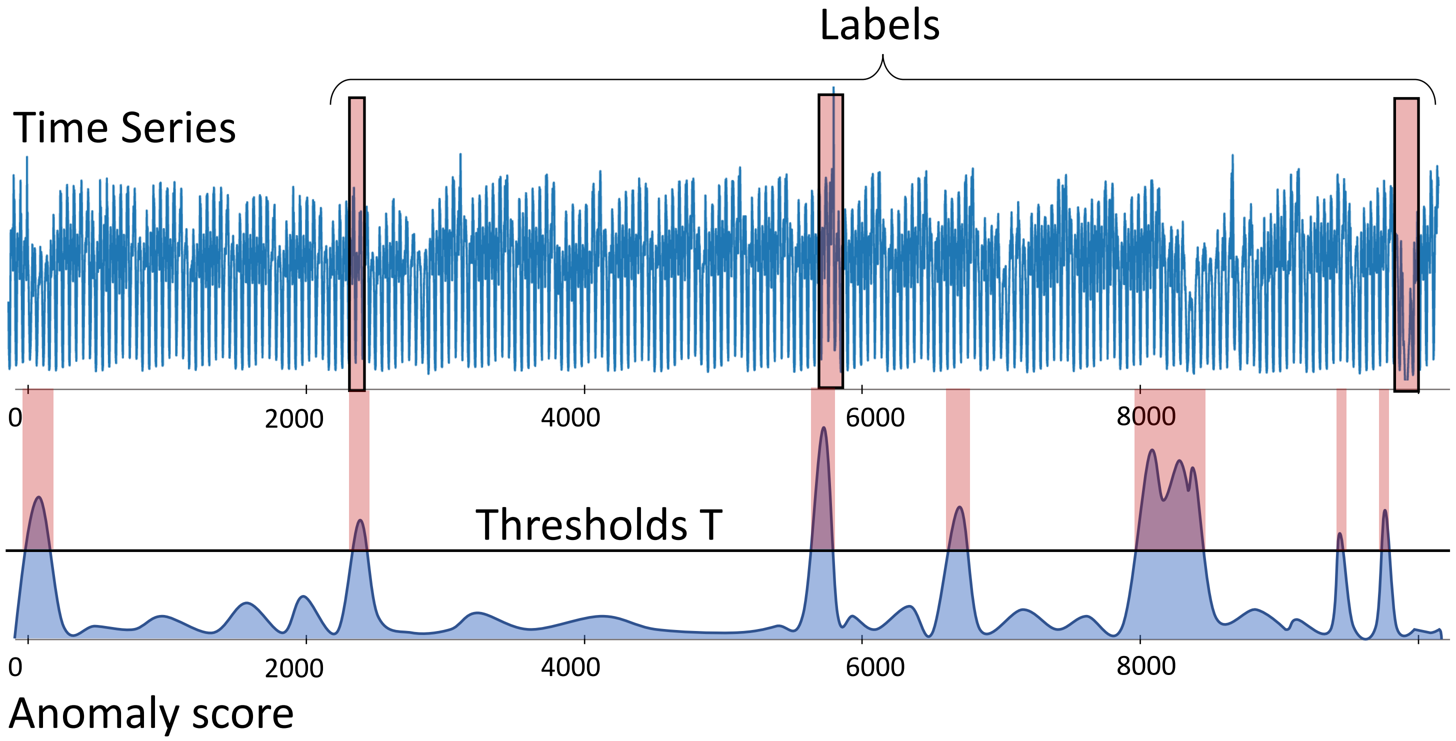




Evaluation measures: *AUC-based*

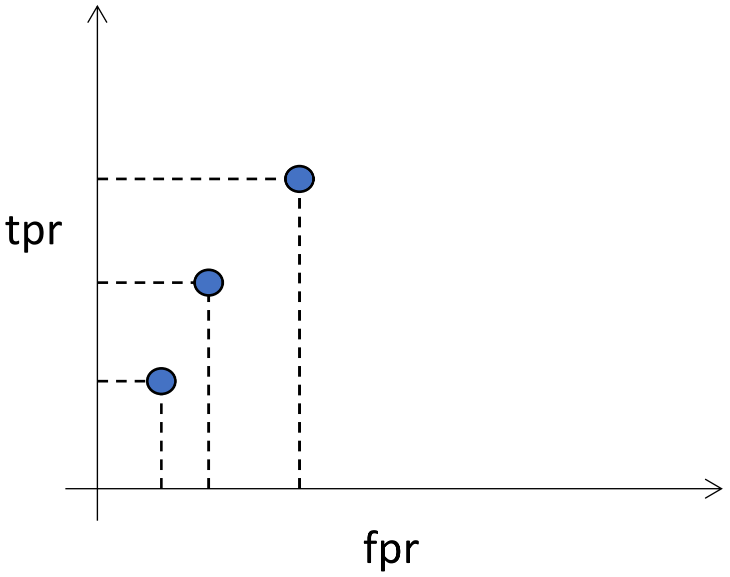
AUC-based Evaluation Measures:

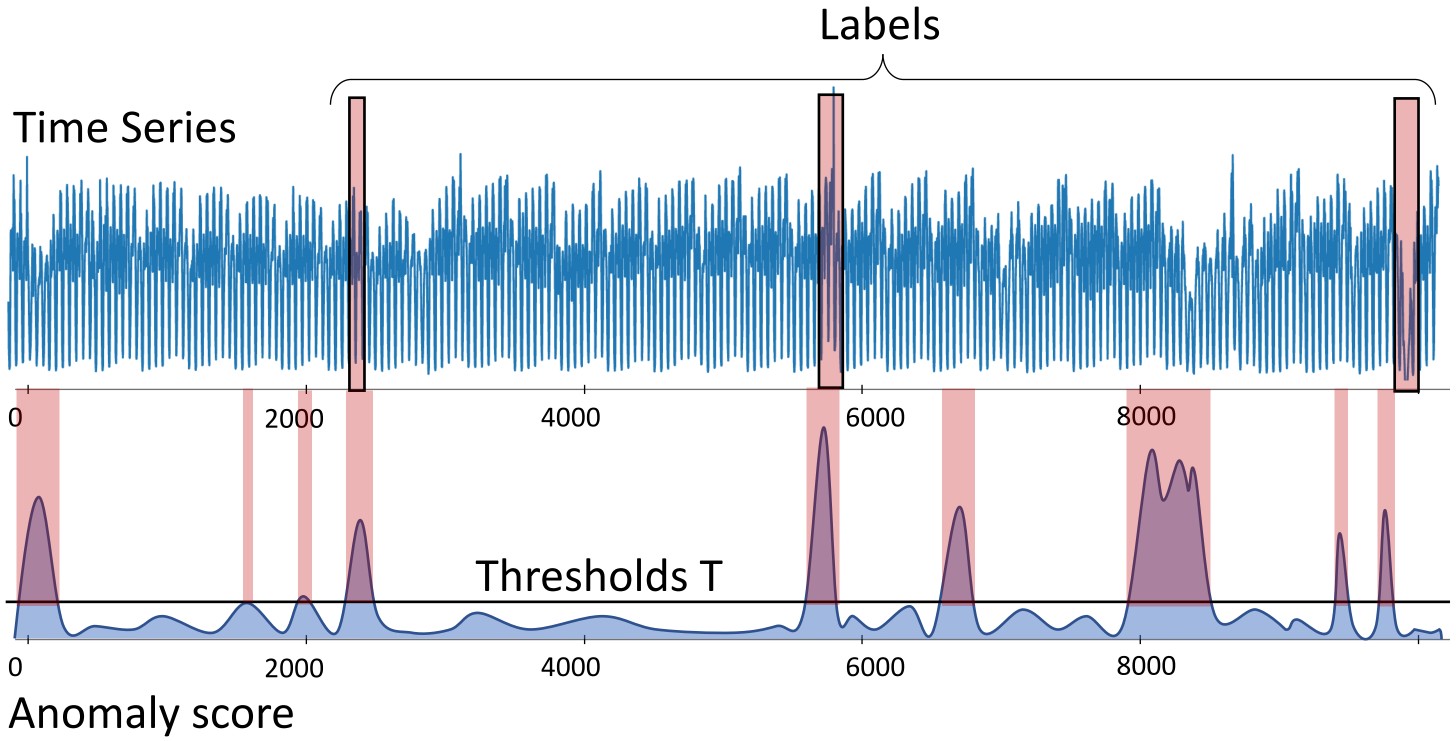




Evaluation measures: *AUC-based*

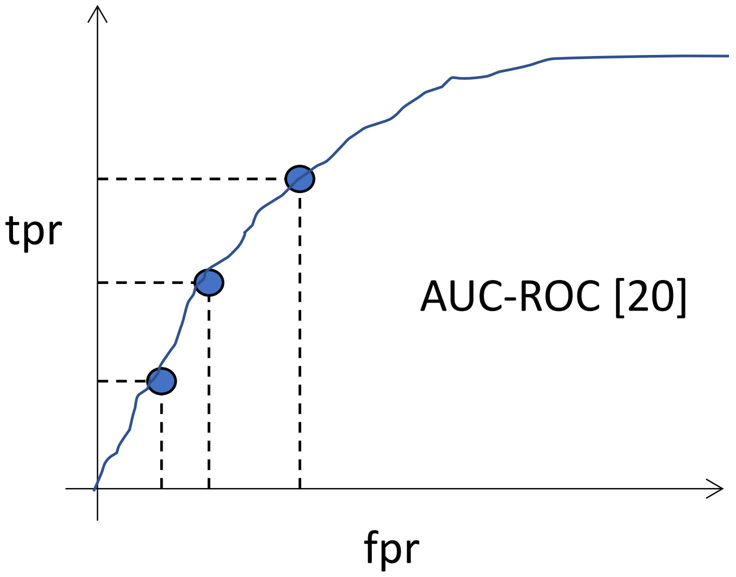
AUC-based Evaluation Measures:

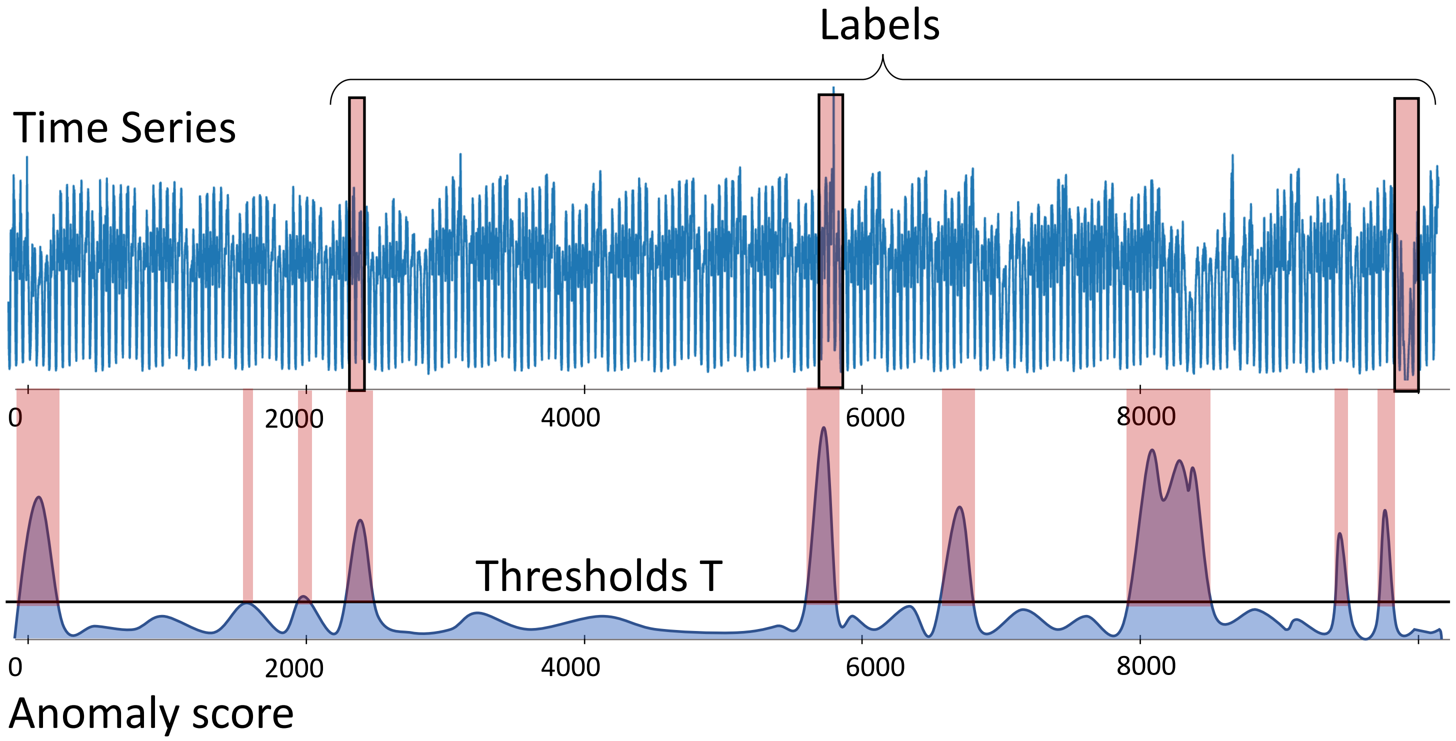




Evaluation measures: *AUC-based*

AUC-based Evaluation Measures:

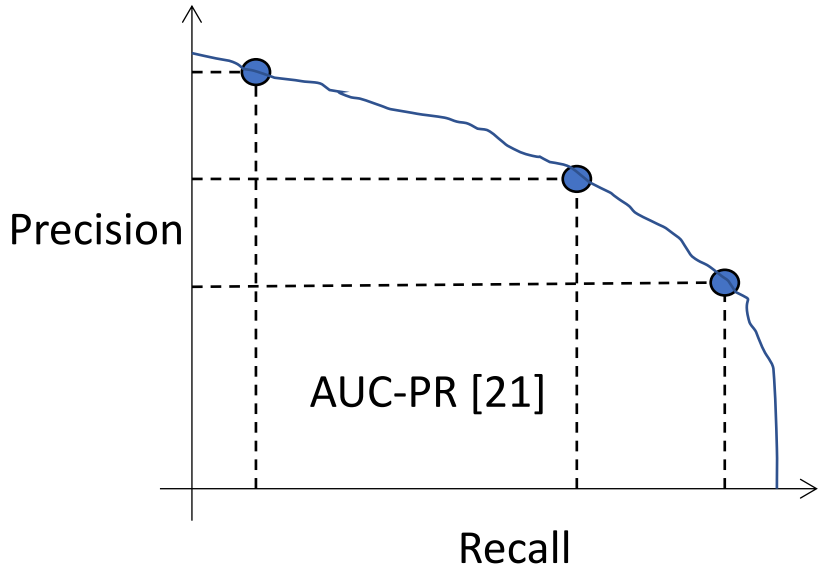


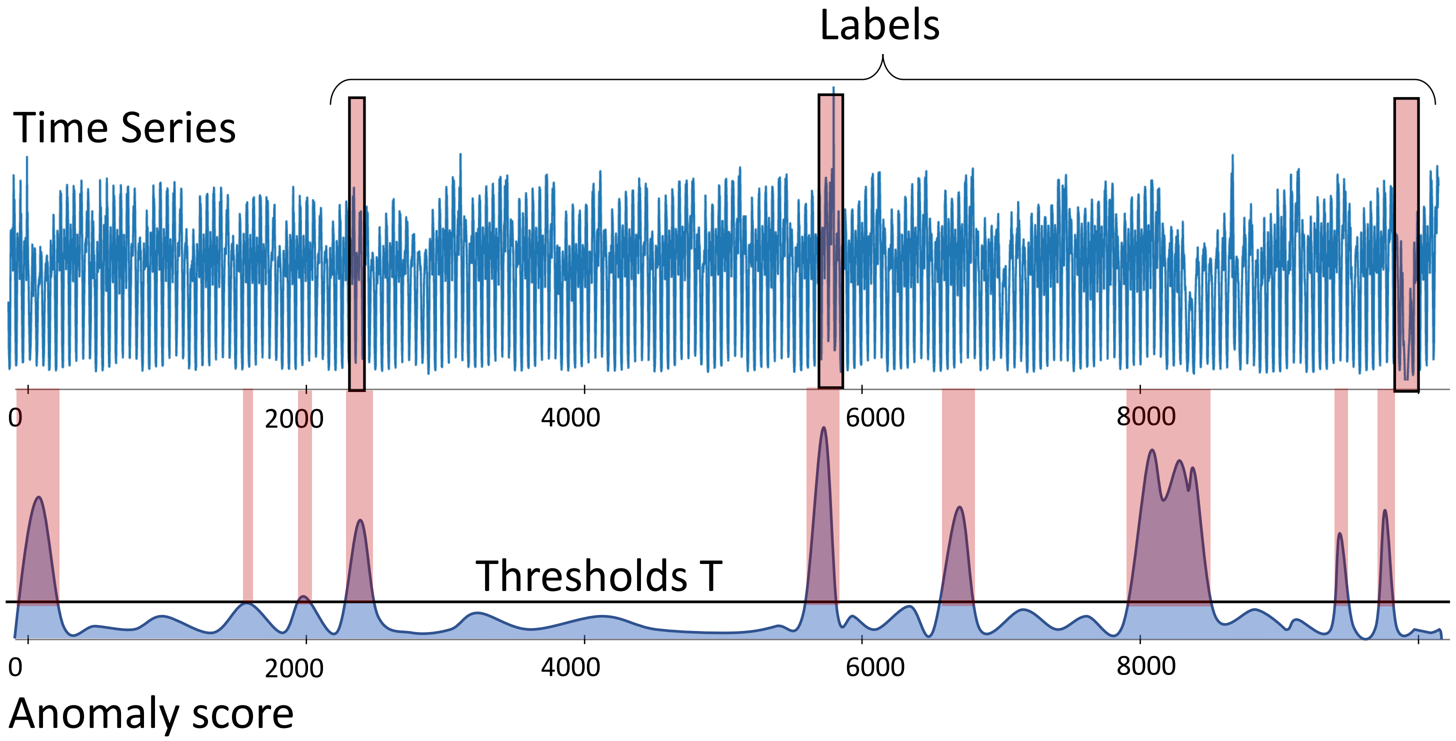


[20] Tom Fawcett. 2006. An introduction to ROC analysis. Pattern Recognition Letters 27, 8 (2006), 861–874.

Evaluation measures: *AUC-based*

AUC-based Evaluation Measures:



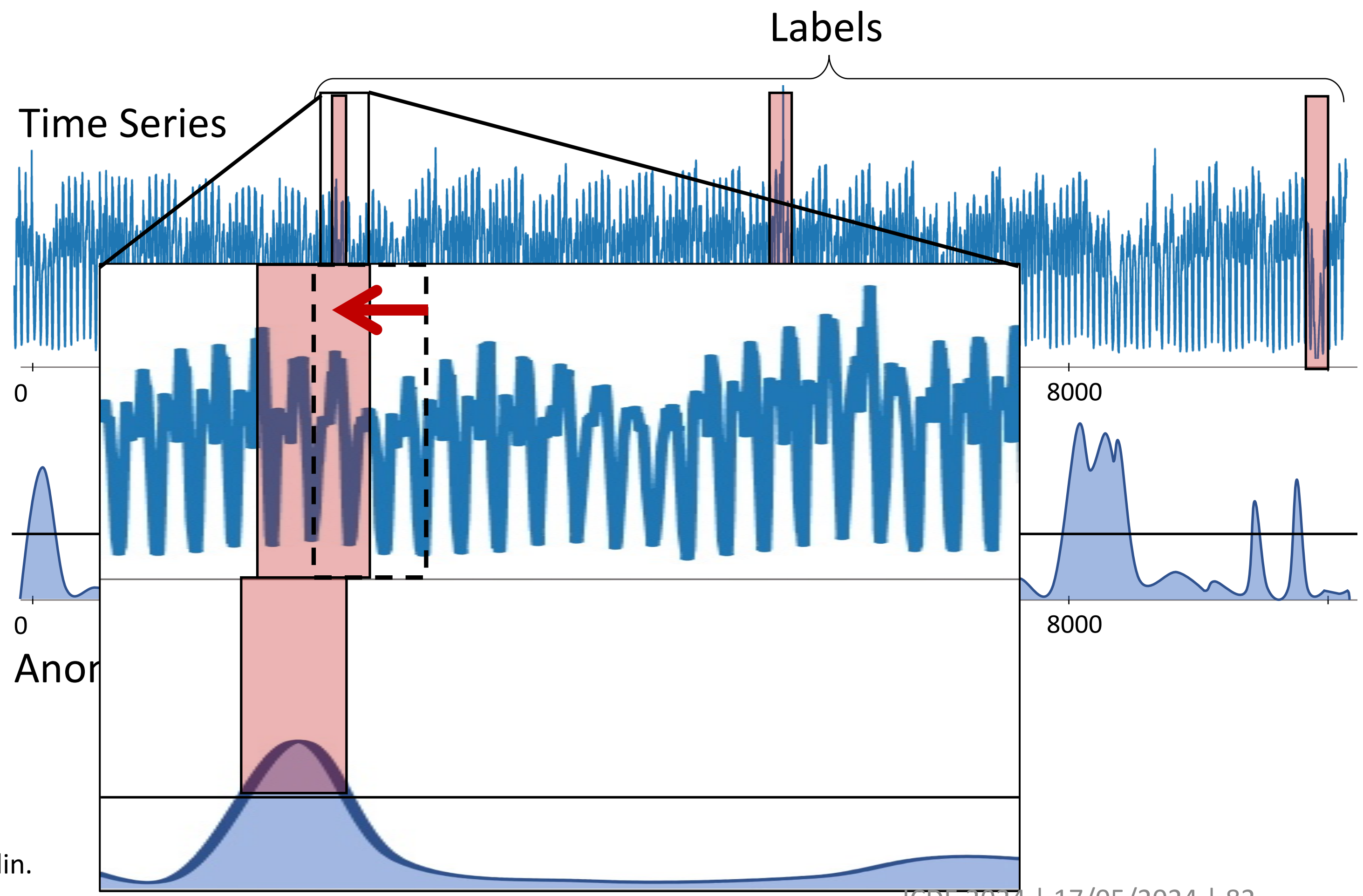


[21] Jesse Davis and Mark Goadrich. 2006. The Relationship between Precision-Recall and ROC Curves. In Proceedings of the 23rd International Conference on Machine Learning (ICML '06).

Evaluation measures: *Labeling issue*

Labeling can be an issue for time series [22]:

- Misalignment can lead to significant changes of accuracy values.

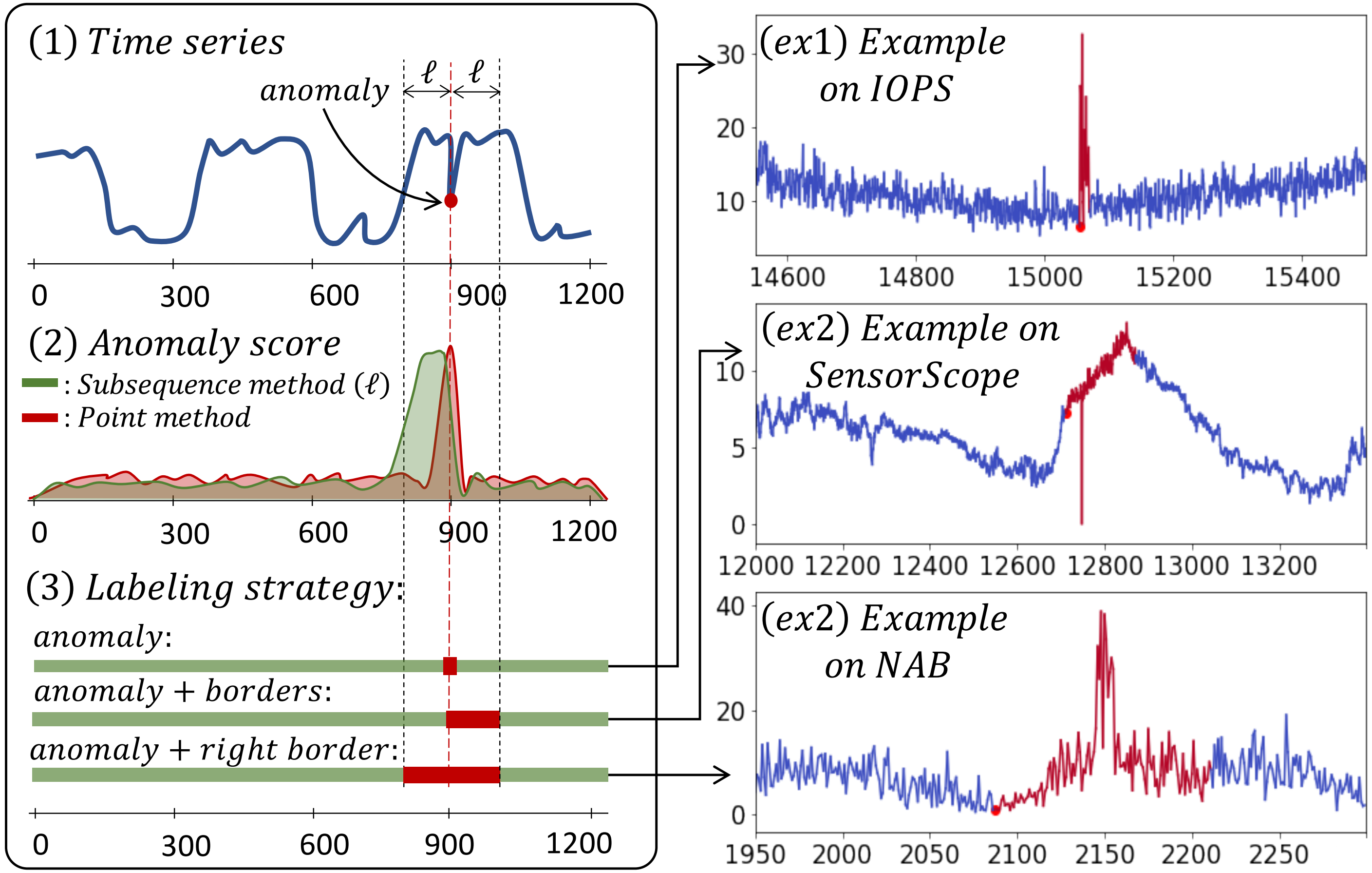


[22] J. Paparrizos, P. Boniol, T. Palpanas, R. S. Tsay, A. Elmore, and M. J. Franklin. Volume under the surface: a new accuracy evaluation measure for time-series anomaly detection. Proc. VLDB Endow. 15, 11 (2022), 2774–2787.

Evaluation measures: *Labeling issue*

Labeling can be an issue for time series [22]:

- Misalignment can lead to significant changes of accuracy values.
- This is a real issue because of:
  - Different Labeling strategies between domains and applications
  - Methods that produce misaligned anomaly scores.



Evaluation measures: *Labeling issue*

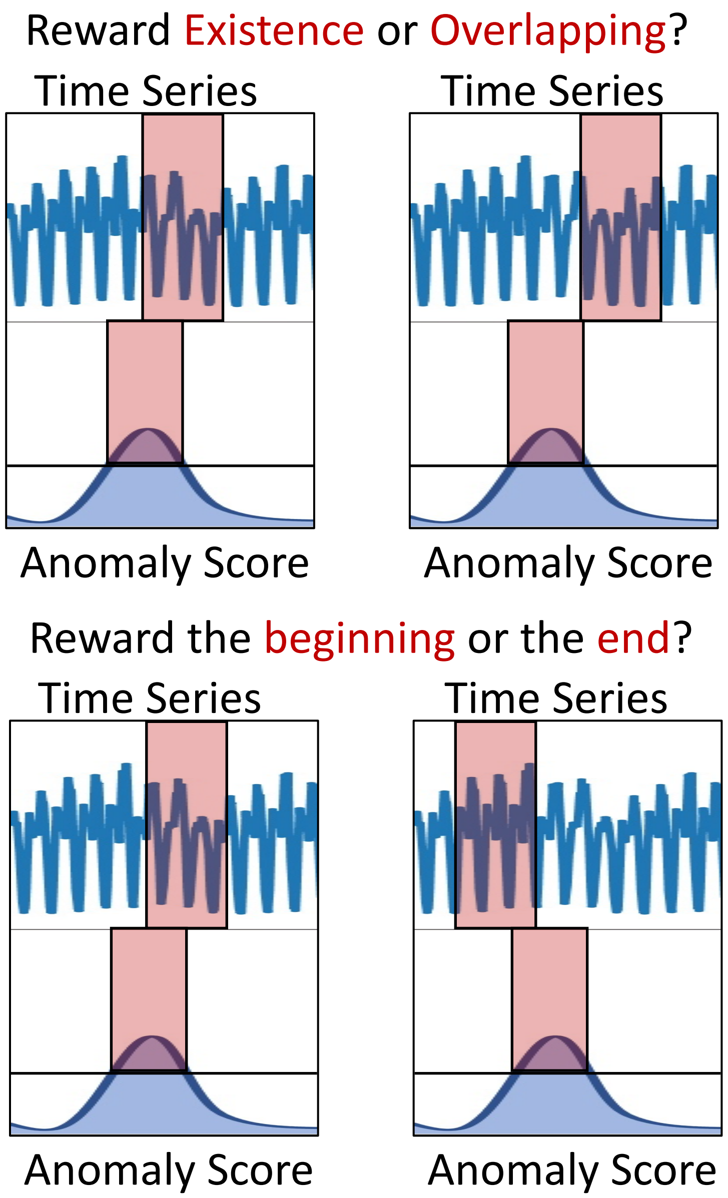
Existing solutions:

- Range Precision and Recall [23]:

- $Recall_T(R,P) = \frac{\sum_{i=1}^{N_r} Recall_T(R_i,P)}{N_r}$
- $Recall_T(R\_i,P) = \alpha * ExistenceR(R_i,P) + (1-\alpha) * OverlappingR(R\_i,P)$

- $Precision_T(R,P) = \frac{\sum_{i=1}^{N_p} Precision_T(R,P_i)}{N_p}$
- $Precision_T(R,P_i) = CardinalityFactor(P_i,R) * \sum_{j=1}^{N_r} w\left(P_i, P_i \cap R_j, \delta\right)$

- Functions $w(\,), \delta(\,)$ are tunable functions to represent the overlap size and position respectively.

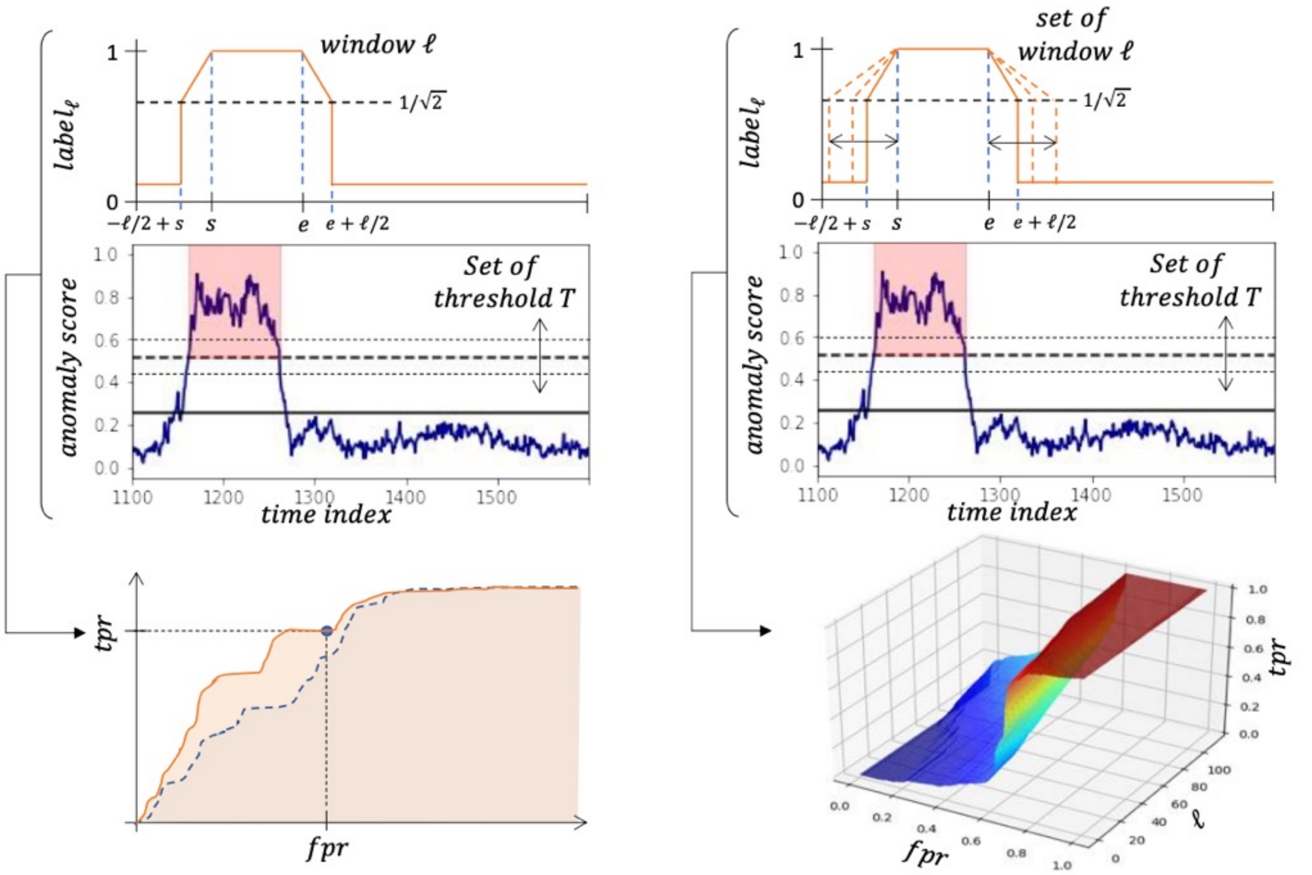


[23] Nesime Tatbul, Tae Jun Lee, Stan Zdonik, Mejbah Alam, and Justin Gottschlich. 2018. Precision and Recall for Time Series. In Advances in Neural Information Processing Systems, Vol. 31.

Evaluation measures: *Labeling issue*

Existing solutions:

- Volume Under the Surface [22] (VUS):
- Modify the labels with buffer regions at the beginning and at the end of an anomaly
- We vary the buffer size (as well as the threshold) and we obtain a surface
- We use the volume under the surface (VUS) as accuracy



(c) R-AUC-based Accuracy measure (ex: R-AUC-ROC, R-AUC-PR)

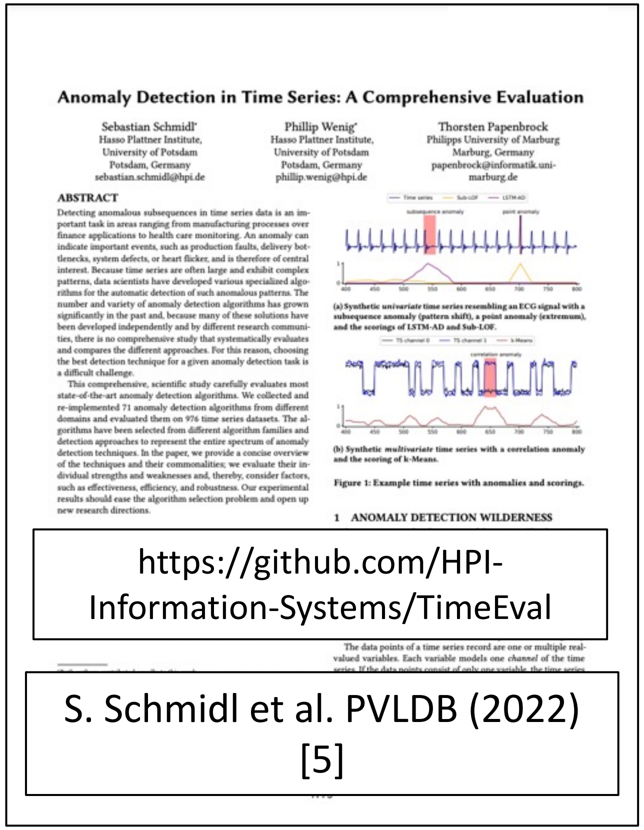
(d) VUS-based Accuracy measure (ex: VUS-ROC, VUS-PR)

[22] J. Paparrizos, P. Boniol, T. Palpanas, R. S. Tsay, A. Elmore, and M. J. Franklin. Volume under the surface: a new accuracy evaluation measure for time-series anomaly detection. Proc. VLDB Endow. 15, 11 (2022), 2774–2787.

Conclusion and Open Problems

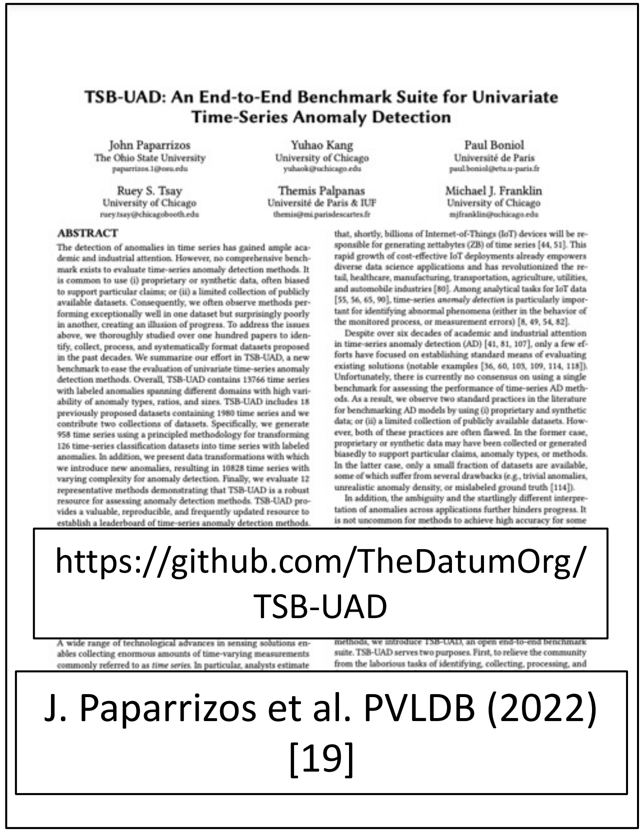
Conclusion and Open Problems

If you are interested in anomaly detection in time series…



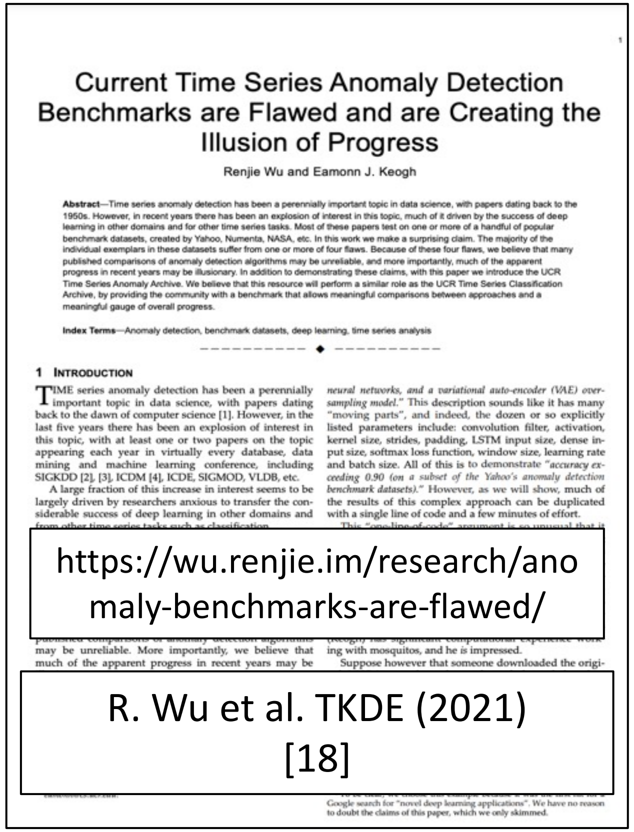
https://github.com/HPI-Information-Systems/TimeEval

S. Schmidl et al. PVLDB (2022) [5]



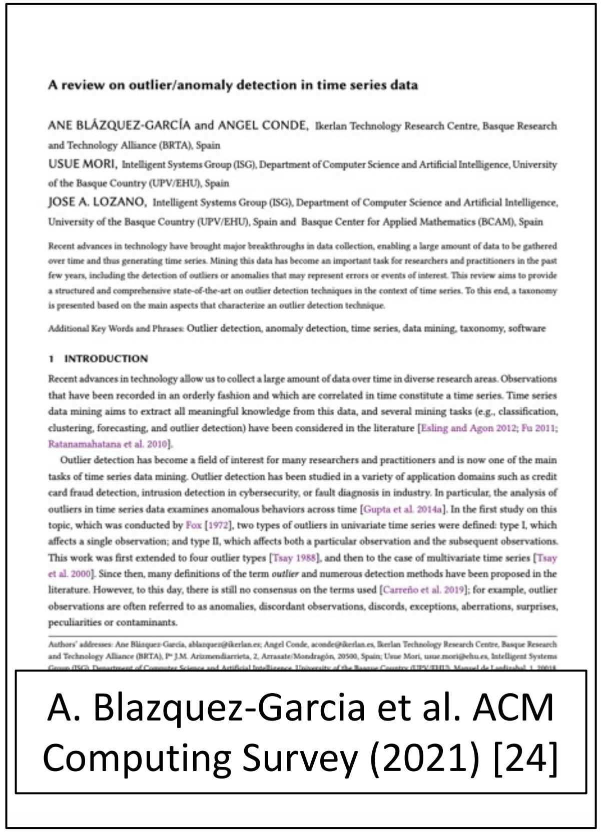
https://github.com/TheDatumOrg/TSB-UAD

J. Paparrizos et al. PVLDB (2022) [19]



https://wu.renjie.im/research/anomaly-benchmarks-are-flawed/

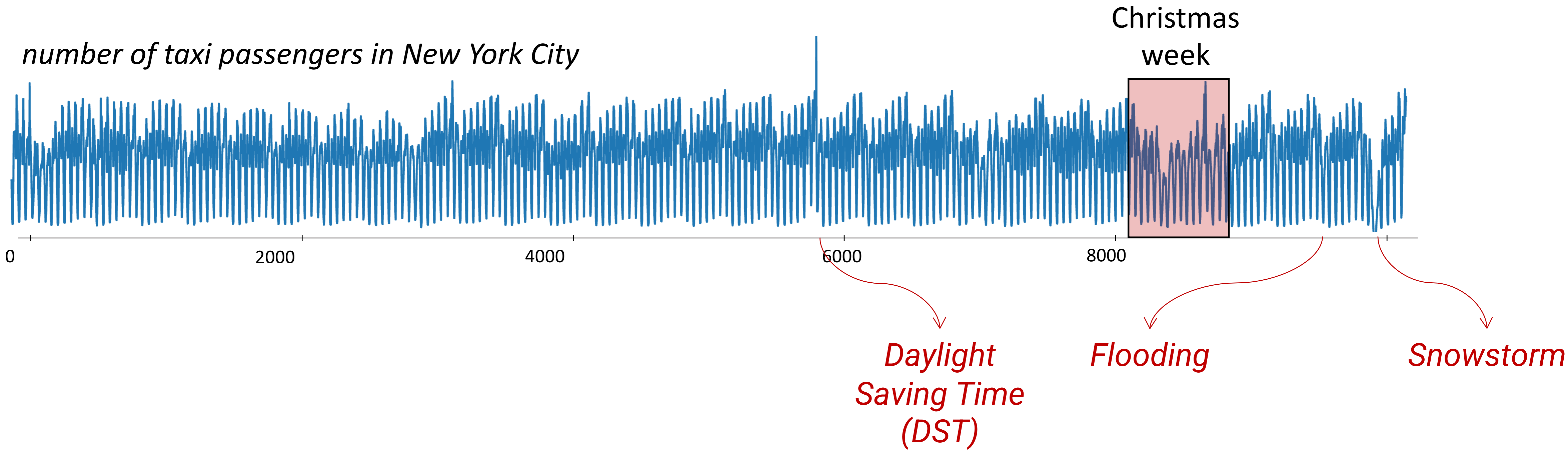
R. Wu et al. TKDE (2021) [18]



A. Blazquez-Garcia et al. ACM Computing Survey (2021) [24]

Conclusion and Open Problems

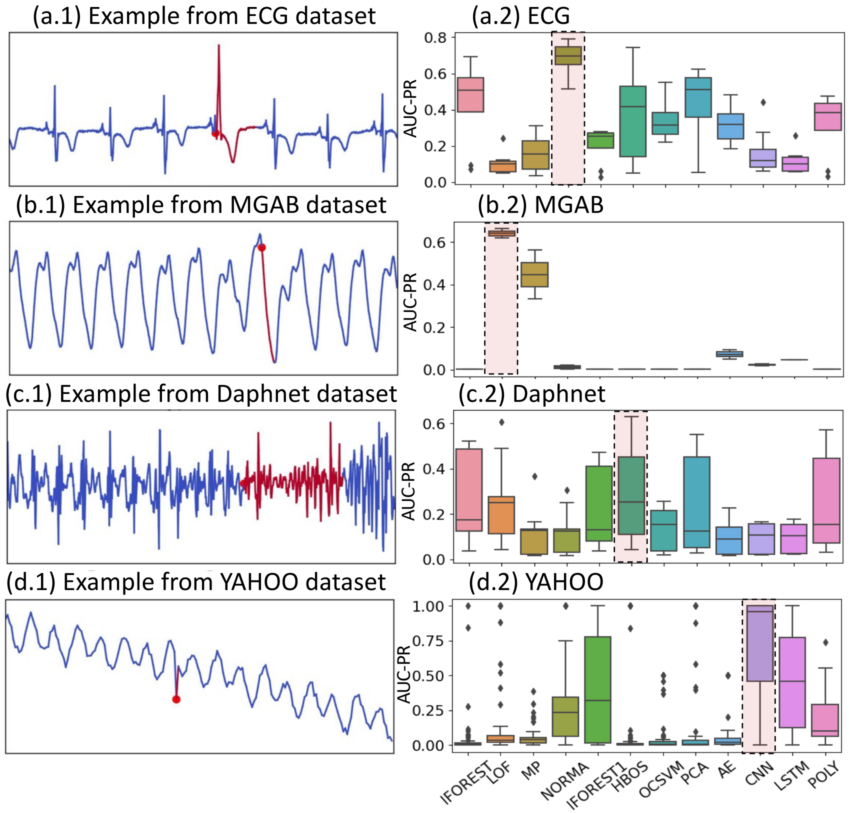
Context-aware Unsupervised Anomaly Detection



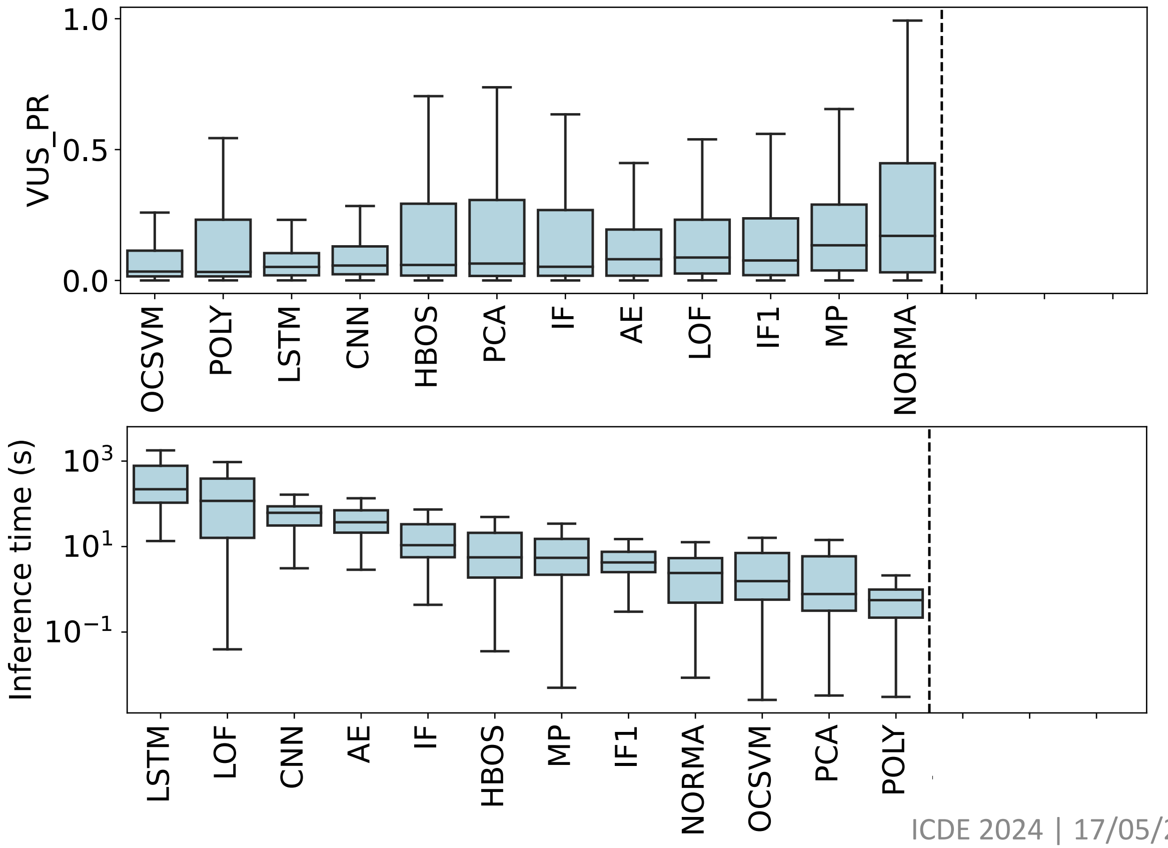
Conclusion and Open Problems

Model selection for anomaly detection

Methods ranking changes
significantly between datasets [19]



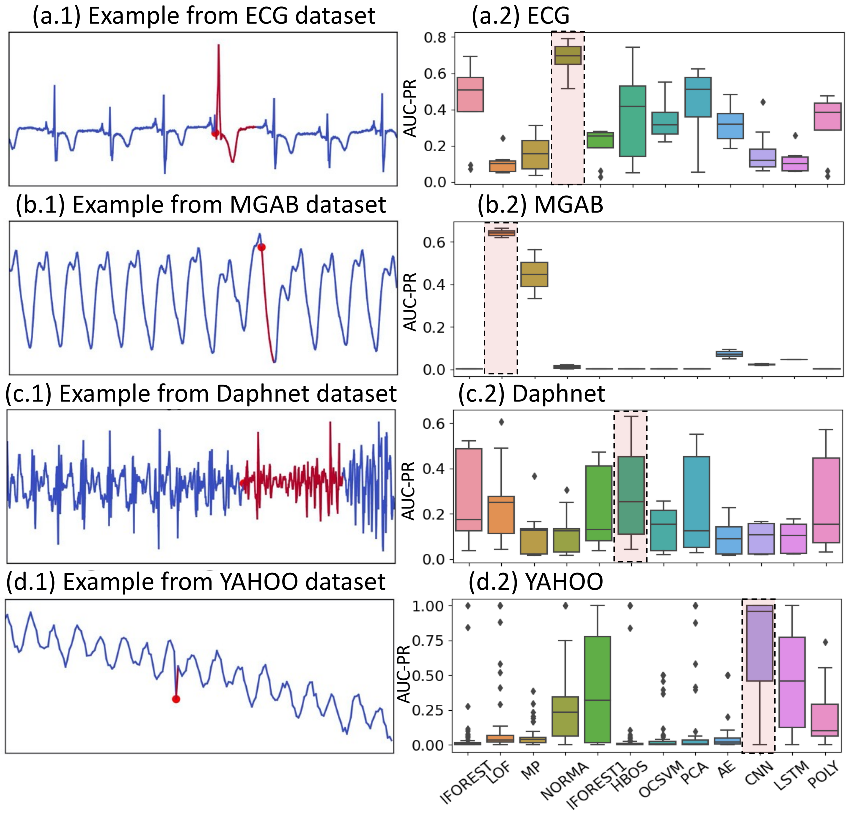
Results over TSB-UAD



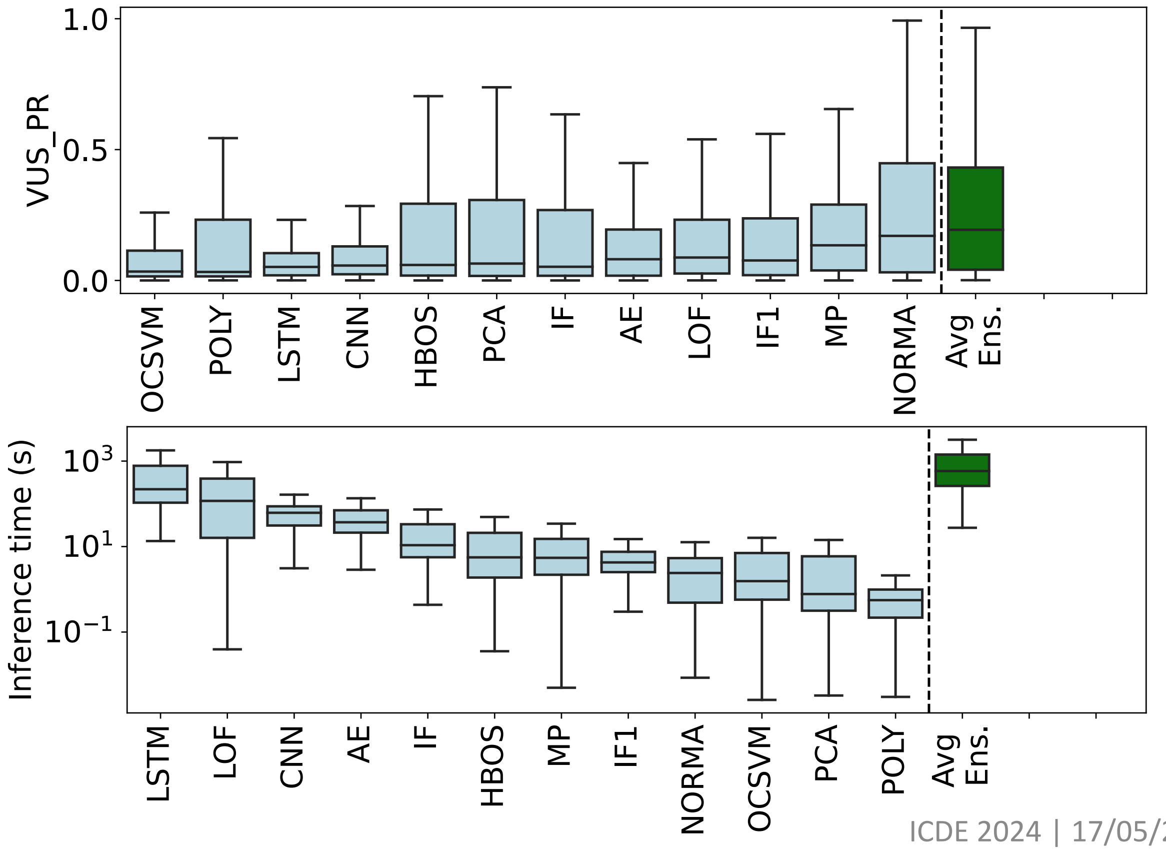
Conclusion and Open Problems

Model selection for anomaly detection

Methods ranking changes
significantly between datasets [19]



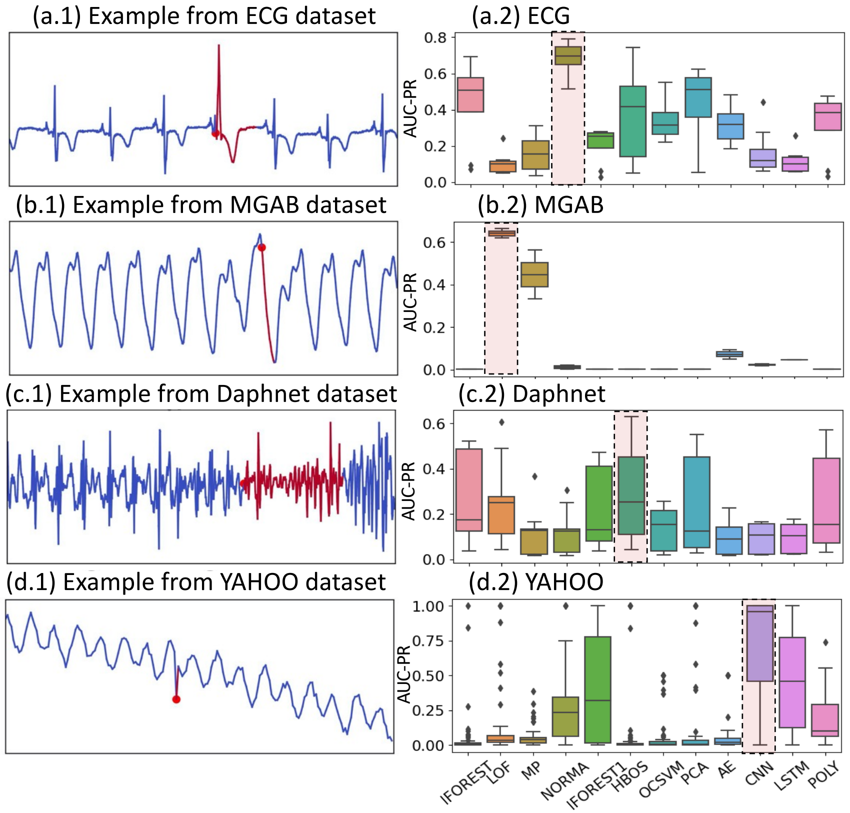
Can *Ensembling* methods solve the
problem?



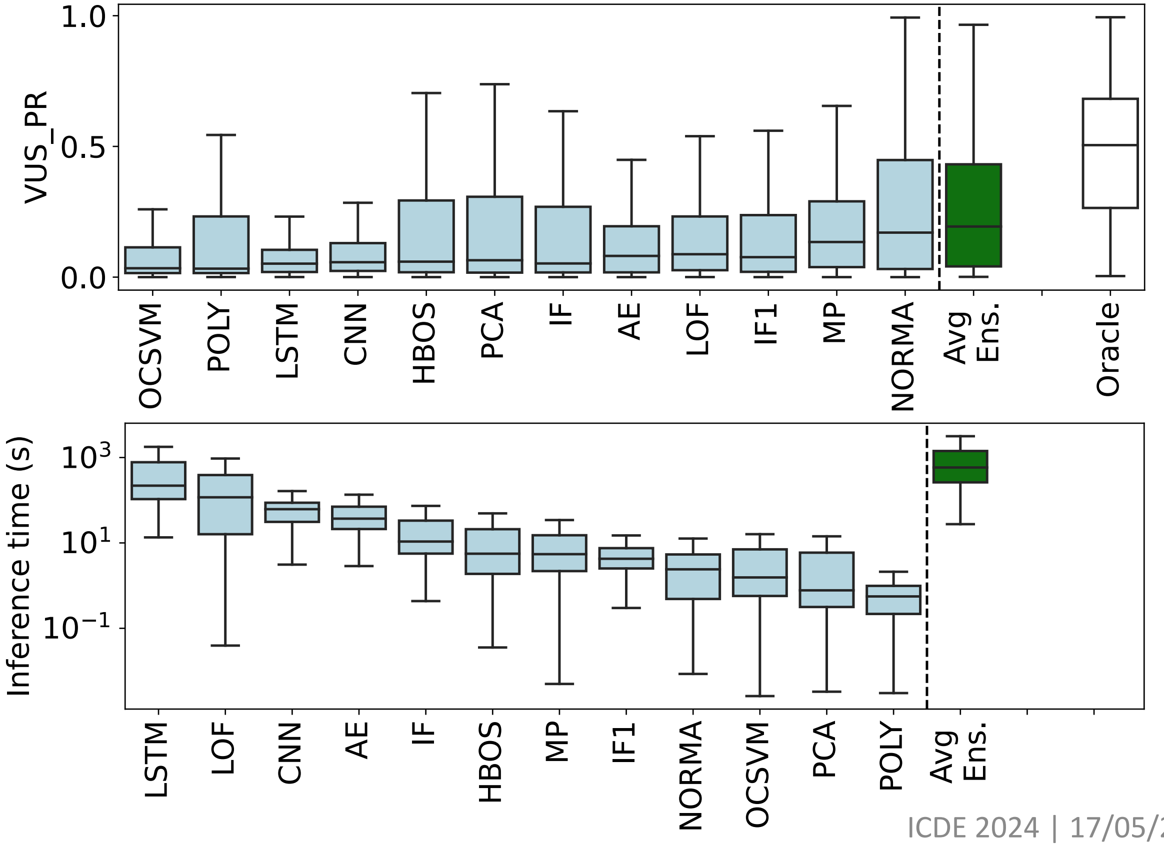
Conclusion and Open Problems

Model selection for anomaly detection

Methods ranking changes significantly between datasets [19]



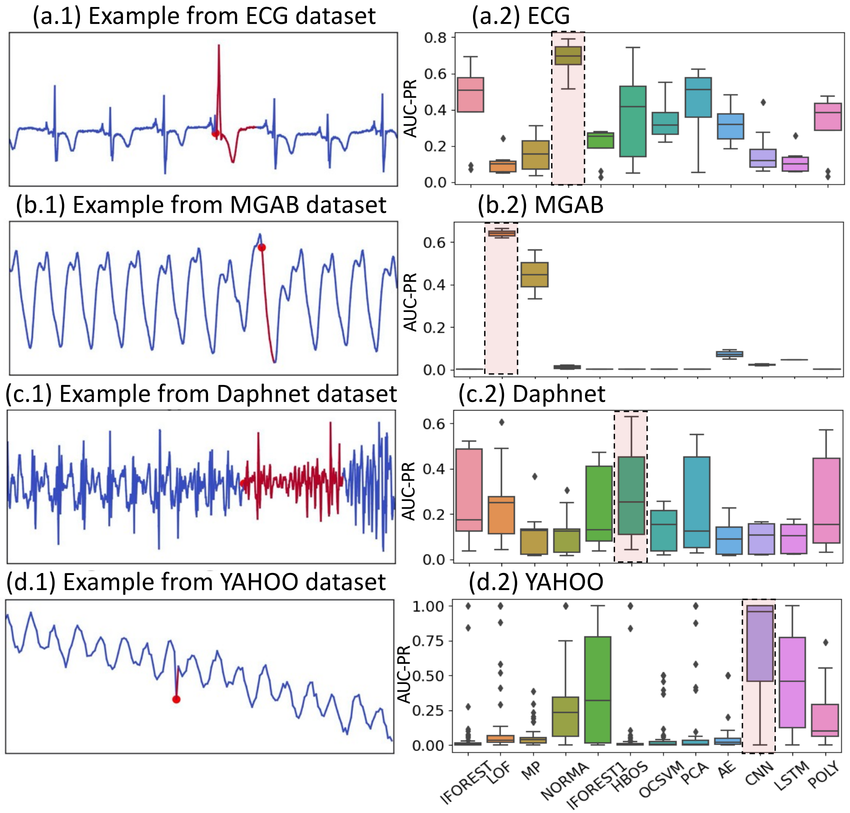
Can *automatic model selection solve the problem*?



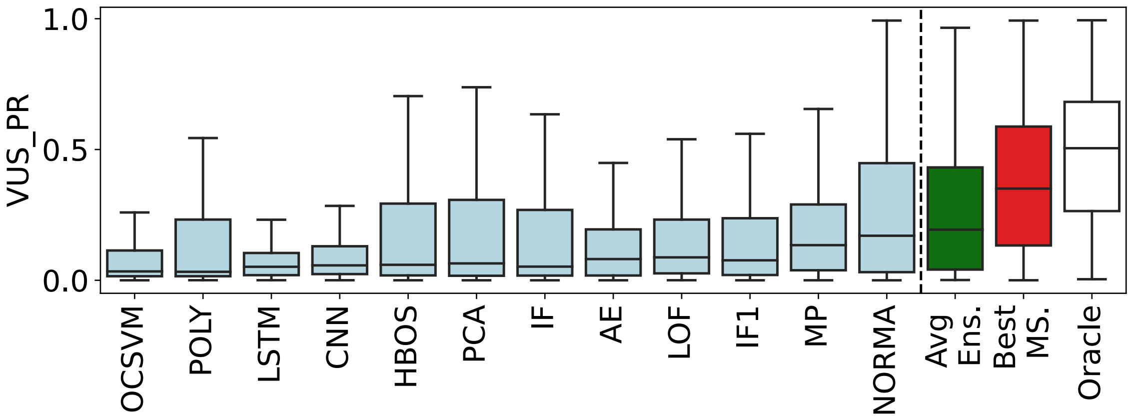
Conclusion and Open Problems

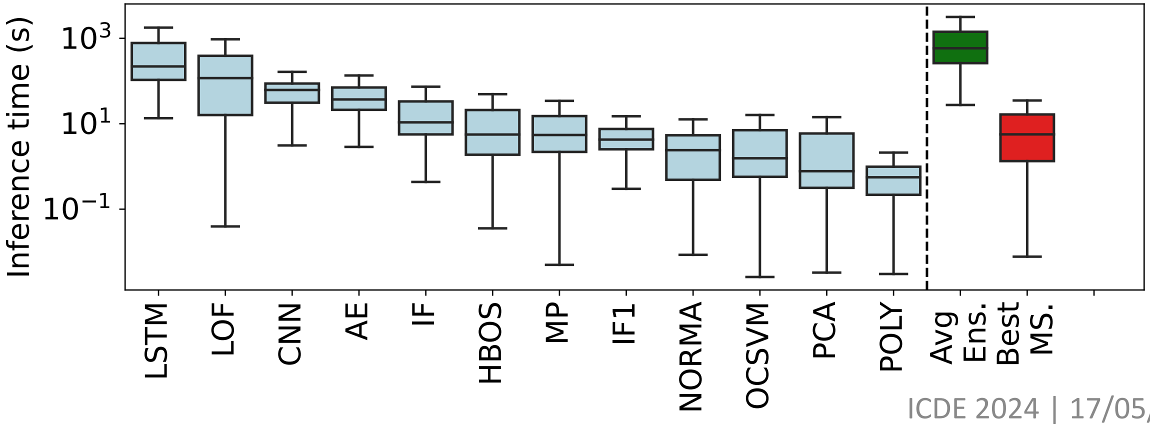
Model selection for anomaly detection

Methods ranking changes significantly between datasets [19]



Can *automatic model selection solve the problem*?

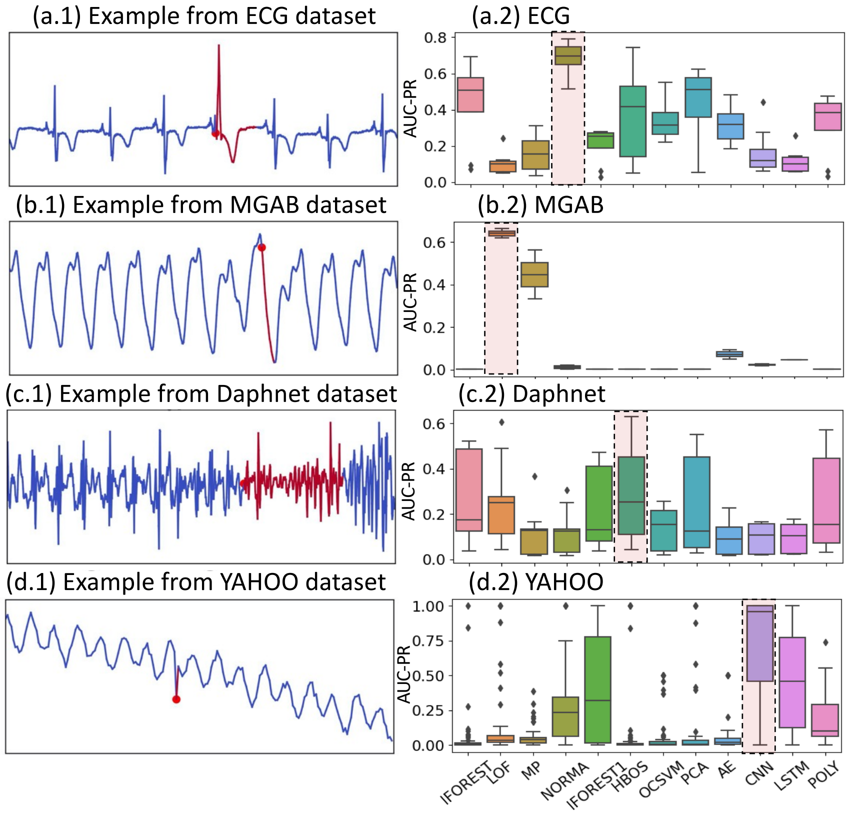




Conclusion and Open Problems

Model selection for anomaly detection

Methods ranking changes significantly between datasets [19]



Can *automatic model selection solve the problem*?

Choose Wisely [29]

An experimental evaluation of model selection for time series anomaly detection

VLDB 2023 *ICDE 2024*

[29] Emmanouil Sylligardos, Paul Boniol, John Paparrizos, Panos Trahanias, and Themis Palpanas. 2023. Choose Wisely: An Extensive Evaluation of Model Selection for Anomaly Detection in Time Series. Proc. VLDB Endow. 16, 11 (July 2023), 3418–3432.

References

[1] N. Laptev, S. Amizadeh, and Y. Billawala. 2015. **S5 - A Labeled Anomaly Detection Dataset**, version 1.0(16M).

[2] Markus Thill, Wolfgang Konen, and Thomas Bäck. 2020. **MGAB: The Mackey-Glass Anomaly Benchmark**.

[3] Pawel Benecki, Szymon Piechaczek, Daniel Kostrzewa, and Jakub Nalepa. 2021. **Detecting Anomalies in Spacecraft Telemetry Using Evolutionary Thresholding and LSTMs**. In Proceedings of the Genetic and Evolutionary Computation Conference Companion (Lille, France) (GECCO '21)

[4] Scott David Greenwald. 1990. **Improved detection and classication of arrhythmias in noise-corrupted electrocardiograms using contextual information**. Thesis. Massachusetts Institute of Technology.

[5] Sebastian Schmidl, Phillip Wenig, and Thorsten Papenbrock. 2022. **Anomaly detection in time series: a comprehensive evaluation**. Proc. VLDB Endow. 15, 9 (May 2022), 1779–1797.

[6] Chin-Chia Michael Yeh, Yan Zhu, Liudmila Ulanova, Nurjahan Begum, Yifei Ding, Hoang Anh Dau, Diego Furtado Silva, Abdullah Mueen, and Eamonn J. Keogh. 2016. **Matrix Prole I: All Pairs Similarity Joins for Time Series.** In ICDM.

[7] Yan Zhu, Zachary Zimmerman, Nader Shakibay Senobari, Chin-Chia Michael Yeh, Gareth Funning, Abdullah Mueen, Philip Brisk, and Eamonn Keogh. 2016. **Matrix Profile II: Exploiting a Novel Algorithm and GPUs to Break the One Hundred Million Barrier for Time Series Motifs and Joins.** In Proceedings of the International Conference on Data Mining (ICDM), 739–748.

[8] Yue Lu, Renjie Wu, Abdullah Mueen, Maria A. Zuluaga, and Eamonn Keogh. 2022. **Matrix Profile XXIV: Scaling Time Series Anomaly Detection to Trillions of Datapoints and Ultra-fast Arriving Data Streams.** In Proceedings of the 28th ACM SIGKDD.

[9] C. -C. M. Yeh, N. Kavantzas and E. Keogh, **Matrix Profile VI: Meaningful Multidimensional Motif Discovery**, 2017 IEEE International Conference on Data Mining (ICDM), New Orleans, LA, USA, 2017, pp. 565-574, doi: 10.1109/ICDM.2017.66. Data Mining (KDD '22).

[10] Paul Boniol, Michele Linardi, Federico Roncallo, Themis Palpanas, Mohammed Meftah, and Emmanuel Remy. 2021. **Unsupervised and scalable subsequence anomaly detection in large data series.** The VLDB Journal 30, 6 (Nov 2021), 909–931.

[11] F. T. Liu, K. M. Ting and Z. -H. Zhou, **Isolation Forest**, 2008 Eighth IEEE International Conference on Data Mining, Pisa, Italy, 2008, pp. 413-422

[12] Markus Goldstein and Andreas Dengel. 2012. **Histogram-based outlier score (hbos): A fast unsupervised anomaly detection algorithm**. KI-2012: poster and demo track 9 (2012).

[13] Paul Boniol and Themis Palpanas. 2020. **Series2Graph: graph-based subsequence anomaly detection for time series.** Proc. VLDB Endow. 13, 12 (August 2020), 1821–1834.

[14] Ali Abdul-Aziz, Mark R Woike, Nikunj C Oza, Bryan L Matthews, and John D lekki. 2012. **Rotor health monitoring combining spin tests and data-driven anomaly detection methods**. Structural Health Monitoring (2012).

[15] Pankaj Malhotra, Lovekesh Vig, Gautam Shro, and Puneet Agarwal. 2015. **Long Short Term Memory Networks for Anomaly Detection in Time Series**. (2015).

[16] M. Munir, S. A. Siddiqui, A. Dengel, and S. Ahmed. 2019. **DeepAnT: A Deep Learning Approach for Unsupervised Anomaly Detection in Time Series**. IEEE Access 7 (2019), 1991–2005.

[17] Mayu Sakurada and Takehisa Yairi. 2014. **Anomaly Detection Using Autoencoders with Nonlinear Dimensionality Reduction**. In Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis (Gold Coast, Australia QLD, Australia) (MLSDA'14).

[18] R. Wu and E. Keogh, **Current Time Series Anomaly Detection Benchmarks are Flawed and are Creating the Illusion of Progress** in IEEE Transactions on Knowledge & Data Engineering, vol. 35, no. 03, pp. 2421-2429, 2023.

[19] John Paparrizos, Yuhao Kang, Paul Boniol, Ruey S. Tsay, Themis Palpanas, and Michael J. Franklin. 2022. **TSB-UAD: an end-to-end benchmark suite for univariate time-series anomaly detection**. Proc. VLDB Endow. 15, 8 (April 2022), 1697–1711.

[20] Tom Fawcett. 2006. **An introduction to ROC analysis**. Pattern Recognition Letters 27, 8 (2006), 861–874.

[21] Jesse Davis and Mark Goadrich. 2006. **The Relationship between Precision-Recall and ROC Curves**. In Proceedings of the 23rd International Conference on Machine Learning (ICML '06).

[22] John Paparrizos, Paul Boniol, Themis Palpanas, Ruey S. Tsay, Aaron Elmore, and Michael J. Franklin. 2022. **Volume under the surface: a new accuracy evaluation measure for time-series anomaly detection**. Proc. VLDB Endow. 15, 11 (July 2022), 2774–2787.

[23] Nesime Tatbul, Tae Jun Lee, Stan Zdonik, Mejbah Alam, and Justin Gottschlich. 2018. **Precision and Recall for Time Series**. In Advances in Neural Information Processing Systems, Vol. 31.

[24] Ane Blázquez-García, Angel Conde, Usue Mori, and Jose A. Lozano. 2021. **A Review on Outlier/Anomaly Detection in Time Series Data**. ACM Comput. Surv. 54, 3, Article 56 (April 2022), 33 pages.

[25] Paul Boniol, John Paparrizos, Themis Palpanas, and Michael J. Franklin. 2021. **SAND: streaming subsequence anomaly detection**. Proc. VLDB Endow. 14, 10 (June 2021), 1717–1729.

[26] Schneider, J., Wenig, P. & Papenbrock, T. **Distributed detection of sequential anomalies in univariate time series**. The VLDB Journal **30**, 579–602 (2021).

Thank you for attending!

Any Questions?