An Automated Forecasting Framework based on Method Recommendation for Seasonal Time Series

André Bauer, Marwin Züfle, Johannes Grohmann, Norbert Schmitt,
Nikolas Herbst, and Samuel Kounev

Chair of Software Engineering, University of Würzburg

11th ACM/SPEC International Conference on Performance Engineering

https://se.informatik.uni-wuerzburg.de/

Motivation

- Fast living and changing requirements
  - In the course of digitalization, new services and applications will arise
  - The amount of connected devices will increase
- Cloud Computing allows flexibility
  - Complexity exceeds human capacity
  - “Naive” resource allocation
    - Consumes 2% of the world‘s electricity
    - Emits as much CO_2 as the airline industry[1]

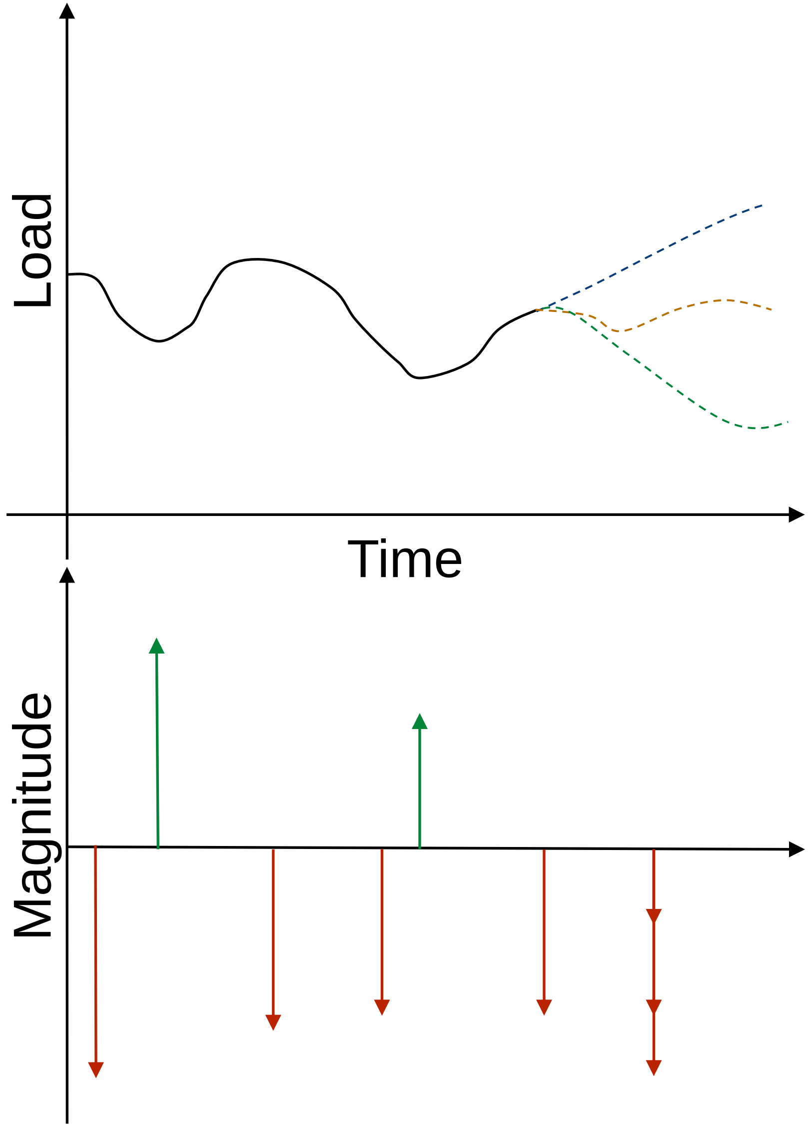
[1] https://e360.yale.edu/features/energy-hogs-can-huge-data-centers-be-made-more-efficient

Problem Visualization

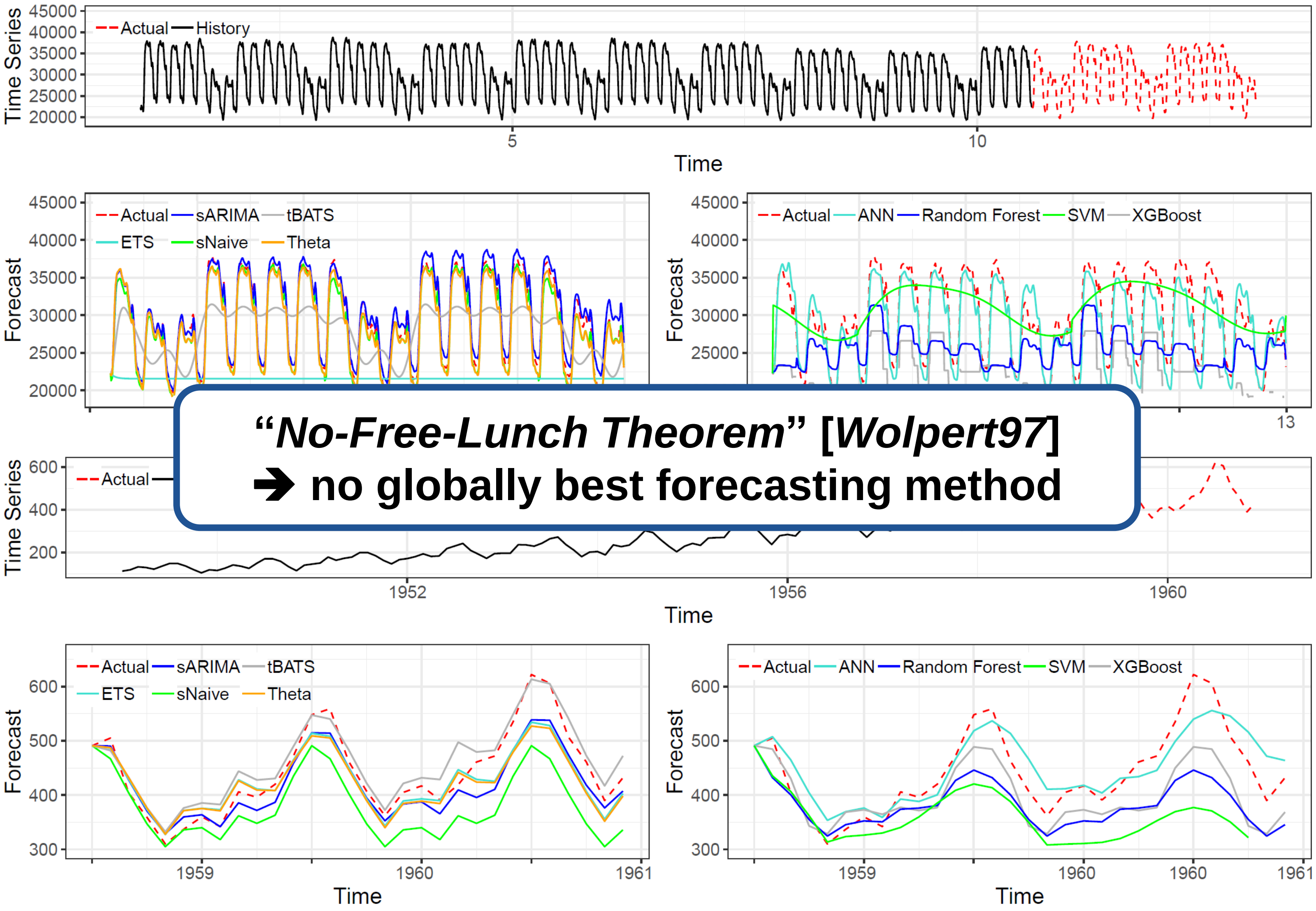
- ➢ Resources shall be supplied
  - Automatically and
  - In advance

- ➢ Adapt the amount of resources
  - When?
  - How much?

Reliable Time Series Forecasting Method



No-Free-Lunch



"*No-Free-Lunch Theorem*" [*Wolpert97*]
➔ no globally best forecasting method

Related Work – Hybrid Forecasting

Ensemble Forecasting	Forecaster Recommendation	Decomposition Forecasting
- Historically first hybrid forecasting task	- Learning a rule set from a set of time series	- Combining advantages of different methods
- Weighting results from several methods	- System guesses best forecasting method	- Decomposition of time series into its components
- Linear combination of these weighted results	- Expert system or machine learning techniques	- Executing several methods one after another
[*Bates69, Clemen89, Boulegane19, more*]	[*Collopy92, Wang09, Montero20, more*]	[*Zhang03, Züfle17, Bauer20, more*]

Contributions

Research Question I

How to build a hybrid forecasting framework for seasonal time series?

Research Question II

What are suitable time series characteristics for the recommendation?

Research Question III

What are suitable approaches for the recommendation system?

Contribution I

Automated framework for seasonal time series based on decomposition.

Contribution II & III

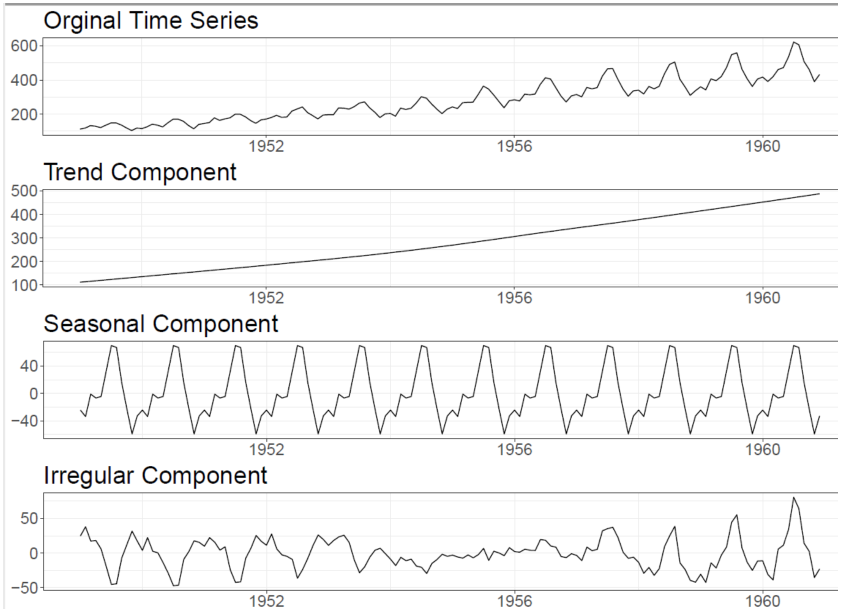
New characteristics and three recommendation approaches.

Contribution IV

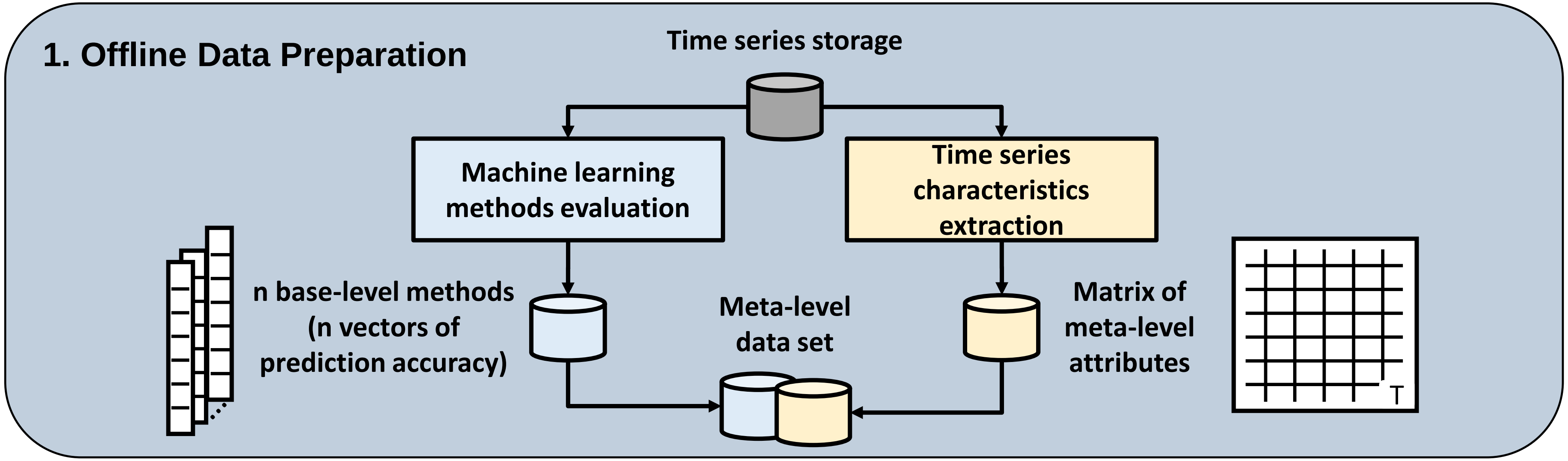
Evaluation of the approaches and discussion.

Forecasting Framework – Underlying Idea

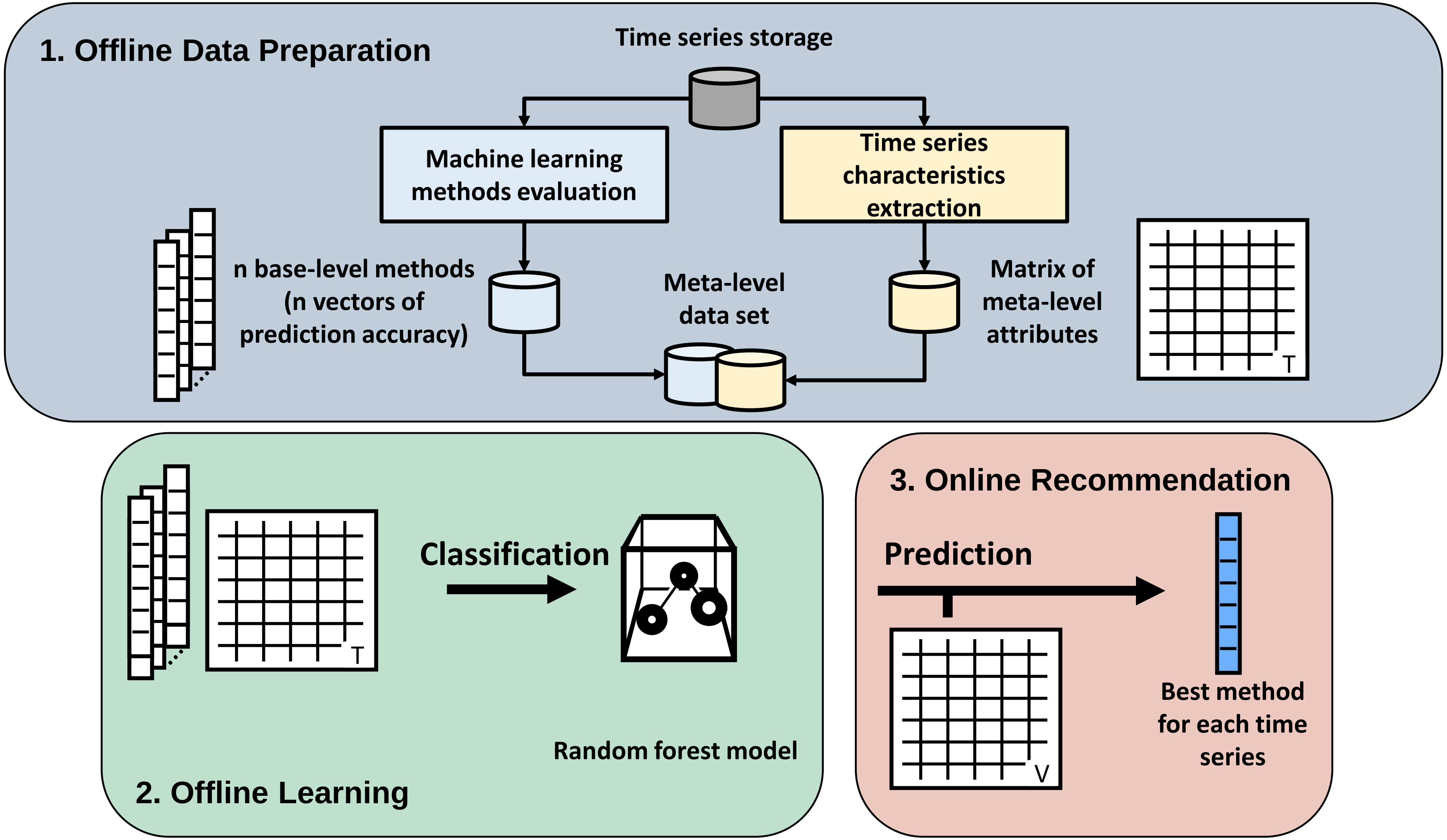
- Simplifying time series with logarithm transformation
- Extracting time series features
  - Fourier Terms
  - Season
  - Trend
  - Remainder
- Consulting recommendation system for best suitable machine learning method
- Forecasting time series
  - Forecasting each feature with individual methods
  - Train and predict machine learning method with features



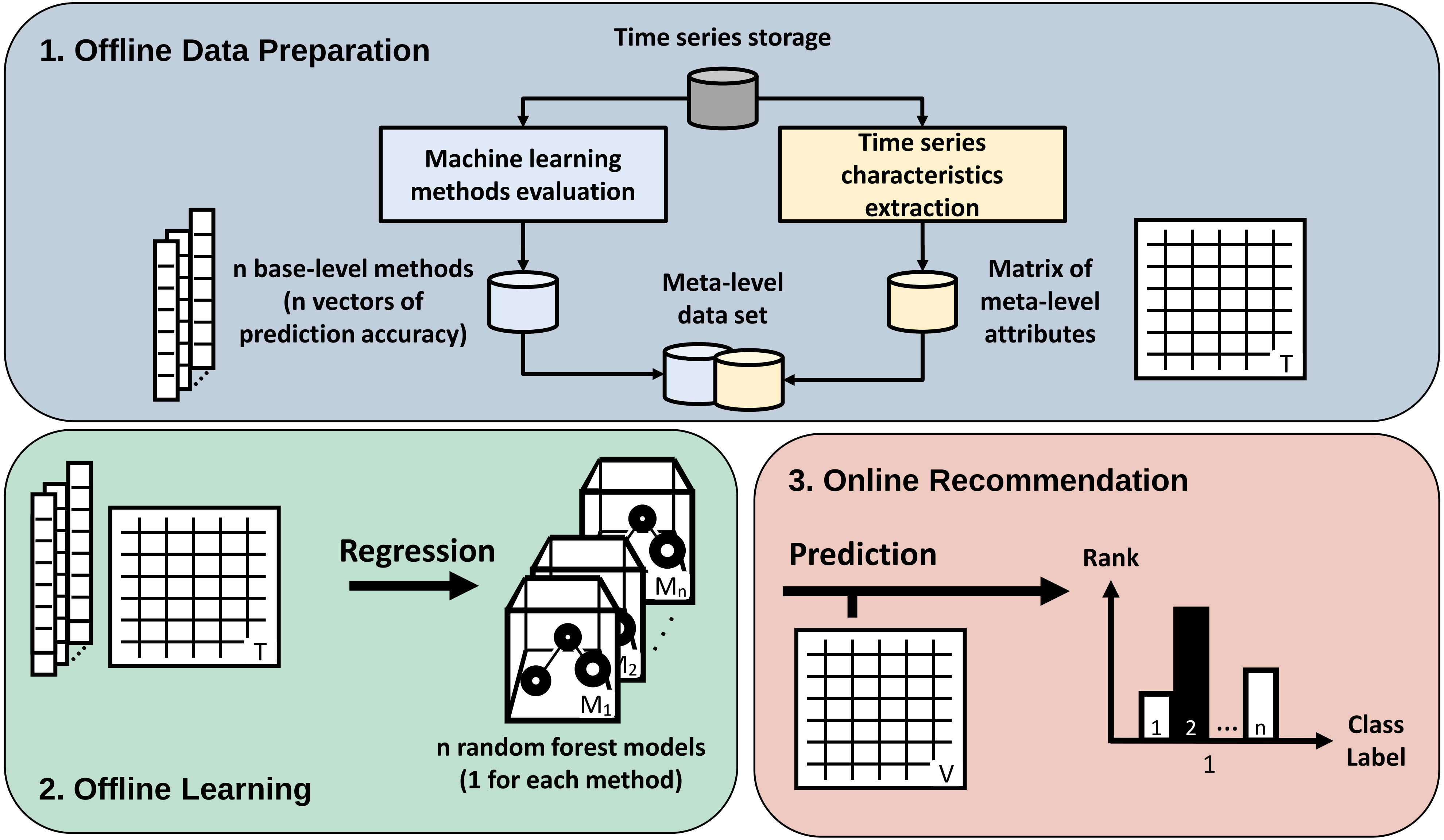
Recommendation Offline Training



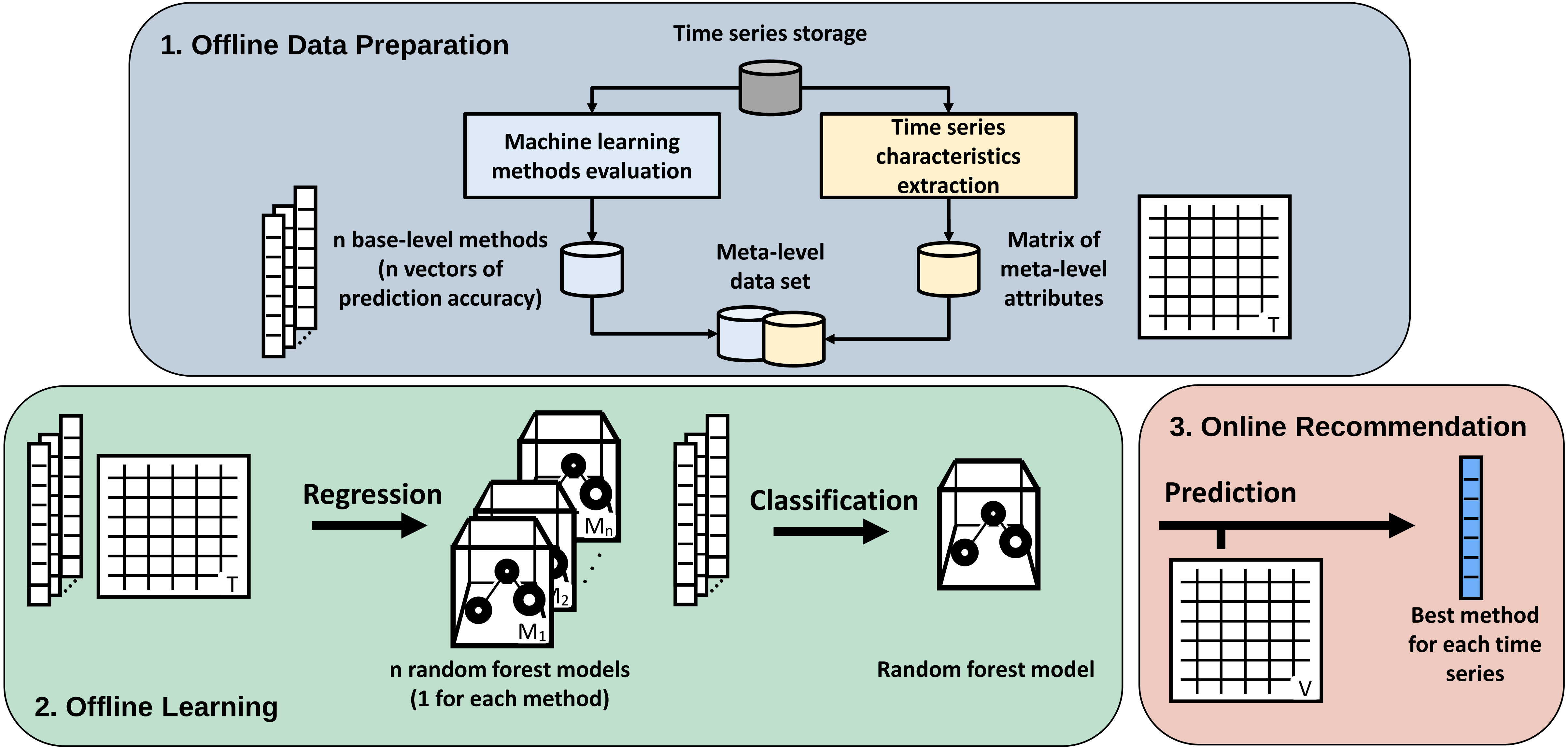
Recommendation System A_C



Recommendation System A_R



Recommendation System A_H



Evaluation Design

- Machine learning methods:
  - Catboost [Prokhorenkova18]
  - Cubist [Quinlan93]
  - Evtree [Grubinger14]
  - Feed-forward NN [Hyndman14]
  - Random forest [Breiman11]
  - Rpart [Breiman93]
  - SVR [Drucker97]
  - XGBoost [Chen16]
- Evaluation measures:
  - Rank $\in \{1, 2, 3, 4,5,6,7,8\}$
  - Accuracy degradation compared to the best method
- Compared recommendation approaches:
  - S* best method a-posteriori
  - S_L lowest average degradation
  - S_B most often the best method
  - A_C classification approach
  - A_R regression approach
  - A_H hybrid approach
  - State-of-the art forecasting methods
- Data set
  - 150 seasonal time series from various sources
  - 100 splits (100 train, 50 test); from train data set 10,000 new time series are created

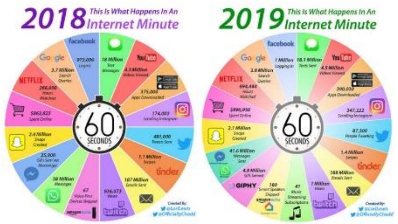
Evaluation

Degradation	S*	S_L	S_B	A_C	A_R	A_H
Mean	1.000	1.409	1.249	1.235	1.172	**1.159**
Median	1.000	1.045	1.076	**1.016**	1.035	1.032
SD	0.000	3.674	0.427	2.458	1.382	**0.382**

MAPE	A_C	A_R	A_H	ETS	tBATS	sARIMA
Mean	24.40	**23.26**	23.68	56.96	36.28	28.12
Median	12.31	13.07	13.18	14.47	**10.83**	13.00
SD	50.31	40.41	**38.52**	136.22	98.68	64.72

- In terms of recommendation, the hybrid version has the lowest mean and standard deviation
- Our approach has a lower mean forecast error and standard deviation than the forecasting methods

Autonomic Forecasting Method Selection Takeaways



- Forecasting is an important task for many decision making fields, for instance, Cloud Computing

- There is no globally best forecasting method (“*No-Free-Lunch Theorem*”)
- We propose an automated forecasting approach based on decomposition and method recommendation
- Our experimental results show that…
  - Our recommendation approaches perform almost equally
  - The whole forecasting approach outperforms existing state-of-the-art forecasting methods

References

- **Wolpert97**: D. H. Wolpert and W. G. Macready, “No free lunch theorems for optimization,” IEEE Trans. on Evolutionary Computation, vol. 1, no. 1, 1997.
- **Bates69**: J. M. Bates and C. W. Granger, “The combination of forecasts,” Journal of the Operational Research Society, vol. 20, no. 4, 1969.
- **Clemen89**: R. T. Clemen, “Combining forecasts: A review and annotated bibliography,” Int. Journal of Forecasting, vol. 5, no. 4, 1989.
- **Boulegane2019**: Boulegane, D., Bifet, A., and Madhusudan, G. (2019). “Arbitrated Dynamic Ensemble with Abstaining for Time-Series Forecasting on Data Streams”. In: 2019 IEEE International Conference on Big Data (Big Data). IEEE, pp. 1040–1045.
- **Collopy92**: F. Collopy and J. S. Armstrong, “Rule-based forecasting: Development and validation of an expert systems approach to combining time series extrapolations,” Management Science, vol. 38, no. 10, 1992.
- **Wang09**: X. Wang, K. Smith-Miles, and R. Hyndman, “Rule induction for forecasting method selection: Meta-learning the characteristics of univariate time series,” Neurocomputing, vol. 72, no. 10-12, 2009.
- **Montero20**: Montero-Manso, Pablo, et al. "FFORMA: Feature-based forecast model averaging." International Journal of Forecasting 36.1 (2020): 86-92.
- **Zhang03**: G. P. Zhang, “Time series forecasting using a hybrid arima and neural network model,” Neurocomputing, vol. 50, 2003.
- **Züfle17**: M. Züfle, A. Bauer, N. Herbst et al., “Telescope: a hybrid forecast method for univariate time series,” in Int. Work-Conference on Time Series, 2017.
- **Bauer20**: A. Bauer, M. Züfle, N. Herbst, S. Kounev, and V. Curtef. Telescope: An automatic feature extraction and transformation approach for time series forecasting on a level-playing field. In Proceedings of the 36th International Conference on Data Engineering (ICDE), April 20-24, 2020.

References – Cont'd

- **Prokhorenkova18**: Liudmila Prokhorenkova, Gleb Gusev, Aleksandr Vorobev, Anna Veronika Dorogush, and Andrey Gulin. 2018. CatBoost: unbiased boosting with categorical features. In Advances in Neural Information Processing Systems. 6638–6648.
- **Quinlan93**: J Ross Quinlan. 1993. Combining instance-based and model-based learning. In Proceedings of the tenth international conference on machine learning. 236–243.
- **Grubinger14**: Thomas Grubinger, Achim Zeileis, and Karl-Peter Pfeiffer. 2014. evtree: Evolutionary Learning of Globally Optimal Classification and Regression Trees in R. Journal of Statistical Software, Articles 61, 1 (2014), 1–29.
- **Hyndman14**: Rob J Hyndman and George Athanasopoulos. 2014. Forecasting: principles and practice. OTexts, Melbourne, Australia.
- **Breiman11**: Leo Breiman. 2001. Random forests. Machine learning 45, 1 (2001), 5–32.
- **Breiman93**: Leo Breiman, JosephHFriedman, R. A. Olshen, and C. J. Stone. 1983. Classification and Regression Trees.
- **Drucker97**: Harris Drucker, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. 1997. Support vector regression machines. In Advances in neural information processing systems. 155–161.
- **Chen16**: Tianqi Chen and Carlos Guestrin. 2016. Xgboost: A scalable tree boosting system. In ACM SIGKDD 2016. ACM, 785–794.

Thank you for your attention!