INTERNATIONALE
HOCHSCHULE

Aeroconf 2024

DANIEL LAKEY
PROF. TIM SCHLIPPE

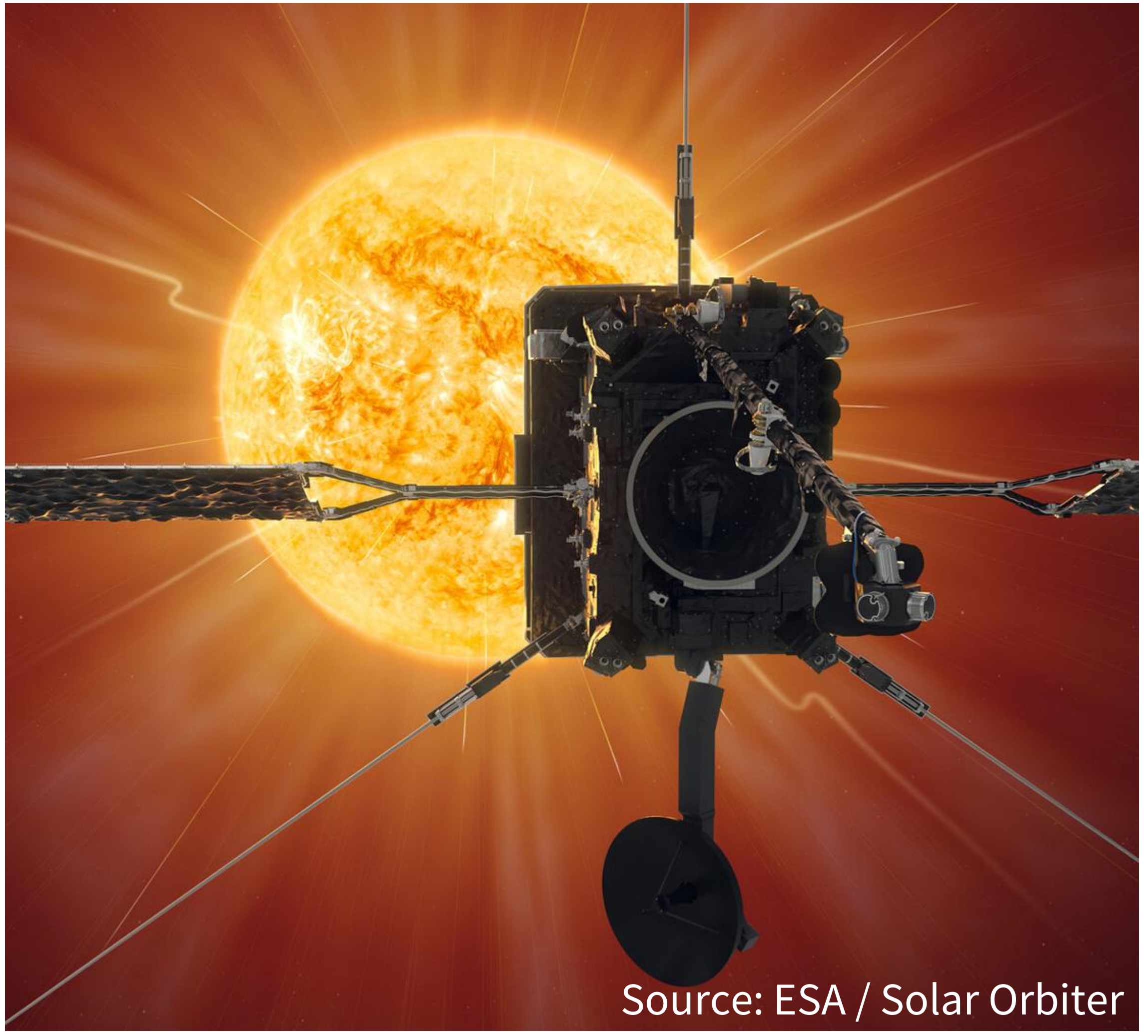
A COMPARISON OF DEEP LEARNING ARCHITECTURES FOR SPACECRAFT ANOMALY DETECTION

Big Sky, MT, 7th March 2024

Session: **12.05 Automation and Machine Learning Applications in Spacecraft Operations**

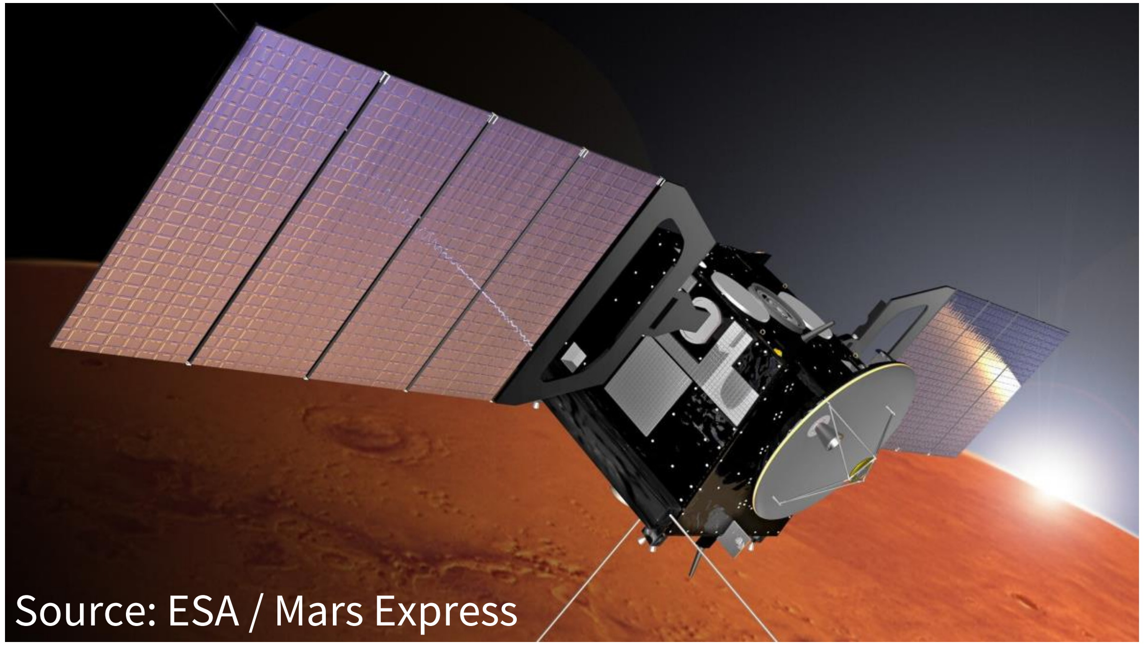
IEEE AEROSPACE CONFERENCE

PROBLEM MOTIVATION : SPACECRAFT ARE COMPLICATED



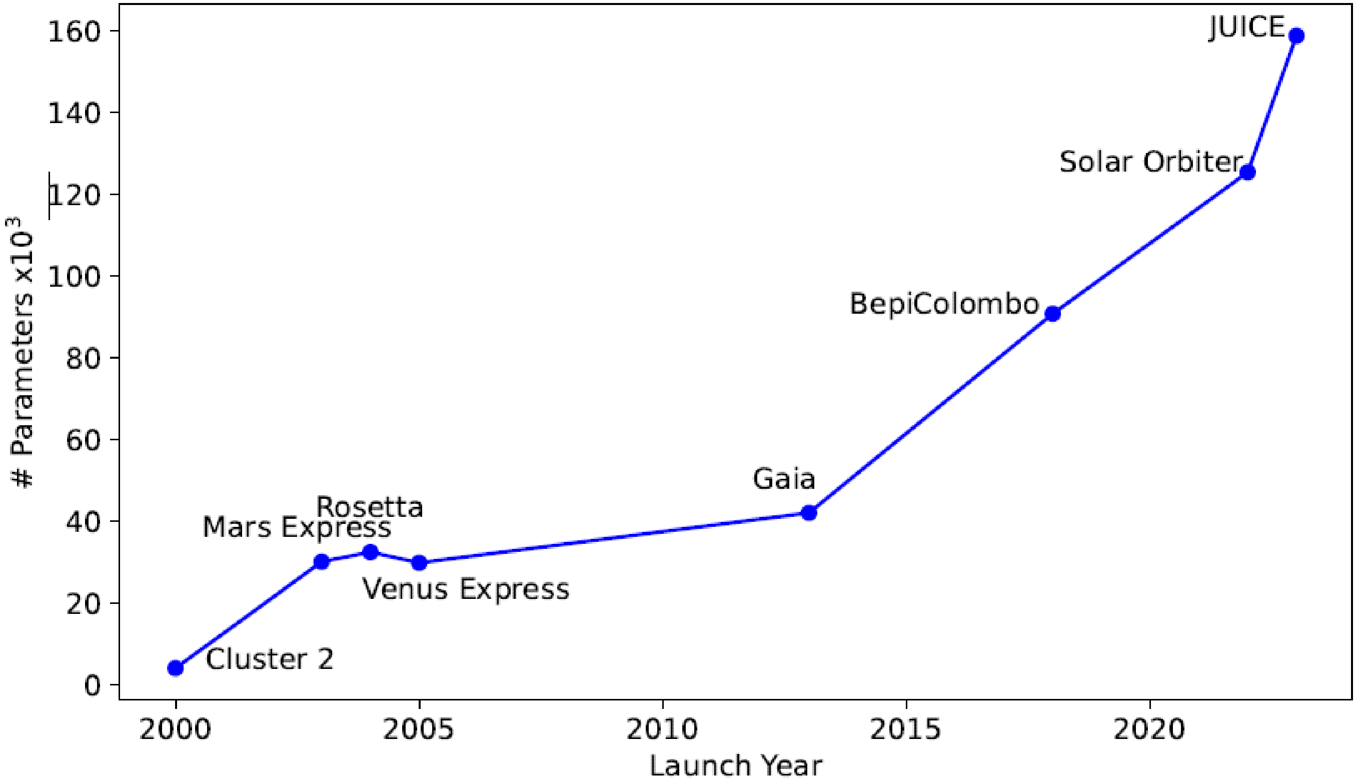
Source: ESA / Solar Orbiter

- Spacecraft are complicated
- Monitoring by expert engineers
- High/Low Limits for small number of TM
- Pressure to “use AI” in operations



Source: ESA / Mars Express

PROBLEM MOTIVATION: IT’S NOT GETTING EASIER

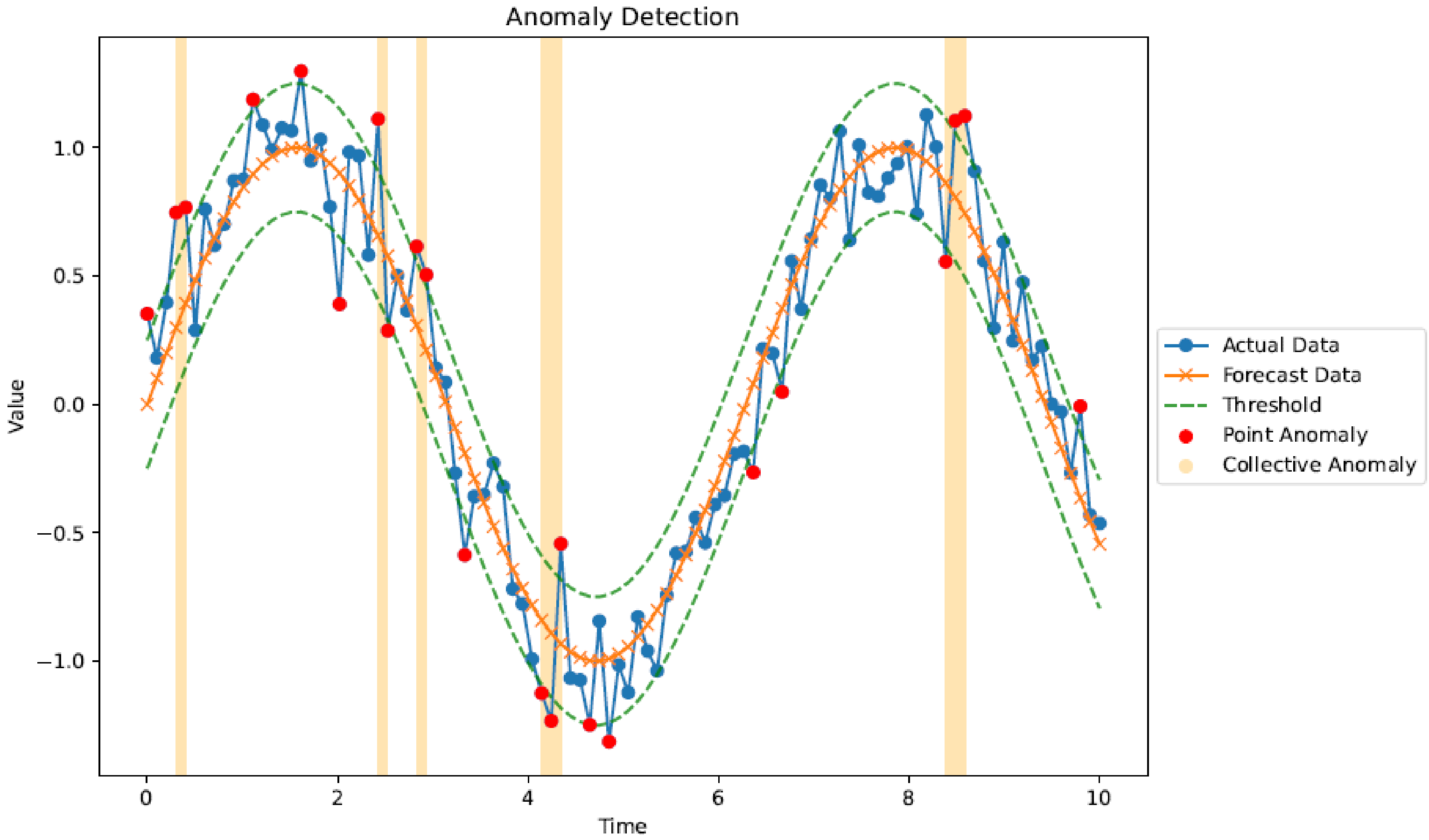


Increasing complexity of spacecraft over time

- ➔ **What anomaly detection approaches are out there?**
- ➔ **What performance do they have?**
- ➔ **Can we do better?**

CURRENT APPROACHES

- Forecasting & Threshold is by far the most common existing approach
  - Hundman et al. (2018), Heras & Donati (2014), Pilastre et al. (2020), Baireddy et al. (2021), Yang et al. (2021), and Wang et al. (2022)…



AVAILABLE DATA SETS

- Public data sets of real spacecraft data are hard to find
- Data sets with well-labelled anomalies even harder to find
- Hundman, *et al* (2018) provide labelled anomalies from MSL/SMAP, anonymised data and benchmark performance of LSTM-based forecasting & threshold approach: “Telemanom”
  - ➔ **We use this data set in our study**
  - Known limitations to dataset :
    - no anomalies in training set
    - high density of anomalies in test set
    - no indication of time span, or relation between telemetry channels

DEEP LEARNING APPROACHES FOR FORECASTING

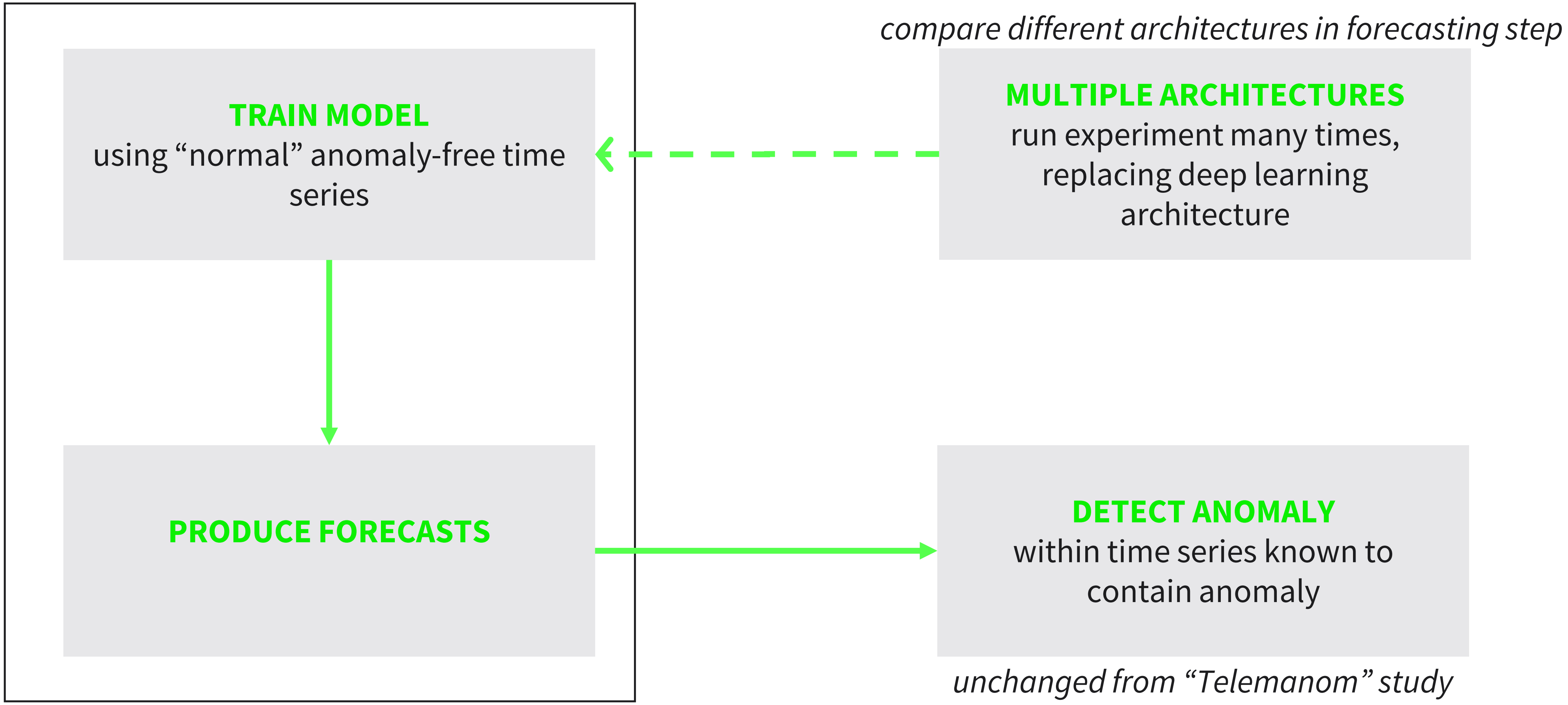
- Many deep-learning architectures are proposed for forecasting-based anomaly detection
  - (Walther et al., 2023, Schmidl et al., 2022)

- Popular deep learning architectures from the literature:

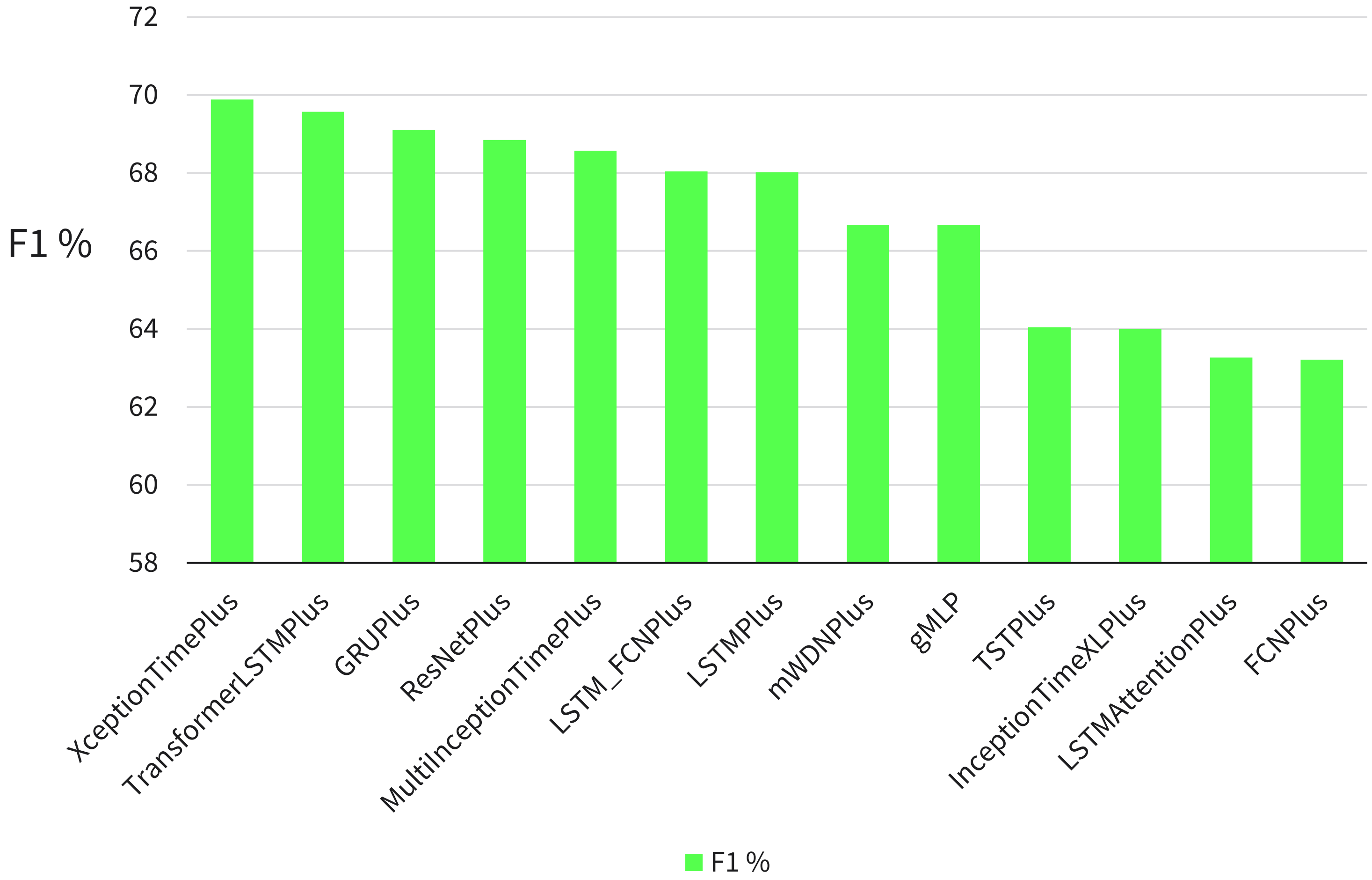
Previously used for spacecraft anomaly detection	Not used for spacecraft anomaly detection
Multi-Layered Perceptron (MLP) (Liu et al. (2021a), Bernal-Mencia et al., 2021)	Multilevel Wavelet Decomposition Network (mWDN) (Wang et al., 2018, Michau et al., 2022)
Long Short Term Memory (LSTM) (Hundman et al., 2018)	Transformer (Vaswani et al., 2023, Zerveas et al., 2021)
Gated Recurrent Unit (GRU) (Xiang & Lin, 2021)	InceptionTime (Fawaz et al., 2020)
	XceptionTime (Rahimian et al., 2019)

IDEA: TRY DIFFERENT FORECASTING ARCHITECTURES

- 81 models x 13 architectures = 1053 models trained!

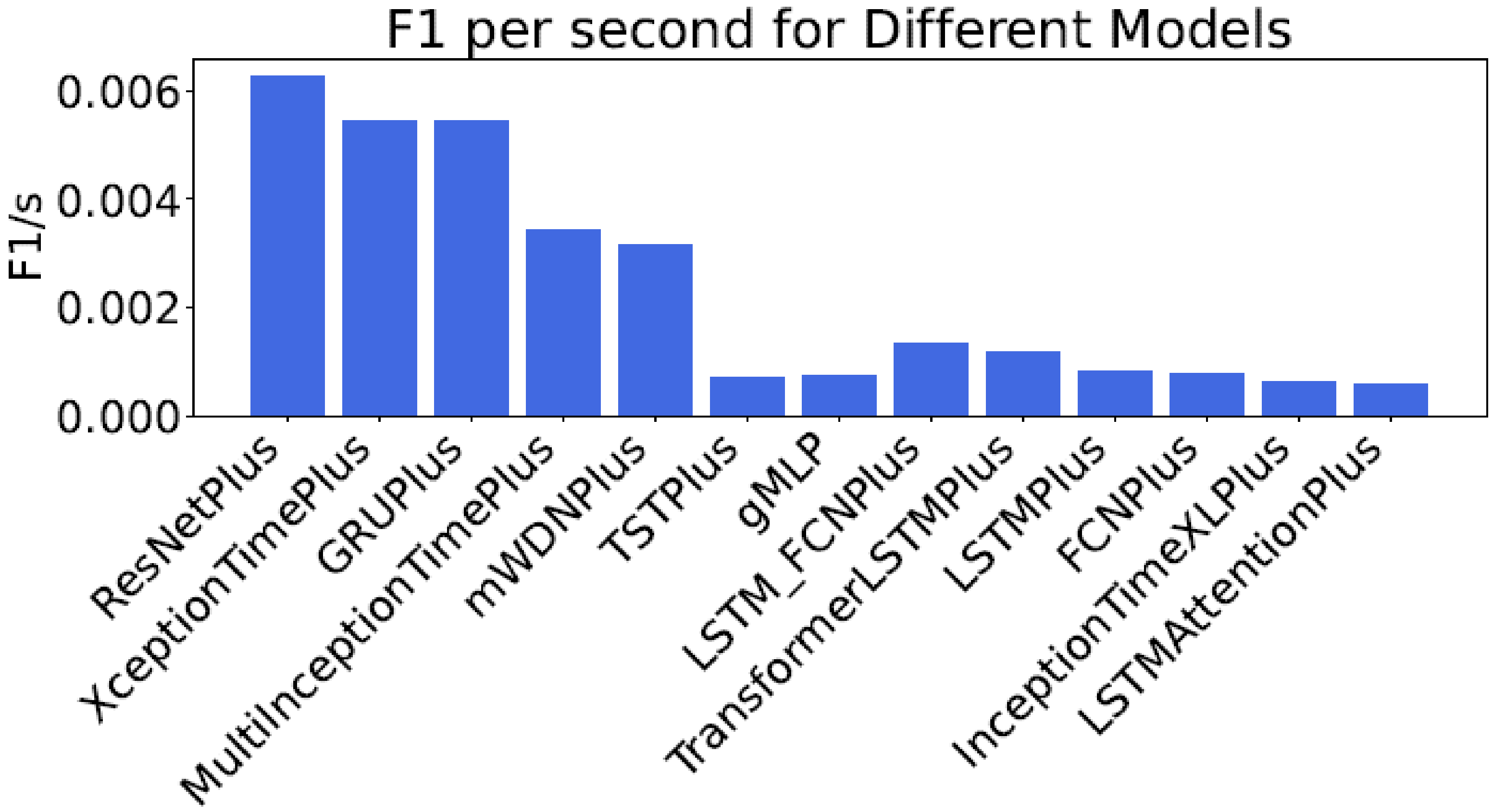


BEST PERFORMING ARCHITECTURE OF FORECAST AND THRESHOLD

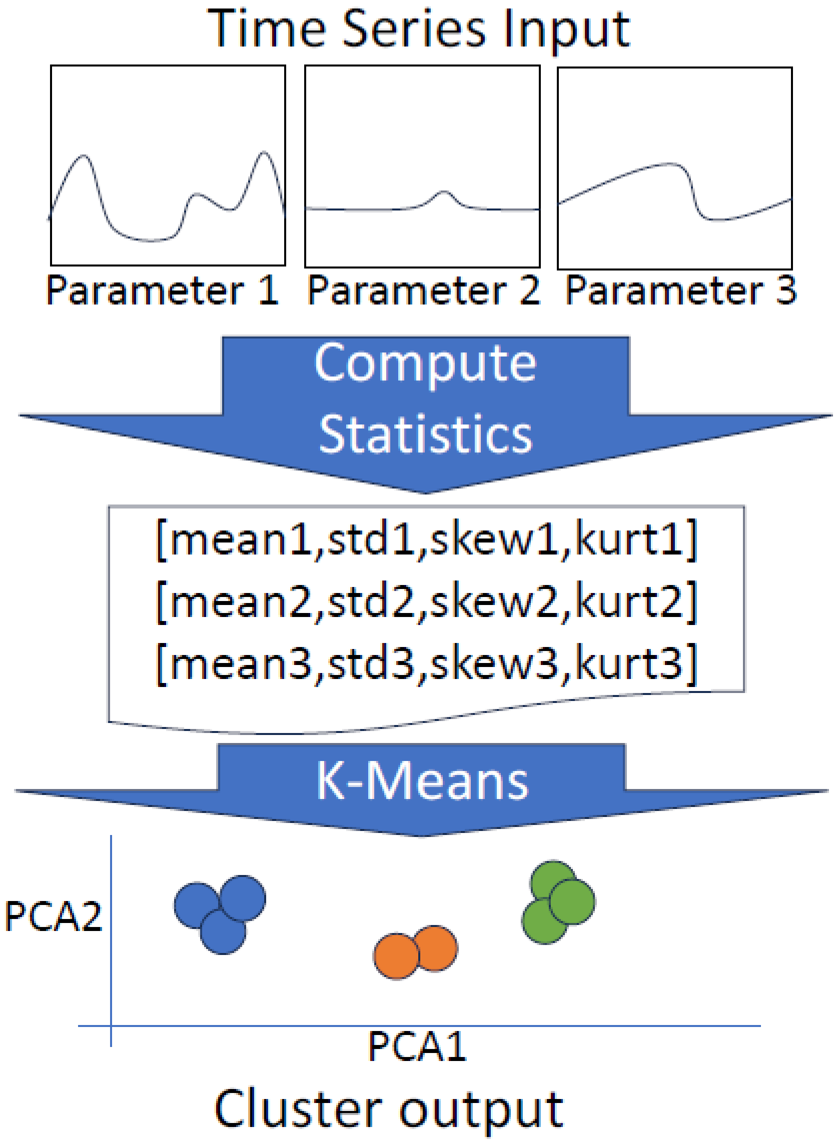


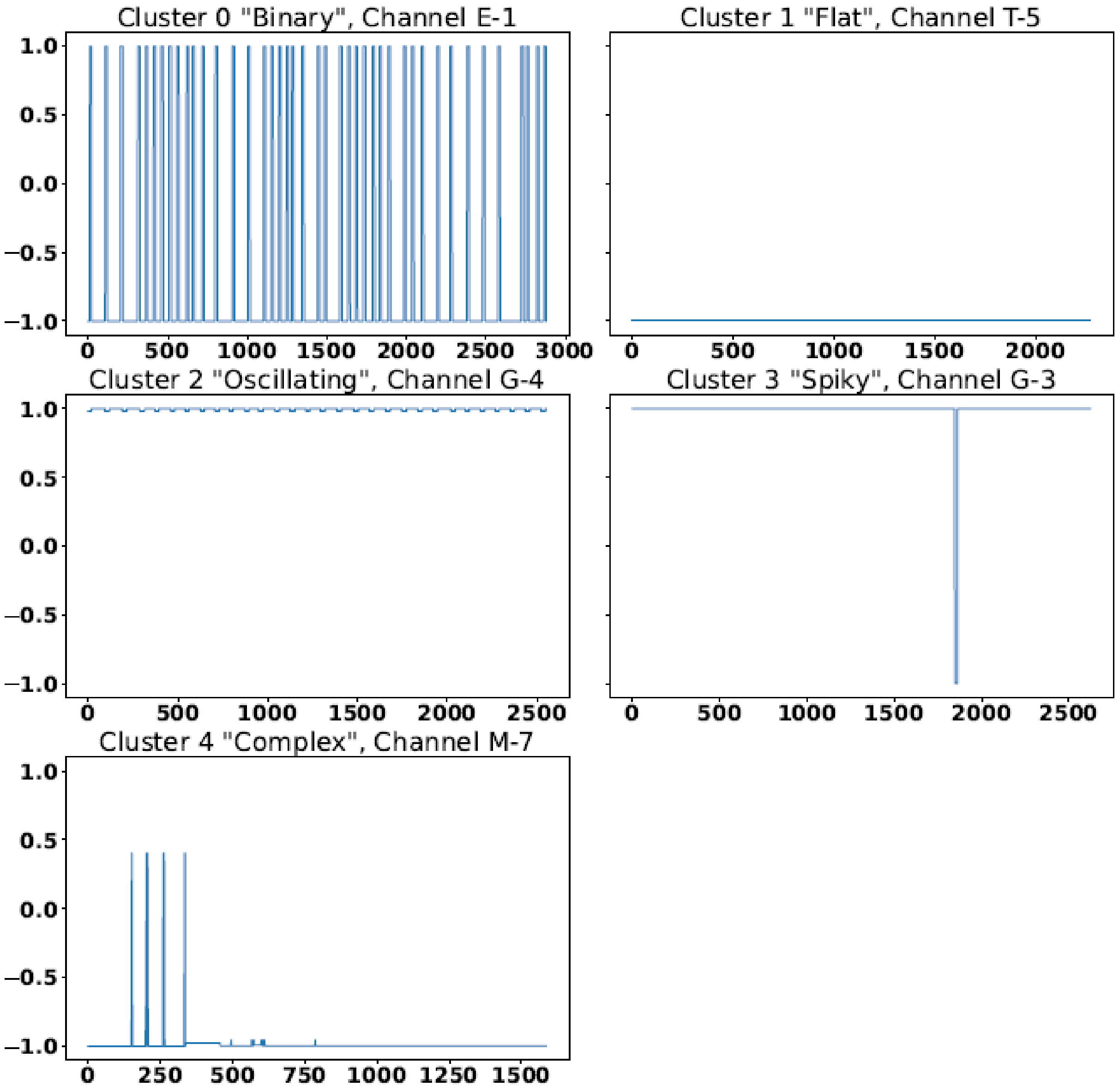
TRAINING EFFICIENCY

Training time can be significant for spacecraft data!



TELEMETRY DATA CLUSTERING





RESULTS: CLUSTERING VS ONE-FOR-ALL

Architecture	Data type cluster	F1 score (%)
MultiInceptionTimePlus	Binary	65
gMLP	Complex	100
XceptionTimePlus	Flat	86
XceptionTimePlus	Oscillating	72
gMLP	Spiky	100
Average		***85***

Architecture	Data type cluster	F1 score (%)
XceptionTimePlus	n/a	70
TransformerLSTMPlus	n/a	70
GRUPlus	n/a	69
ResNetPlus	n/a	69
MultiInceptionTimePlus	n/a	69
LSTM_FCNPlus	n/a	68
LSTMPlus	n/a	68
mWDNPlus	n/a	67
gMLP	n/a	67
TSTPlus	n/a	64
Average		***68***

➔ Clustering ensemble beats one-for-all

CONCLUSIONS

1. **The choice of deep learning architecture has big impact on anomaly detection performance**

– CNN-based XceptionTimePlus and hybrid TransformerLSTMPlus, **F1 score ≈ 70%** vs 84% in baseline

2. **Certain architectures work better with certain types of data**

– Cluster-based, per-data type approach, averaged **F1 score ≈ 85%** vs 84% in baseline

– Unsupervised clustering based on simple statistical measurements is fast and effective

3. **Training time matters!**

– Absolute performance needs to be balanced against training time

FUTURE WORK

- Integrate clustering with forecast and threshold pipeline to create a single ensemble model, and tune hyper-parameters per type

- Data set was very imbalanced. Forthcoming “ESA Anomalies Dataset” hopes to address this.

THANK YOU

Supported by

INTERNATIONALE HOCHSCHULE

DAAD Deutscher Akademischer Austauschdienst German Academic Exchange Service

Daniel Lakey
daniel.lakey@iu-study.org / daniel.lakey@cgi.com

REFERENCES

Baireddy, S., Desai, S. R., Mathieson, J. L., Foster, R. H., Chan, M. W., Comer, M. L., & Delp, E. J. (2021). Spacecraft time-series anomaly detection using transfer learning. In 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 1951–1960.
DOI: 10.1109/CVPRW53098.2021.00223.

Bernal-Mencia, P., Doerksen, K., & Yap, C. (2021). Machine learning for early satellite anomaly detection. Proceedings of the 35th Small Satellite Conference. URL: https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=5002&context=smallsat, [Accessed 27-10-2023].

Fawaz, H. I., Lucas, B., Forestier, G., Pelletier, C., Schmidt, D. F.,Weber, J.,Webb, G. I., Idoumghar, L., Muller, P.-A.,& Petitjean, F. (2020). InceptionTime: Finding AlexNet for time series classification. Data Mining and Knowledge Discovery, 34(6):1936–1962. DOI: 10.1007/s10618-020-00710-y.

Heras, J. & Donati, A. (2014). Enhanced telemetry monitoring with novelty detection. AI Magazine, 35(4):37–46. DOI: 10.1609/aimag.v35i4.2553.

REFERENCES

Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018). Detecting spacecraft anomalies using LSTMs and nonparametric dynamic thresholding. DOI: 10.48550/ARXIV.1802.04431.

Karim, F., Majumdar, S., Darabi, H., &Chen, S. (2018). LSTM fully convolutional networks for time series classification. IEEE Access, 6:1662–1669. DOI: 10.1109/access.2017.2779939.

Liu, H., Dai, Z., So, D. R., & Le, Q. V. (2021a). Pay attention to mlps. arXiv. DOI: 10.48550/arXiv.2105.08050.

Michau, G., Frusque, G., & Fink, O. (2022). Fully learnable deep wavelet transform for unsupervised monitoring of high-frequency time series. Proceedings of the National Academy of Sciences, 119(8). DOI: 10.1073/pnas.2106598119.

REFERENCES

Pilastre, B., Boussouf, L., D’Escrivan, S., & Tourneret, J.-Y. (2020). Anomaly detection in mixed telemetry data using a sparse representation and dictionary learning. Signal Processing, 168:107320. DOI: 10.1016/j.sigpro.2019.107320.

Rahimian, E., Zabihi, S., Atashzar, S. F., Asif, A., & Mohammadi, A. (2019). Xceptiontime: A novel deep architecture based on depthwise separable convolutions for hand gesture classification. arXiv. DOI: 10.48550/ARXIV.1911.03803.

Schmidl, S., Wenig, P., & Papenbrock, T. (2022). Anomaly detection in time series: A comprehensive evaluation. Proc. VLDB Endow., 15(9):1779–1797. DOI: 10.14778/3538598.3538602.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). Attention is all you need. arXiv. DOI: 10.48550/arXiv.1706.03762.

Walther, D., Viehweg, J., Haueisen, J., & Mäder, P. (2023). A systematic comparison of deep learning methods for EEG time series analysis. Frontiers in Neuroinformatics, 17. DOI: 10.3389/fninf.2023.1067095.

REFERENCES

Wang, J., Wang, Z., Li, J., & Wu, J. (2018). Multilevel wavelet decomposition network for interpretable time series analysis. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD ’18, page 2437–2446, New York, NY, USA. Association for Computing Machinery. DOI: 10.1145/3219819.3220060.

Wang, Y., Gong, J., Zhang, J., & Han, X. (2022). A deep learning anomaly detection framework for satellite telemetry with fake anomalies. International Journal of Aerospace Engineering, 2022:1–9. DOI: 10.1155/2022/1676933.

Wang, Z., Yan, W., & Oates, T. (2016). Time series classification from scratch with deep neural networks: A strong baseline. arXiv. DOI: 10.48550/arXiv.1611.06455.

Xiang, G. & Lin, R. (2021). Robust anomaly detection for multivariate data of spacecraft through recurrent neural networks and extreme value theory. IEEE Access, 9:167447–167457. DOI: 10.1109/ACCESS.2021.3136505.