GREEN
FORMANCE

Martin PACZONA
martin@greenformance.com
A-8330 Feldbach
www.greenformance.com

Physics-Informed Machine Learning and Hybrid Modelling

Conceptual-Model-Guided Physics-Inspired Feature Engineering: A 3D-Printer Case Study

INTRODUCTION

Power-only monitoring provides limited observability: relevant internal device states remain latent, and different physical mechanisms may produce similar load signatures. This paper introduces conceptual-model-guided physics-inspired feature engineering (CPFE), a lightweight representation-level strategy that formalizes the derivation of interpretable features from a conceptual model of the monitored device and its operation.

Research Question

Can CPFE improve machine-state estimation from active-power measurements compared to baseline and automated features, while enabling interpretation through traceable domain concepts?

RELATED WORK

Physics-informed and knowledge-guided approaches integrate prior domain structure through feature engineering and preprocessing, training objectives and constraints, and architecture-level design, forming the basis for hybrid modeling approaches [1,2]. In the energy domain, NILM infers states from electrical measurements under limited observability, especially in low-frequency power-only settings [3].

Automated feature engineering explores large transformation spaces to construct useful representations, while theory-guided data science advocates the controlled incorporation of prior knowledge into data-driven pipelines [2,4]. Building on this research, the present work formalizes a structured and traceable derivation of features from a conceptual model of the device.

METHOD

Conceptual-Model-Guided Physics-Inspired Feature Engineering (CPFE)

CPFE grounds feature generation in an explicit conceptualization of the observed system, the available raw data, and the intended prediction task. This conceptualization is structured through the meta-model shown in Fig. 1, which organizes system elements, domain concepts, and their relationships. Physical principles, constraints, and operational concepts are translated into domain concepts, which are then parameterized and operationalized into proxy features computable from the raw data stream. This creates a transparent abstraction chain from the observed system and its measurements to implemented features.

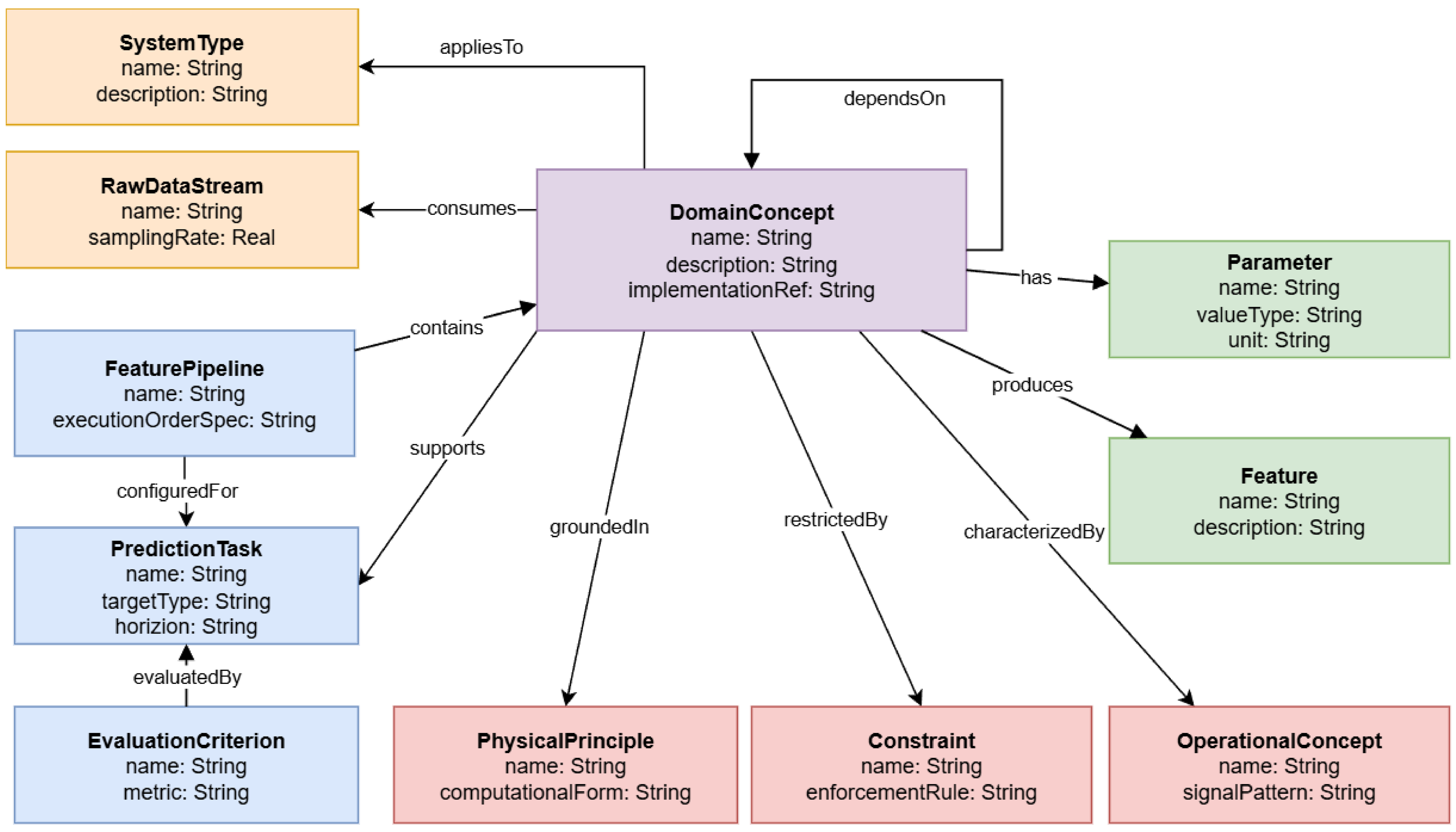

Figure 1: Meta-model of the proposed feature conceptualization method.

3D PRINTER CASE STUDY

The proposed CPFE methodology is demonstrated on a filament 3D printer using active-power measurements at 10 s resolution for state estimation. Training labels were obtained from the printer's G-code interface. In the present implementation, nine domain concepts (visible in Figure 3) are represented by 26 features. The equation of the feature *stored_heat_proxy_rc* that represents thermal state with gradual decay over time is shown in (1).

$$stored\_heat\_proxy\_rc_k = \alpha_{RC}\, x_{k-1} + (1-\alpha_{RC})\, \Delta P_k$$
$$\alpha_{RC} = e^{-\Delta t / \tau_{RC}} \qquad (1)$$

RESULTS

Comparison of the feature sets

For the controlled comparison, the same Random Forest hyperparameters have been used for all feature sets. The proposed core-physics representation achieved the best overall performance, reaching 90.6% accuracy also 0.738 macro-F1 and 0.766 balanced accuracy (see Figure 2). It outperformed both a compact baseline and automatically extracted tsfresh features. Misclassification of the rare WAIT class is likely driven by its short duration and indistinct power signature in 10 s data, suggesting weak class separability as the primary limiting factor.

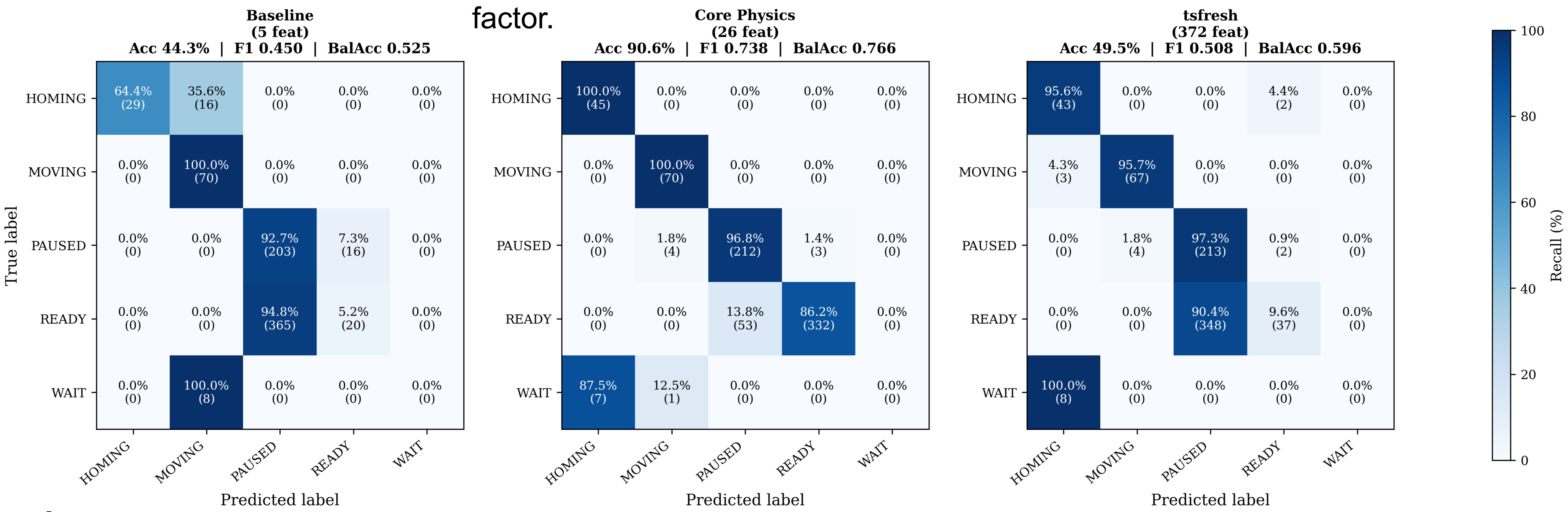

Figure 2: Confusion matrices for the compared feature sets, using identical Random Forest hyperparameters.

Figure 3 shows the aggregated feature importance by domain of the proposed feature set with *baseline-to-active decomposition, thermal load accumulation*, *multi-timescale envelope*, and *operating variability* as the top domain concepts.

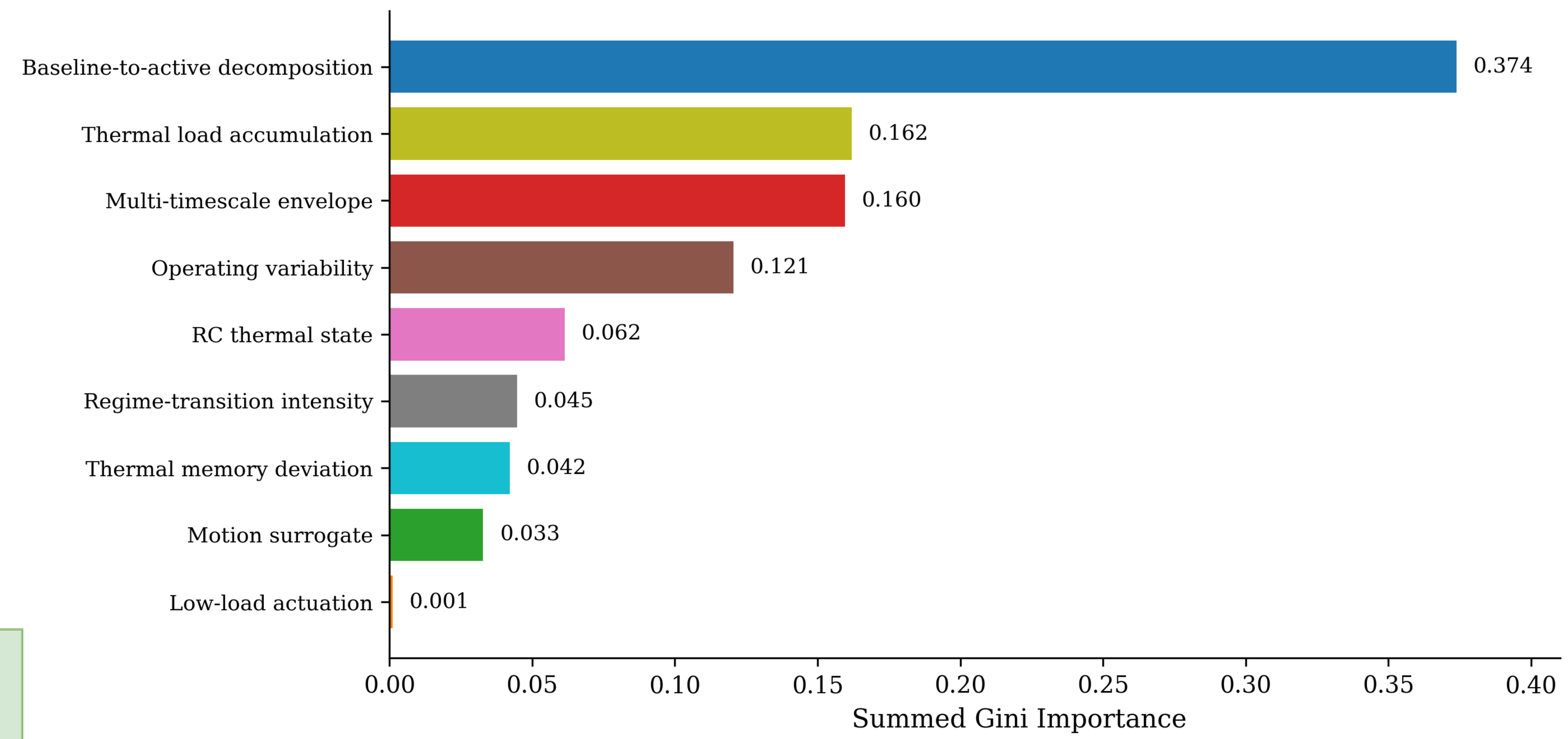

Figure 3: Aggregated feature importance by domain concept for the proposed approach.

CONCLUSION AND OUTLOOK

In this 3D-printer case study, the proposed core-physics representation outperformed both a compact baseline and the automatic feature extraction using tsfresh. Overall, the findings suggest that conceptual-model-guided and traceable feature construction offers a practical complement to physics-informed learning approaches that incorporate physics into the loss function or model architecture, particularly in limited-observability settings. Future work will address transferability across devices, evaluation across model architectures, automated concept-to-feature generation, and integration with end-to-end physics-inspired machine learning.

References

[1] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang. Physics-informed machine learning. Nature Reviews Physics, 3(6):422–440, 2021.
[2] A. Karpatne, G. Atluri, J. H. Faghmous, M. Steinbach, A. Banerjee, A. Ganguly, S. Shekhar, N. Samatova, and V. Kumar. Theory-guided data science: A new paradigm for scientific discovery from data. IEEE Transactions on Knowledge and Data Engineering, 29(10):2318–2331, 2017.
[3] G. Tanoni, E. Principi, and S. Squartini. Non-intrusive load monitoring in industrial settings: A systematic review. Renewable and Sustainable Energy Reviews, 202:114703, 2024.
[4] J. M. Kanter and K. Veeramachaneni. Deep feature synthesis: Towards automating data science endeavors. In Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA), pages 1–10. IEEE, 2015.