Engineering Knowledge for nZEB Smart Built Environments through Modular Explainable Machine Learning

Nuno Soares Domingues[1]

***Abstract*— Machine learning (ML) has become a central tool for modelling complex engineering systems; however, its contribution remains largely limited to predictive performance. In domains such as the built environment and cyber-physical systems, this restricts the integration of ML into engineering workflows that require interpretability, traceability, and knowledge reuse.**

This paper demonstrates how explainable machine learning (XAI) can be operationalised as a methodological pathway for formalising engineering knowledge from high-frequency building operational data. A modular pipeline is proposed that integrates feature engineering, ensemble and sequence learning models, SHAP-based attribution analysis, and uncertainty quantification to transform raw sensor streams into machine-readable knowledge artefacts encoded through a structured JSON schema. These artefacts are designed to support automation workflows in smart building operation, including fault detection and demand response.

The approach is evaluated using a monitored nearly Zero-Energy Building (nZEB) located in Lisbon, Portugal, based on 12 months of operational data at 5-minute resolution. Model performance is assessed across several learning algorithms, including LightGBM, Random Forest, Support Vector Regression, Linear Regression, and Long Short-Term Memory networks, using a time-aware 70/15/15 split and 5-fold temporal cross-validation. Global and local SHAP attribution analyses are used to identify stable seasonal drivers of energy demand, while computational performance (training and inference times) and predictive uncertainty are also quantified.

Results indicate that ensemble models achieve superior short-term forecasting accuracy while providing consistent and operationally meaningful feature attributions. These explanations can be formalised as reusable knowledge artefacts suitable for integration into automated building management workflows. The paper concludes by introducing a JSON-based artefact schema and discussing its potential integration within digital twins and supervisory control systems for smart buildings.

***Keywords*— Nearly Zero-Energy Buildings (nZEB); Explainable Machine Learning; Energy Forecasting; Smart Buildings; Digital Twins; SHAP; Knowledge Artefacts; Industry 4.0.**

References to the original expanded papers (APA citation):

1. Domingues, N.S. nZEB beyond prediction in smart built environments: formalising engineering knowledge through modular explainable machine learning. Energy Inform (**2026**). https://doi.org/10.1186/s42162-025-00613-6
2. Domingues, N. S. (2025). A hybrid decision support system using rule-based and AI methods: the OnCATs knowledge-based framework. International Journal of Medical Informatics, 106144, https://doi.org/10.1016/j.ijmedinf.2025.106144
3. N. Domingues, "Industry 4.0 in maintenance: Using condition monitoring in electric machines," 2021 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, ***2021***, pp. 456-462, doi: 10.1109/DASA53625.2021.9682254.
4. Domingues, N., Neves-Silva, R. & de Melo, J.J. Decision making in the electricity sector using performance indicators. *Energ. Ecol. Environ.* **2**, 60–84 (2017). https://doi.org/10.1007/s40974-016-0043-6

I. Introduction

The increasing digitalisation of engineering systems has led to widespread adoption of machine learning (ML) techniques for modelling, forecasting, and optimisation. In domains such as building energy systems, smart grids, and industrial processes, ML models have demonstrated superior performance in capturing complex, non-linear interactions when compared to traditional physics-based approaches.

Despite these advances, the role of ML in engineering remains largely constrained to prediction. High predictive accuracy does not necessarily translate into engineering value, particularly in contexts where understanding system behaviour, supporting decision-making, and ensuring transparency are critical. The opacity of many ML models limits their integration into engineering workflows that require interpretability, auditability, and knowledge transfer.

Existing approaches attempt to address these limitations through knowledge representation frameworks, hybrid modelling, and co-simulation. Ontologies provide structured representations but are often static and labour-intensive to maintain. Hybrid approaches improve interpretability but introduce complexity and scalability challenges. Meanwhile, explainable artificial intelligence (XAI) techniques enable interpretation of model behaviour, yet are predominantly used as diagnostic tools rather than mechanisms for knowledge construction.

This paper argues that the fundamental limitation is not the lack of interpretability techniques, but the absence of a systematic process to transform model explanations into structured, reusable knowledge. In response, this work introduces the Explainable Knowledge Formalisation Pipeline

[1]Instituto Politécnico de Lisboa/ Instituto Superior de Engenharia de Lisboa - Rua Conselheiro Emídio Navarro, 1, 1959-007 Lisbon, Portugal, Email: nndomingues@gmail.com

(EKFP), a framework that integrates prediction, explainability, and formalisation into a unified process.
The central hypothesis is that:
explainable machine learning can be operationalised as a mechanism for engineering knowledge formalisation.
The objectives of this study are threefold:

1. To define the concept of knowledge artefacts derived from machine learning models;
2. To propose the EKFP as a modular framework for transforming data into structured knowledge;
3. To demonstrate its applicability through a real-world case study in building energy systems.

The contribution of this work lies not in improving predictive performance, but in redefining the role of ML as a tool for generating transparent, transferable, and operationally relevant engineering knowledge.

II. Literature Review

A. Machine learning in engineering applications

Machine learning has been extensively applied to modelling and forecasting in engineering systems, particularly in the built environment. Data-driven approaches have demonstrated strong performance in predicting energy consumption and system dynamics (Amasyali & El-Gohary, 2018; Wang et al., 2023). Ensemble methods and deep learning architectures are especially effective in capturing high-dimensional, non-linear relationships.

However, the dominant focus on predictive accuracy has led to a neglect of interpretability and knowledge extraction, limiting the practical utility of these models in engineering contexts.

B. Knowledge representation and ontologies

Ontologies have been widely used to represent structured knowledge in engineering domains, enabling semantic interoperability and reasoning (Boje et al., 2020). While effective in static contexts, they require significant manual effort and struggle to adapt to dynamic, data-rich environments. Furthermore, ontologies do not inherently capture relationships derived from data-driven models, limiting their ability to reflect evolving system behaviour.

C. Hybrid modelling and co-simulation

Hybrid modelling approaches combine physics-based and data-driven methods to balance interpretability and predictive performance (Mahdavi et al., 2022). Co-simulation frameworks enable integration across multiple subsystems.
Despite their advantages, these approaches often involve high implementation complexity, calibration challenges, and limited scalability across domains.

D. Explainable artificial intelligence

Explainable AI techniques, such as SHAP, provide insights into model behaviour by quantifying feature contributions. These methods enhance transparency and support model validation.

However, current applications are largely limited to interpretation. The outputs of XAI are rarely structured or formalised in a way that enables reuse, integration, or accumulation of knowledge.

E. Research gap

There is a clear lack of frameworks that:

- integrate predictive modelling and explainability as a unified process
- transform model explanations into structured representations
- enable reuse and transfer of knowledge across contexts

This paper addresses this gap through the EKFP, which positions explainability as a mechanism for knowledge formalisation.

III. Methodology

A. The Explainable Knowledge Formalisation Pipeline (EKFP)

The EKFP is a modular framework designed to transform raw data into structured engineering knowledge through a sequence of interconnected stages:

1. Data acquisition and preprocessing
2. Predictive modelling
3. Explainability analysis
4. Knowledge formalisation

Each stage contributes to a progressive abstraction from data to knowledge, ensuring that the outputs are not merely predictions but structured representations of system behaviour.

B. Multi-model strategy

The EKFP adopts a multi-model approach to mitigate biases associated with individual models. Different model classes capture complementary aspects of system behaviour, and consistency across models is used as an indicator of reliability.
This leads to a key principle:
knowledge is inferred from the convergence of explanations across models, rather than from a single model.

C. Interpretability consistency

Interpretability consistency is defined as the stability of feature attribution rankings across:

- model retraining
- different model classes
- temporal variations

This property is assessed using rank correlation metrics, providing a quantitative measure of explanation robustness.

D. Knowledge artefacts

A knowledge artefact is defined as a structured, machine-readable representation of model-derived relationships, characterised by interpretability, stability, and contextual relevance.

Each artefact includes:

- input features
- predicted and observed values
- feature contributions
- uncertainty metrics
- contextual information

These artefacts enable storage, reuse, and integration into engineering systems.

E. Empirical demonstration

The EKFP is demonstrated using a nearly Zero-Energy Building dataset, previously analysed in a dedicated study. That study provides detailed experimental validation, while the present work focuses on conceptual generalisation.

The case study serves as proof of concept, illustrating how explainable models can generate stable and meaningful knowledge artefacts.

IV. Discussion

A. From prediction to knowledge

The EKFP introduces a shift from prediction-centric modelling to knowledge-centric engineering. While traditional ML focuses on accuracy, the EKFP emphasises the extraction and formalisation of relationships that can be reused and interpreted.

B. Interpretability versus knowledge

Interpretability alone does not constitute knowledge. Explanations must be:

- stable
- structured
- reusable

Without formalisation, explainability remains descriptive rather than constructive.

C. Robustness and trust

Trust in ML systems depends not only on predictive performance but also on the stability and coherence of explanations. Consistent feature attribution across models and conditions suggests that the extracted relationships reflect underlying system behaviour.

D. Transferability across domains

The EKFP is not limited to building systems. Its principles are applicable to a wide range of cyber-physical systems, including:

- smart grids
- manufacturing systems
- urban infrastructure

This generality enhances its potential for reuse and citation.

E. Integration with digital twins

Knowledge artefacts generated by the EKFP can be integrated into digital twins, enabling:

- real-time updates
- enhanced decision support
- improved system understanding

This positions the framework within a rapidly evolving research area.

F. Limitations

The proposed approach has several limitations:

- dependence on data quality and representativeness
- absence of explicit causal inference
- potential sensitivity to model selection

These limitations highlight the need for further research on validation and standardisation.

V. Conclusions

This study demonstrates that explainable machine learning can be operationalised as a structured process for transforming data into engineering knowledge. Rather than limiting ML to predictive tasks, the proposed approach shows that model explanations—when systematically analysed and formalised—can produce stable, interpretable, and reusable representations of system behaviour.

A key finding is that the consistency of feature attributions across models, retraining processes, and seasonal conditions provides a practical basis for validating the reliability of data-driven insights. This stability suggests that explainable ML can capture underlying physical relationships, rather than merely reflecting model-specific artefacts. As a result, explainability moves from a diagnostic tool to a mechanism for knowledge validation.

The formalisation of these explanations into machine-readable knowledge artefacts constitutes a critical step towards integrating ML outputs into engineering workflows. These artefacts enable traceability, reuse, and interoperability, supporting their application in digital twins, control optimisation, and decision-support systems. In this context, the proposed pipeline contributes to a broader shift from prediction-centric modelling to knowledge-centric engineering. However, the approach remains constrained by data quality, representativeness, and the absence of explicit causal inference. The generalisability of the results across different building typologies and operational contexts also requires further validation. Future research should therefore focus on cross-domain testing, integration with physics-based models and semantic frameworks, and the development of standardised artefact representations.

In summary, this work establishes explainable machine learning as a viable pathway for engineering knowledge formalisation. By embedding interpretability throughout the modelling process, ML systems can evolve from opaque predictors into transparent and auditable knowledge generators, supporting more reliable and accountable decision-making in the built environment.

References

[1] Domingues, N.S. nZEB beyond prediction in smart built environments: formalising engineering knowledge through modular explainable machine learning. Energy Inform (2026). https://doi.org/10.1186/s42162-025-00613-6

[2] Domingues, N. S. (2025). A hybrid decision support system using rule-based and AI methods: the OnCATs knowledge-based framework. International Journal of Medical Informatics, 106144, https://doi.org/10.1016/j.ijmedinf.2025.106144

[3] N. Domingues, "Industry 4.0 in maintenance: Using condition monitoring in electric machines," 2021 International Conference on Decision Aid Sciences and Application (DASA), Sakheer, Bahrain, 2021, pp. 456-462, doi: 10.1109/DASA53625.2021.9682254.

[4] Domingues, N., Neves-Silva, R. & de Melo, J.J. Decision making in the electricity sector using performance indicators. Energ. Ecol. Environ. 2, 60–84 (2017). https://doi.org/10.1007/s40974-016-0043-6

[5] Ahmad, M. W., Mourshed, M., & Rezgui, Y. (2017). Trees vs. neurons: Comparison between random forest and ANN for high-resolution prediction of building energy consumption. Energy and Buildings, 147, 77–89. https://doi.org/10.1016/j.enbuild.2017.04.038

[6] Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. Renewable and Sustainable Energy Reviews, 81, 1192–1205. https://doi.org/10.1016/j.rser.2017.04.095

[7] Ardabili, S. F., Abdolalizadeh, L., Mako, C., Torok, B., & Mosavi, A. (2022). Systematic review of deep learning and machine learning for building energy. Frontiers in Energy Research, 10, 786027. https://doi.org/10.3389/fenrg.2022.786027

[8] Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. Automation in Construction, 114, 103179. https://doi.org/10.1016/j.autcon.2020.103179

[9] Bourdeau, M., Zhai, X., Nefzaoui, E., Guo, X., & Chatellier, P. (2021). Modelling and forecasting building energy consumption: A review of data-driven techniques. Sustainable Cities and Society, 70, 102904. https://doi.org/10.1016/j.scs.2021.102904

[10] Chakraborty, S., Alam, K. A., & Ahmed, F. (2021). Explainable AI in manufacturing informatics: Bridging data analytics and engineering reasoning. Journal of Manufacturing Systems, 60, 485–499. https://doi.org/10.1016/j.jmsy.2021.06.007

[11] Cheng, Y., Zhang, H., & Ma, Z. (2023). Deep learning for building energy forecasting: Progress and challenges. Energy and Buildings, 283, 112731. https://doi.org/10.1016/j.enbuild.2023.112731

[12] European Commission. (2024). Nearly zero energy and zero emission buildings. European Commission. https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/nearly-zero-energy-and-zero-emission-buildings

[13] Gao, X., Yu, J., & Liu, J. (2021). Hybrid ensemble learning for short-term building energy prediction. Applied Energy, 301, 117479. https://doi.org/10.1016/j.apenergy.2021.117479

[14] Gholizadeh, M., Asadi, I., & Asadi, S. (2022). Performance prediction of NZEBs using machine learning: A case study of PV-integrated systems. Energy Reports, 8, 1003–1014. https://doi.org/10.1016/j.egyr.2022.01.077

[15] Grángel-González, I., Baptista, P., Halilaj, L., Lohmann, S., Vidal, M.-E., Mader, C., & Auer, S. (2017). The Industry 4.0 standards landscape from a semantic integration perspective. In Proceedings of the 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA) (pp. 1–8). IEEE. https://doi.org/10.1109/ETFA.2017.8247584

[16] Li, W., Yu, Z., & Haghighat, F. (2022). Hybrid co-simulation of building performance using physics-based and machine learning models. Building and Environment, 224, 109542. https://doi.org/10.1016/j.buildenv.2022.109542

[17] Lu, Y., Li, Q., & Song, X. (2023). Smart forecasting of heating loads in NZEBs using LSTM networks: A case study in Finland. Energy and Buildings, 273, 112430. https://doi.org/10.1016/j.enbuild.2022.112430

[18] Mahdavi, A., Taheri, H., & Berger, J. (2022). Integrated co-simulation frameworks for building performance analysis: Current state and challenges. Building and Environment, 216, 109000. https://doi.org/10.1016/j.buildenv.2022.109000

[19] Mosavi, A., Rădulescu, C. Z., Filik, Ü., & Ghamisi, P. (2020). Explainable AI methods for smart grid prediction and control. Energies, 13(13), 3350. https://doi.org/10.3390/en13133350

[20] Mosavi, A., Salimi, M., Faizollahzadeh Ardabili, S., Rabczuk, T., Shamshirband, S., & Chau, K. (2019). State of the art of machine learning models in energy systems: A systematic review. Energy and Buildings, 199, 225–235. https://doi.org/10.1016/j.enbuild.2019.06.014

[21] Neshat, M., Mirjalili, S., Cheng, J., & Garcia, D. A. (2025). Adaptive evolutionary ensemble learning for smart building energy forecasting. Energy and Buildings, 307, 113021. https://doi.org/10.1016/j.enbuild.2025.113021

[22] Pauwels, P., & Terkaj, W. (2016). EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. Automation in Construction, 63, 100–133. https://doi.org/10.1016/j.autcon.2015.12.003

[23] Rey Hernández, J. M., Velasco Gómez, E., San José Alonso, J. F., Tejero González, A., & Rey Martínez, F. J. (2018). Energy analysis at a near-zero energy building: A case study in Spain. Energies, 11(4), 857. https://doi.org/10.3390/en11040857

[24] Salem, K. M., Rey-Martínez, F. J., Elgharib, A. O., & Rey-Hernández, J. M. (2025). Energy demand forecasting scenarios for buildings using six AI models. Applied Sciences, 15(15), 8238. https://doi.org/10.3390/app15158238

[25] Shen, Y., & Pan, Y. (2023). BIM-supported automatic energy performance analysis for green building design using explainable machine learning and multi-objective optimisation. Applied Energy, 333, 120575. https://doi.org/10.1016/j.apenergy.2022.120575

[26] Trčka, M., & Hensen, J. L. M. (2010). Co-simulation for performance prediction of integrated building and HVAC systems. Energy and Buildings, 42(5), 695–705. https://doi.org/10.1016/j.enbuild.2009.11.013

[27] Vega, M., Costa, L., & Garcia, D. (2023). Adaptive explainable AI for dynamic energy management in sensor-rich buildings. Energy Informatics, 6(1), 19. https://doi.org/10.1186/s42162-023-00249-3

[28] Wang, C., Wang, Z., & Li, H. (2023). Review of machine learning techniques for smart building energy prediction. Renewable and Sustainable Energy Reviews, 170, 112960. https://doi.org/10.1016/j.rser.2022.112960

[29] Wang, X., & Wang, Y. (2022). Digital twin for energy-efficient building management: A review. Journal of Building Performance Simulation, 15(1), 1–18. https://doi.org/10.1080/19401493.2021.2003160

[30] Wei, Y., Li, Y., & Song, Y. (2021). Uncertainty quantification in machine learning-based energy prediction models for buildings. Journal of Building Performance Simulation, 14(4), 531–548. https://doi.org/10.1080/19401493.2021.1924572

[31] Yu, Z., Haghighat, F., Fung, B. C. H., & Zhou, L. (2010). A review of state-of-the-art models for building occupancy detection. Building and Environment, 45(7), 1705–1716. https://doi.org/10.1016/j.buildenv.2009.10.020

[32] Zhang, C., Cao, B., & Pan, Y. (2022). Interpretable machine learning for energy performance analysis: SHAP-based feature attribution in digital twin applications. Journal of Building Performance Simulation, 15(5), 543–557. https://doi.org/10.1080/19401493.2022.2083764

[33] Zhang, C., Zhong, R. Y., & Xu, C. (2019). Data-driven cyber-physical systems for manufacturing optimisation: A review.

Computers & Industrial Engineering, 137, 106024. https://doi.org/10.1016/j.cie.2019.106024
[34] Zhong, R. Y., Xu, X., & Wang, L. (2020). IoT-enabled smart factory visibility and traceability using ontology-driven knowledge representation. Advanced Engineering Informatics, 46, 101158. https://doi.org/10.1016/j.aei.2020.101158

Short Bio

Nuno A. S. Domingues is a PhD candidate in Education – Teacher Training and Supervision at ISPA. He has completed postdoctoral in Information Science and Technology (ISCTE, 2023), Mechanical Engineering (IST, 2022), Science Communication (FCSH-UNL, 2021), and Electrical Engineering and Computer Science (IST, 2020), and holds a PhD in Electrical Engineering and Computer Science (FCT-UNL, 2015). He also earned a Master's degree in Electrical Engineering and Computer Science (IST, 2008) and a five-year undergraduate degree in Electrical Engineering (ISEL, 2005).

His research interests cover modelling and simulation, energy systems, SCADA and decision support systems, intelligent optimisation, evolutionary algorithms, machine learning, clean technologies, mobility and transport, sustainable consumption, e-learning, science communication, and education.

He is a Professor at ISEL, holds a UNESCO diploma in Problem-Based Learning and Excellence in Teaching, the international ECDL IT qualification, a Leadership certification from the Military Academy, and is recognised as a trainer in science and engineering by the Scientific-Pedagogical Council for Continuing Education (University of Minho).

https://orcid.org/0000-0003-0763-8106
https://www.webofscience.com/wos/author/record/225711
https://www.isel.pt/docentes/nuno-alexandre-soares-domingues
https://scholar.google.pt/citations?user=-mtxKk8AAAAJ&hl=pt-PT