

DATA - DRIVEN ASSESSMENT OF CIRCULAR ENERGY EFFICIENCY: AN EXPLAINABLE AI APPROACH

DOI: 10.26341/issn.2241-4010-2026-1c-10

Ivan Ciric

University of Nis, Serbia

ivan.ciric@masfak.ni.ac.rs

Milica Tasic

University of Nis, Serbia

milica.tasic@masfak.ni.ac.rs

Ana Đorđević (formerly Kitić)

University of Nis, Serbia

ana.kitic@masfak.ni.ac.rs

Zlatan Car

Catholic University of Zagreb, Croatia

zlatan.car@unicath.hr

Abstract

This paper presents an operational, closed-loop-inspired framework for assessing circular energy efficiency in district heating systems using real Supervisory Control and Data Acquisition (SCADA) data. Inefficiencies are characterized through regime-dependent “wastage signatures” (elevated supply/return temperatures, low ΔT , and control-induced mismatches).

As a baseline of expected consumption for the Lamella L17 substation (Faculty of Mechanical Engineering, University of Nis), we train ensemble models - Random Forest and Least-Squares Boosting on time-aligned measurements (outdoor temperature, secondary-side supply temperature, and delivered heat), augmented with temporal features and short-horizon lags. Evaluation with a chronological train/test split shows that Random Forest provides the most robust performance (RMSE = 3.85, MAE = 1.67, R^2 = 0.894), particularly during abrupt regime changes.

An explainability layer (permutation importance) is then applied to identify dominant drivers and flag high-deviation episodes as candidates for inefficient operation. The proposed workflow links baseline forecasting, interpretable diagnostics, and intervention-oriented feedback to support continuous efficiency improvement and more reliable heat delivery.

Key words: district heating, circular energy efficiency, machine learning, explainable artificial intelligence

Introduction

Circular energy efficiency in thermal energy systems can be interpreted as the systematic identification and reduction of resource dissipation across the full value chain - generation, the distribution network, and end-use - while maximizing the utilization of useful heat,

including waste heat and low-temperature sources. In district heating (DH), circularity is typically reflected in operational symptoms such as unnecessarily high supply temperatures, unfavorable return temperatures, low temperature difference (ΔT), and excessive mass flows, which jointly increase distribution and pumping losses and may degrade service quality. Because these inefficiencies are strongly regime-dependent (e.g., seasonal transitions, load ramps, partial-load operation, and network constraints), an actionable circular-efficiency assessment requires methods that are accurate, computationally feasible, and transparent to operational stakeholders.

In practice, DH operators already have extensive measurement infrastructure (e.g., SCADA, substation heat meters, temperature sensors, and in some cases building-level signals). However, translating these streams into continuous improvement actions is not straightforward. Raw measurements can confirm that a deviation occurred (e.g., rising return temperature or collapsing ΔT), but they do not, on their own, explain whether the deviation is demand-driven, caused by control logic, driven by hydraulic limitations, or linked to end-user behavior. Moreover, high-frequency or irregularly sampled operational data often exhibit missingness, regime switching, and delayed effects, making purely physics-based supervisory diagnostics difficult to maintain at scale. Consequently, the methodological challenge is not only to predict and detect deviations, but also to produce interpretable, regime-aware evidence that supports operational decisions, intervention design, and subsequent evaluation.

The recent DH literature provides the necessary building blocks for such an assessment, spanning reduced-order and data-driven representations that enable scalable deployment (Kudela et al., 2020; Rein et al., 2020), short-term forecasting as an operational baseline (Maryniak et al., 2024; Potočník et al., 2021), explainability methods to diagnose the drivers of inefficiency (Neubauer et al., 2025), and optimization/control strategies to test and implement interventions (Schindler et al., 2023; Vesterlund and Toffolo, 2017; Stock et al., 2024). In parallel, the building-energy community has demonstrated learning-based approaches that are relevant at the consumer interface of DH, including heat-load prediction for buildings connected to DH networks (Hajri et al., 2025), reinforcement-learning-based heating control motivated by cost reduction (Kannari et al., 2023), and adaptive modeling concepts for long-term operation under changing conditions (Chen et al., 2024). These strands are directly relevant to circularity because the main circular “symptoms” in DH are ultimately produced at the interface between network operation and end-use behavior.

However, these components are often treated in isolation. Forecasting is frequently evaluated primarily through error metrics, Explainable Artificial Intelligence (XAI) is used for model interpretation rather than operational diagnosis, and optimization/control is addressed as a standalone objective rather than as part of a measurable efficiency-improvement loop. Similarly, consumer-side learning approaches (e.g., reinforcement learning control or building load prediction) are often discussed as local improvements, without explicit linkage to network-side temperature regimes, return-temperature constraints, or system-level Key Performance Indicators (KPIs) that define circular efficiency in DH. As a result, the field still lacks a unified, operationally oriented workflow that links baseline modeling, interpretable loss signatures, and intervention evaluation into a transparent closed loop suitable for continuous improvement (Kudela et al., 2020; Neubauer et al., 2025; Medina Cruz et al., 2025). Recent applied studies comparing machine-learning methods for DH load

prediction and addressing irregular time-series and explainability themes further support the operational relevance of closed-loop, interpretable assessment workflows (Tasić et al., 2024; Tasić et al., 2025; Tasić et al., 2025b).

COST Action CA22124 - ECO4ALL (Circular Economy that Works for Consumers) is a European research initiative that promotes sustainable development by strengthening the role of consumers in the circular economy. The Action focuses on sustainable consumption, resource efficiency, waste prevention, and consumer protection through evidence-based and cross-country assessment of circular strategies from legal, social, economic, and environmental perspectives. Within the broader context of the ECO4ALL Action, this work contributes to the understanding of how circular-economy principles can be operationalized in energy systems in a way that is meaningful for multiple stakeholders along the value chain, including system operators and end users. ECO4ALL emphasizes the role of consumers, transparency, and data-informed decision-making as prerequisites for effective circularity, moving beyond abstract policy objectives toward measurable and actionable improvements. In district heating systems, operational inefficiencies such as elevated return temperatures, low temperature differentials, and excessive mass flows represent not only technical losses but also a failure to align network operation with end-use behaviour and consumer-level engagement. By developing an interpretable, data-driven workflow for identifying, explaining, and evaluating circular-efficiency losses in DH operation, this study directly supports the ECO4ALL objectives of evidence-based assessment and quantitative measurement of inefficiencies and their drivers, thereby contributing to the design of circular strategies that are transparent, verifiable, and capable of supporting informed interventions at both the system and consumer interfaces.

Related work

Prior work covers complementary building blocks that can be assembled into a closed-loop assessment logic for district heating: scalable/deployable models, forecasting baselines for deviation detection, interpretable diagnostics, and intervention-oriented methods. In addition, related work from building energy management contributes useful concepts at the demand interface, particularly for data-driven load estimation and learning-based control.

A practical prerequisite for continuous assessment is deployability: models must be accurate yet computationally light, robust under data imperfections, and maintainable as operating conditions shift. Kudela et al. (2020) discuss how machine learning can be integrated efficiently into district-heating predictive workflows, with attention to implementation aspects that determine whether a model remains usable outside of a research setting (Kudela et al., 2020). From an assessment perspective, this is significant because a closed-loop efficiency workflow cannot rely on models that are too fragile, too expensive to re-train, or too opaque to validate. In parallel, Rein et al. (2020) show that network control problems become difficult to solve online when full dynamics are included and demonstrate how reduced-order modelling supports tractable optimization while preserving key dynamic constraints (Rein et al., 2020). Together, these contributions motivate a pragmatic stance: circular-efficiency monitoring should build on models that can operate at substation and network scales in near real time, even if they intentionally trade some fidelity for reliability and speed.

Accurate forecasting of heat load and/or heat production is a cornerstone of efficiency assessment because it enables the separation of normal demand variability from operationally induced losses (e.g., overheating, over-supply, and inadequate tracking of demand). Maryniak et al. (2024) demonstrate the use of a neural network for forecasting daily heat production in a district heating plant, emphasizing model validation and the use of performance indicators as a basis for comparing alternative approaches. In the presented case, the authors report an effectiveness of 85% for the trained network (Maryniak et al., 2024), which is relevant to the assessment layer because a stable forecast allows observed deviations to be interpreted as signals of inefficient operation rather than as consequences of inherently unpredictable demand. Potočník et al. (2021) further strengthen this argument by focusing on short-term demand forecasting in district heating, reporting that, for a 48-hour forecast horizon, prediction errors below approximately 3% can be achieved in their scenario (Potočník et al., 2021). Such accuracy levels are particularly important when circular efficiency is operationalized through indicators such as: (1) excess heat delivered relative to an optimum, (2) unnecessarily high temperature regimes, and (3) losses driven by mismatches between supply and demand.

At the end-use interface, forecasting and baseline estimation are equally relevant because building demand is the main driver of supply-temperature decisions and network mass-flow patterns. Hajri et al. (2025) specifically address heat-load prediction for buildings connected to DH networks using data-driven modelling (Hajri et al., 2025). This line of work supports circular-efficiency assessment in two ways: first, it can refine what “normal” demand behaviour looks like for a given building or consumer group; second, it helps localize inefficiency by distinguishing network-driven anomalies from genuine demand changes. In a closed-loop setting, such consumer-level baselines are valuable for diagnosing whether high supply temperatures or poor ΔT arise from network control settings, substation heat exchanger behaviour, or demand-side disturbances.

Although data-driven models can achieve high predictive accuracy, assessing circular efficiency requires more than a “good predictor”: it is necessary to explain which factors - and under which operating regimes - generate resource wastage. This is particularly important in DH systems, where decisions affect costs, emissions, hydraulic conditions, and end-user comfort (Stock et al., 2024). Neubauer et al. (2025) explicitly emphasize interpretability for multi-step heating load forecasting by combining SHapley Additive exPlanations (SHAP)-based attribution with temporal mechanisms, enabling the identification of contributions from individual features and past time steps to the forecast (Neubauer et al., 2025). From an assessment standpoint, the key value of this approach is that XAI makes it possible to link inefficiencies (e.g., episodes of excess supply or unfavourable temperature regimes) to actionable drivers and to form a transparent diagnostic of circular efficiency, rather than relying solely on aggregate accuracy metrics. In operational environments, this diagnostic role is critical: when a KPI indicates deterioration, operators need to understand whether the driver is weather sensitivity, control hysteresis, lagged thermal inertia, abnormal flow conditions, or a shift in consumer-side behaviour.

A closely related practical challenge is non-stationarity: DH systems and connected buildings evolve over seasons, retrofits, occupancy changes, tariff structures, and control updates. Chen et al. (2024) address dynamic and adaptive modelling in building energy

systems through lifelong learning concepts, aiming to sustain model performance under changing conditions (Chen et al., 2024). While their focus is broader than DH alone, the implication for circular-efficiency assessment is direct: a closed-loop workflow must remain valid as the system drifts, otherwise KPI baselines and explainability conclusions become unreliable. Lifelong or adaptive learning strategies can therefore be interpreted as enablers of continuous circular-efficiency monitoring, reducing the need for disruptive re-calibration cycles and supporting more stable decision support.

In literature, optimization and control are often treated as primary objectives; however, within a circular-efficiency framework they can be viewed as a downstream layer—interventions designed only after it has been identified where and why resource wastage occurs. Rein et al. (2020) analyse optimal control of DH networks using a reduced-order model, emphasizing that network dynamics as constraints in the optimization problem make online computation of the optimum difficult without model reduction (Rein et al., 2020). In an assessment context, this work is relevant because it demonstrates how models can be made sufficiently fast to support systematic intervention testing and evaluation of their effects on efficiency indicators. Schindler et al. (2023) consider optimization-driven operation of DH networks Mixed-Integer Linear Programming (MILP), focusing on the integration and coordination of sources in real systems (Schindler et al., 2023). This approach aligns well with circular logic because it enables the evaluation and selection of strategies that increase the utilization of available resources (including local and low-carbon options), while explicitly managing trade-offs among energy, costs, and operational constraints. A common conclusion is that reliable assessment (data-driven and transparent) and optimization should function as a unified chain: (1) measurement/forecasting, (2) identification of losses and their causes, and (3) selection and evaluation of interventions (Rein et al., 2020; Neubauer et al., 2025; Schindler et al., 2023).

At the consumer-control layer, learning-based strategies can also be interpreted as intervention mechanisms. Kannari et al. (2023) study reinforcement-learning-based heating control driven by energy cost considerations (Kannari et al., 2023). Within a DH circular-efficiency narrative, this is relevant because it exemplifies how advanced control can shift operational regimes toward lower losses - provided that the reward structure and constraints reflect DH-specific circular objectives (e.g., penalizing high return temperature, incentivizing higher ΔT , or discouraging excessive mass flows). This highlights a key methodological point for closed-loop assessment: the same measurable KPIs used to diagnose inefficiency should, where possible, be aligned with the objectives and constraints of downstream control and optimization layers, ensuring that interventions target the actual sources of dissipation rather than merely improving local cost or comfort metrics.

In DH systems, circular efficiency largely manifests through return temperature, ΔT , and flow rates, as these determine distribution losses, pumping energy, and the feasibility of integrating low-temperature sources. Tol and Madessa (2024) focus on return-temperature reduction at end-user sites, emphasizing that improvements often rely on actions that affect return temperature and system hydraulics (Tol and Madessa, 2024). Tol and Madessa (2025) further review return-temperature reduction strategies and their role in overall efficiency improvement (Tol and Madessa, 2025). Within a circular assessment approach, these findings suggest that efficiency metrics should include indicators related to return temperature and ΔT ,

as they represent direct levers for reducing wastage and increasing resource utilization. Importantly, these indicators are interpretable to operators and can be connected to specific intervention types (substation control tuning, balancing, retrofits, or consumer guidance), making them well suited for a closed-loop monitoring framework.

At the network level, Stock et al. (2024) address the optimization of DH network separation to improve the utilization of heat-source potentials, particularly when temperature levels are mismatched or source availability is unfavourable (Stock et al., 2024). This is directly connected to circular efficiency because separation and reconfiguration can enable local use of waste heat or low-temperature sources, reduce transport losses, and better align system temperature with source temperature. From a planning perspective, Vesterlund and Toffolo (2017) examine optimization of DH network expansion by balancing capital costs of piping against pumping costs (Vesterlund and Toffolo, 2017), which is relevant to assessment because it supports evaluation of the long-term effects of infrastructure decisions on resource-efficiency indicators (losses, pumping energy, and source availability). These planning and topology decisions also motivate why operational assessment must be regime-aware: what is “efficient” during peak winter operation may differ from shoulder seasons, when low-temperature sources or local separation strategies become more advantageous.

Finally, recent applied studies support the feasibility and relevance of operator-oriented data-driven assessment. Tasić et al. (2024) compare machine-learning approaches for DH load prediction, providing evidence that practical predictive baselines can be built and evaluated in DH contexts (Tasić et al., 2024). Tasić et al. (2025) analyse the effect of sliding-window techniques on decision-tree time-series forecasting, which is directly relevant to real operational data where regime changes and temporal context strongly affect model performance (Tasić et al., 2025). Tasić et al. (2025b) further position irregular time series and explainability as central themes for industrially relevant energy modelling (Tasić et al., 2025b). Collectively, these studies motivate the present work’s emphasis on a closed-loop-inspired workflow: not only producing forecasts but using them to construct interpretable loss signatures and a feedback mechanism that supports systematic intervention design and evaluation over time.

In summary, the literature provides strong individual components - forecasting, explainability, optimization/control, consumer-side modelling, and adaptive learning - but the operational gap remains the integration of these components into a transparent, measurable loop for continuous circular-efficiency improvement. The framework developed in this paper addresses this gap by explicitly linking baseline modelling with KPI/EEI wastage signatures and explainable attribution, thereby enabling intervention selection and post-intervention assessment in a form that can be communicated to operators and aligned with end-user-oriented objectives.

Methodology

This paper addresses the assessment of circular energy efficiency based on real operational data collected from the SCADA system of the Faculty of Mechanical Engineering in Nis, with a focus on the district heating substation - lamella L17. The analysis relies on time-synchronized measurements including outdoor air temperature, secondary-side supply water temperature, and a signal of delivered thermal power, from which the consumed thermal

energy per time step is calculated. The dataset represents an authentic record of system operation under real exploitation conditions and inherently contains nonlinearities, delays, and regime changes characteristic of practical district heating systems.

The primary objective of the study is not limited to energy consumption forecasting, but rather to the definition of a reliable baseline representation of system behaviour. This baseline describes the expected energy consumption under given external and temporal conditions and serves as a reference for further efficiency analysis. Accordingly, energy consumption is modelled as a time-series regression function of outdoor temperature, secondary-side-supply temperature, and temporal features describing time of day and day of week. To capture the dynamic nature and thermal inertia of the system, lagged values of previously consumed energy are also included in the model.

Based on this set of input variables, ensemble machine learning models are trained, with preference given to models that demonstrate the best generalization performance on a test set obtained through chronological data splitting. The selected model is not interpreted as a physical representation of the system, but as a statistically learned approximation of the normal, expected operating regime of the L17 substation. In this sense, the model output is treated as a reference baseline prediction.

Model validation is performed using standard regression metrics. The primary evaluation metric is the Root Mean Squared Error (RMSE),

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_i - \hat{E}_i)^2},$$

which emphasizes larger deviations and enables the identification of operating regimes with potentially significant energy excess. In addition, the Mean Absolute Error (MAE),

$$MAE = \frac{1}{N} \sum_{i=1}^N |E_i - \hat{E}_i|,$$

and the coefficient of determination R^2 ,

$$R^2 = 1 - \frac{\sum (E_i - \hat{E}_i)^2}{\sum (E_i - \bar{E})^2},$$

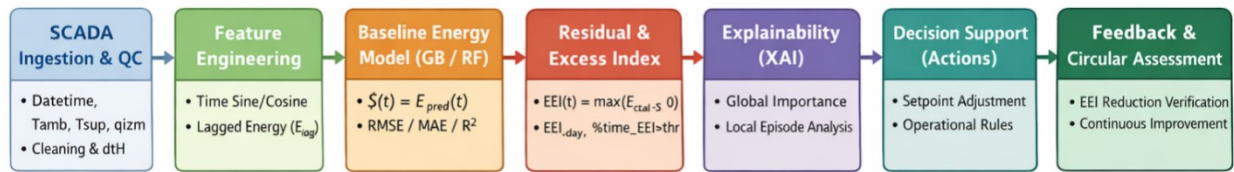
are used to assess overall prediction accuracy and the proportion of variance in measured energy consumption explained by the model. The difference between measured and predicted energy consumption is interpreted as an indicator of deviation from expected system behaviour. Rather than being directly classified as a fault or loss, this deviation is treated as an analytical signal indicating potentially inefficient or atypical operating regimes. Emphasis is placed on positive deviations, where actual consumption exceeds the baseline prediction, as these cases represent relevant candidates for circular energy efficiency analysis.

To make such deviations interpretable and operationally meaningful, an explainability layer based on Explainable Artificial Intelligence (XAI) methods is introduced. At the global level, the relative contribution of individual input variables to the baseline model prediction is analysed to identify dominant drivers of energy consumption. At the local level, the analysis focuses on specific time instances with the largest deviations, examining how variations in

each input variable influence the model output. This enables interpretation of concrete operational situations associated with increased energy consumption.

The central scientific contribution of this work lies in the definition of a reliable baseline prediction and the application of explainable artificial intelligence for interpreting the learned system behaviour. Elements related to the assessment of circular energy efficiency and potential operational interventions are presented as guidelines derived from the analysis results and illustrated through the proposed block diagram, rather than as explicitly implemented control strategies.

Figure 1. Closed-loop SCADA - XAI for circular energy efficiency



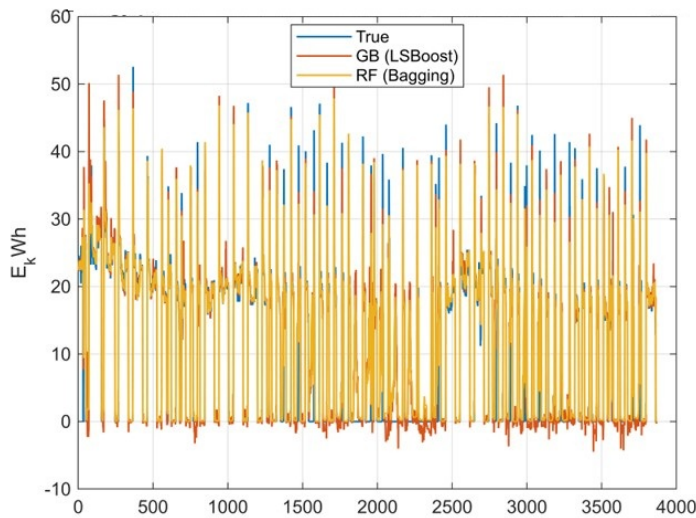
This approach enables flexible application of methodology in accordance with system-specific characteristics and operator requirements, without compromising the primary research objective.

In this way, a closed analytical loop is established in which real SCADA data are used to learn baseline system behaviour, identify deviations, and explain their underlying causes. The proposed approach demonstrates how modern machine learning and XAI techniques can support the assessment of circular energy efficiency in real district heating systems, as illustrated by the case study of the Lamella L17 substation at the Faculty of Mechanical Engineering in Nis.

Results and discussion

The presented results refer to the prediction of consumed thermal energy for the heating substation (lamella) L17, based on time-synchronized SCADA measurements. In total, 12,894 valid samples were used, and the dataset was chronologically split into 9,025 samples for training (TRAIN) and 3,869 samples for testing (TEST). The typical sampling interval is approximately 0.250 h (about 15 minutes), and a maximum historical lag of 4 was introduced, meaning that the input feature set includes previous signal values up to four - time steps back.

Two machine learning models were compared: (1) Least-Squares Boosting (LSBoost) and (2) Random Forest in a bagging regime (RF). For Least-Squares Boosting, hyperparameters were tuned on a meta-validation set, yielding the following best configuration: cycles = 300, learning rate = 0.200, maxSplits = 10, minLeaf = 10, with RMSE(meta) = 2.6693. Subsequently, an RF model was trained as a “baseline” for comparison.

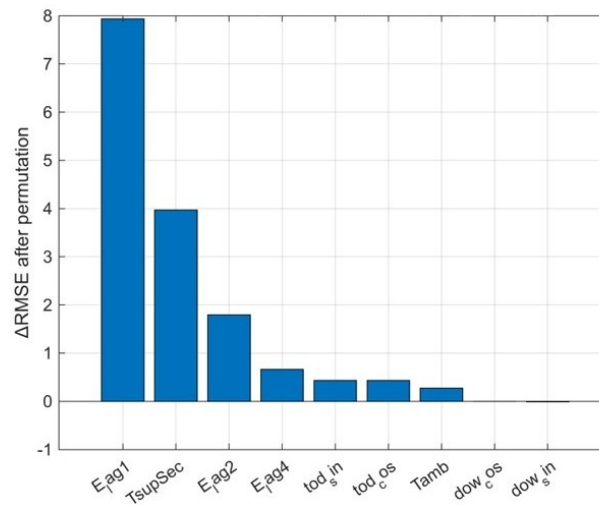
Figure 2. Comparative overview of model performance on the test set

From the energy prediction plot (Fig.2) on the test set, it can be observed that both models follow the overall level of the signal; however, RF behaves more robustly in segments where pronounced peaks and rapid changes occur (which is typical for energy values computed over discrete sampling intervals). This suggests that the bagging ensemble smooths noise and local anomalies more effectively, whereas boosting may be more sensitive to specific training patterns in the presence of abrupt spikes.

Table 1. Numerical comparison of model performance (RF vs. LSBoost): RMSE, MAE, and R^2

MODELS	RMSE	MAE	R^2	BEST RMSE	BEST MAE	BEST R^2
RF	3.8537	1.6696	0.89446	3.8537	1.6696	0.89446
LSBOOST	4.0364	1.8435	0.88422			

The numerical evaluation on the TEST split, which is presented in Table 1, confirms the advantage of the Random Forest model. In practical terms, the obtained $R^2 \approx 0.895$ indicates that the model explains approximately 89-90% of the variance of the target variable on the test set, which is a very solid level of accuracy for real operational data that contain noise, regime changes, and potential measurement errors. The difference between RMSE and MAE is expected: RMSE is higher because it penalizes larger errors (e.g., peaks) more strongly, whereas MAE better reflects the typical absolute error in day-to-day operation. In addition to predictive accuracy, model interpretation was performed using permutation importance on the test set.

Figure 3. XAI analysis of predictor importance: permutation test and change in model error

In the feature-importance histogram (Fig. 3), the dominant predictors are the thermal energy consumption lagged by one time step, followed by the secondary-side supply water temperature, then the thermal energy consumption lagged by two time steps and the thermal energy consumption lagged by four time steps, while the time-of-day components (sine and cosine) and the outdoor air temperature are notably smaller but still measurably relevant; the day-of-week components (sine and cosine) are almost negligible. This ranking is consistent with energy time-series behaviour: autoregressive features typically carry the most information (energy in the next step is strongly related to energy in previous steps), while the secondary-side supply temperature represents a direct operational “signature” of heat delivery (secondary supply temperature is closely linked to thermal load). The smaller contribution of the outdoor temperature and time-of-day components indicates that part of the meteorological and scheduling influence is already captured through the recent consumption history (lags), yet additional information from the daily cycle and ambient conditions remains beneficial.

Overall, the results show that, for L17 under the observed conditions, RF is a more reliable choice than LSBoost: it yields higher accuracy (lower RMSE and MAE, higher R^2) and tracks varying operating regimes more stably.

The XAI analysis further confirms that consumption dynamics are primarily driven by the system’s short-term inertia and by the operational state of the secondary circuit, which is valuable for operators because it highlights which variables should be prioritized when interpreting deviations and potential losses. In this section, it is important to emphasize the connection between the obtained results and the last two blocks of the block diagram presented in the previous section - Decision Support (Actions) and Feedback & Circular Assessment - in order to complete the closed-loop logic of circular energy efficiency. Within the *Decision Support (Actions)* block, results derived from the baseline predictive model, residual/EEI indicators, and XAI analyses will be translated into a set of practical operational guidelines and decision rules. The goal is for elevated EEI episodes not to be treated merely as “statistical deviations,” but as clear indicators of potentially suboptimal operating regimes. In this sense, future work will define procedures for identifying typical patterns (e.g., frequent EEI episodes in specific time windows, associations between EEI and secondary-side

operating temperatures, or sudden localized deviations), followed by corresponding interventions (setpoint correction, heating-curve adjustment, night-regime optimization, verification of sensors and actuators, etc.). This ensures that the analytical layer (ML + KPI + XAI) gains direct operational applicability in control and monitoring.

The *Feedback & Circular Assessment* block represents the final step that closes the loop and completes the concept of circular efficiency. After interventions are implemented, a repeated evaluation using the same KPI/EEI metrics and comparison against the baseline will be performed to quantitatively confirm that “excess” consumption (wastage) has been reduced and that the operating regime has stabilized. This feedback mechanism enables iterative improvement: interventions are not treated as one-off actions, but are verified, adjusted, and systematically refined through continuous monitoring. In this way, future research will enable full closed-loop realization: prediction - wastage detection (EEI) - explanation (XAI) - operational action - effect verification, thereby completing, both methodologically and practically, the assessment and enhancement of circular energy efficiency.

Conclusions

This study advances a practical pathway for operationalizing *circular energy efficiency* in district heating by treating circularity not as an abstract sustainability principle, but as a measurable, continuously assessable property of real system operation.

The core contribution is the establishment of a data-driven baseline learned from time-synchronized SCADA measurements for the Lamella L17 substation, which represents “expected” behaviour under prevailing external and temporal conditions, rather than a physics-based surrogate. This baseline framing is essential because it enables the separation of routine variability from potentially avoidable excess, and it provides a consistent reference against which circular-efficiency deviations can be monitored in a repeatable manner. Beyond baseline forecasting, the work strengthens the bridge from predictive analytics to circular-efficiency assessment by interpreting prediction deviations as an analytical signal for atypical or potentially inefficient regimes - specifically focusing on *positive deviations* where measured consumption exceeds the baseline estimate. This perspective is important for operational environments: deviations are not automatically treated as “faults,” but as *structured indicators* that can trigger targeted diagnostic steps. In this way, the baseline model becomes a scalable screening mechanism for identifying candidate episodes of resource dissipation (e.g., oversupply behaviour, transient anomalies, or control-driven mismatches) that merit deeper inspection.

Interpretability is addressed through an explicit explainability layer that supports operational trust and actionability. By combining global and local explainability analyses, the paper demonstrates how the learned baseline can be *explained* in terms of dominant drivers, and how high-deviation instances can be examined in a structured manner to connect operational context to observed excess. This is a key element for circular-efficiency assessment because it transforms model output from a “black-box prediction” into an evidence-based diagnostic tool that can be communicated to practitioners and used for prioritization.

At the same time, the paper deliberately positions the intervention layer as a *forward-looking operationalization*, not as a fully implemented control strategy. The block-diagram logic clarifies how the analytics stack (baseline + deviation indicator + explainability) can feed into decision support and subsequent verification, while remaining adaptable to site-specific constraints and operator requirements.

This is aligned with ECO4ALL’s emphasis on linking efficiency metrics to consumer-relevant outcomes: a transparent monitoring and diagnostic process is a prerequisite for fairer heat delivery, reduced overheating/underheating risk, and more accountable operation—benefits that directly affect end-user comfort and costs.

Several limitations define the next research steps. First, the analysis is currently centered on a single substation and a particular operational context; expanding to multiple substations and seasons would strengthen external validity and enable comparative circular-efficiency benchmarking. Second, the present feature set focuses on readily available SCADA signals; incorporating additional hydraulic and temperature-performance indicators (e.g., return temperature behaviour, ΔT signatures, and flow-related proxies where available) would provide a richer mapping between deviations and physically meaningful loss mechanisms. Third, operational deployment requires explicit governance of data quality (sensor drift, missingness, and regime labelling), as well as clear thresholds for actionable alerts versus benign variance.

Future work will therefore focus on completing the closed-loop circular assessment: translating explainability-guided deviation patterns into formalized operational playbooks (Decision Support) and then verifying the effectiveness of interventions through repeated KPI/EEI evaluation (Feedback & Circular Assessment).

This progression - baseline prediction which include wastage detection, explainability, action and verification provide a concrete, scalable mechanism for continuous improvement and for demonstrating, with operational evidence, that circular-efficiency measures yield measurable benefits for system performance and end users.

References

- Chen S., Ge W., Liang X., Jin X., Du Z., (2024), “Lifelong learning with deep conditional generative replay for dynamic and adaptive modeling towards net zero emissions target in building energy system”, “Applied Energy”, vol. 353 (Part B), Art. no. 122189, doi: 10.1016/j.apenergy.2023.122189.
- Hajri A., Garay-Martinez R., Macarulla A. M., and Ben Sassi M. A., “Data-driven model for heat load prediction in buildings connected to district heating networks,” Energy, vol. 329, Art. no. 136684, Aug. 15, 2025, doi: 10.1016/j.energy.2025.136684.
- Kannari L., Kantorovitch J., Piira K., Piippo J., (2023), “Energy Cost Driven Heating Control with Reinforcement Learning”, “Buildings”, vol. 13(2), Art. no. 427, doi: 10.3390/buildings13020427.
- Kudela L., Chýlek R., Pospíšil J., (2020), “Efficient Integration of Machine Learning into District Heating Predictive Models”, “Energies”, vol. 13, p. 6381, doi: 10.3390/en13236381.
- Maryniak A., Banaś M., Michalak P., Szymiczek J., (2024), “Forecasting of Daily Heat Production in a District Heating Plant Using a Neural Network”, “Energies”, vol. 17, p. 4369, doi: 10.3390/en17174369.

-
- Medina Cruz J. R., Quispe Huayta J. E., Carpio Cutipa M. Á., Molina Cruz A. D., Lopez Palomeque F., (2025), “The sustainable development of rural economies within the territorial scope of the Geopark located between the Colca Valley and volcanoes in the Arequipa region, Peru”, “Sustainable Futures”, vol. 9, 100716, doi: 10.1016/j.sft.2025.100716
- Neubauer A., Brandt S., Kriegel M., (2025), “Explainable multi-step heating load forecasting: Using SHAP values and temporal attention mechanisms for enhanced interpretability”, “Energy and AI”, vol. 20, p. 100480, doi: 10.1016/j.egyai.2025.100480.
- Potočník P., Škerl P., Govekar E., (2021), “Short-term heat demand forecasting in a district heating system: A comparative study of state-of-the-art methods”, “Energy and Buildings”, vol. 234, p. 110673, doi: 10.1016/j.enbuild.2020.110673.
- Rein M., Mohring J., Damm T., Klar A., (2020), “Optimal control of district heating networks using a reduced order model”, “Optimal Control Applications and Methods”, 41, 1352-1370, doi: 10.1002/oca.2610
- Schindler M., Gnam L., Puchegger M., Medwenitsch K., Jasek P., (2023), “Optimization-Based Operation of District Heating Networks: A Case Study for Two Real Sites”, “Energies”, vol. 16, p. 2120, doi: 10.3390/en16052120.
- Stock J., Xhonneux A., Müller D., (2024), “Optimisation of district heating network separation for the utilisation of heat source potentials”, “Energy”, vol. 303, p. 131872, doi: 10.1016/j.energy.2024.131872.
- Tasić M., Ćirić I., Ignjatović M., Stojiljković D., (2024), “Comparison of Various Machine Learning Methods for Heat Load Prediction in District Heating System”, in “Proceedings of the XVII International Conference on Systems, Automatic Control and Measurements (SAUM 2024)”, Niš, Serbia, pp. 169-172, doi: 10.46793/SAUM24.169T.
- Tasić M., Ćirić I., Jovanović V., Ignjatović M. G., Mitrović D., (2025), “The Effect of Sliding Window Technique on Decision Tree Time Series Forecasting in District Heating System”, in “Proceedings of the 17th International Conference on Accomplishments in Mechanical and Industrial Engineering (DEMI 2025)”, Banja Luka, Bosnia and Herzegovina, May 29-30, 2025, pp. 163-170.
- Tasić M., Ćirić I., Jovanović V., Simonović M., Ignjatović M., (2025b), “Smart Energy Production and Industry 5.0: XAI and Irregular Time Series Forecasting”, in “Proceedings of ICPES 2025”, pp. 496-505, doi: 10.46793/ICPES25.498T.
- Tol H. I., Madessa H. B., (2024), “Return-Temperature Reduction at District Heating Systems: Focus on End-User Sites”, “Energies”, vol. 17, p. 4901, doi: 10.3390/en17194901.
- Tol H. I., Madessa H. B., (2025), “Enhancing District Heating System Efficiency: A Review of Return Temperature Reduction Strategies”, “Applied Sciences”, 15(6), 2982, doi: 10.3390/app15062982.
- Vesterlund M., Toffolo A., (2017), “Design Optimization of a District Heating Network Expansion, a Case Study for the Town of Kiruna”, “Applied Sciences”, vol. 7, p. 488, doi: 10.3390/app7050488.