Neural Network-Driven Optimization: Enhancing Speed in Energy Systems

Theresa Reinhardt, Viktor Wesselak, Christoph Schmidt, Sebastian Voswinckel and Rohith Bala Krishnan

in.RET Institute for Renewable Energy Technology, University of Applied Sciences, 99734 Nordhausen, Germany

Abstract

This work explores the use of neural network surrogate models to approximate outputs such as installed renewable capacities, storage sizing and system costs, thereby reducing simulation runtimes while maintaining predictive accuracy. Training datasets are generated to ensure broad coverage of input scenarios, including economic variables like electricity prices and capital expenditures, as well as climatic factors influencing renewable generation. The surrogate models are integrated with outputs from the oemof framework, which represents the regional energy system through graph-based optimisation. Performance assessment employs metrics such as root mean square error, mean absolute percentage error, and coefficient of determination, with attention to prediction accuracy across typical and extreme operational conditions. A data-driven surrogate based on a feedforward multilayer perceptron was developed to emulate the optimization results of the energy system model. The surrogate provides accurate capacity predictions with an order-of-magnitude lower computational cost, enabling extensive scenario and sensitivity analyses.

Keywords: decarbonisation, energy system, energy system modelling, open source, optimization, neural network, multi-layer perceptrons, convolutional neural networks, recurrent neural networks

1. Introduction

The increasing complexity of regional energy system models has amplified the computational burden associated with scenario analysis. Long simulation times constrain the number of design alternatives that can be tested and, consequently, limit the breadth of decision support offered to stakeholders. In situations where multiple input parameters need to be varied, ranging from electricity and fuel prices to annual full-load hours and projected demand levels, the space of possible scenarios expands quickly. This can make brute-force simulation approaches inefficient or even infeasible within reasonable timeframes (Shirzadi, Lau and Stylianou, 2025). The essential motivation here is to find a method that preserves predictive accuracy while substantially reducing computational expense. A promising pathway to accomplish this is to train surrogate models capable of learning and replicating the behaviour of the original high-fidelity simulations. These surrogate models are based on neural networks that directly map selected scalar inputs, representative of scenario characteristics, to target outputs like installed renewable generation capacity, storage sizing needs, or aggregated system cost. Rather than re-running time-intensive optimisations for each variation in input assumptions, the trained network can generate approximations almost instantaneously once deployed (Kalogirou, 2001).

Artificial neural networks (ANNs) have seen broad application in energy forecasting because they can capture nonlinear interactions between inputs and outputs that are difficult to represent with explicit analytical equations. Unlike traditional modelling techniques that require predefining fixed mathematical relationships, which is especially problematic when system behaviours shift nonlinearly over time, ANNs learn these mappings from data through iterative parameter adjustments. Experience suggests that properly configured ANNs reach high correlation between predicted and actual values even with comparatively noisy or uncertain training sets (Kalogirou, 2001). For this particular application, different architectures come under scrutiny,

10.18086/swc.2025.05.09 Available at http://proceedings.ises.org

each with potential advantages depending on the specific statistical structure of the energy model outputs. Feedforward multilayer perceptrons offer simplicity and generally fast inference but may need large amounts of training data for complex systemic behaviours (Zhang). Convolutional networks are more specialised towards structured spatial-temporal patterns but can also serve time-series-related estimation in hybrid forms. Recurrent structures like RNNs or LSTMs incorporate feedback loops enabling them to carry temporal dependencies, an advantage when historical sequences matter, but their computational demands and tendency to overfit in small datasets pose challenges unless mitigated by architectural tweaks or regularisation strategies (Rahman *et al.*, 2021).

The choice among these is not only a question of accuracy but also processing speed and interpretability. Comparative evaluations have shown stark differences: while simple linear regression executes predictions extremely quickly, deeper or recurrent neural networks may require much longer runtimes both during training and inference (Zhang). For operational integration in planning processes where rapid iteration is valued over marginal gains in precision, these runtime considerations can tip decisions toward lighter architectures. In fact, model construction benefits from stripping away irrelevant variables before training starts. Input parameters without measurable effect on variation in outputs should be excluded; their retention only adds noise and slows convergence without improving generalisation (Kalogirou, 2001).

Determining this relevance can rely on contribution factor analyses embedded within ANN frameworks themselves or alternatively on prior domain expertise about which scenario descriptors truly matter for capacities and costs. Given that underlying resource availabilities such as wind speeds or solar irradiation patterns are inherently variable across years and locations, a model trained solely on deterministic values risks being misled by atypical data points. The solution often lies in either integrating synthetic variability during sample generation or selecting training examples that span worst-case through best-case boundaries rather than clustering densely around average conditions (Medina-Santana and Cárdenas-Barrón, 2022). Failing to do so may result in optimistic projections underestimating necessary reserve margins or storage capacities. Neural network predictions should therefore be accurate enough to preserve trust when replacing longer simulations while remaining transparent about uncertainties inherent in approximation-based approaches.

2. Description of the Thuringia Energy System Model

The Thuringia energy system model describes a multi-sectoral representation of energy generation, storages and demand-side dynamics within a regional boundary. Its spatial scope is confined to the state of Thuringia, yet it also incorporates constraints and opportunities arising from interconnection with external regions. Core inputs to the model include technology-specific investment costs (CapEx), operating and maintenance costs (OpEx), and conversion efficiencies. These parameters are primarily drawn from established reference datasets such as national studies that provide harmonized assumptions for all German regions (Oberdorfer et al., 2021). The model also embeds regional potentials, for example limiting onshore wind capacity to a fixed land-use share or mapping photovoltaic deployment along suitable transport corridors. The supply side encompasses multiple renewable energy options. Wind power potential in Thuringia is geographically constrained, while solar photovoltaics, both rooftop and free-field systems, are distributed based on land availability and irradiation levels. Renewable generation profiles are developed through historical weather datasets that translate resource quality into annual full-load hour estimations for each technology. Dispatchable generation units fuelled by natural gas or biogas serve as flexible backup capacity within their allowed emission limits. In parallel, sector coupling technologies like heat pumps or electrolyzers establish links between electricity and heating or hydrogen vectors, enhancing utilization of surplus renewable output. Electricity demand is constructed from sector-specific load profiles including residential, commercial, industrial, and public services consumption patterns (Oberdorfer et al., 2021). These are aggregated across the year at an hourly resolution to capture intraday variability. Seasonal effects are evident in heating-related loads peaking in colder months and reduced daytime net demand when solar generation is high.

At the system level, balancing supply and demand across these time slices requires adequate dispatchable resources or storage solutions. Storage technologies incorporated in the model include lithium-ion batteries for short-term balancing and pumped hydro or hydrogen-based storage for longer-term shifts. Batteries mitigate intra-day mismatches between solar output peaks and evening demand ramps, whereas hydrogen infrastructure offers seasonal flexibility through electrolyzer production during prolonged renewable surpluses. The model optimizes both installed capacities and operational schedules of these assets to minimize total system cost while meeting reliability standards and renewable penetration targets.

The oemof framework, central to modelling the Thuringia energy system, provides a flexible environment for integrating diverse energy sector components into a unified computational representation. Its foundation rests on a graph-based structure composed of buses and components connected by directed edges that indicate energy flows, conversions and consumption. This arrangement is particularly amenable to regional-scale studies because each bus can represent a physical or logical node, such as a substation, district heating plant, or energy storage unit, while edges capture the transfer of resources between them. In this setting, input parameters such as projected electricity prices, technology-specific feedin-profiles, and regionally differentiated demand profiles are associated with relevant nodes or flow connections, ensuring that variations in these drivers are faithfully propagated throughout the model simulation. Within this architecture, the solph library functions as the bridge between the qualitative network description and quantitative optimisation routines. It translates the component-and-bus arrangement into a set of linear equations solvable through optimisers (Krien et al., 2020). The optimiser then determines whether this energy is dispatched immediately to meet demand sinks, stored in intermediate nodes like batteries or pumped hydro reservoirs, or curtailed due to network constraints captured through capacity limits on edges.

The energy system model can alternatively be implemented through the web-based graphical user interface available at https://ensys.hs-nordhausen.de (in.RET EnSys, 2025), which enables intuitive model configuration and execution. The system layout can be constructed using a drag-and-drop approach, allowing the user to visually assemble, parameterize, and link components, thereby facilitating transparent system design.

3. Neural Network Architectures

Selecting neural network architectures appropriate for predicting outcomes in the Thuringia energy system model requires balancing computational efficiency against the ability to capture complex, nonlinear relationships between scalar inputs and target outputs. The inputs available for training comprise variables such as electricity prices, annual full-load hours across diverse technologies, regional demand values for urban and rural zones, and cost parameters like CapEx and OpEx. These scalars feed into networks designed to forecast outputs including installed generation capacities, total storage sizing, and aggregated system costs. Among the options considered, multi-layer perceptrons (MLPs) form the most straightforward architecture. Structurally, an MLP consists of dense layers where each neuron is connected to all neurons in adjacent layers. Their inference speed is attractive for rapid scenario testing; they can approximate continuous nonlinear mappings with reasonable accuracy given adequate hidden layer width and depth (Rahman *et al.*, 2021). However, when input–output relationships involve strong temporal dependencies or spatial correlations, such as capacity outcomes influenced by seasonal resource patterns, MLPs may require substantial training datasets to avoid underfitting. This increase in data demand has implications for the initial dataset generation phase since oemof simulations would need to produce more varied combinations of input parameters. Feedforward neural networks (FFNNs) essentially form a subset of MLP-type designs where information flows strictly forward from inputs to outputs without feedback loops (Medina-Santana and Cárdenas-Barrón, 2022). In energy-system prediction tasks where historical sequences are not directly part of the scalar feature set, FFNNs remain effective; they are simpler to train and less prone to instability during optimisation. Their operation can be formally described by nested compositions of activation functions over weighted sums of inputs:

$$y = f(W_n f(W_{n-1} \dots f(W_1 x + b_1) + \cdots + b_{n-1}) + b_n)$$

where W_i are weight matrices, b_i biases, and f denotes nonlinear activation functions such as ReLU or sigmoid (Mohammadi *et al.*, 2024). Choosing between activations affects learning stability: ReLU generally accelerates convergence but sigmoid might be preferable when input magnitudes vary subtly across low ranges. When the structure of inputs includes ordered or cyclical patterns, like full-load hours varying month-to-month or seasonally adjusted demand coefficients, recurrent neural networks (RNNs) offer advantages through their internal state mechanisms that preserve prior timestep information during inference (Rahman *et al.*, 2021). This enables modelling temporal persistence such as how a capacity plan responds cumulatively to successive high-demand periods. Standard RNNs may suffer from vanishing or exploding gradients when sequences extend far back; long short-term memory (LSTM) variants mitigate this with gated cell structures that regulate information flow over time (Medina-Santana and Cárdenas-Barrón, 2022). For Thuringia's case, LSTMs could be beneficial if datasets embed synthetic sequences representing inter-year variabilities or extreme climatic events. Convolutional neural networks (CNNs), though traditionally associated with spatial data like images, have practical application here when scalar features are reshaped into pseudo-spatial matrices capturing multi-technology relationships over parameter dimensions (Mohammadi *et al.*, 2024). A convolutional layer applies local filters across these structured inputs:

$$h_{ij} = \sum_{m,n} x_{i+m,j+n} \cdot w_{mn} + b$$

where x represents input feature maps, w convolutional kernels, and b bias terms. Such CNN-based processing can detect localised coupling effects, for example between solar full-load hour patterns and PV-related CapEx estimates, more efficiently than fully connected alternatives without requiring explicit manual feature engineering. Hybrid architectures also emerge as appealing candidates when predictions must account for both static scenario descriptors and evolving operational constraints. Combining CNN modules for feature extraction with recurrent units like GRUs or LSTMs lets the network first identify localised parameter correlations then propagate temporal patterns arising from constraints such as ramping limits on dispatchable generation. This configuration does raise computational load during training but could improve generalisation over mixed-input datasets containing both snapshot-like scalars and series-derived aggregates. Back-propagation algorithms used in training these architectures vary according to convergence speed requirements and avoidance of local minima traps. Options include gradient descent with momentum, scaled conjugate gradient, or Levenberg-Marquardt optimisation depending on dataset size and noise characteristics (Rahman *et al.*, 2021). When training on simulation-derived datasets where certain input ranges are sparsely populated, for instance extremely high electricity price scenarios, it can be advantageous to adopt adaptive learning rate methods that slow learning in well-represented regions while stepping faster through poorly covered zones. The quantity of training data required depends markedly on architecture choice. Simpler FFNNs often plateau in performance after several thousand samples due to diminishing returns on additional diversity (Shirzadi, Lau and Stylianou, 2025). In contrast, deeper CNN-RNN hybrids may continue improving accuracy with larger datasets because they exploit fine-grained structural variations present in expanded sample sets. Given that generating each sample entails a full oemof run, choosing architectures must weigh incremental accuracy gain against simulation time costs incurred during dataset creation.

The energy system model developed for Thuringia utilizes a surrogate modeling approach based on a MLP architecture. To ensure optimal model configuration, Bayesian hyperparameter optimization was employed. This method constructs a probabilistic surrogate model of the objective function that maps hyperparameter settings to model performance metrics. By iteratively updating this probabilistic model, it strategically selects new hyperparameter combinations expected to yield performance improvements. This guided search process balances exploration of the hyperparameter space with the exploitation of promising regions, thus achieving a well-generalizing MLP surrogate with minimal computational overhead.

4. Evaluation of Prediction Performance

Assessing how well the surrogate neural networks replicate the outputs of Thuringia's baseline optimisation model requires a robust and multi-faceted evaluation strategy. The patterns of deviation across different operational contexts must be understood, especially because scalar input parameters such as electricity prices, annual full-load hours, and region-specific demands drive strongly non-linear interactions within the original oemof simulations (Jiang *et al.*, 2022). The first step involves determining which performance indicators best capture predictive quality for this application. Root mean square error (RMSE) is particularly informative for capacity predictions, since large discrepancies could directly impair downstream feasibility analyses if acted upon without correction (Zouloumis *et al.*, 2021). Mean absolute percentage error complements RMSE by expressing deviations in relative terms, which is useful when comparing across technologies with very different capacity scales, for example, solar PV installations in MW versus battery storage in MWh (Zhang). Coefficient of determination (R^2) serves as another key metric, reflecting the proportion of variance explained by the model and therefore indicating broader reliability rather than case-specific precision. Using these measures highlights how architecture choice influences predictive behaviour. Multi-layer perceptrons maintain competitive RMSE values when inputs lack strong sequential components; however, their accuracy can deteriorate if seasonal variability encoded in full-load hours exerts major influence on outputs (Rahman *et al.*, 2021).

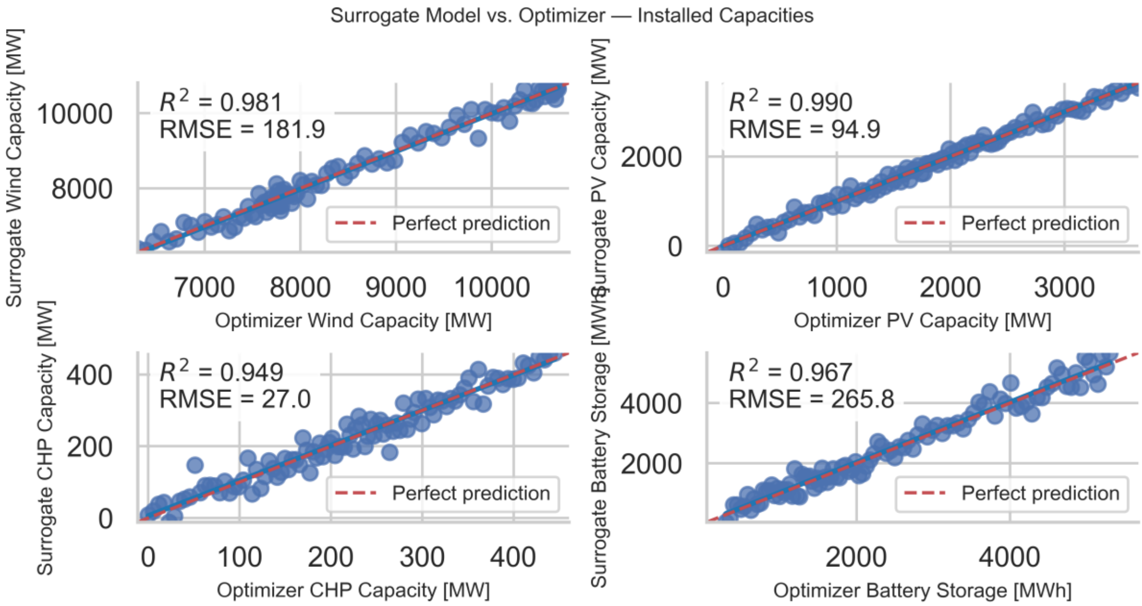

Figure 1: Surrogate model predictions vs. optimizer results for installed capacities of four key technologies.

Figure 1 compares the optimizer results and the surrogate model predictions for installed capacities of wind, photovoltaics (PV), combined heat and power (CHP), and battery storage. The surrogate model is implemented as a feedforward multilayer perceptron trained to approximate the outcomes of a detailed energy system optimization model.

The training dataset consists of 3,500 optimizer runs, generated using Latin Hypercube Sampling to ensure a well-distributed and representative coverage of the 25-dimensional design space (corresponding to 25 installable technologies). The model was trained using 80 % of the data for training and 20 % for validation. The MLP, implemented in PyTorch, comprises two hidden layers with 128 neurons each, ReLU activation functions, and early stopping to prevent overfitting. This network configuration was chosen to balance model complexity and generalization capability while maintaining low computational cost.

The performance of the surrogate model is summarized in Figure 1. For wind capacity, the surrogate model achieved a coefficient of determination of R^2=0.981 and a root mean square error (RMSE) of 181.9 MW across a capacity range of 6.3–10.7 GW. For PV capacity, R^2=0.99 and RMSE=94.9 MW, indicating good alignment with the optimizer outputs despite larger relative uncertainty at low capacity levels. CHP capacity exhibited R^2=0.949 and RMSE=27 MW, which corresponds to a very good approximation relative to its smaller capacity range (0–445 MW). For battery storage capacity, the model achieved R^2=0.967 and RMSE=265.8 MWh. Normalized over the 0–5000 MWh capacity range, this corresponds to an error of approximately 5.6 %**,** which is well within acceptable limits for scenario and planning studies.

Across all technologies, the surrogate model accurately reproduces both the magnitude and trend of the optimized capacity allocations, with only minor deviations from the perfect prediction line. This indicates that the MLP surrogate is able to capture the underlying nonlinear relationships in the energy system model with high fidelity.

5. Conclusion

This study demonstrates the potential of surrogate modeling based on multilayer perceptrons to approximate the results of complex energy system optimization models with high accuracy and significantly reduced computational effort. By training the surrogate on 3,500 optimizer runs generated via Latin Hypercube Sampling, the model successfully reproduced capacity allocations with R^2 values above 0.85 and normalized errors below 6 %. This confirms the capability of MLP surrogates to capture essential system dynamics despite the high dimensionality of the input space.

The proposed approach enables rapid scenario exploration and uncertainty analyses that would be computationally prohibitive with full optimization runs. This opens up new opportunities for large-scale planning, policy analysis, and robust system design, where computational constraints have traditionally limited the scope of analysis.

Future work will focus on extending the surrogate framework from predicting scalar capacity values to forecasting time series outputs such as generation profiles, storage operation trajectories, and system dispatch patterns. This will allow the surrogate to capture not only system-level investment decisions but also their temporal operational dynamics, further enhancing its applicability in energy system planning and operation studies.

6. Acknowledgments

This work is a part of our main project Zero Carbon Cross Emission (ZO.RRO II) and supported by the Thuringian Ministry of Environment, Energy, Nature Conservation and Forestry.

7. References

in.RET EnSys (2025) – Web-based graphical user interface for energy system modeling. Retrieved October 22, 2025, from https://ensys.hs-nordhausen.de

Jiang, J. *et al.* (2022) ‘Surrogate model assisted multi-criteria operation evaluation of community integrated energy systems’, *Sustainable Energy Technologies and Assessments*, 53, p. 102656. Available at: https://doi.org/10.1016/j.seta.2022.102656.

Kalogirou, S.A. (2001) ‘Artificial neural networks in renewable energy systems applications: A review’, *Renewable and Sustainable Energy Reviews*, 5, pp. 373–401. Available at: www.elsevier.com/locate/rser.

Krien, U. *et al.* (2020) ‘Oemof.solph, a model generator for linear and mixed-integer linear optimisation of energy systems’, *Software Impacts*, 6, p. 100028. Available at: https://doi.org/10.1016/j.simpa.2020.100028.

Medina-Santana, A.A. and Cárdenas-Barrón, L.E. (2022) ‘Optimal design of hybrid renewable energy systems considering weather forecasting using recurrent neural networks’, *Energies*, 15, p. 9045. Available at: https://doi.org/10.3390/en15239045.

Mohammadi, S. *et al.* (2024) ‘Surrogate modeling for solving OPF: A review’, *Sustainability*, p. 9851. Available at: https://doi.org/10.3390/su16229851.

Oberdorfer, A. *et al.* (2021) ‘Transition pathways towards a carbon-neutral Thuringia’. *Digitale Bibliothek Thüringen*. Available at: https://doi.org/10.22032/dbt.53108.

Rahman, M.M. *et al.* (2021) ‘Prospective methodologies in hybrid renewable energy systems for energy prediction using artificial neural networks’, *Sustainability*, 13, p. 2393. Available at: https://doi.org/10.3390/su13042393.

Shirzadi, N., Lau, D. and Stylianou, M. (2025) ‘Surrogate modeling for building design: Energy and cost prediction compared to simulation-based methods’, p. 2361. Available at: https://doi.org/10.3390/buildings15132361.

Zhang, G. (2021) ‘Surrogate modelling for multi-carrier distribution networks’. Master's thesis TU Delft, Available at: https://resolver.tudelft.nl/uuid:9a65a8f5-7e1c-4b51-a1ce-ff65880ed735

Zouloumis, L. *et al.* (2021) ‘Dynamic simulation-based surrogate model for the dimensioning of building energy systems’, *Energies*, 14, p. 7141. Available at: https://doi.org/10.3390/en14217141.