Energy Proceedings
Vol 61, 2025
ISSN 2004-2965

Integrating Deep Learning with Modelica for a CO_2 Heat Pump System: A Hybrid Modeling Case Study in Oslo#

Ge Song[1*], Qian Zhang[2], Natasa Nord[1]
1 Energy and Process Technology, Norwegian University of Science and Technology, Norway
2 Department of Energy and Petroleum Engineering, University of Stavanger, Stavanger, Norway
(Ge Song: ge.song@ntnu.no)

ABSTRACT

Modelica is a powerful language for modeling thermal systems. However, the heat pump models available in open source Modelica libraries often lack sufficient accuracy for CO_2 heat pump systems. Unlike conventional heat pumps that use condensers, CO_2 heat pumps employ gas coolers, introducing significant thermodynamic differences not well captured in standard libraries. While commercial libraries offer more accurate CO_2 heat pump models, they are often costly, computationally intensive, and require substantial development effort—especially when the research focus is on the overall heating system rather than the detailed behavior of the heat pump itself. This study presents a deep learning approach to develop a CO_2 heat pump model based on real system data. The trained model was integrated into a Modelica-based thermal system simulation, enabling a hybrid modeling framework. A case study using a school building located in Oslo demonstrated the effectiveness of the approach. The hybrid deep learning-based heat pump model, when combined with the Modelica system model, improved overall simulation accuracy compared to a traditional Modelica-only model. Furthermore, the proposed framework is flexible and can be applied to other industrial thermal systems, offering a scalable solution for data-driven thermal system modeling.

Keywords: Deep learning, Modelica, CO_2 heat pump, thermal system, Hybrid modeling

NONMENCLATURE

Abbreviations	
DNN	Deep Neural Network
FMU	Functional Mock-up Unit
MSE	Mean Squared Error
MAE	Mean Absolute Error

1. INTRODUCTION

The transition toward sustainable energy solutions has intensified the demand for innovative and intelligent heating system technologies, particularly in the building sector. Heat pumps have emerged as a promising low-carbon alternative, but their performance is highly dependent on the design of advanced control algorithms. To enable such optimized control, an accurate, reliable, and computationally efficient model of the heating system is urgently needed. This necessity is especially pronounced for real-time applications, large-scale simulations, and predictive control schemes.

International standards such as the IPMVP [1], ASHRAE Guideline 14-2002 [2], and the CIBSE Operational Performance TM-63 outline methodologies for the development and verification of building thermal energy models. These frameworks classify modeling approaches into three main types: white-box (physics-based), black-box (data-driven), and grey-box (a combination of both). The primary difference among these categories lies in the extent to which physical laws and system knowledge are embedded within the model structure and simulation environment [3]

White-box models are physics-based and rely on detailed energy balance equations—typically heat and mass transfer formulations—to represent system dynamics. Their internal logic is transparent and interpretable, allowing dynamic and modular interaction between sub-components [4]. Building white-box models requires comprehensive parameters, including building envelope characteristics, HVAC system configurations, internal gains, equipment specifications, occupancy profiles, thermal zones, geographical location, and local meteorological conditions. Popular platforms for developing white-box models include TRNSYS and Modelica. TRNSYS offers a highly modular structure well-suited for simulating dynamic thermal systems, allowing users to customize components for a

This is a paper for the Resilient-Applied Energy Symposium and Forum: Resilient energy systems (Resilient2025), Sep. 23-25, 2025, Västerås, Sweden.

wide range of applications. Modelica, on the other hand, supports multi-domain modeling through an object-oriented language and libraries such as Buildings and IDEAS, enabling seamless integration of thermal, control, and electrical subsystems. Both platforms are widely adopted in co-simulation frameworks to facilitate advanced energy system studies, including model predictive control, optimization strategies, and the development of digital twin environments [5]. However, white-box modeling often entails high development complexity and computational overhead due to its reliance on extensive physical data and equation-solving.

In contrast, black-box models are data-driven and rely on historical input–output relationships without explicitly incorporating physical laws. These models can range from simple linear regression techniques [6] to advanced deep learning architectures [7, 8]. Due to their minimal reliance on physical parameters, black-box models are particularly advantageous in scenarios where detailed system information is unavailable or computational resources are limited. In the domain of building energy systems, they are commonly employed to enhance energy management strategies and optimize control algorithms [9]. However, their performance heavily depends on the quality and quantity of available data, and their lack of interpretability makes it difficult to capture underlying physical interactions.

In this study, a full year of operational data from the CO_2 heat pump system was collected, providing a sufficiently large and representative dataset for developing a reliable black-box model. Since constructing a high-fidelity CO_2 heat pump model in Modelica typically requires access to commercial libraries—which may involve significant cost and complexity—the modeling strategy adopted here focuses on replacing only the heat pump component with a black-box model. This surrogate model, developed using a deep neural network (DNN) trained on the collected operational data, was seamlessly integrated with the remaining system components—including the building model, water tank model, borehole model, and a district heating module (primarily used to supplement domestic hot water)—all modeled in Modelica and coupled through the Functional Mock-up Unit (FMU) standard. This hybrid framework allowed for faster simulation while retaining physical accuracy in the overall system, enabling effective analysis of flexibility strategies such as demand response and load shifting.

2. METHODOLOGY

2.1 Heating system

Fig. 1 Voldsløkka school building

This study is based on a Norwegian demonstration project located at Voldsløkka School in Oslo, as illustrated in Fig. 1. The facility spans an area of approximately 11,100 m² and employs a CO_2-based ground-source heat pump system to provide space heating. Thermal energy is extracted from a borehole field and upgraded through two gas coolers—a high-temperature and a low-temperature unit—before being distributed throughout the building. A 400L buffer tank is connected to the high-temperature gas cooler to store excess heat.

2.2 White box system model

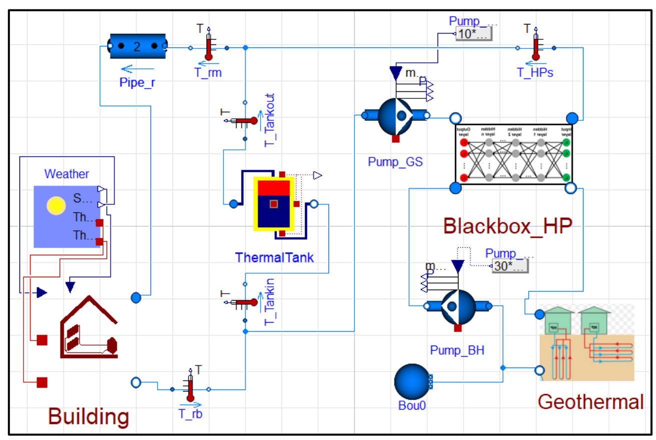

Fig. 2 White box system model in Modelica

A detailed Modelica model was developed using the IDEAS library in Dymola environment to capture dynamic interactions among four key components: building model, water tank, CO_2 heat pump, and borehole model as shown in Fig. 2. These subsystems were interconnected to reflect the operational logic of the real system, with heat and cooling flows determined by real-time demand and resource availability.

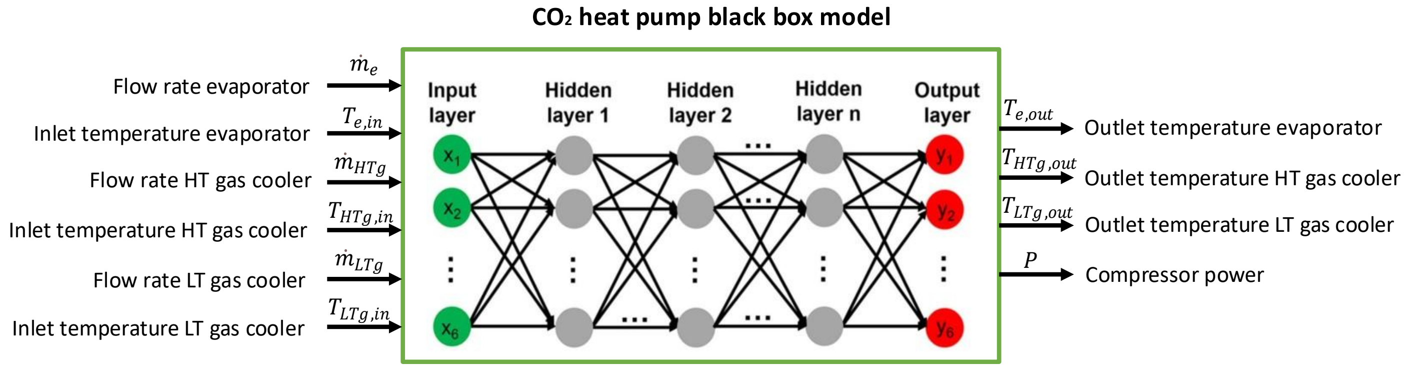

Fig. 3 Overview of the black box model of the CO_2 heat pump system.

2.3 Black box CO_2 heat pump model

A deep learning CO_2 heat pump model based on DNN was developed to simulate the performance of a CO_2 heat pump as shown in Fig. 3. Key input features include the inlet temperature and mass flow rate at the high-temperature gas cooler $\dot{m}_{HTg}$, $T_{HTg,in}$, low-temperature gas cooler $\dot{m}_{LTg}$, $T_{LTg,in}$, and evaporator $\dot{m}_e$, $T_{e,in}$, while the outputs are the outlet temperatures of these three components $T_{HTg,out}$, $T_{LTg,out}$, $T_{e,out}$ and the power use of the compressor P. The DNN architecture comprises three hidden layers with 128, 128, and 64 neurons (each using ReLU activation) followed by an output layer with four neurons (with ReLU activation). All input features were scaled to a [0,1] range using min–max normalization. Approximately 90% of the cleaned dataset was used for model development, with this subset further divided into 70% for training and 30% for validation, while the remaining 10% was held out for testing. The model was trained using the Adam optimizer (learning rate = 0.001) to minimize the mean squared error (MSE) loss, and early stopping (monitoring validation loss) was employed to prevent overfitting. The model performance was evaluated by computing the mean absolute error (MAE) on the test set and by visually comparing the predicted outputs against the measured values, confirming that the DNN could accurately capture the behavior of the heat pump.

A neural network was developed using Python Keras library to model the CO_2 heat pump system. To enable integration with the Modelica-based Dymola environment, the trained Keras model (saved as an .h5 file) was first imported into MATLAB using the import Keras Network function. The resulting network was then loaded into Simulink, where the solver was set to fixed step to allow compatibility with FMU export. An FMU containing the neural network was generated from Simulink and subsequently imported into Dymola, enabling co-simulation between the black-box neural network and the white-box physical model components.

3. RESULTS

3.1 Black box model validation

Fig.4 illustrates the training and validation curves of the DNN-based CO_2 heat pump model. In the left plot, both the training and validation losses MSE decrease rapidly during the initial epoch and then converge to stable values around 0.0012. The close alignment

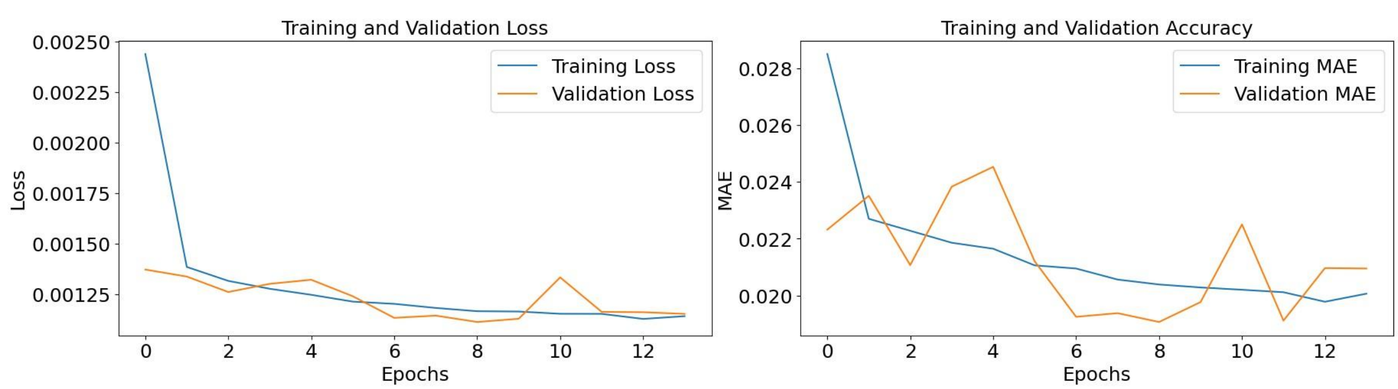

Fig. 4 Training and validation mean squared error (MSE) loss and mean absolute error (MAE) curves of the ANN-based CO_2 heat pump model

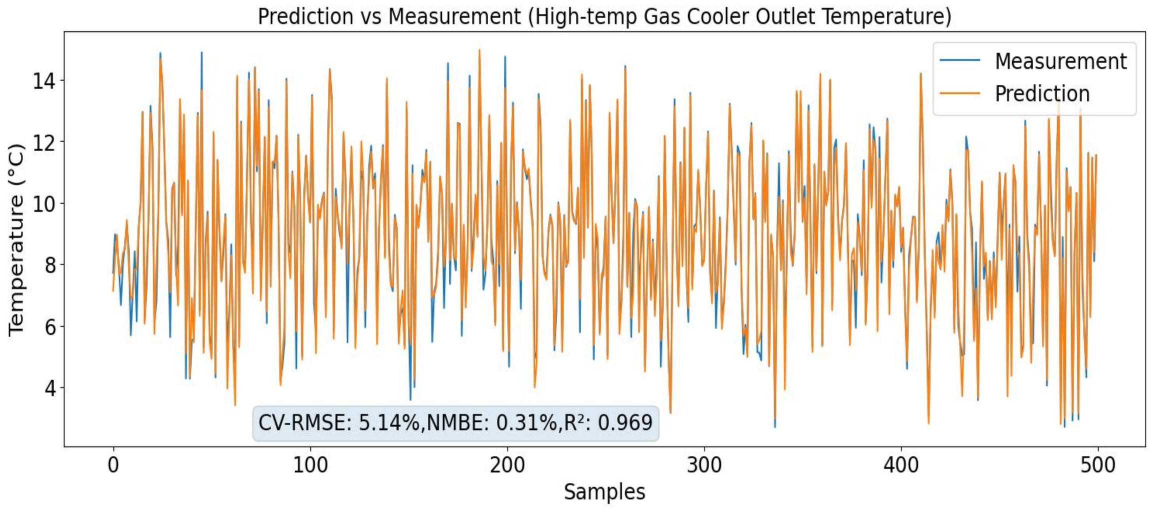

Fig. 5 Comparison of measured and predicted outlet temperature of the high-temperature gas cooler

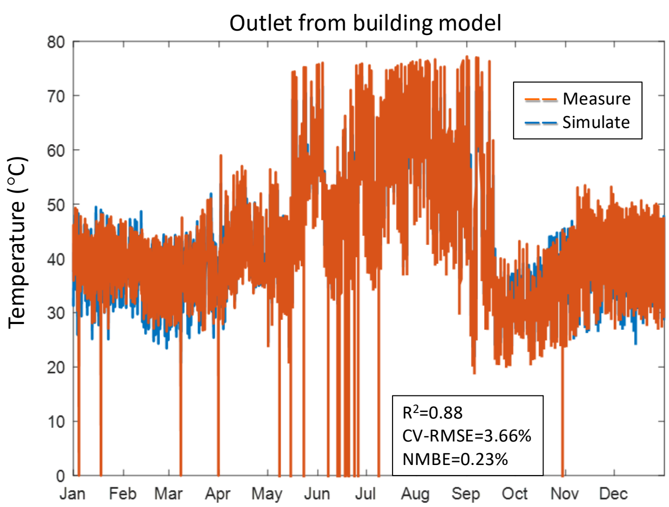

Fig. 6 Validation of the building return temperature with the DNN-based CO_2 heat pump model integrated in Modelica

between training and validation loss indicates that the network learns effectively without evident overfitting and generalizes well to unseen data. The right plot in Fig.4 presents the MAE curves, which also decline steadily and stabilize around 0.019–0.021. Although minor fluctuations appear in the validation MAE, its overall trend closely follows that of the training MAE. These results confirmed that the DNN successfully captures the relationship between the inlet temperatures and mass flow rates of the gas coolers and evaporator, and the corresponding outlet temperatures and compressor power, achieving low and consistent prediction errors across both training and validation sets.

Fig.5 presents a comparison between the measured and predicted outlet temperature of the high-temperature gas cooler over 500 validation samples. The predicted values closely follow the measured data—CV-

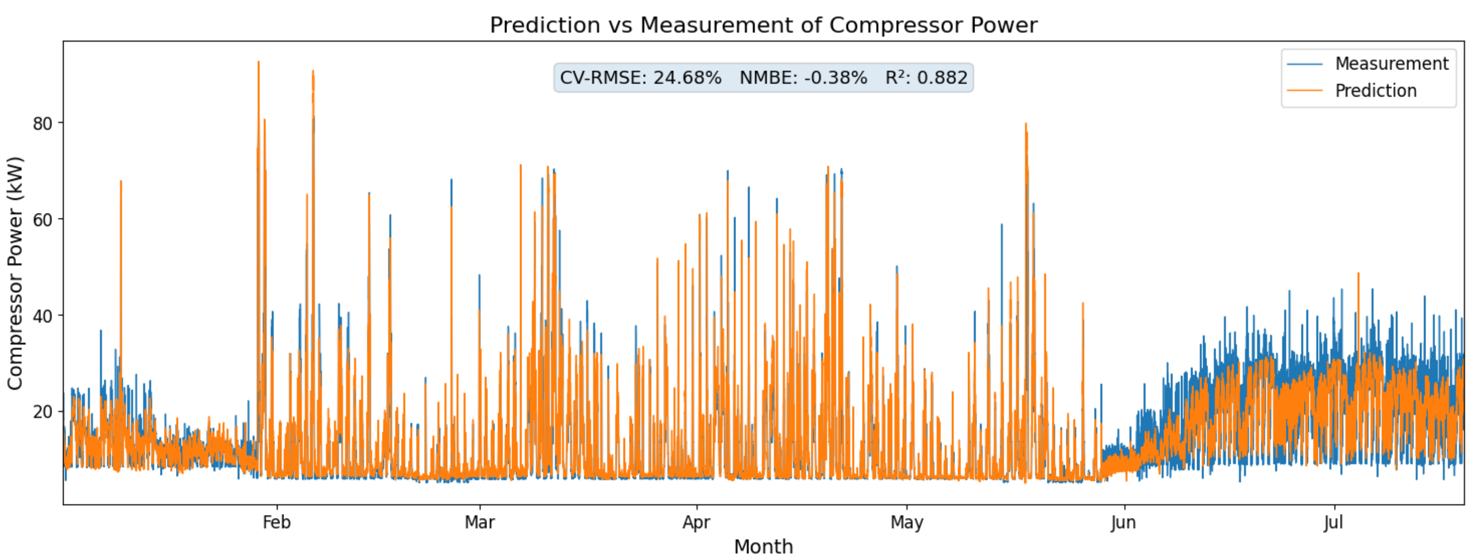

Fig. 7 Validation of the compressor power of the DNN-based CO_2 heat pump model integrated in Modelica

RMSE = 5.14%, NMBE = 0.31%, and R^2 = 0.969—indicate that the model achieves high accuracy, with very low systematic bias and strong agreement between predicted and measured values. These results confirmed that the DNN-based approach provided reliable predictions of the CO_2 heat pump performance under varying operating conditions.

3.2 Black box model combines white box model validation

Fig. 6 and 7 present the validation results of the DNN-based CO_2 heat pump model integrated into the system-level building simulation in Modelica. For the outlet (return) temperature, the simulated results closely followed the measured data throughout the year, successfully capturing both seasonal and short-term variations. The statistical indicators (R^2 = 0.88, CV-RMSE = 3.66%, NMBE = 0.23%) demonstrate good agreement between simulation and measurement, with low normalized errors and negligible bias. Similarly, for the compressor power, the predicted values reproduce the measured trends across half of the year, reflecting the short-term operational fluctuations. Although the CV-RMSE is relatively higher (24.68%) due to the inherently dynamic nature of compressor behavior, the overall agreement remains satisfactory, with R^2 = 0.882 and NMBE = –0.38% indicating negligible bias. These results confirm that coupling the DNN-based CO_2 heat pump model with the system-level building model provides a reliable representation of both thermal and electrical performance under real operating conditions.

4. CONCLUSIONS

In this study, an DNN was developed to model the performance of a CO_2 heat pump and subsequently integrated into a whole-building system in Modelica. The DNN, trained with inlet temperatures and mass flow rates as inputs and outlet temperatures and compressor power as outputs, demonstrated fast convergence and stable performance, as confirmed by the MAE and MSE curves. Validation against measured data showed high predictive accuracy, with low CV-RMSE and NMBE values and R^2 consistently above 0.95 for individual component predictions. When coupled with the building energy model, the DNN-based CO_2 heat pump maintained reliable performance, accurately reproducing the return water temperature over a full year with R^2 = 88.29%, CV-RMSE = 3.66%, and NMBE = 0.23%. These results confirmed that the proposed deep learning approach not only captures the detailed dynamics of the CO_2 heat pump but also integrated seamlessly into system-level simulations, providing a robust and efficient tool for building energy analysis and optimization.

ACKNOWLEDGEMENT

This work was collaboratively prepared as part of the cooperative projects "ChiNoZEN" (Key technologies and demonstration of combined cooling, heating and power generation for low-carbon neighbourhoods/buildings with clean energy) and "ARV" (Climate Positive Circular Communities). The authors gratefully acknowledge financial support from the Research Council of Norway (ENERGIX, project no. 304191), the Ministry of Science and Technology of China (MOST, project no. 2019YFE0104900), and the European Union's Horizon 2020 research and innovation programme (grant agreement no. 101036723, ARV project).

REFERENCE

[1] J. Cowan, International performance measurement and verification protocol: Concepts and Options for Determining Energy and Water Savings-Vol. I, International Performance Measurement & Verification Protocol, 1 (2002).
[2] A.G. ASHRAE, Guideline 14-2002: Measurement of Energy and Demand Savings, ASHRAE, Atlanta, (2002).
[3] K. Arendt, M. Jradi, H.R. Shaker, C. Veje, Comparative analysis of white-, gray-and black-box models for thermal simulation of indoor environment: Teaching building case study, in: Building Performance Analysis Conference and SimBuild: Co-organized by ASHRAE and IBPSA-USA, ASHRAE, 2018, pp. 173–180.
[4] G.R. Ruiz, C.F. Bandera, T.G.-A. Temes, A.S.-O. Gutierrez, Genetic algorithm for building envelope calibration, Applied energy, 168 (2016) 691–705.
[5] A. Shahcheraghian, H. Madani, A. Ilinca, From white to black-box models: A review of simulation tools for building energy management and their application in consulting practices, Energies, 17 (2) (2024) 376.
[6] G. Ciulla, A. D'Amico, Building energy performance forecasting: A multiple linear regression approach, Applied Energy, 253 (2019) 113500.
[7] E. Mocanu, P.H. Nguyen, M. Gibescu, W.L. Kling, Deep learning for estimating building energy consumption, Sustainable Energy, Grids and Networks, 6 (2016) 91–99.
[8] Y. Sun, F. Haghighat, B.C. Fung, A review of the-state-of-the-art in data-driven approaches for building energy prediction, Energy and Buildings, 221 (2020) 110022.
[9] C. J Hopfe, R. S McLeod, T. Rollason, Opening the black box: Enhancing community design and decision making processes with building performance simulation, in: Building Simulation 2017, IBPSA, 2017, pp. 920–929.