Science and Technology for the Built Environment, (2025) **31**, 533–550

ISSN: 2374-4731 print / 2374-474X online
DOI: 10.1080/23744731.2025.2460374

Taylor & Francis
Taylor & Francis Group

Development and performance evaluation of empirical models compatible with building energy modeling engines for unitary equipment

SHAHZAD YOUSAF[1,2]*, CRAIG R. BRADSHAW[1,2], RUSHIKESH KAMALAPURKAR[2]** and OMER SAN[3]

[1]*Department of Mechanical Science and Engineering, University of Illinois, Urbana-Champaign, IL, USA*
[2]*Center for Integrated Building Systems, Department of Mechanical and Aerospace Engineering, Oklahoma State University, Stillwater, OK, USA*
[3]*Department of Mechanical, Aerospace and Biomedical Engineering, University of Tennessee, Knoxville, TN, USA*

Modeling of the air conditioner (AC) and heat pump (HP) systems is critical for maintaining thermal comfort and regulating building energy use. They are designed to meet the load of extreme climatic and occupancy conditions. This makes them oversized for the majority of their operational hours, resulting in part-load operation. Their part-load operation affects efficiency and energy consumption. Many of the equipment models currently used in building energy models (BEM) require large quantities of experimental data and/or do not capture modern equipment behavior. This paper discusses models compatible with BEM for AC and HP and evaluates their efficacy. First, six distinct models are identified in the literature and assessed for their predictive capabilities of cooling/heating capacities and coefficient of performance (COP). This evaluation is conducted using high-fidelity data from 6 equipment pieces: one roof-top air conditioner (RTU), three residential 2-speed split-systems, and two variable-speed heat pump systems. Drawing from insights gained from previous models, a new empirical model is then proposed to effectively capture the behavior of contemporary equipment while minimizing the need for extensive testing data. All models are ranked based on mean absolute percentage error (MAPE) as well as the coefficients employed by the model. The proposed model, with 20 coefficients, is found to be superior compared with the legacy models, predicting heating/cooling capacity and COP below 5% for the experimental units and sensible heat ratio (SHR) below 5% for all the units in AC operation.

1. Introduction

Space heating and cooling are almost everywhere in the United States (US) and are becoming increasingly common in other developed countries in the world. The most common method to cool/heat the indoor air is the electricity-driven vapor compression cycle (VCC). This makes air conditioning an energy-intensive process. In the US, approximately 76% of the total energy consumed in a residential home is dedicated to space and water heating/cooling applications (U.S. Energy Information Administration 2015). Continuous worldwide initiatives aim to enhance building energy efficiency and promote sustainability.

Building energy models (BEMs) play an important part because of their ability to simulate the thermal behavior of buildings and predict their energy consumption. The thermal behavior of a building is often challenging to model because of the heat transfer processes that occur at the building envelope and air conditioning and heat pump (ACHP) systems. BEM is a crucial tool used to integrate energy-efficient technologies and design principles into both new and existing buildings (Feng, Lu and Wang 2019; Shen, Braham, and Yi 2018; Jung et al. 2023) to achieve significant energy and cost savings. While BEM has diverse applications, for example in ensuring compliance with energy codes, supporting green certification, qualifying for incentives, enabling real-time building control, improving building performance, and conducting fault detection studies (Fan and Xia 2018; Yang, Gao, and You 2023; Gholamibozanjani et al. 2018; Zhang and Hong 2017), the accuracy and reliability of the BEM is highly dependent upon an ACHP system model (Joe, Im and Dong 2020) because of the high energy use by these systems in buildings.

Received May 7, 2024; accepted January 13, 2025
Shahzad Yousaf, MSc, Student Member ASHRAE, is a Graduate Research Assistant. **Craig R. Bradshaw, PhD**, Full Member ASHRAE, is an Associate Professor. **Rushikesh Kamalapurkar, PhD**, is an Associate Professor. **Omer San, PhD**, is an Associate Professor.

*Corresponding author e-mail: syousaf3@illinois.edu
**Current affiliation: Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, USA.

Since the ACHP systems are designed to meet the load requirements for extreme conditions, for that reason they are normally oversized, and as a result, run at part load (Cetin et al. 2019). The part load operation is achieved by adjusting the output of AC and HP components, such as the compressor and fan, to match the actual heating or cooling load of a building, while operating at less than full capacity. While the majority of the equipment operates at part load, there is a rapid drop in their efficiency as the equipment operation moves away from the optimal operating point (Doty 2010; Ahmadisedigh and Gosselin 2022). Therefore, there is a pressing demand for the development of an ACHP model that not only aligns with BEM tools but also has the capability to replicate the performance of advanced, highly efficient variable speed systems.

Modeling of the ACHP system is essential for analyzing energy usage and indoor thermal comfort quality control. Typically, physics and data-based modeling approaches are employed for ACHP systems. In the first approach, the system model is formulated by employing thermodynamic laws to link the inputs and outputs which are referred to as physics-based models and are also categorized as white box models. For physics-based models, it is essential to have complete knowledge of the processes and the system. For instance, Domanski and Didion (1983), Herbas et al. (1993), Rossi (1995), Shen (2006), Bell (2012) presented physics-based models. These models are computationally expensive and require iterative techniques to be implemented to find system outputs. Such models are used to optimize the heat pump system on the level of refrigerant level, for example. The second approach is the black-box models which are developed from system performance data. Inputs and outputs are linked using regression or machine learning techniques. Unlike their physics-based counterparts, these models do not necessitate an extensive understanding of the system, yet they offer remarkable accuracy and faster response times (O'Neill and O'Neill 2016). The aim of developing such models is to integrate them with BEM tools to calculate building-specific seasonal performance factor and/or compare specific heat pump units using a seasonal coefficient of performance factor.

In literature, artificial neural network (ANN) is largely used to model heat pump systems among the black-box models when the available data is large either in terms of input information or the number of data points. For instance, ANN is used to predict the steady-state cooling capacity of an AC system by Yousaf et al. (2022) using 40 inputs to the model with 10 hidden neurons. A large amount of field data with 37 input variables is used to model the ground source heat pump system's long-term behavior using ANN by Puttige et al. (2022). Similarly, Ye et al. (2020) uses field measurements from retrofit residential housing to predict energy consumption using 9 inputs and 40 neurons in the hidden layer. Kizilkan (2011) formulated a variable speed heat pump system model to investigate the compressor frequency for capacity control using ANN with better energy savings using 80 data points with 6 inputs. Lower inputs of 4 but with large data of 38071 data points, Belman-Flores et al. (2013) evaluated the energetic performance of a variable speed heat pump system. Ledesma and Belman-Flores (2014) identified the best performance zones for heat pump systems using only 4 inputs by utilizing a data set of 54654 data points. Not only a large amount of data is required, but in addition, the modeler's proficiency in selecting and utilizing an appropriate training dataset while minimizing overfitting is the primary determinant of success when employing ANN models (Lee et al. 2019). This is due to the absence of reliable guidelines for determining the optimal training structure, including the number of neurons in the input, output, and hidden layers (Yousaf, Shafi, and Ahmad 2018; Shafi et al. 2006). Selecting the ultimate trained model often hinges on the expertise of the researcher. Additionally, implementing ANNs poses a unique challenge due to the substantial number of model parameters generated during training. To illustrate, consider a scenario where a heat pump model utilizes just six inputs (e.g., indoor and outdoor dry/wet bulb temperatures, compressor and fan data) and employs a single hidden layer with merely ten neurons to predict three outputs (such as cooling capacity, COP, and sensible heat ratio). In this case, the training process yields a total of 103 model parameters. The exact count may vary, influenced not only by the volume and quality of the training data but also by the proficiency of the model developer. These challenges underline why ANN and machine-learning black-box models may not be suitable for integrating the heat pump model with BEMs.

The second category in black-box models are empirical models. The coefficients in the polynomials are mostly evaluated using data fitting techniques. These models find applications in ACHP systems and their individual components to streamline the modeling process, reduce the computational burden, and mitigate the risk of overfitting. These models are characterized by explicit polynomial expressions and possess predefined input variables in terms of number and type. For instance, an empirical model is formulated to represent all types of heat exchangers for their effectiveness and charge estimation along pressure drop using polynomials by Park, Guo, and Rasmussen (2021). A second-order polynomial function is developed using inlet pressure, the mass flow rate of refrigerant, inlet quality, temperature of the secondary fluid, and heat exchanger length. In another study for heat exchangers, an empirical model is developed to calculate the effectiveness of the heat exchanger as a function of the number of transfer units (NTU), fluid inlet temperatures, and surface temperatures of heat exchangers by Lim, Cheon, and Jeong (2018). The scope of empirical modeling extends to compressor performance as well, where parameters like compressor capacity and power consumption are estimated. This is achieved by a cubic polynomial equation in Air conditioning, heating, and refrigeration institute (AHRI) well-known 10-coefficient mapping model as a function of evaporating and condensing temperatures (AHRI-540 2020). More complex phenomena like vapor injection ((Tello-Oquendo et al. 2019)) and variable speed operation (Shao et al. 2004) in compressors are also modeled using empirical modeling. These exemplary cases, drawn from the literature, underscore the widespread adoption of empirical models within the realm of air conditioning and refrigeration system components. A second-order polynomial as a function of supply air flow rate is used to calculate the fan power consumption in Shao, Yang, and Zhang (2010). Empirical correlations are not limited to modeling individual ACHP components but are also indispensable in developing comprehensive system-level models. For example, a linear model using a single variable is employed to estimate the energy requirements of both a variable-capacity air source heat pump (ASHP) and a ground source heat pump (GSHP) by considering the outside temperature as an independent variable by Safa (2012). Single and multiple input variable empirical models are formulated for ASHP and GSHP systems to predict the power consumption by Peskova (2023). They investigated different models with one, two, and three independent variables selected on the basis of Spearman and Pearson coefficients from the data. The independent variables selected for this purpose are outdoor dry bulb temperature, supply air temperature, and refrigerant temperature of the evaporator. It is reported that models having all three inputs are performing significantly better in prediction of the power consumption. Simple polynomial-based heat pump model correlations are developed for a dual, water and air, source heat pump to predict the condenser capacity along with compressor power consumption through data generated from a computer simulation by Marchante-Avellaneda et al. (2019). The independent variables selected are compressor frequency, fan frequency, pressure ratio of the compressor, inlet and exit temperatures of the secondary fluid across the evaporator and condenser along with the relative humidity of the air. An empirical model for the heat pump is provided by Afjei and Wittwer (1995). The model uses bi-quadratic polynomials based on fluid entering temperature to the evaporator and exit temperature to the condenser to calculate the heating capacity as well as compressor work. A heat pump model is given by Morrison (1994) based on catalog data provided compressor manufacturer for the evaporation and condensation temperatures. They provided bi-quadratic polynomials for the heating capacity and power consumption of the compressor. The component models are closed through energy conservation laws. A steady-state empirical model is developed by Filliard, Guiavarch, and Peuportier (2009) to predict the heat pump performance. Separate models are formulated for full load, part-load, and defrost operation for the heat pump. The model uses indoor and outdoor temperatures to calculate the capacity and COP using a second-order polynomial equation in the full load model. The part load model predictions are carried out by the introduction of part-load factor in the model while in the defrost model, a degradation coefficient is applied to the heating capacity. They validated the full load model through manufacturer data while part-load and defrost models were not validated due to data unavailability. A reference model is formulated by Kim et al. (2010) for a heat pump for fault detection in cooling mode to predict seven dependent features of the heat pump by using dry bulb temperature for indoor and outdoor along with indoor humidity ratio as independent variables. They formulated polynomial-based models upto 4th order and found that higher-order models are accurate but may worsen the interpolation for small data sets. Using source and sink temperatures as input to the model, a simple polynomial-based heat pump model is presented in Gupta and Irving (2013). The regression-based model developed from the test results of the heat pump is intended to be simple enough to be used in a spreadsheet to provide predictions for monthly loads and energy consumption. The difference between the manufacturer's provided performance and *in situ* performance of the heat pump is evaluated for COP using polynomial models involving different independent variables by Chesser et al. (2021). The independent variables selected for different models are outside dry bulb temperature, rain precipitation, dew point air temperature, mean sea level pressure, relative humidity, and wind speed.

Air-conditioners and heat pump models that leverage inputs commonly available from manufacturers and employ polynomials of lower degree, requiring fewer model coefficients, are highly advantageous for building energy modelers Yousaf et al. (2024). Using lower-degree polynomials not only reduces the volume of training data required but also helps in avoiding over-fitting—a prevalent issue in data driven models, like ANN-based models, due to the lack of standardized architectures (Yousaf, Shafi, and Ahmad 2018). Over-fitting compromises model generalizability, often occurring when researchers, in pursuit of high accuracy, inadvertently use higher-degree polynomials that increase the number of coefficients. It is essential to select model coefficients judiciously to maintain consistent and reliable performance across both training and test datasets. On this criterion, an air-conditioning model is formulated by Brandemuehl, Gabel, and Andresen (1993) for fixed-speed equipment as a function of indoor and outdoor temperatures along with indoor air flow rate. The model can predict the operational performance of the system other than rated conditions by adopting correction factors in indoor and outdoor temperatures along with an indoor air flow rate. A similar model for the heat pump is presented by Winkelmann et al. (1993) but with a lower number of model coefficients. A relatively simple model for capacity calculation is presented by Nyika et al. (2014). For power calculation, their model separately calculates the power of each component of the system while demanding more component-specific data as well as training data. Their model can be used for both cooling and heating applications. Using Brandemuehl, Gabel, and Andresen (1993)'s model as a baseline, Hjortland (2018) formulated a variable speed unitary equipment model with higher training data requirements. A variable speed cooling-only model is presented by Cheng, Braun, and Horton (2021). The model takes sensible load as input from a BEM and predicts total cooling capacity and power consumption of the unit.

Despite extensive research on modeling unitary equipment, there is no clear consensus on the most suitable black-box model for a diverse range of equipment. Additionally, we have identified critical limitations in the existing models that necessitate a shift in our research focus.

- ANN and machine learning models lack a standardized design and are susceptible to modifications dependent on the characteristics and quantity of the training data.
- A significant number of literature models rely on input variables that are challenging to obtain in real-world settings, rendering them impractical for widespread

application. This is more pronounced in machine learning-based models.

- Many of the models currently available do not comprehensively capture the intricacies of modern heat pump equipment and their variable operating conditions.
- A significant limitation seen in some models is the omission of part-load operation of the equipment, which has considerable significance in ensuring the correctness of the model's predictions.

In light of these limitations, we are focusing on finding a model in literature or formulating a new one that better aligns with the practical demands of contemporary energy modeling. Formulating a model is akin to cooking a recipe; while all the ingredients are available, mastering the art lies in understanding the precise proportions, timing, and quantities of each ingredient required. In (Yousaf et al. 2023), the authors use experimental data from a ten-ton rooftop unit (RTU) and 3 state-of-the-art high-efficiency split heat pumps with a capacity range of 3 to 5 tons to study the influential model inputs and proposed two sets of critical model inputs for empirical and semi-empirical models. Based on the model input recommendations from (Yousaf et al. 2023), we identify models with such inputs that can be integrated into an existing BEM, smoothly. Our next steps involve evaluating the performance of these selected models, assessing their suitability for real-world applications, and exploring opportunities for further refinement. This approach allows us to not only critique the limitations of existing models but also advance the field by identifying and enhancing models that show promise in accurately representing real-world behavior. Our research focuses on the following:

- Selecting BEM engines compatible empirical models from the literature, including Brandemuehl, Gabel, and Andresen (1993), Winkelmann et al. (1993), Nyika et al. (2014), Hjortland (2018), Cheng, Braun, and Horton (2021), and EnergyPlus (2022). Specifically regarding EnergyPlus, for cooling, we used the EnergyPlus model *Coil:Cooling:DX:SingleSpeed* and for heating, we used the EnergyPlus model "*Coil:Heating:DX:SingleSpeed* and will be referred to as EnergyPlus model in the latter sections.
- Conducting a rigorous evaluation of these selected models using high-fidelity experimental data obtained from state-of-the-art unitary equipment, including high-efficiency commercial units of 10 tons (35 kW) and 5 residential units of capacities ranging from 3.5 to 5 tons (12 to 17.6 kW).
- Identifying the strengths and weaknesses of the evaluated models in accurately capturing the complex behavior of air conditioners and heat pumps.
- Proposing an improved empirical model for variable speed equipment based on the lessons learned from the evaluation of legacy models.
- Enhancing the accuracy and efficiency of empirical models for air conditioners and heat pumps to enable more precise predictions and enhance energy efficiency in buildings.

Through this comprehensive evaluation and improvement process, our work aims to contribute to the advancement of empirical modeling techniques for unitary air conditioners and heat pumps, facilitating more accurate and energy-efficient designs.

2. Legacy models

This section provides a brief description of empirical models used to model the performance of unitary equipment. Due to space limitations, detailed descriptions of legacy models are skipped here and can be found in the original references. Legacy models, compatible with BEM, identified from the literature have few inputs and outputs that are common across all models. These are identified as standard inputs and outputs as shown in Figure 1. Standard inputs include outdoor dry bulb temperature (ODT), indoor wet bulb temperature (IWB), and indoor dry bulb temperature (IDT). The standard outputs are heating/cooling capacity and the sensible heat ratio (SHR) (cooling mode only). In addition to standard inputs and outputs, indoor air supply, compressor

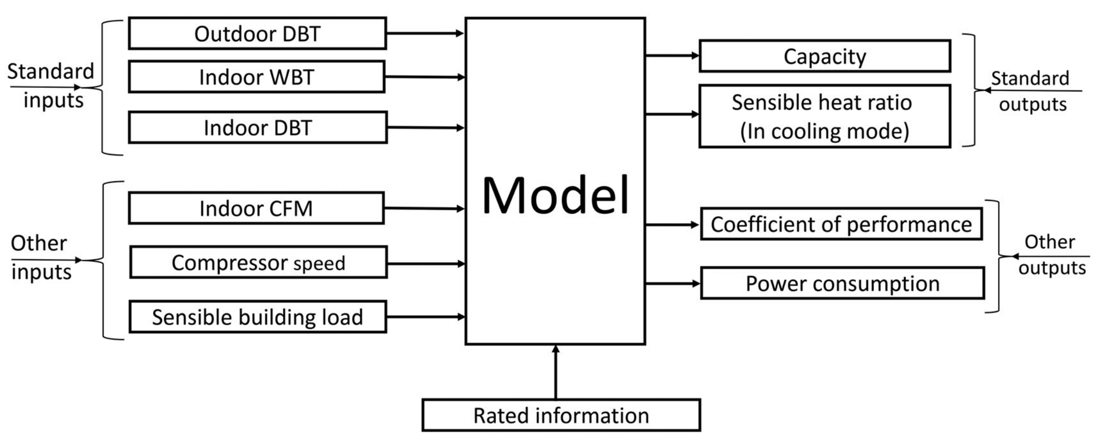

Fig. 1. General form of legacy models.

speed, and building sensible load are sometimes used as inputs while coefficient of performance (COP) or power are the outputs as well. Apart from conventional inputs, the performance of the equipment at the standard rated conditions is also required. AHRI standards define standard rating conditions as a comparison for performance characteristics AHRI (2008). These performance characteristics include heating/cooling capacity, and COP. The industry has accepted two standard rating conditions, one each for cooling and heating purposes to compare the equipment performance. The standard rated conditions are shown in Table 1.

At the standard rating conditions, cooling capacity along with COP is recorded with the highest compressor and fan speeds. Legacy models use correction factors to the rated performance of the equipment to predict cooling capacity and COP at various operating conditions. Details of correction factors for the capacity and COP are provided in Table 2. Another metric, bypass factor (BF) is also recorded at the rated conditions (Cheng, Braun, and Horton 2021). The BF represents the amount of air traveling through the coil that does not directly make contact with the surface of the cooling coil. All the legacy models use the bypass factor approach (BFA) to calculate the SHR at different operating conditions, except for the model presented in Cheng, Braun, and Horton (2021).

3. The proposed model

The current work proposes a modification to the Brandemuehl model (Brandemuehl, Gabel, and Andresen 1993)) for the estimation of equipment performance at variable speed. The proposed model uses model inputs recommended by Yousaf et al. (2023) for black box models.

Figure 2 shows the inputs as well as the outputs of the model in cooling operation. The proposed modification uses correction factors for temperature, air flow rates, and compressor speed but with fewer data point requirements to train

Table 1. Standard rating conditions for cooling and heating AHRI (2008).

AHRI Standard Rating conditions	Air Entering Indoor Unit		Air Entering Outdoor Unit	
	Dry bulb °F (°C)	Wetbulb °F (°C)	Dry bulb °F (°C)	Wetbulb °F (°C)
Cooling (A-Test)	80 (26.7)	67 (19.4)	95 (35)	75 (23.9)
Heating (H1-Test)	70 (21.1)	60 (15.6)	47 (8.33)	43 (6.11)

Table 2. Correction factors and the degree of polynomials used in legacy models.

Author	Model Coefficients	Correction Factors				SHR
		CFM	Temperature	Part load	Compressor speed	
Brandemuehl	16	Linear	Bi-quadratic	–	–	BFA
EnergyPlus	18	Quadratic	Bi-quadratic	Cubic for power	–	BFA
Nyika	21	Quadratic	Bi-quadratic	Bi-quadratic	–	BFA
Cheng	20	–	Bi-quadratic	Degree 4	–	–
Winkelmann	17	Linear	Bi-quadratic	–	–	–
Hjortland	26	Quadratic	Bi-quadratic	–	Cubic	BFA

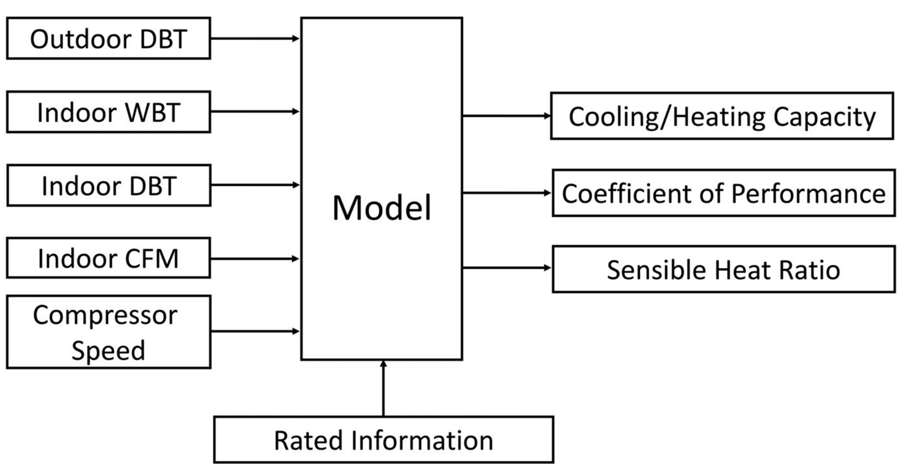

Fig. 2. Proposed model inputs and outputs (SHR only in cooling mode).

the model coefficients. These correction factors account for differences in ambient temperatures, indoor air flow rate, and compressor speed. The differences in ambient temperatures are modeled using a bi-quadratic function of ODT and IWB while the air flow rate and compressor speeds are modeled using a linear correction factor. The model predicts the capacity using,

$$\dot{Q}_{tot} = f_{cap,temp} \cdot f_{cap,\dot{V}} \cdot f_{cap,\Omega} \cdot \dot{Q}_{rat}, \quad (1)$$

and the COP is predicted using,

$$COP = f_{cop,temp} \cdot f_{cop,\dot{V}} \cdot f_{cop,\Omega} \cdot COP_{rat}. \quad (2)$$

The model makes use of the BFA to calculate the SHR as follows in the cooling operation of HP.

$$SHR = \frac{h(T_{ID}, \omega_{adp}) - h_{adp}}{h_{ID} - h_{adp}}, \quad (3)$$

where,

$$h_{adp} = h_{ID} - \frac{h_{ID} - h_{sup}}{1 - BF}. \quad (4)$$

The BF is given by,

$$BF = \frac{h_{sup} - h_{adp}}{h_{ID} - h_{adp}}. \quad (5)$$

The model coefficients are trained using performance data at various temperatures and compressor speeds. Table 3 shows the correction factors for the capacity and COP prediction. Correction factors with a total of 20 coefficients for cooling operations are listed in Table 4.

The proposed model is used for heating operations by using indoor dry bulb temperature instead of indoor wet bulb temperature, similar to the legacy models. In the heating mode operations, SHR is not calculated and the model outputs are the heating capacity and COP. For both heating and cooling, heat pump power consumption ($\dot{W}$) is obtained directly by dividing the capacity by the COP as given below,

$$\dot{W} = \frac{COP}{\dot{Q}_{tot}}. \quad (6)$$

4. Methodology

This section provides details about model training and testing methodology along with the performance metrics adopted for the evaluation of the various models. Furthermore, detailed information about the collection of unitary equipment data is provided along with an uncertainty analysis of the data collected for training and model evaluation.

4.1. Model training and testing

Models presented in this study predict the performance of the equipment by correcting for the changes in operating conditions to the rated performance using correction factors. The correction factors utilized by models are for capacity and COP prediction. Coefficients of correction factors are obtained from unitary equipment data using regression methods. The cost function used to get the model coefficients is given as,

$$Error = \sum_{i=1}^{n} \left(\frac{X_{Pred} - X_{\text{Exp}}}{X_{\text{Exp}}} \right)^2. \quad (7)$$

whereas X_{Pred} and X_{Exp} are predicted and experimental values for capacity or COP.

For training purposes, the unitary equipment data is divided into training and testing data sets for AC and HP as shown in Figure 3. The Nyika model uses the highest amount of model coefficients in the power correction factor which is 16. So, for model training, 17 data points are selected from the data set of each unit, manually. Manual selection of training data points is preferred over algorithmic random selection due to a consistent use of similar conditions representative of similar operations in all data sets. IWB and ODT vary from their minimum to their maximum values in the training data set for each unit. The training data sets also included all the fan speeds as well as compressor stages/speeds. The training data set provided us with model coefficients which are used to predict the capacity

Table 3. Correction factors and the degree of polynomials used in the proposed model.

Author	Model Coefficients	Correction Factors				SHR
		CFM	Temperature	Part load	Compressor speed	
Proposed model	20	Linear	Bi-quadratic	–	Linear	BFA

Table 4. Proposed model correction factors.

Correction Factors	Equation to be fit
$f_{cap,temp}$	$a_0 + a_1T_{iwb} + a_2T_{iwb}^2 + a_3T_{odt} + a_4T_{odt}^2 + a_5T_{iwb}T_{odt}$
$f_{cap,\dot{V}}$	$b_0 + b_1(\frac{\dot{V}}{\dot{V}_{rat}})$
$f_{cap,\Omega}$	$k_0 + k_1\Omega$
$f_{cop,temp}$	$c_0 + c_1T_{iwb} + c_2T_{iwb}^2 + c_3T_{odt} + c_4T_{odt}^2 + c_5T_{iwb}T_{odt}$
$f_{cop,\dot{V}}$	$d_0 + d_1(\frac{\dot{V}}{\dot{V}_{rat}})$
$f_{cop,\Omega}$	$l_0 + l_1\Omega$

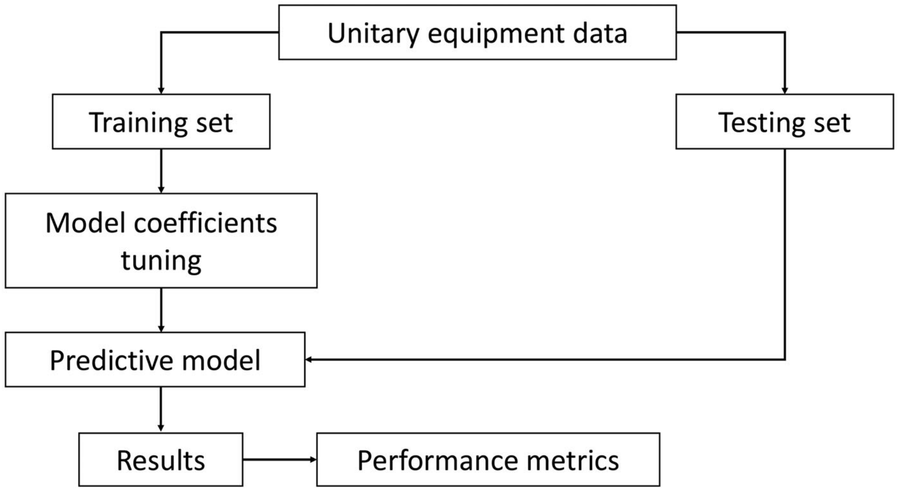

Fig. 3. Model training and testing methodology.

and COP on the test data set. The complete process of model training and testing is shown in Figure 3.

4.2. Performance metrics

To quantify the accuracy of the model prediction, the selected performance metric is the MAPE. It is useful when comparing models with varying error magnitudes, as it provides a percentage-based error metric, which can be easier to interpret. Additionally, it provides a holistic view of the model's predictive capabilities. It is calculated as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_{\exp} - X_{pred}}{X_{\exp}}\right|.100\%. \quad (8)$$

4.3. Unitary equipment data

To evaluate the performance of models in both cooling and heating modes, experimental data is recorded on one commercial and 5 residential pieces of unitary equipment, as per ASHRAE standard 37 ASHRAE (2009). The commercial unit is a 10-ton (35.2 kW) fixed-speed RTU AC, referred to as Unit-1, and is tested in in-house psychrometric chambers by varying conditions of indoor and outdoor temperatures. Indoor air flow rate and compressor speeds are kept fixed and the equipment performance is captured in only cooling mode.

Among the residential split systems, five units are examined, with three of them being 2-speed systems, operating with only two compressor and fan speeds. These systems undergo testing in both full and part-load operations across various operational and environmental conditions. The three residential 2-speed systems, designated as Unit-2 (3.5 tons/12.3 kW), Unit-3 (4 tons/14 kW), and Unit-4 (5 tons/17.5 kW), are tested in cooling and heating modes, with further details on fan and compressor speeds provided in Table 5.

Unit-2, Unit-3, and Unit-4 undergo extensive testing in both heating and cooling modes, spanning a wide range of temperature conditions within an external psychrometric chamber. Additionally, Unit-5 (4 tons/14 kW) and Unit-6 (5 tons/17.5 kW) are variable speed units, also tested in both heating and cooling modes. Operational and temperature range specifics for these variable speed systems are detailed in Table 5.

Unit-5 undergoes testing within an in-house psychrometric chamber, while experimental data for Unit-6 is sourced from Shoukas et al. (2022). Figure 4 illustrates the experimental setup, featuring the 10-ton RTU and Unit 5 under test in the in-house psychrometric chambers. For in-house testing, steady-state data is recorded for a minimum of 15 min, with sensors capturing data at a sampling frequency of 1 Hz. Figure 5 provides insight into the placement of pressure, temperature, and mass flow sensors used in the in-house data collection.

4.3.1. Uncertainty analysis

In the current work, uncertainty is estimated for the aforementioned units using a propagation analysis using the method developed in Taylor and Kuyatt (1994). Table 6 lists the accuracy specifications for each sensor, as well as a brief description of how it is used. An EES Klien and Alvardo (2000) script is written to calculate the uncertainty propagation for all the unitary equipment data. Uncertainty in the cooling capacity for the fixed speed RTU ranged from ±0.74 kW (±0.21 ton) to ±1.2 kW (0.35 ton). For the 12.3 kW (3.5 tons) unit, the uncertainty range is ±0.33 kW (±0.10 ton) to ±0.73 kW (±0.21 ton) while in the residential units with capacities of 14(4) and 17.6(5) kW(tons), the uncertainty range is ±0.38 kW (±0.11 ton) to ±0.75 kW (±0.2 ton) and ±0.47 kW (±0.13 ton) to ±0.83 kW (±0.23 ton) respectively. In the variable speed heat pump, the air side uncertainty in capacity is found to in between ±0.26 kW (±0.08 ton) to ±0.22 kW (±0.06 ton).

4.4. Model ranking

To evaluate the performance of the models, two key metrics are considered: MAPE and the number of coefficients in the

Table 5. Details of the units along with the experiment data ranges.

Unit no.	Rated capacity, Tons (kW)	Type	Mode	ODT range, °F (°C)	IWBT range, °F (°C)	IDT range, °F (°C)	Supply air, CFM (m^3/hr)	Compressor speed (RPM) /stage (Hi-Low)	No. of data points
Unit-1	10 (35.16)	RTU	Cooling	67–115 (19.4–46.1)	47–77 (8.3–25)	56–92 (13.3–33.3)	3600 (6114)	2700	31
Unit-2	3.5 (12.3)	Split sytem	Heating	0–60 (−17.8–15.6)	55–60 (12.7–15.5)	60–80 (15.5–26.6)	800–1750 (1360–2970)	Hi-Low	35
			Cooling	55–125 (12.8–51.7)	57–72 (13.9–22.2)	75–80 (23.9–26.7)			52
Unit-3	4 (14)		Heating	0–60 (−17.8–15.6)	55–60 (12.7–15.5)	60–80 (15.5–26.6)	1130–1750 (1920–2970)	Hi-Low	35
			Cooling	55–125 (12.8–51.7)	57–72 (13.9–22.2)	75–80 (23.9–26.7)			54
Unit-4	5 (17.5)		Heating	0–60 (−17.8–15.6)	55–60 (12.7–15.5)	60–80 (15.5–26.6)	1500–2000 (2548–3398)	Hi	30
			Cooling	55–125 (12.8–51.7)	57–72 (13.9–22.2)	75–80 (23.9–26.7)			50
Unit-5	4 (14)		Heating	17–60 (−8.3–15.5)	55–60 (12.7–15.5)	60–75 (15.5–23.9)	700–1600 (1188–2715)	6000–1400	39
			Cooling	67–115 (19.4–46.1)	55–72 (12.8–22.2)	65–80 (18.3–26.7)		3725–1400	31
Unit-6	5 (17.5)		Heating	−7–60 (−21.6–15.5)	–	60–76 (15.5–24.4)	615–1402 (1044–2380)	7000–1800	41
			Cooling	55–120 (12.7–48.9)	46–75	68–82 (20–27.8)	872–1456 (1480–2470)	4248–1827	44

model. The models are individually assessed based on their MAPE values, where lower values indicate better accuracy and hence accurate formulation of the model. An average MAPE is obtained for all the heat pump performance metrics for the units predicted by every model. The average score is separate for heating and cooling applications.

The second metric to assess the models is the number of coefficients in the model. A model with higher accuracy but the least number of coefficients is desirable. The number of the model coefficient represents the amount of training data it would require. Additionally, more coefficients in the model can over-fit the model on the training data set, which might lead to poor prediction on the test data set. To rank all the models, a criterion is chosen that combines both the MAPE and the number of coefficients. The final rank score, *S*, is a weighted sum of MAPE, having a weight of 90%, and the number of model coefficients with a weight of 10%. This is given as:

$$S = \sigma.\frac{1}{n}\sum_{i=1}^{n} MAPE + \rho.C, \tag{9}$$

where n is the number of performance metrics predicted, C is the number of model coefficients. At the same time, σ (set to 0.9) and ρ (set to 0.1) represent the weightage of MAPE and the number of model coefficients. To determine the weights assigned to MAPE and the number of coefficients, we considered the relative importance of each criterion in accurately assessing model performance. Given the fundamental significance of accuracy in predictive modeling, MAPE was assigned a higher weight of 90%, while the number of coefficients, although important for model simplicity and training data requirements, was given a weight of 10% to avoid overly penalizing models with slightly higher complexity. This ranking methodology enabled a holistic evaluation of the models, taking into account both accuracy (MAPE) and number of model coefficients (C) in determining their overall performance.

5. Results and discussion

All the models are trained and tested as per the methodology laid out in Figure 3. Each model is trained on the same data and tested on the same test data set as well. The coefficients of the model are obtained through applying methods of least squares. Once the coefficients are obtained, the model is updated and a predictive model is complete. The predictive model is used to predict the unseen test data. Individual model performance is assessed by an error metric, MAPE. All the models are coded in Python programming language. This section provides the results of all the models evaluated for cooling and heating operations. All the percentages given in this section are percent MAPEs.

5.1. Cooling mode predictions

This section presents the results of the model evaluation for cooling operations. Figure 6 shows the performance of all

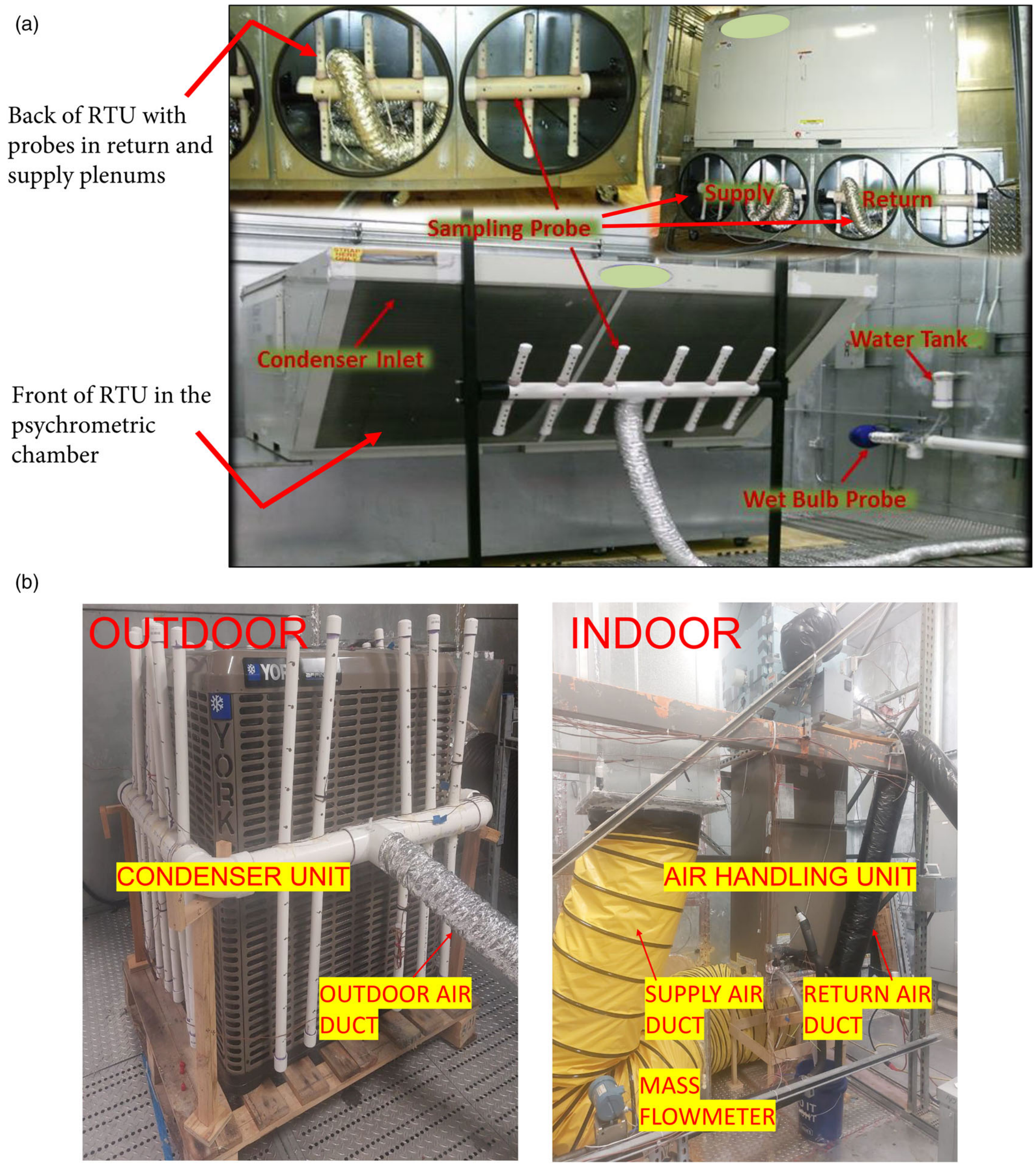

Fig. 4. Test units under test at in-house psychrometric chambers.

the evaluated models for the cooling capacity, Figure 7 provides details about the COP predictions, while Figure 8 shows the MAPEs for SHR being predicted for different units included in this study.

The evaluation of MAPE values across various models provides a comprehensive comparison of prediction accuracy for cooling capacity, COP, and SHR metrics. The EnergyPlus model performs well for fixed-speed and 2-speed units, with MAPE values under 5%, but struggles with variable-speed units, particularly Unit-5, where the cooling capacity MAPE is 16.6%. The Brandemuehl model exhibits higher errors, especially for variable-speed units, with a 17.27% MAPE for Unit-5's cooling capacity. The Nyika model shows higher COP prediction errors for Units 2 and 3, with MAPEs of 8.13% and 15.9%, respectively. The Hjortland model is consistent across all units, maintaining MAPE values below 5% for most metrics, except for Unit-3's COP. The Cheng model performs well for cooling capacity predictions but has significant inaccuracies in COP predictions, notably with Unit-2 and Unit-5 showing MAPEs of 7% and 11%, respectively. The proposed model stands out with high accuracy across all performance metrics, maintaining MAPE values well under 5% for most units, except for a 7.32% MAPE for Unit-3's cooling capacity. To further clarify what the MAPEs represent in terms of model accuracy, Figure 9 is provided for cooling capacity, illustrating the measured versus predicted cooling capacities for each model across the different units evaluated in this study. By directly comparing measured and predicted values, these plots provide a visual interpretation of model accuracy

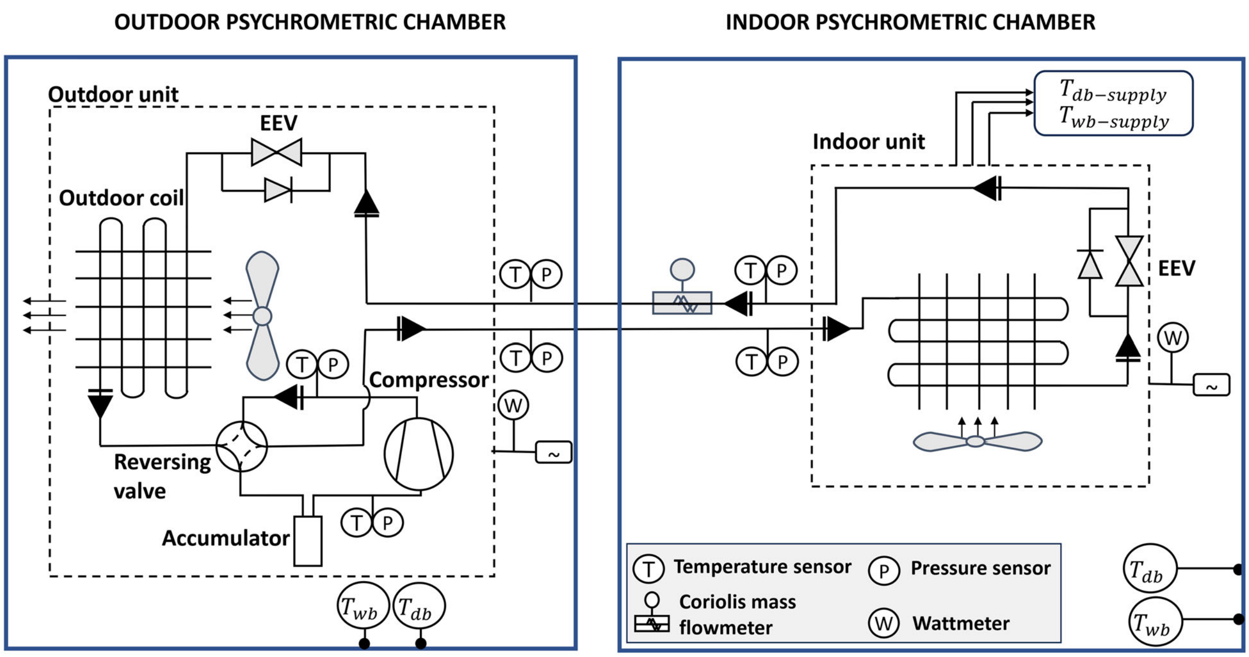

Fig. 5. Schematic showing locations of key sensors in the split heat pump system installed in a pair of psychrometric chambers.

Table 6. Sensor accuracy and usage for variable speed residential units.

Sensor	Accuracy	Use
Thermocouples	0.9 °F (0.5 °C)	Air side temperature
RTD	0.2 °F (0.11 °C)	Refrigerant side temperature
Pressure sensor	±0.06% of F.S	Refrigerant side pressure
Power meter	±0.1% of F.S	Indoor and outdoor unit power
Coriolis flow meter	±0.1% of rate	Refrigerant flow rate
Dew point hygrometer	±.27°F (.15°C)	Dew point

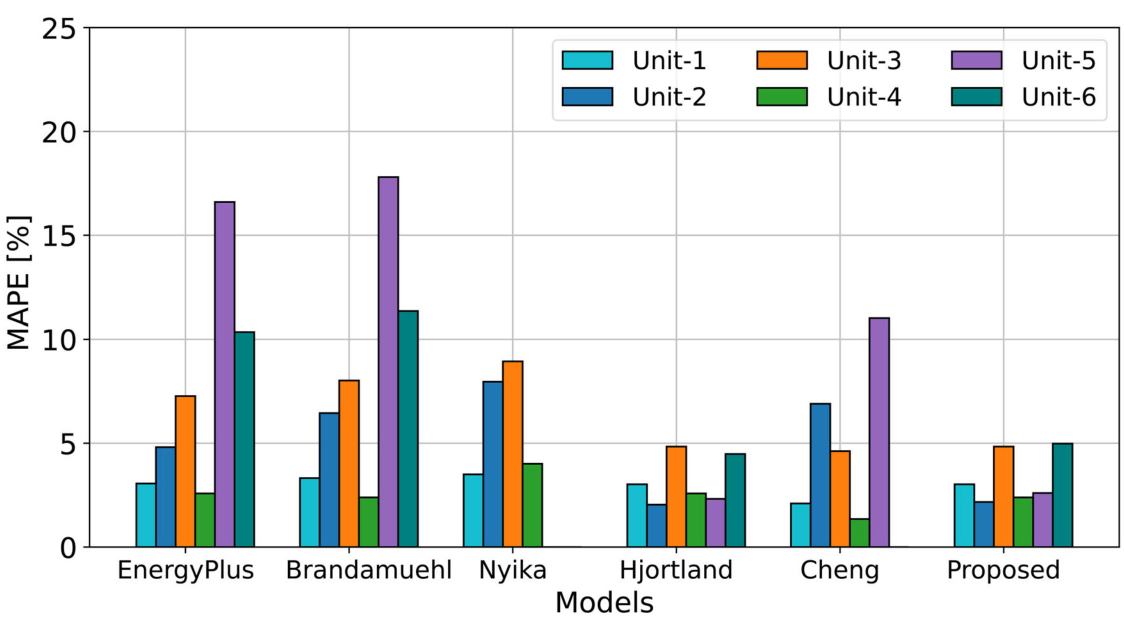

Fig. 6. Model performance comparison for total cooling capacity prediction.

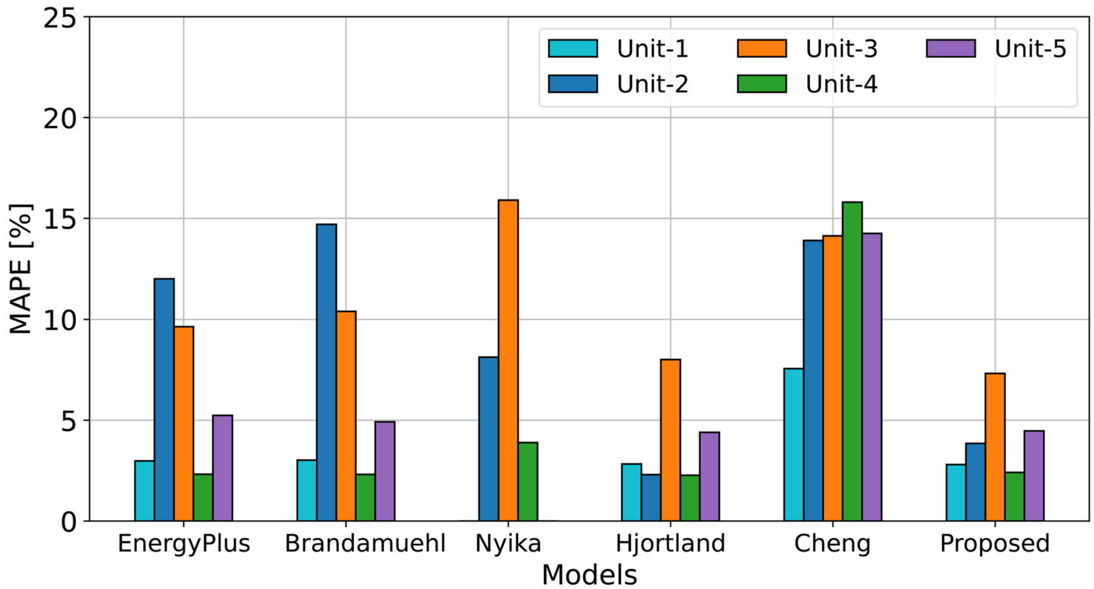

Fig. 7. Model performance comparison for COP prediction, cooling mode.

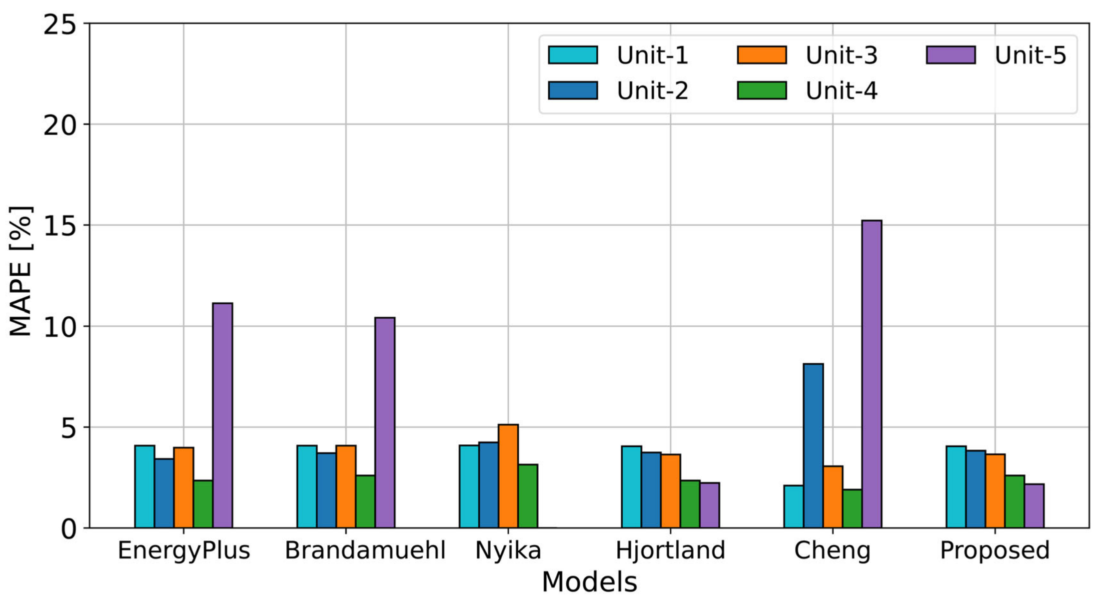

Fig. 8. Model performance comparison for SHR prediction.

beyond the MAPE values, allowing readers to assess how closely each model captures the actual system performance across varying operational conditions.

Key observations include the consistent accuracy of fixed-speed units across all models, significant challenges in modeling variable-speed units, and the superior performance of the proposed model. Addition of a linear correction factor for compressor speed enhances accuracy of the proposed model on the test data, making it a robust and effective tool for building energy modeling.

5.2. Heating mode predictions

The selected models are evaluated on three residential 2-speed units, and 2 variable-speed heat pump systems. Legacy as well as the currently proposed model are assessed through MAPE, similar to the cooling mode of operation. For HP, the Cheng model is not used because it applies only to AC. Another model, the Winkelmann model applies only to HP and is added to the list of models to be evaluated. For HP, only the total heating capacities and COP are to be predicted. Figures 10 and 11 show the MAPEs for heating capacity and COP, respectively. The EnergyPlus model shows low MAPE values (under 5%) for heating capacity in 2-speed systems but struggles with variable-speed systems, particularly Unit-5 (MAPE of 20.6%) while maintaining the COP predictions remain within 5%. Brandemuehl's model performs similarly, with low MAPE for Unit-4 but higher errors for other 2-speed and variable-speed systems. Nyika's model shows higher MAPE values

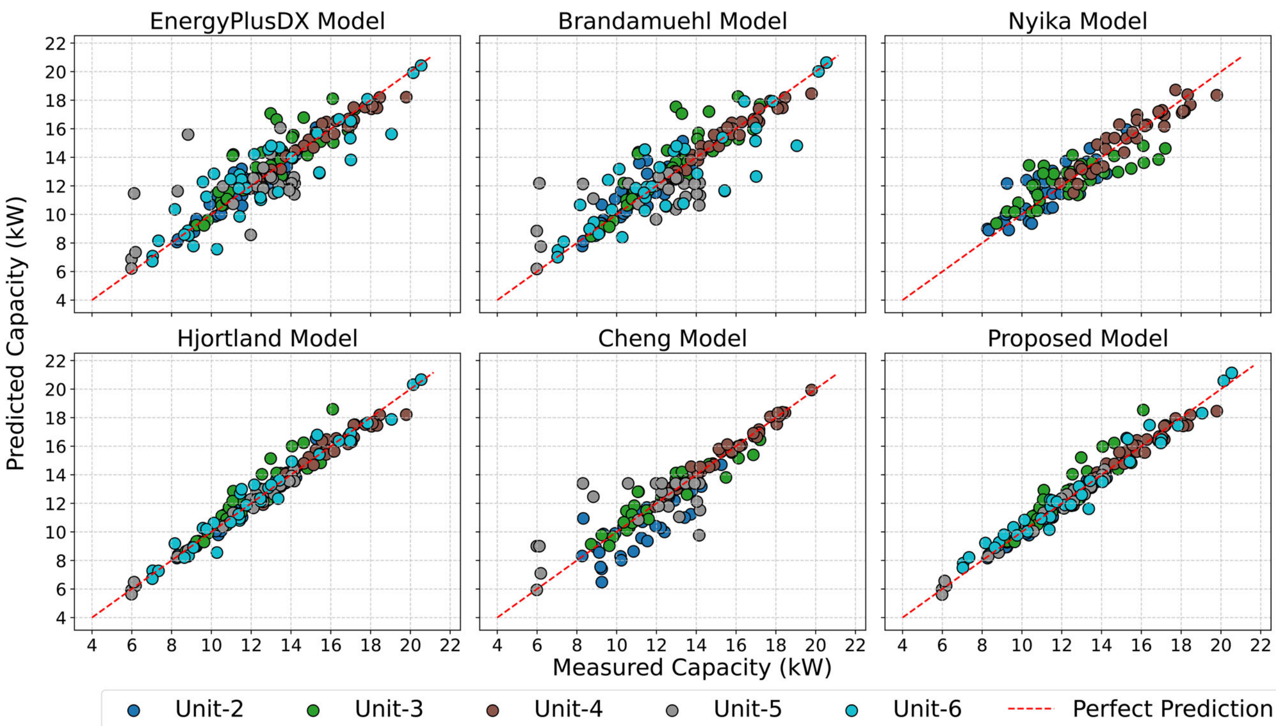

Fig. 9. Model performance in predicting the cooling capacity for two speed and variable speed units.

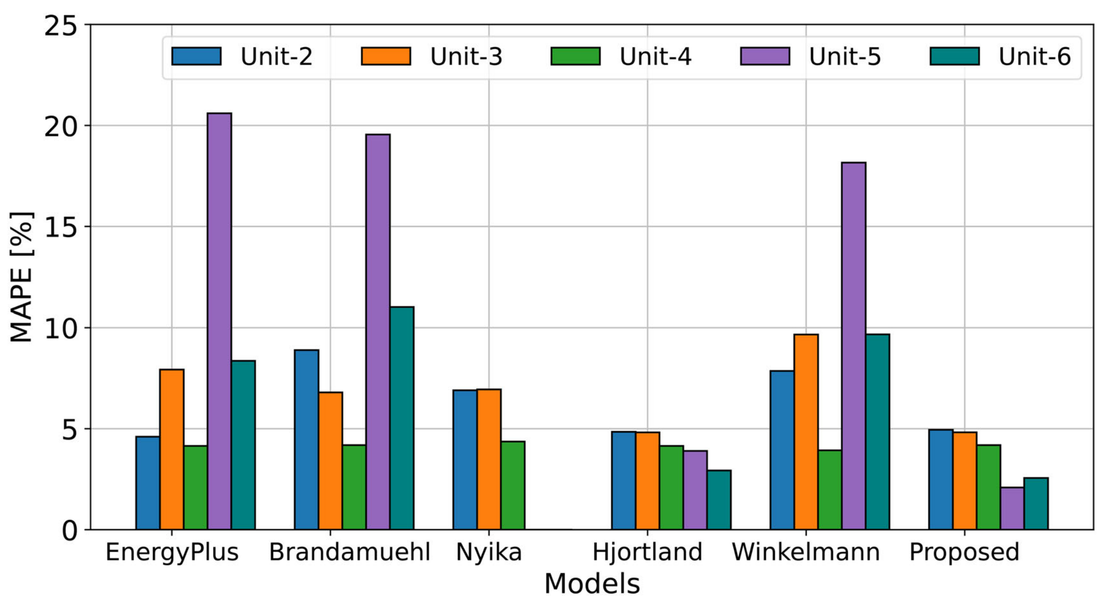

Fig. 10. Model performance comparison for heating capacity prediction.

for COP predictions due to its use of more coefficients trained on limited data, indicating overfitting in the predictions. The Hjortland model maintains MAPE values under 5% for most units, except for Unit-5's COP. The Winkelmann model predicts low MAPE for Unit-4 but has higher errors for other units in both heating capacity and COP predictions. The proposed model consistently outperforms others, maintaining MAPE values under 5% for all units in predicting heating capacity and COP, demonstrating its robustness and effectiveness. To enhance understanding of the MAPEs and their implications for model accuracy, we include Figure 12 showing measured versus predicted heating capacities across the different units assessed. These plots offer a direct visual comparison, highlighting how each model performs in capturing the actual heating capacity across varied conditions. This additional insight complements the MAPEs by demonstrating each model's accuracy in replicating real system behavior.

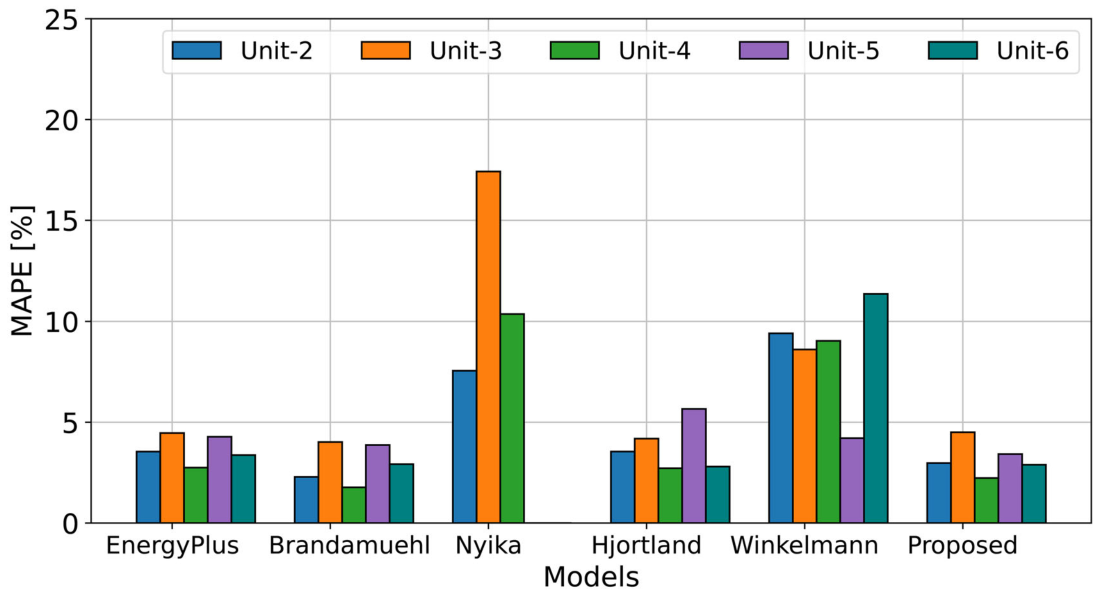

Fig. 11. Model performance comparison for COP prediction, heating mode.

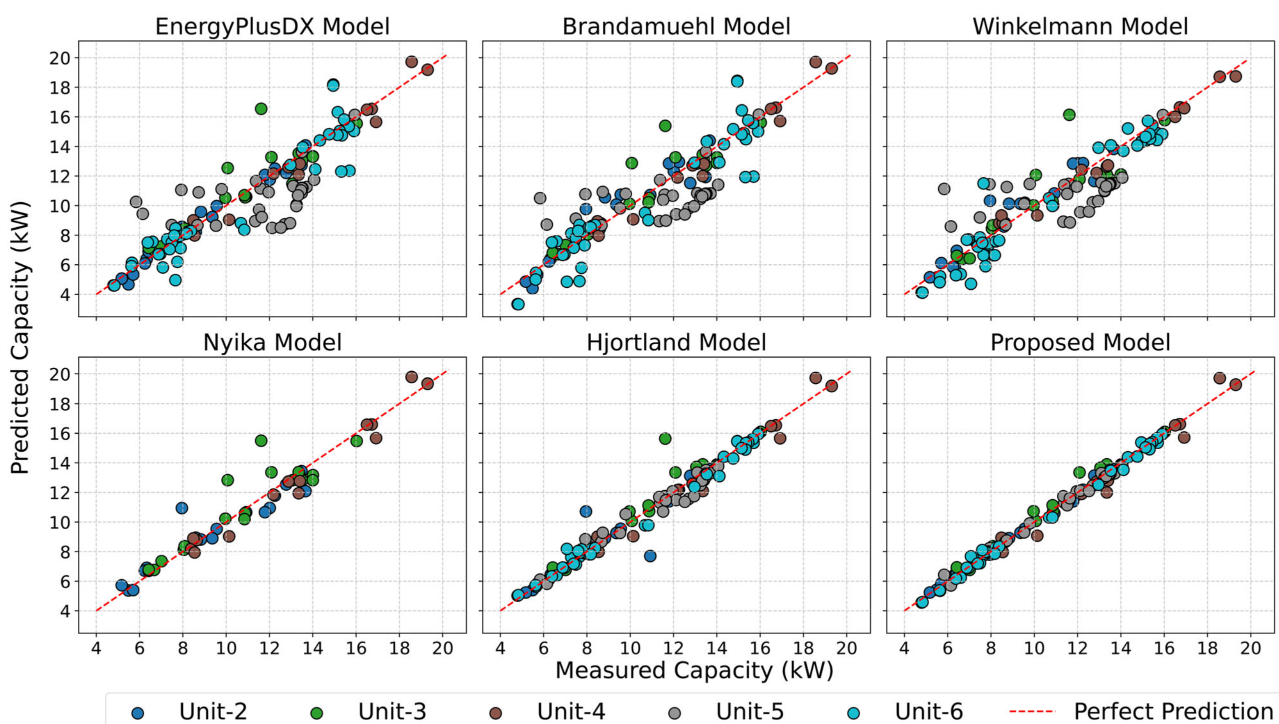

Fig. 12. Model performance in predicting the heating capacity for two speed and variable speed units.

Key observations from the evaluation include the consistent and accurate predictions for 2-speed systems (Units 2 and 4), which consistently show low MAPE values across all evaluated models. This indicates that these models are well-calibrated for 2-speed systems. In contrast, variable-speed systems (Units 5 and 6) present significant challenges, as evidenced by the higher MAPE values, highlighting the difficulty in accurately modeling the complexities of variable-speed operations. Additionally, models with a higher number of coefficients, such as the Nyika model, tend to show higher COP prediction errors when trained on limited data, suggesting that models with fewer coefficients may offer more accurate predictions in such scenarios. Notably, the proposed model demonstrates superior performance by consistently maintaining low MAPE values across all metrics and units, underscoring its robustness and

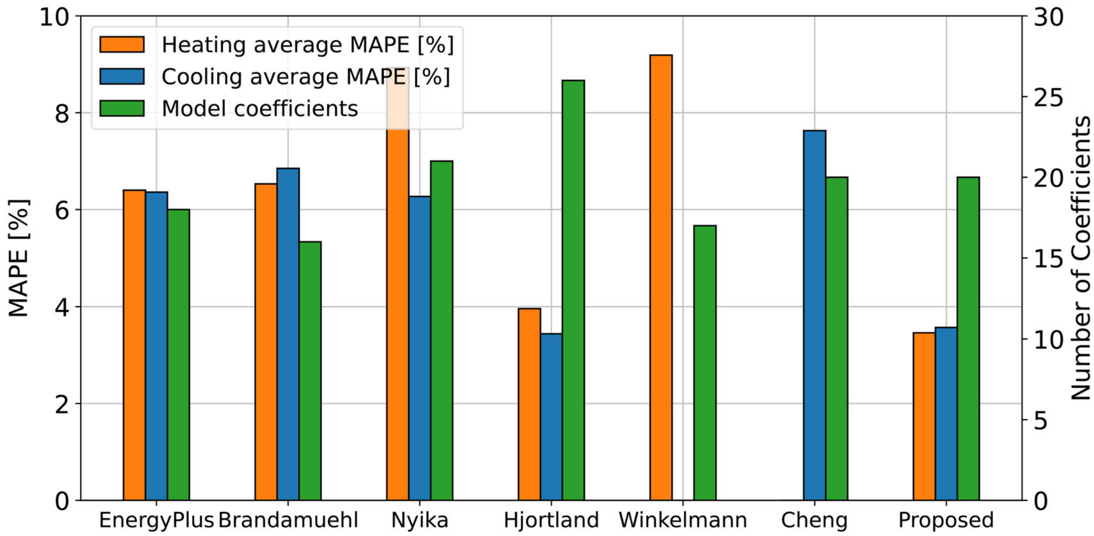

Fig. 13. Averaged MAPE for cooling and heating operation in all units.

Table 7. Heating and cooling scores (S) of the models (lower is better).

Model	Heating score	Cooling score
EnergyPlus DX	7.56	7.52
Brandamuehl	7.48	7.77
Nyika	10.14	7.74
Hjortland	6.16	5.70
Winkelmann	9.97	N/A
Cheng	N/A	8.87
Proposed	5.11	5.21

effectiveness in practical applications for building energy modeling.

5.3. Model ranking

As per the ranking criterion described in Section 4.4, all the evaluated models are ranked based on their predictive performance along with the number of coefficients. The number of coefficients in the model shows the amount of training data required.

Figure 13 shows the average MAPE of evaluated models in both heating and cooling operations. This figure also shows the coefficients associated with each of the models. The EnergyPlus model exhibits similar predictive capabilities for both cooling (6.36%) and heating (6.40%) operations along with its 18 coefficients while Brandamuehl's model with 16 coefficients has slightly higher cooling (6.85%) and heating (6.53%) MAPEs than EnergyPlus's model. The third model evaluated and shown in the figure is Nyika's model with an average heating MAPE of 8.93% and cooling MAPE of 6.27% which is achieved using 21 model coefficients. The fourth model on the list is Hjortland's model having the highest number of model coefficients of 26 achieving a MAPE of 3.96% and 3.44% in heating and cooling operation respectively in all the units. The heating-only model, Winkelmann's model, has a MAPE of 9.19% with only 17 model coefficients. On the list, the cooling-only model is Cheng's model with a MAPE of 7.63% and 20 model coefficients. The last among these models is the currently proposed model performance. The model predicts the heating and cooling operation of different units with a MAPE of 3.46% and 3.57%. The current model has the same number of coefficients of (20) as several other models evaluated but the predictive capabilities are much improved.

Table 7 shows individual score of each model while Figure 14 shows the final rank received by each model as per Equation (9). The proposed model received rank 1 while Hjortland's model stands 2nd in both heating and cooling operations. Brandamuehl's model rank is 3rd in heating while in cooling it is 5th. Similarly, EnergyPlus's model has a better performance in cooling as evident from its 3rd position as opposed to its 4th heating position. The heating and cooling-only models, Winkelmann and Cheng models, are in 5th and 6th position. Whereas, Nyika's model has 4th position in cooling and 6th in heating.

A comparison of the proposed model with the Hjortland model shows that both models achieve similar MAPE values. However, there are significant differences in their construction. The Hjortland model utilizes higher-order polynomials, which can closely fit the training data but are more prone to overfitting. This overfitting can lead to poorer performance on new, unseen data. Additionally, higher-order polynomials require a larger dataset to accurately train the model, which may not always be available. In contrast, the proposed model employs lower-order polynomials, which are less susceptible to overfitting and perform more reliably with smaller datasets. This makes our model not only more robust but also easier to implement in practical applications where extensive data collection may not be feasible. Consequently, we recommend our model over the Hjortland model due to its balance between simplicity, accuracy, and practical usability.

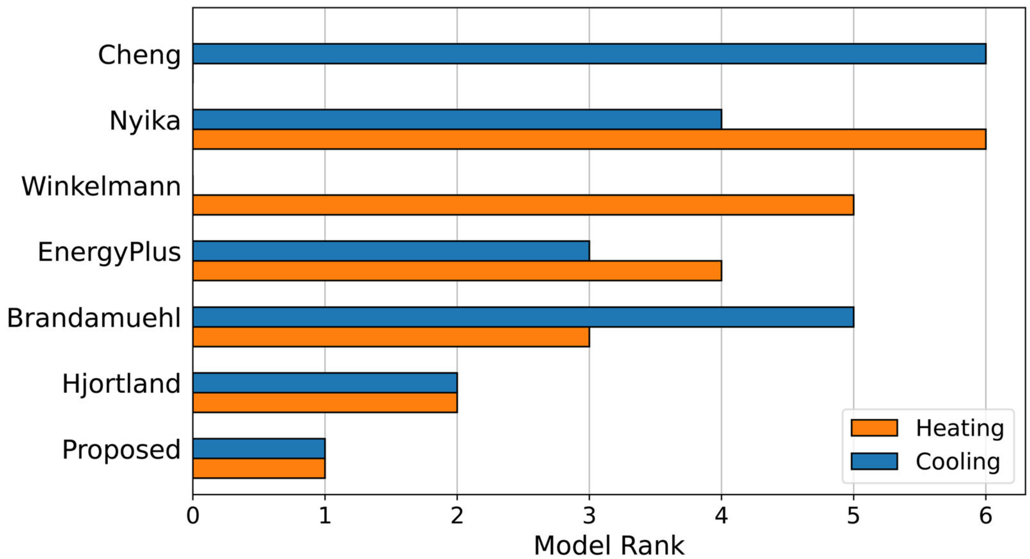

Fig. 14. Final ranks of models for AC and HP (lower is better).

6. Concluding remarks

This work evaluated 6 empirical models from the literature for different operating conditions to identify the best-performing model to capture AC and HP operation of unitary equipment. Additionally, a new model is proposed from the lessons learned during the legacy model evaluation phase. The models are trained and tested on high-fidelity experimental data obtained from different types of unitary equipment. They included a fixed-speed commercial AC unit, 3 different split-type 2-speed residential heat pumps, and two variable-speed heat pump systems. MAPE is used as an error metric to quantify the predictive capabilities of all the models at various operating conditions. All the models are ranked based on their predictive abilities as determined by the two performance metrics selected. From the results, some conclusions drawn are as follows:

- All models demonstrated comparable performance when applied to Unit-1, a fixed-speed commercial unit. Similarly, their performance remained consistent for Unit-4, where only high-speed compressor data was available. This suggests that legacy models excel in predicting full-load data. Notably, the models' predictive capabilities are not as strong in scenarios that involve a mix of compressor, fan, and temperature conditions.
- EnergyPlus and Brandamuehl's models exhibit competitive performance. EnergyPlus quadratic formulation for the supply air correction factor improved its accuracy prediction compared to the original Brandamuehl model.
- The heating-only model, Winkelmann's model, and the cooling-only model, Cheng's model, demonstrate less accuracy relative to other models.
- The proposed model with its 20 coefficients demonstrates superior predictive accuracy compared to existing models, achieving MAPEs consistently under 5% for both cooling and heating operations across various units. Hjortland's model follows closely in performance but with a higher number of coefficients (26).
- The comparative results highlight the usefulness of the proposed model, which uses fewer training data points while achieving improved predictive performance across both cooling and heating operations, making it a potential improvement in unitary equipment performance prediction.

Overall, the proposed model has better predictive capabilities both in the cooling and heating modes of heat pumps operated at part-load conditions. The model is trained on 17 data points and tested on the remaining data points with 10 model coefficients for predicting capacity and COP, each. To have a fair comparison with the legacy model the training data points are kept the same. The proposed formulation, although, has the following limitations as well,

- The model requires the performance of the unit at the rated conditions. Without such information, the model would not be useful.
- Caution should be exercised when using the model coefficients beyond the training data range, as the model operates as a black-box formulation. Additionally, the selection of the training dataset significantly impacts the model's predictive capabilities, requiring careful consideration.

In conclusion, our study synthesized and consolidated current research on state-of-the-art approaches, unveiling the potential for developing a new empirical model. By doing so, our findings serve as a catalyst, paving the way for significant contributions to the digital transformation of the built and energy industries. These advancements hold promising prospects for enhanced efficiency, optimization, and innovation in these sectors, ultimately driving sustainable and impactful developments.

Nomenclature

Variables

BF = Bypass factor, (-)
h = enthalpy, (BTU/lb)
Ω = compressor speed, (RPM)
ω = humidity ratio, (-)
$\dot{Q}$ = capacity, (BTU/hr)
S = rank score, (-)
T = temperature, (°F)
$\dot{V}$ = supply air flow rate, (CFM)
$\dot{W}$ = power, (W)

Subscripts

adp = apparatus dew point
cap = capacity
cop = coefficient of performance
f = correction factor
ID = indoor
rat = rated
sup = supply
temp = temperature
tot = total

Acronyms

AC = air conditioner
ACHP = air conditioning and heat pump
AHRI = air conditioning, heating, and refrigeration institure
ANN = artificial neural network
ASHP = air source heat pump
ASHRAE = American society of heating, refrigeration, and air conditioning engineers
BEMs = building energy models
BF = bypass factor
BFA = bypass factor approach
COP = coefficient of performance
EES = engineering equation solver
F.S = full scale reading
GSHP = ground source heat pump
HP = heat pump
IDT = indoor dry bulb temperature
IWB = indoor wet bulb temperature
MAPE = mean absolute percentage error
NTU = number of transfer units
ODT = outdoor dry bulb temperature
RTD = resistance temperature detector
RTU = roof top unit
SHR = sensible heat ratio
VCC = vapor compression cycle

Constants

$a_0, a_1, \ldots, a_5$ = regression coefficients
b_0, b_1 = regression coefficients
C = number of model coefficients
$c_0, c_1, \ldots, c_5$ = regression coefficients
d_0, d_1 = regression coefficients
k_0, k_1 = regression coefficients
l_0, l_1 = regression coefficients
ρ = coefficients weightage
σ = accuracy weightage

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work has been funded by the Oklahoma Center for the Advancement of Science and Technology (OCAST) with grant number AR21-042.

Notes on contributors

Shahzad Yousaf, MSc, Student Member ASHRAE, is a Graduate Research Assistant.

Craig R. Bradshaw, PhD, Full Member ASHRAE, is an Associate Professor.

Rushikesh Kamalapurkar, PhD, is an Associate Professor.

Omer San, PhD, is an Associate Professor.

References

Afjei, T., and D. Wittwer. 1995. Yearly Utilization Model YUM WP/Holz. *Benutzerhandbuch mit Beispielen. Ein Simulationsprogramm zur Berechnung des Jahresenergieverbrauchs und Nutzungsgrads bivalenter Wärmepumpen-Heizsysteme mit Holzkessel*. INFEL und BEW.

Ahmadisedigh, H., and L. Gosselin. 2022. Combined heating and cooling networks with part-load efficiency curves: Optimization based on energy hub concept. *Applied Energy* 307:118245. 10.1016/j.apenergy.2021.118245

AHRI. 2008. Performance rating of unitary air-conditioning & air-source heat pump equipment. *AHRI Standard* 210:240.

AHRI-540 2020. *Standard for performance rating of positive displacement refrigerant compressors and compressor units*. Arlington, VA: Air Conditioning, Heating, and Refrigeration Institute.

ASHRAE. 2009. *ANSI/ASHRAE standard 37-2009: Methods of testing for rating electrically driven unitary air-conditioning and heat pump equipment*.

Bell, I. 2012. ACHP. Available at: https://achp.sourceforge.net/_static/ACHP.pdf.

Belman-Flores, J. M., S. E. Ledesma, M. G. Garcia, J. Ruiz, and J. L. Rodríguez-Muñoz. 2013. Analysis of a variable speed vapor compression system using artificial neural networks. *Expert Systems with Applications* 40 (11):4362–9. 10.1016/j.eswa.2013.01.015

Brandemuehl, M. J., S. Gabel, and I. Andresen. 1993. *HVAC 2 Toolkit: Algorithms and subroutines for secondary HVAC system energy calculations*. ASHRAE.

Cetin, K. S., M. H. Fathollahzadeh, N. Kunwar, H. Do, and P. C. Tabares-Velasco. 2019. Development and validation of an HVAC on/off controller in EnergyPlus for energy simulation of residential and small commercial buildings. *Energy and Buildings* 183:467–83. 10.1016/j.enbuild.2018.11.005

Cheng, L., J. E. Braun, and W. T. Horton. 2021. A methodology for mapping the performance of variable-speed residential cooling equipment using load-based testing. *International Journal of Refrigeration* 132:133–44. 10.1016/j.ijrefrig.2021.09.001

Chesser, M., P. Lyons, P. O'Reilly, and P. Carroll. 2021. Air source heat pump in-situ performance. *Energy and Buildings* 251:111365. 10.1016/j.enbuild.2021.111365

Domanski, P., and D. Didion. 1983, May. Computer modeling of the vapor compression cycle with constant flow area expansion device. *Final Report, National Bureau of Standards*, Washington, DC.

Doty, S. 2010. Part-load HVAC efficiency. *Energy Engineering* 107 (3):6–28. 10.1080/01998591009709874

EnergyPlus. 2022. EnergyPlus. Available at: https://energyplus.net/assets/nrel_custom/pdfs/pdfs_v22.1.0/EngineeringReference.pdf.

Fan, Y., and X. Xia. 2018. Building retrofit optimization models using notch test data considering energy performance certificate compliance. *Applied Energy* 228:2140–52. 10.1016/j.apenergy.2018.07.043

Feng, K., W. Lu, and Y. Wang. 2019. Assessing environmental performance in early building design stage: An integrated parametric design and machine learning method. *Sustainable Cities and Society* 50:101596. 10.1016/j.scs.2019.101596

Filliard, B., A. Guiavarch, and B. Peuportier. 2009. Performance evaluation of an air-to-air heat pump coupled with temperate air-sources integrated into a dwelling. In *Proceedings of Building Simulation 2009, 11th International Building Performance Simulation Association Conference, Glasgow, UK*, July 2008, 2266–2273. International Building Performance Simulation Association.

Gholamibozanjani, G., J. Tarragona, A. d Gracia, C. Fernández, L. F. Cabeza, and M. M. Farid. 2018. Model predictive control strategy applied to different types of building for space heating. *Applied Energy* 231:959–71. 10.1016/j.apenergy.2018.09.181

Gupta, R., and R. Irving. 2013. Development and application of a domestic heat pump model for estimating CO2 emissions reductions from domestic space heating, hot water, and potential cooling demand in the future. *Energy and Buildings* 60:60–74. 10.1016/j.enbuild.2012.12.037

Herbas, T. B., E. C. Berlinck, C. A. T. Uriu, R. P. Marques, and J. A. R. Parise. 1993. Steady-state simulation of vapor-compression heat pumps. *International Journal of Energy Research* 17 (9):801–16. 10.1002/er.4440170903

Hjortland, A. L. 2018. Automated fault detection, diagnostics, impact evaluation, and service decision-making for direct expansion air conditioners. PhD thesis, Purdue University.

Joe, J., P. Im, and J. Dong. 2020. Empirical modeling of direct expansion (Dx) cooling system for multiple research use cases. *Sustainability* 12 (20):8738. 10.3390/su12208738

Jung, D. E., S. Kim, S. Han, S. Yoo, H. Jeong, K. H. Lee, and J. Kim. 2023. Appropriate level of development of in-situ building information modeling for existing building energy modeling implementation. *Journal of Building Engineering* 69:106233. 10.1016/j.jobe.2023.106233

Kim, M., S. H. Yoon, W. V. Payne, and P. A. Domanski. 2010. Development of the reference model for a residential heat pump system for cooling mode fault detection and diagnosis. *Journal of Mechanical Science and Technology* 24 (7):1481–9. 10.1007/s12206-010-0408-2

Kizilkan, Ö. 2011. Thermodynamic analysis of variable speed refrigeration system using artificial neural networks. *Expert Systems with Applications* 38 (9):11686–92. 10.1016/j.eswa.2011.03.052

Klien, S. A., and F. L. Alvardo. 2000. EES-engineering equation solver. Available at: https://www.fchartsoftware.com/ees/.

Ledesma, S., and J. M. Belman-Flores. 2014. Application of artificial neural networks for generation of energetic maps of a variable speed compression system working with R1234yf. *Applied Thermal Engineering* 69 (1-2):105–12. 10.1016/j.applthermaleng.2014.04.050

Lee, D., R. Ooka, S. Ikeda, and W. Choi. 2019. Artificial neural network prediction models of stratified thermal energy storage system and borehole heat exchanger for model predictive control. *Science and Technology for the Built Environment* 25 (5):534–48. 10.1080/23744731.2018.1557464

Lim, H., S.-Y. Cheon, and J.-W. Jeong. 2018. Empirical analysis for the heat exchange effective-ness of a thermoelectric liquid cooling and heating unit. *Energies* 11 (3):580. 10.3390/en11030580

Marchante-Avellaneda, J., J. M. Corberán, E. Navarro-Peris, and B. Llopis. 2019. Characterization of the performance for water-to-water and air-to-water heat pumps: Development of simple and compact polynomial correlations. In *Congreso Nacional de Ingeniería Termodinámica (CNIT), Albacete, Spain*, June 12–14, 934–945. Universidad de Castilla-La Mancha.

Morrison, G. L. 1994. Simulation of packaged solar heat-pump water heaters. *Solar Energy* 53 (3):249–57. 10.1016/0038-092X(94)90631-9

Nyika, S., and S. O. Holloway, et al. 2014. Generalized performance maps for variable-speed, ducted, residential heat pumps. *ASHRAE Transactions* 120:80.

O'Neill, Z., and C. O'Neill. 2016. Development of a probabilistic graphical model for predicting building energy performance. *Applied Energy* 164:650–8. 10.1016/j.apenergy.2015.12.015

Park, D., F. Guo, and B. P. Rasmussen. 2021. A method of mapping heat exchanger as simple polynomials. *International Refrigeration and Air Conditioning Conference*, Paper 2235. Available at: https://docs.lib.purdue.edu/iracc/2235.

Peskova, M. 2023. Data-driven models for estimating heat pump power consumption. MSc Thesis. KTH, School of Industrial Engineering and Management.

Puttige, A. R., S. Andersson, R. Östin, and T. Olofsson. 2022. Modeling and optimization of hybrid ground source heat pump with district heating and cooling. *Energy and Buildings* 264: 112065. 10.1016/j.enbuild.2022.112065

Rossi, T. M. 1995. Detection, diagnosis, and evaluation of faults in vapor compression equipment. PhD thesis. Purdue University.

Safa, A. A. 2012. Performance analysis of a two-stage variable capacity air source heat pump and a horizontal loop coupled ground source heat pump system. MSc Thesis. Ryerson University.

Shafi, I., J. Ahmad, S. I. Shah, and F. M. Kashif. 2006. Impact of varying neurons and hidden layers in neural network architecture for a time frequency application. 2006 IEEE International Multitopic Conference, 188–193. 10.1109/INMIC.2006.358160

Shao, S., W. Shi, X. Li, and H. Chen. 2004. Performance representation of variable-speed compressor for inverter air conditioners based on experimental data. *International Journal of Refrigeration* 27 (8):805–15. 10.1016/j.ijrefrig.2004.02.008

Shao, L.-L., L. Yang, and C.-L. Zhang. 2010. Comparison of heat pump performance using fin-and-tube and microchannel heat exchangers under frost conditions. *Applied Energy* 87 (4):1187–97. 10.1016/j.apenergy.2009.08.021

Shen, P., W. Braham, and Y. Yi. 2018. Development of a lightweight building simulation tool using simplified zone thermal coupling for fast parametric study. *Applied Energy* 223:188–214. 10.1016/j.apenergy.2018.04.039

Shen, B. 2006. Improvement and validation of unitary air conditioner and heat pump simulation models at off-design conditions. PhD thesis. Golden, CO: Purdue University.

Shoukas, G., E. Bonnema, G. Paranjothi, R. Faramarzi, and L. Klun. 2022. Performance assessment of high-efficiency variable-speed air-source heat pump in cold climate applications. National Renewable Energy Lab. (NREL), Golden, CO (United States).

Taylor, B. N., and C. E. Kuyatt. 1994. *Guidelines for evaluating and expressing the uncertainty of NIST measurement results*. Vol. 1297. Gaithersburg, MD: US Department of Commerce, Technology Administration, National Institute of Standards and Technology.

Tello-Oquendo, F. M., E. Navarro-Peris, F. Barceló-Ruescas, and J. Gonzálvez-Maciá. 2019. Semi-empirical model of scroll compressors and its extension to describe vapor-injection compressors. Model description and experimental validation. *International Journal of Refrigeration* 106:308–26. 10.1016/j.ijrefrig.2019.06.031

U.S. Energy Information Administration 2015. Use of energy in homes. Available at: https://www.eia.gov/energyexplained/use-of-energy/homes.php.

Winkelmann, F. C., B. E. Birdsall, W. F. Buhl, K. L. Ellington, E. Erdem, J. J. Hirsch, and S. Gates. 1993. *DOE-2 supplement, Version 2.1E*. Berkeley, CA: Lawrence Berkeley National Laboratory.

Yang, S., H. O. Gao, and F. You. 2023. Building electrification and carbon emissions: Integrated energy management considering the dynamics of the electricity mix and pricing. *Advances in Applied Energy* 10:100141. 10.1016/j.adapen.2023.100141

Ye, K. K., G. Demirezen, A. S. Fung, and E. Janssen. 2020. The use of artificial neural networks (ANN) in the prediction of energy consumption of air-source heat pump in retrofit residential housing. *IOP Conference Series: Earth and Environmental Science* 463 (1):012165. 10.1088/1755-1315/463/1/012165

Yousaf, S., C. Bradshaw, R. Kamalapurkar, and O. San. 2022. Physics-informed machine learning-based reduced-order model of unitary equipment. In *19th International Refrigeration and Air Conditioning Conference*, West Lafayette, IN, July 10–14, Indiana: Purdue University Press.

Yousaf, S., C. R. Bradshaw, R. Kamalapurkar, and O. San. 2023. Investigating critical model input features for unitary air conditioning equipment. *Energy and Buildings* 284:112823. 10.1016/j.enbuild.2023.112823

Yousaf, S., C. R. Bradshaw, R. Kamalapurkar, and O. San. 2024. A gray-box model for unitary air conditioners developed with symbolic regression. *International Journal of Refrigeration* 168: 696–707. 10.1016/j.ijrefrig.2024.10.008

Yousaf, S., I. Shafi, and J. Ahmad. 2018. Calculation of pressure loss coefficients in combining flows of a solar collector using artificial neural networks. *International Journal of Advanced Computer Science and Applications* 9 (9):555–559. 10.14569/IJACSA.2018.090969

Zhang, R., and T. Hong. 2017. Modeling of HVAC operational faults in building performance simulation. *Applied Energy* 202:178–88. 10.1016/j.apenergy.2017.05.153

Appendix A: Proposed model coefficients

This section provides coefficients for the proposed model obtained through the training methodology outlined in 4.1 for both heating and cooling mode of operation. These can be used to obtain the correction factors for different operating conditions and prediction can be made. The model coefficients are given in Table A1.

Table A1. Proposed model coefficients obtained through linear regression analysis.

Unit #	Operation	Coefficient	0	1	2	3	4	5	6	7	8	9
Unit 1	Cooling	a	2661.087	−82.8078	0.93714	144.0744	−1.38306	2.25628	−0.00534	0.013223	0.19014	−0.15652
		b	920.7129	4.107646	0.504585	13.77742	−0.22221	−0.06961	0.349565	−0.70387	0.357803	−0.35319
Unit 2	Cooling	a	6518.087	7.331101	−0.27682	−179.343	−2.45865	11.27099	−0.00016	−5.3E-05	−0.1527	−0.67173
		b	−0.75795	−0.01755	−0.00054	−0.05752	0.000506	0.000852	−0.39395	0.455121	−0.30928	−1.1523
	Heating	a	−263.234	−7.79162	−0.06494	−0.97037	0.044722	0.058597	−0.00531	−0.00113	−0.05678	0.537028
		b	1510.174	46.3063	0.20337	18.82242	−0.82628	−1.19005	−0.25683	−0.06706	−0.00202	0.000136
Unit 3	Cooling	a	10147.55	−8.25274	−0.55408	−456.519	−0.7408	21.1538	0.130237	0.095142	0.000152	0.000401
		b	1456.016	18.73174	0.168179	−73.4304	−0.12929	2.135672	0.000162	0.000287	−0.00014	−0.4947
	Heating	a	2070.011	49.24579	0.209451	−9.64075	−0.07593	0.244741	0.211657	0.027295	−0.00019	0.002594
		b	1771.585	35.35882	−0.10351	−5.31041	−0.46169	−0.40993	0.166878	0.042983	0.003147	−0.00042
Unit 4	Cooling	a	619.0808	−2.38064	−0.06275	−12.4925	0.158198	0.441122	0.859535	0.254281	0.069287	−0.06753
		b	−291.851	−0.53835	−0.06821	6.966938	−0.03856	−0.18306	−0.00146	0.001913	−0.00116	2.006225
	Heating	a	29.90054	3.097655	0.05966	9.254803	0.018691	−0.21627	−0.00288	−4.8E−05	−1.03631	−1.05985
		b	−1.93915	−0.03454	−0.00015	0.015333	0.000311	0.000282	−0.12201	−0.02142	2.010991	2.010991