CHEUNG, Howard; WANG, Shengwei. A comparison of the effect of empirical and physical modeling approaches to extrapolation capability of compressor models by uncertainty analysis: A case study with common semi-empirical compressor mass flow rate models. International Journal of Refrigeration, 2018, 86: 331-343.
This is the Pre-Published Version.
https://doi.org/10.1016/j.ijrefrig.2017.11.020

A comparison of the effect of empirical and physical modeling approaches to extrapolation capability of compressor models by uncertainty calculation: a case study with common semi-empirical compressor mass flow rate models

Howard Cheung, Shengwei Wang*
Department of Building Services Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong
* Corresponding author: beswwang@polyu.edu.hk

Abstract

Some semi-empirical compressor models are claimed to be more accurate at extrapolation conditions than their empirical counterparts which has a long history of industrial applications due to their uses of physical principles, but it is unknown how much improvement the principles can bring to the modeling of extrapolation scenarios quantitatively. This paper studies the effect of the number of empirical coefficients and physical principles on model accuracy and uncertainty by comparing the estimation of five regression models of compressor mass flow rates. The choice of model training data follows the industrial norm, and model accuracy and uncertainty are calculated. The quantitative results show that the use of neither empirical coefficients nor physical principles guarantees good accuracy and reliability. If a coefficient is redundant to explain the behavior of the phenomenon, regardless of its empirical or physical origin, it should be removed to reduce model inaccuracy in extrapolation scenarios.

Keywords

compressor; modeling; empirical models; extrapolation; uncertainty

Highlights

- Models with different number of empirical coefficients are compared
- Changes of model performance with the use of empirical coefficients is quantified
- Redundant coefficients reduce model reliability regardless of their physical origins
- Extrapolation effect is not only dependent on the number of physical rules involved

Nomenclature

c_p	Isobaric heat capacity [$Jkg^{-1}K^{-1}$]
c_v	Isochoric heat capacity [$Jkg^{-1}K^{-1}$]
e_{is}	Compression exponent
f	Rotational frequency [Hz]
h	Threshold in numerical methods [dimensionless]
$\boldsymbol{H}$	Hessian matrix [unit varies]
j_{lev}	Jacobian leverage [dimensionless]
$\boldsymbol{J}$	Jacobian matrix [unit varies]
$\dot{m}_{comp}$	Compressor mass flow rate [kgs^{-1}]
n	Number of data points [dimensionless]
P	Pressure [Pa]
q	Number of regression coefficients in a model [dimensionless]
s	Specific entropy [$Jkg^{-1}K^{-1}$]
$t_{n_{train}-q,0.95}$	Student's t value for a 95% confidence interval of estimates from equations with n training data points and q coefficients [dimensionless]
T	Temperature [K]
v	Specific volume [m^3kg^{-1}]
V_d	Displacement volume [m^3]

Greek

β	Regression coefficient [unit varies]
γ	Adjustment factor in Richardson Extrapolation [dimensionless]

Γ	Threshold multiplier in Richardson Extrapolation [dimensionless]
Δ	Uncertainty [unit varies]
η_{vol}	Volumetric efficiency [dimensionless]
θ	Arbitrary functions [unit varies]
λ	Order of accuracy of numerical methods [dimensionless]
ρ	Density [kgm^{-3}]
Ψ	Arbitrary variable [unit varies]

Accents

$\widehat{\ }$	estimated

Subscripts

dew	dewpoint
dis	compressor discharge
disc	discretization
EOS	equation of state
exp	expanded
input	input
it	iteration
I	Model I
II	Model II
III	Model III
IV	Model IV
model	model random error
num	numerical method
output	output
rat	rated
suc	compressor suction
train	training data

1. Introduction

Various types of compressor models have been used in the computational models of vapor compression systems and other applications involving prediction of compressor performance [1]–[4]. They should be designed to be accurate and reliable to give correct prediction consistently for reliable engineering designs. In practice, the models are mostly semi-empirical: the empirical part of the model is formulated by best-fitted curves without acknowledging any physics behind the compression mechanism, while some physical rules are used to explain the compression mechanism and to formulate the rest of the model. They contain unknown coefficients that must be predicted with experimental observations of the compressor operation before their applications. The prediction of these coefficients is usually conducted by regression with some performance data from compressor calorimeter experiments. These performance data are called training data.

Some researchers aimed at improving the accuracy and the speed of the models and recommended using more empirical coefficients in the compressor models. Rasmussen and Jakobsen [1] described different types of polynomial models for compressor modeling, including the 10-coefficient cubic polynomial for compressor mass flow rate and power consumption in ANSI/AHRI Standard 540-2004 [5] and quadratic polynomials to estimate compressor isentropic efficiency and volumetric efficiency. Shao et al. [6] used a quadratic polynomial of compressor rotational frequency to adjust the 10-coefficient polynomial to model the performance of variable-speed compressors. Shen [7] used an adjustment multiplier calculated from experimental data to tune any bias in the 10-coefficient polynomial in ANSI/AHRI Standard 540-2004 [5]. Yang et al. [8] used multiple empirical linear equations to form a neural network to model compressor volumetric efficiency.

Other researchers argued that the use of empirical coefficients in the compressor models reduced the accuracy of the models to predict compressor performance at some operating conditions. They claimed that the use of empirical coefficients might not obey physics and reduced the accuracy of the models to predict compressor performance under conditions that are different from that of the training data (i.e. extrapolation). They tried to improve the models' extrapolation capability by including more physical rules and create semi-empirical compressor models. For example, Jähnig et al. [9] introduced pressure drop at the compressor suction in its adiabatic

compression model to model the mass flow rate of reciprocating compressors. Kim and Bullard [10] used Newton's law of cooling to model compressor heat loss. Winandy et al. [11] used under- and over- compression mechanisms to adjust the power consumption estimated by adiabatic compression. Navarro et al. [12] considered various types of flow leakages in compressors as flows across valves to estimate compressor mass flow rate. Duprez et al. [13] suggested the use of compressor suction valve mechanism instead of the compression mechanism to model compressor mass flow rate. Aprora [14] and Zakula et al. [15] included mathematical formula representing re-expansion and back leakage losses in their compressor model. Cheung and Braun [16] considered air-side natural convection and refrigerant-side forced convection to model the heat loss of compressor. Negrão et al. [17] simplified the models of leakages and heat loss by modeling their effect to volumetric efficiency with a linear relationship with pressure ratio only.

To justify if physical rules improve the accuracy of the compressor models, some researchers moved on to test their models at extrapolation conditions. Both Jähnig et al. [9] and Li [18] validated their models model's ability to extrapolate by examining the accuracy of the model with data in addition to their training data. Aute et al. [19], [20] did similar research with multiple semi-empirical and empirical models and found that the uses of more empirical coefficients in compressor models and the current industrial practice to select training data points in the industry did not inhibit model accuracy at extrapolation significantly. Cheung et al. [21] examined how extrapolation reduced the accuracy of the estimation of compressor mass flow rate and power consumption of the 10-coefficient compressor model [5].

While it is well known that the reliability of predicted models can be calculated by uncertainty of the models [22]–[24], the literature on compressor modeling focused exclusively on improving the accuracy of the models, including the accuracy of the models at extrapolation. They did not study how the use of empirical coefficients or introduction of new physical principles in the models affected the uncertainty and the reliability of the prediction. For example, it is unknown if the use of equations of state (EoS) for refrigerant property calculation, which usually contains more empirical coefficients than the compressor models being studied, introduce extra uncertainties of EoS and reduce the reliability of the prediction. It is also unknown if the addition of an extra cubic term to a compressor model overfits the training data and hence reduce the

reliability of the prediction. While Aute et al. [19], [20] and Cheung et al. [21] calculated the uncertainty of the 10-coefficient compressor model [5] which models the compressor power consumption, they did not move forward to investigate the change of uncertainties due to a change in model structure and numerical stability. In general, the change of the extrapolation capability of a compressor model by introducing extra empirical coefficients or physical principles in its modeling process is not well understood.

This paper studies the change of compressor model accuracy and reliability with number of empirical coefficients and physical principles by calculating the uncertainties of various semi-empirical compressor models and comparing their estimates with experimental data. Compressor mass flow rate models having different model structure are selected for the comparison. Data of two compressors with a comprehensive experimental data across their operating ranges are used to train the models to conduct the comparison. Their model output uncertainties are calculated primarily based on the uncertainty calculation method in with adjustment for the use of nonlinear equations and numerical methods in the models [21]. The changes of the model output uncertainties and accuracy with the uses of empirical coefficients and physical principles in the models are studied. Recommendations on the number of empirical coefficients and physical principles in semi-empirical models to model compressors with reasonable model reliability are given based on the results.

2. General calculation method of model uncertainty

Cheung et al. [21] introduced a scheme to calculate the uncertainty of the estimated compressor power consumption of the AHRI 10-coefficient polynomial [5] to the true value of the compressor power consumption. They described how uncertainty of the polynomial can be calculated from 4 main sources: uncertainty due to inputs, uncertainty due to outputs, uncertainty due to training data and uncertainty due to model random error. They neglected uncertainty due to numerical methods and manufacturing and ageing due to the absence of numerical methods in the study. However, since estimation of regression coefficients of semi-empirical models often involves numerical methods, uncertainty due to numerical methods is needed in this study. Hence this study only neglects the uncertainty due to manufacturing and ageing. The overall uncertainty of the estimated compressor mass flow rate in this study is calculated from the sum of squares of the components as shown in Equation (1).

$$\Delta\hat{m}_{comp} = \sqrt{\begin{matrix}\Delta\hat{m}^2_{comp,input} + \Delta\hat{m}^2_{comp,output} + \Delta\hat{m}^2_{comp,train} \\ +\Delta\hat{m}^2_{comp,model} + \Delta\hat{m}^2_{comp,num}\end{matrix}} \quad \textbf{(1)}$$

Equation (1) shows that the uncertainties can be decoupled into 5 parts: the uncertainty due to inputs, the uncertainty due to outputs, the uncertainty due to training data, the uncertainty due to model random error and the uncertainty due to numerical error. They are presented in Equation (1) in standard uncertainty forms according to ASME Performance Test Codes 19.1-2003 [24]. When they are calculated based on the 95% confidence level as expanded uncertainties, they should be calculated with the Student's *t* value as shown in Equation (2).

$$\Delta\hat{m}_{comp,exp} = t_{n_{train}-q,0.95}\Delta\hat{m}_{comp} \quad \textbf{(2)}$$

The following subsections describe the definition of each component of the uncertainty.

2.1 Uncertainty due to inputs

Uncertainty due to inputs are uncertainty propagated from the input variables to the model. When the models predict their outputs, they acquire input variables that carry uncertainties such as measurement uncertainties from sensors. The uncertainty of the input variables form part of the uncertainty of the predicted variables called uncertainty due to inputs, and its calculation method is originated from Kline and McClintock [25].

2.2 Uncertainty due to outputs

Uncertainty due to outputs is a result of using measured values of model output variables instead of their true values to estimate the coefficients of the models. While the model should be used to estimate the true values of the output variable, it is trained by measured values and hence estimates the measured value of the output variables instead with a difference to the true value of the output variable. This difference is quantified by the uncertainty due to outputs, and their values depend on the magnitude of the model estimation and the uncertainty of the sensor that measures the model output variables used in the training data.

2.3 Uncertainty due to training data

Uncertainty due to training data is the propagation of uncertainties of the variables in the training data through the estimated coefficients to the model prediction results. It differs from the

uncertainty due to inputs that the uncertainty due to inputs propagates to the model prediction result directly without propagating through the estimated coefficients. It is calculated by multiplying the derivatives of the model prediction output with the regression coefficients with the uncertainties of the measurement in the training data. However, regression coefficients in semi-empirical models are usually estimated by implicit numerical solvers that cannot be differentiated analytically, and the uncertainty due to training data is calculated by estimating the derivatives by finite difference method [26].

2.4 Uncertainty due to model random error

Cheung et al. [21] described the uncertainty due to model random error as the uncertainty propagated from the uncertainty between the true values and estimated values of the regression coefficients. If the training data point includes all possible inputs to the model, the uncertainty due to model random error is the accuracy of the model. However, the number of training data points in this case will be infinite and the data set will become impossible to obtain. Hence one can only obtain a sample of the population as the training data to estimate the regression coefficients. The model random error hence consists of uncertainty due to the incomprehensiveness of the training data and the accuracy of the model at the training data points. This uncertainty component is necessary to quantify the effect of extrapolation on model uncertainty as shown in [21].

2.5 Uncertainty due to numerical method

The uncertainty due to numerical method is usually neglected in linear equations because linear equations do not involve any numerical methods. However, estimation of regression coefficients in semi-empirical models often involves implicit numerical methods that iterates to convergence. Hence the estimated regression coefficients carry an uncertainty due to iteration that can be calculated by the Eigenvalue method [27]. If the estimation process involves discretization such as finite difference method, the uncertainty due to discretization of the regression coefficients is also needed and is calculated by Richardson extrapolation [27]. Both the uncertainty due to iteration and the uncertainty due to discretization are propagated to the model prediction by the uncertainty propagation method in [25].

3. Compressor mass flow rate models

In this study, five different models representing different levels of applications of empirical coefficients and physical principles are selected. The selection is limited to models that only use compressor suction temperature, compressor suction pressure and compressor discharge pressure as their model inputs for a fair comparison. Since all of them are all more non-linear than the model evaluated in [21] to account for uncertainty due to superheat correction and nonlinear regression, the mathematical formula of the uncertainty calculation methods of these models are slightly modified from that in [21] and are shown in the Appendix for reference.

3.1 Model I: 10-coefficient polynomial with superheat adjustment [5], [28]

ANSI/AHRI Standard 540-2004 [5] describes a 10-coefficient cubic polynomial to model compressor mass flow rate as shown in Equation (3).

$$\begin{aligned}\hat{\dot{m}}_{comp,rat} = {} & \hat{\beta}_0 + \hat{\beta}_1 T_{dew,suc} + \hat{\beta}_2 T_{dew,dis} + \hat{\beta}_3 T_{dew,suc}^2 + \hat{\beta}_4 T_{dew,suc} T_{dew,dis} \\ & + \hat{\beta}_5 T_{dew,dis}^2 + \hat{\beta}_6 T_{dew,suc}^3 + \hat{\beta}_7 T_{dew,suc}^2 T_{dew,dis} \\ & + \hat{\beta}_8 T_{dew,suc} T_{dew,dis}^2 + \hat{\beta}_9 T_{dew,dis}^3\end{aligned} \quad \textbf{(3)}$$

Regression coefficients in Equation (3) can be estimated by linear regression with compressor calorimeter data at a rated compression suction superheat (difference between the compressor suction temperature and compressor suction dewpoint), and Equation (3) can be used to estimate a rated compressor mass flow rate at the given compressor suction and discharge dewpoint temperature. To estimate the mass flow rate at superheat values other than the rated value, Dabiri and Rice [28] obtained an empirical relation Equation (4) to adjust the rated compressor mass flow rate.

$$\hat{\dot{m}}_{comp,I} = \hat{\dot{m}}_{comp,rat}\left(1 + \hat{\beta}_{10}\theta_1(T_{suc}, P_{suc})\right) \quad \textbf{(4)}$$

$$\theta_1(T_{suc}, P_{suc}) = \frac{\rho_{suc}(T_{suc}, P_{suc})}{\rho_{suc,rat}(P_{suc})} - 1 \quad \textbf{(5)}$$

$\hat{\beta}_{10}$ in Equation (4) was given as 0.75 in [28] though they reported that it ranged between 0.62 and 0.75. To facilitate the uncertainty propagation of the uncertainty of $\hat{\beta}_{10}$ to the uncertainty of the estimated mass flow rate, its value is taken as 0.685 with an uncertainty ±0.065 in this paper.

3.2 Model II: Model with volumetric efficiency estimated by a polynomial [1]

One of the simplest and most common methods to model compressor mass flow rate is to consider the actual mass flow rate as a fraction of mass flow rate of an ideal compression process as Equation (6).

$$\hat{\dot{m}}_{comp} = \hat{\eta}_{vol}\rho_{suc}fV_d \tag{6}$$

Rasmussen and Jakobsen [1] described that the volumetric efficiency in Equation (6) can be modeled by a quadratic polynomial of discharge and suction dew point temperature. By considering the compressor displacement volume and rotational speed to be constant for single-speed compressors, a model of mass flow rate can be written as Equation (7).

$$\begin{aligned}\hat{\dot{m}}_{comp,II} &= \rho_{suc}(T_{suc}, P_{suc})\theta_2\left(T_{dew,suc}, T_{dew,dis}\right) \\ &= \rho_{suc}(T_{suc}, P_{suc})(\hat{\beta}_0 + \hat{\beta}_1 T_{dew,suc} + \hat{\beta}_2 T_{dew,dis} + \hat{\beta}_3 T_{dew,suc}^2 \\ &+ \hat{\beta}_4 T_{dew,suc}T_{dew,dis} + \hat{\beta}_5 T_{dew,dis}^2)\end{aligned} \tag{7}$$

Regression coefficients in Equation (7) can be estimated by linear regression with the ratio of mass flow rate to suction density as the dependent variable of the regression equation.

Model II is a semi-empirical model and is less empirical than Model I due to its lower-order polynomial structure and the use of the physical principle Equation (6) to account for the effect of superheat.

3.3 Model III: Reciprocating compressor mass flow rate model based on adiabatic compression [9]

Jähnig et al. [9] modeled the volumetric efficiency in Equation (6) by adiabatic compression and the pressure drop at the compression suction. The resultant equation is Equation (8).

$$\begin{aligned}&\hat{\dot{m}}_{comp,III} \\ &= \rho_{suc}(T_{suc}, P_{suc})\hat{\beta}_0\left(1 - \hat{\beta}_1\left[\left(\frac{P_{dis}}{P_{suc}(1-\hat{\beta}_2)}\right)^{\left(c_v(T_{suc},P_{suc})/c_p(T_{suc},P_{suc})\right)} - 1\right]\right)\end{aligned} \tag{8}$$

Constrained optimization is used to estimate the regression coefficients in Equation (8) which minimizes the objective function Equation (9).

$$O = \frac{1}{n_{train}\dot{m}_{comp,rat}} \sum_{i=1}^{n_{train}} \left(\hat{\dot{m}}_{comp,i} - \dot{m}_{comp,train,i}\right)^2 \quad \textbf{(9)}$$

where the rated mass flow rate in Equation (9) is the average of the measured mass flow rates in the training data.

The objective function Equation (9) is a dimensionless average of the squared differences between the estimated and measured mass flow rate, and a set of regression coefficients that minimizes Equation (9) is presumed to be accurate to estimate the compressor mass flow rate by Equation (8). Sequential Least Square (SLSQP) method [29], which is an implicit and iterative constrained optimization method, is used to find the regression coefficients that minimizes Equation (9) with initial guesses and constraints in Table 1.

Table 1 Initial guesses and constraints to the regression coefficients in model III

Coefficients	Initial guess	Constraints
$\hat{\beta}_0$	Ratio of measured rated mass flow rate to average compressor suction density in the training data	Greater than 0
$\hat{\beta}_1$	0	Greater than 0
$\hat{\beta}_2$	0	Between 0 and 0.1

The optimization process is set to terminate when Equation (9) change less than 10^{-12} kg^2s^{-2} between iterations. The estimation process can be written as a function in Equation (10).

$$\vec{\hat{\beta}} = \theta_3\left(\vec{T}_{suc,train}, \vec{P}_{suc,train}, \vec{P}_{dis,train}, h\right) \quad \textbf{(10)}$$

Equation (10) shows an extra input h to the regression coefficient estimation process in addition to the training data. This is a threshold used by the SLSQP method to estimate the derivatives of Equation (9) with respect to the regression coefficients by finite difference method [26]. It is determined by conducting a convergence analysis with Equation (10).

Model III is less empirical than Model II because it includes adiabatic compression and suction pressure drop in its modeling process.

3.5 Model IV: Mass flow rate model based on adiabatic compression and back leakage losses

Arora [14], Zakula et al. [15] and Cheung and Braun [16] noted that a back leakage loss term could be added to compressor mass flow rate model. This idea was used to add a new physical consideration to Model III to create a new compressor mass flow rate model, and the resultant model is shown by Equation (11).

$$\hat{\dot{m}}_{comp,IV} = \rho_{suc}(T_{suc}, P_{suc})\hat{\beta}_0 \left(1 - \hat{\beta}_1 \left[\left(\frac{P_{dis}}{P_{suc}(1-\hat{\beta}_2)} \right)^{(c_v(T_{suc},P_{suc})/c_p(T_{suc},P_{suc}))} - 1 \right] - \hat{\beta}_3 \left(\frac{P_{dis}}{P_{suc}} \right) \right) \quad \textbf{(11)}$$

The method to estimate the regression coefficients in Model IV is the same as that of Model III with the exception that $\hat{\beta}_3$ in Equation (11) is restricted to be positive and its initial guess is zero.

Model IV is less empirical than Model III because of its back leakage loss model in addition to all physical principles considered in Model III.

3.6 Model V: Compressor mass flow rate model based on isentropic compression

Zakula et al. [15] proposed another variant of Model III by assuming isentropic compression instead of adiabatic compression to model compressor mass flow rate. To compare Model III and V with the same number of physical principles, the model in Zakula et al. [15] was modified, and the resultant model is shown by Equations (12), (13) and (14).

$$\hat{\dot{m}}_{comp,V} = \rho_{suc} \left[\hat{\beta}_0 \left(1 - \hat{\beta}_1 \left[\left(\frac{P_{dis}}{P_{suc}(1-\hat{\beta}_2)} \right)^{(1/e_{is})} - 1 \right] \right) \right] \quad \textbf{(12)}$$

$$e_{is} = \ln\left(\frac{P_{dis}}{P_{suc}}\right) \Big/ \ln\left(\frac{\rho_{dis,s}}{\rho_{suc}}\right) \quad \textbf{(13)}$$

$$\rho_{suc} = \rho(P_{dis}, s = s(T_{suc}, P_{suc})) \quad \textbf{(14)}$$

The method to estimate the regression coefficients in model IV is the same as that of Model III.

3.7 Summary

To summarize the difference between the models, Table 2 is created.

Table 2 Summary of the characteristics of the compressor mass flow rate models

Model	Number of regression coefficients	Physical principles involved
Model I	11	a. Tuning of mass flow rate by compressor suction density
Model II	6	a. Definition of compressor volumetric efficiency
Model III	3	a. Definition of compressor volumetric efficiency b. Adiabatic compression c. Compressor suction pressure loss
Model IV	4	a. Definition of compressor volumetric efficiency b. Adiabatic compression c. Compressor suction pressure loss d. Compressor back leakage loss
Model V	3	a. Definition of compressor volumetric efficiency b. Isentropic compression c. Compressor suction pressure loss

In Table 2, it can be seen that model I is the most empirical model because it contains 11 regression coefficients to form a cubic polynomial. It is only supported by one physical principle that governs the change of mass flow rate due to a change of compressor suction superheat. Model II is the second most empirical model with a more comprehensive physical principle and fewer regression coefficients than Model I. Models III and V are less empirical than Models I and II with more physical principles and fewer regression coefficients, and most of their regression coefficients are related to the physical principles. Model IV is the least empirical model because it involves more physical principles than other models that it contains more regression coefficients than Models III and V to account for the mechanism in the principles.

4. Experimental data

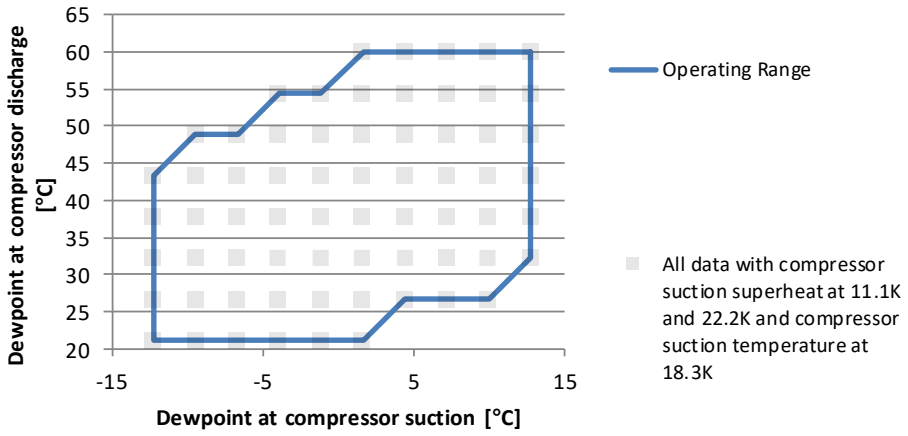
To demonstrate how the reliability of the model outputs changes with the number of empirical coefficients and physical principles in the compressor mass flow rate models, experimental data of compressor performance that are comprehensively tested is needed for verification. However, since it is unclear which data from the industrial catalog are being created by calorimeter tests instead of Model I [19], only data that are documented with calorimeter testing procedure can be used to avoid giving unfair advantages to Model I. In this study, the calorimeter data of two compressors were used [30], [31]. Their specification is shown in

Table 3.

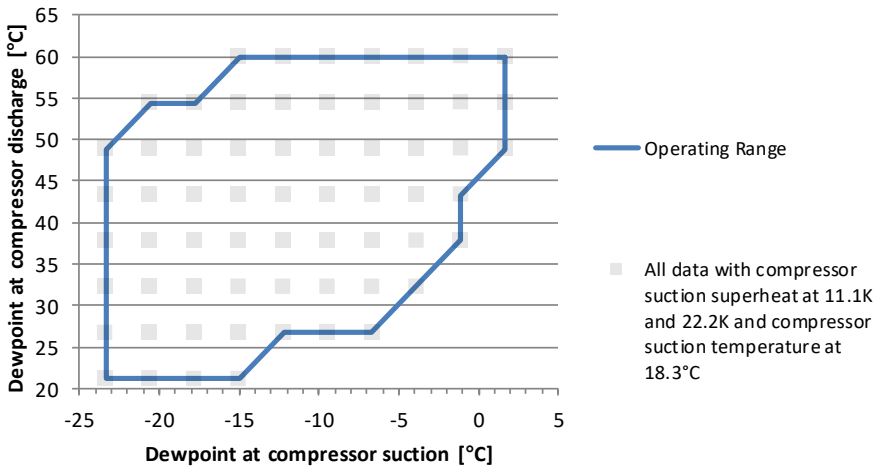
Table 3 Specification of compressors in experimental setups

	Compressor 1	**Compressor 2**
Type	Hermetic scroll	Hermetic scroll
Displacement volume	20.3 cm^3 rev^{-1}	51.0 cm^3 rev^{-1}
Rated power consumption	2.17 kW	3.32 kW
Rated mass flow rate	0.0396 kgs^{-1}	0.0624 kgs^{-1}
Refrigerant	R410A	R404A

The compressors were tested according to ANSI/ASHRAE Standard 23.1-2010 [32] under various compressor suction and discharge dewpoint with compressor suction superheat at 11.1K and 22.2K and compressor suction temperature at 18.3°C as shown in Figure 1.



(a)



(b)

Figure 1 Illustration of compressor calorimeter operating conditions of (a) Compressor 1 and (b) Compressor 2

At each operating condition, its suction and discharge temperature, suction and discharge pressure, refrigerant mass flow rate and power consumption are measured. The apparatus used to conduct the measurement is tabulated in Table 4.

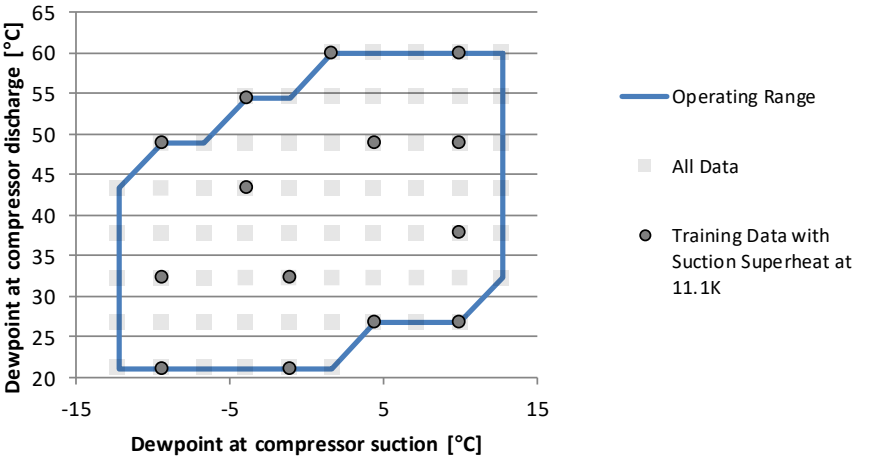
Table 4 Measurement apparatus in the compressor calorimeter test

Sensor	Measurand	Uncertainty
Resistance temperature detector	Compressor suction and discharge temperature	±0.2K
Coriolis mass flowmeter	Refrigerant mass flow rate	±0.1%
Pressure transmitter with full scale at 5,171kPa	Compressor discharge pressure	±0.25% full scale
Pressure transmitter with full scale at 1,378kPa	Compressor suction pressure	±0.25% full scale

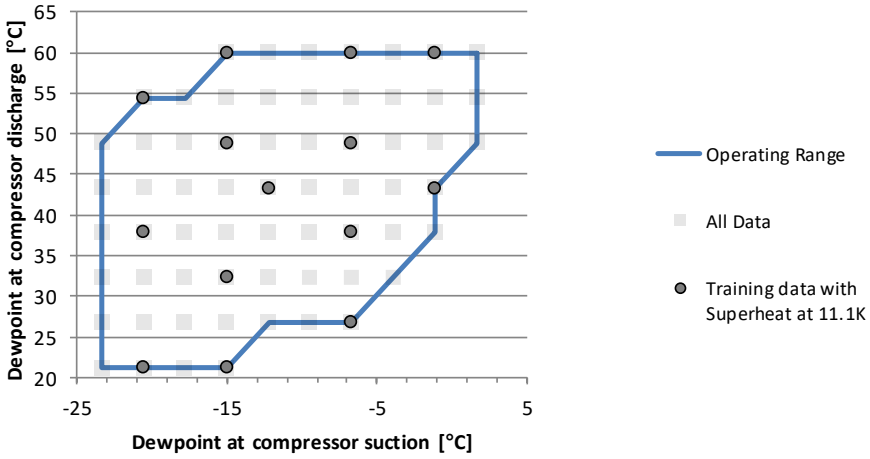
Other details of the tests can be seen in [30], [31].

5. Evaluation methodology

To compare the reliability between models, the models are first trained by training data selected according to the rules of thumb of the compressor industry. The operating conditions of the training data of the two compressors are given in Figure 2.



(a)



(b)

Figure 2 Testing conditions of training data in demonstration scenarios of (a) Compressor 1 and (b) Compressor 2

Figure 2 shows 14 training data points with constant compressor suction superheat at 11.1K scattered across the entire operating range of both compressors. Their compressor suction

superheat was maintained at constant 11.1K to satisfy the training data requirement of constant compressor suction superheat of model I, and only 14 data points were selected because this was the minimum number of data points used by compressor manufacturers to create Model I according to the survey in [19]. The selection of the operating data points in Figure 2 is arbitrary with reference to the results in [19] so as to imitate how the compressor manufacturers choose training data points for compressor models.

After selecting the training data, the regression coefficients of the compressor models are estimated by methods specified in Section 3. The models are then used to estimate the compressor mass flow rates and their uncertainties at all operating conditions in Figure 2. The uncertainties calculated are compared with the accuracy of the models. This examines the impact of the use of empirical coefficients and physical principles to the accuracy and reliability of the model prediction. Further details of the effect of extrapolation on the accuracy of the models can be found in the supplementary materials.

6. Results and discussion

To examine the change of uncertainties of various models, the expanded uncertainty components and the overall uncertainty of various models at all operating conditions are plotted in Figure 3.

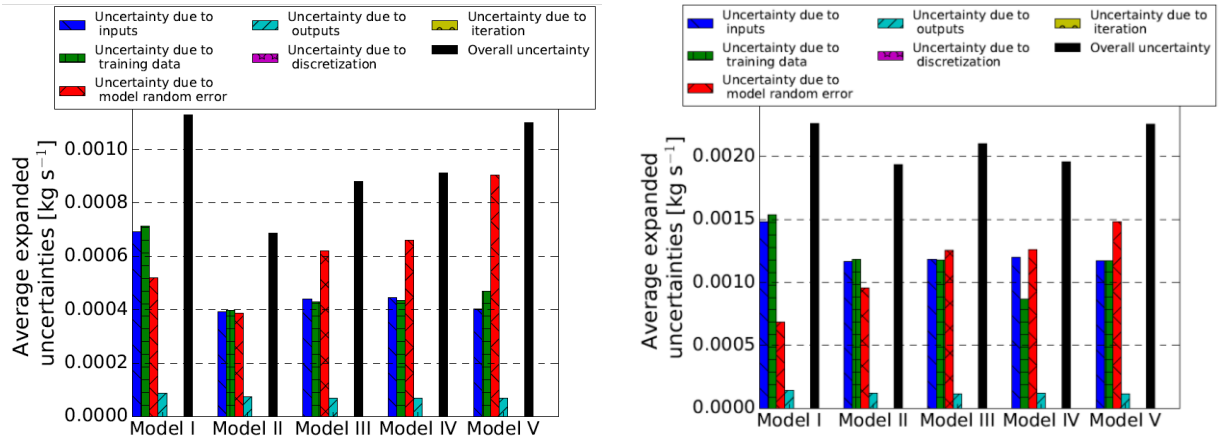


Figure 3 Average expanded uncertainties and their components under all operating conditions of (a) Compressor 1 and (b) Compressor 2

Figure 3 shows that Model I yields the highest uncertainty for both compressors. Although Model I is the most empirical model and should be mostly subjected to extrapolation with a high uncertainty due to model random error among all models, its uncertainty due to model random error is low comparing to other models. This is caused by its high accuracy at training data points. However, it is less reliable than other models with higher overall uncertainty due to its complex polynomial structure and hence high uncertainty due to inputs and training data. In contrast, Model II, which is a bi-quadratic equation, yields a much smaller uncertainty due to inputs and training data for both compressors. Hence the use of too many regression coefficients in compressor models reduces the reliability of its estimation.

Similar phenomenon of model redundancy also appears in models based on physical principles as shown by the uncertainty of Model III and IV in Figure 3. The overall uncertainty of Model IV is higher than that of Model III for Compressor 1 but vice versa for Compressor 2. This can be explained by the model accuracy as quantified in Table 5.

Table 5 Coefficients of determination (R^2) of various models calculated based on data from all operating conditions

Type	Compressor 1	Compressor 2
Model I	0.9988	0.9976
Model II	0.9991	0.9991
Model III	0.9977	0.9983
Model IV	0.9977	0.9987
Model V	0.9982	0.9991

Coefficient of determination (R^2) in Table 5 is commonly used to assess how accurate the estimated mass flow rates are [33]. A smaller R^2 means less accurate prediction, and it is calculated by Equation (15).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(\dot{m}_{comp,i} - \hat{\dot{m}}_{comp,i}\right)^2}{\sum_{i=1}^{n}\left(\dot{m}_{comp,i} - \frac{1}{n}\sum_{i=1}^{n}\dot{m}_{comp,i}\right)^2} \quad \textbf{(15)}$$

Table 5 shows that Model IV is as accurate as Model III for Compressor 1 but is more accurate than Model III for Compressor 2. For Compressor 1, the additional physical rule about back leakage loss in Model IV cannot improve the accuracy of estimation for Compressor 1, and this redundancy increases the uncertainty of its estimation. For Compressor 2, the additional physical rule in Model IV improves the accuracy of the model and reduces the uncertainty due to training data of Model IV for Compressor 2. This shows that physical rules that do not help explaining the data are also redundant for the model and may reduce the reliability of the models.

Figure 3 also shows that nonlinear models with estimated regression coefficients do not suffer from numerical issues. Despite the use of numerical methods to estimate regression coefficients for Models III, IV and V, none of them show significant uncertainties due to discretization or iteration in Figure 3. A more in-depth investigation shows that the maximum uncertainty due to numerical methods is in the order of 10^{-8} kgs^{-1} which is much smaller than other uncertainty components in Figure 3.

Figure 3 and Table 5 illustrate an issue about the calculation method of uncertainty in this study. Model V is only less accurate than Model II for Compressor 1 and is one of the most accurate models for Compressor 2 according to Table 5, but its average uncertainty is only lower than that of Model I according to Figure 3. As shown in Figure 3, the high uncertainty of Model V is primarily because of its high uncertainty due to model random error whereas other uncertainty components of Model V are similar to Model II, III and IV. The reason for its high uncertainty for Compressor 1 lies in its accuracy over the training data points as shown in Table 6.

Table 6 Coefficients of determination (R^2) calculated based on data from operating conditions of the training data

Type	**Compressor 1**	**Compressor 2**
Model I	0.9999	0.9999
Model II	0.9998	0.9995
Model III	0.9991	0.9988
Model IV	0.9991	0.9989
Model V	0.9983	0.9990

Table 6 shows that the accuracy of Model V in Compressor 1 is much lower than other models. This is different from that in Table 5 which shows that Model V is more accurate than Model III and IV when all experimental data points are considered. The low accuracy of Model V over its training data points overestimates the uncertainty due to model random error of Model V in Figure 3. Hence the overall uncertainty of Model V of Compressor 1 is overestimated.

The high uncertainty of Model V for Compressor 2 in Figure 3 can be studied by investigating the uncertainty components of the models over the training data points only as shown in Figure 4.

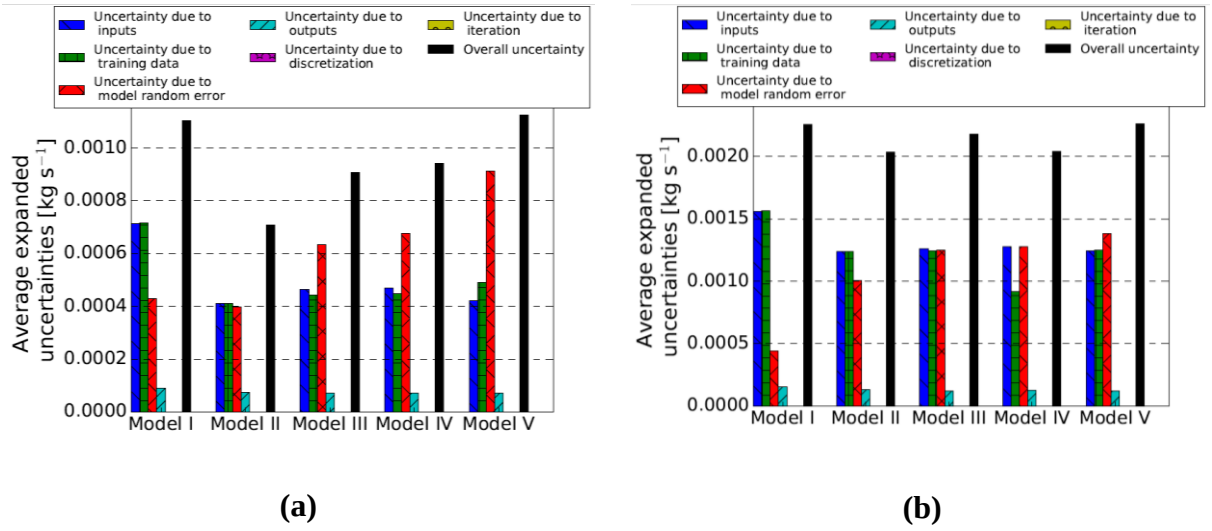


Figure 4 Average expanded uncertainties and their components under training data point conditions of (a) Compressor 1 and (b) Compressor 2

Figure 4 shows that the uncertainties due to model random error of both Models I and V for Compressor 2 are significantly lower than that in Figure 3. This shows that both models are subjected to extrapolation issues in their estimation of mass flow rate of Compressor 2, and more training data at different operating conditions are needed to ensure the reliability of the two models. To verify this claim, the models of Compressor 2 are re-trained with extra data points at the top-left handed corner of the operating range as shown in Figure 5.

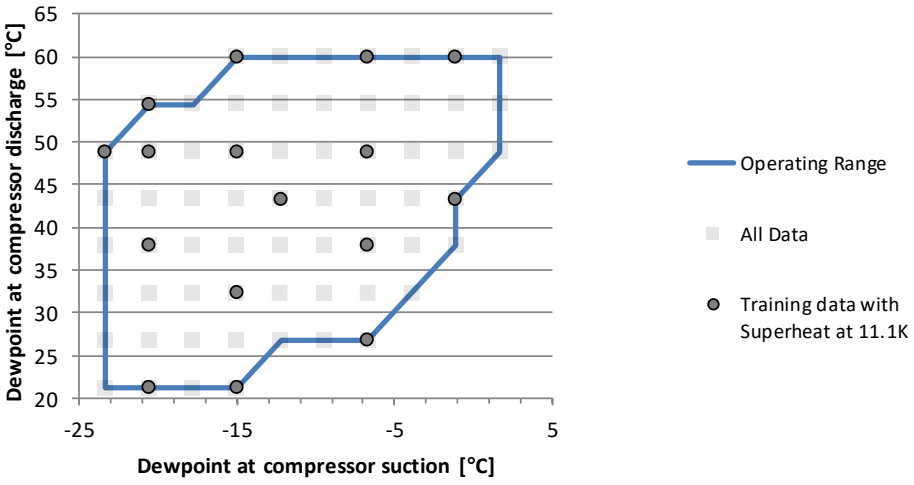


Figure 5 New testing conditions of training data with more data points on the top left-handed corner than that in Figure 2 to re-train models of Compressor 2

The results of the re-training are shown in Figure 6.

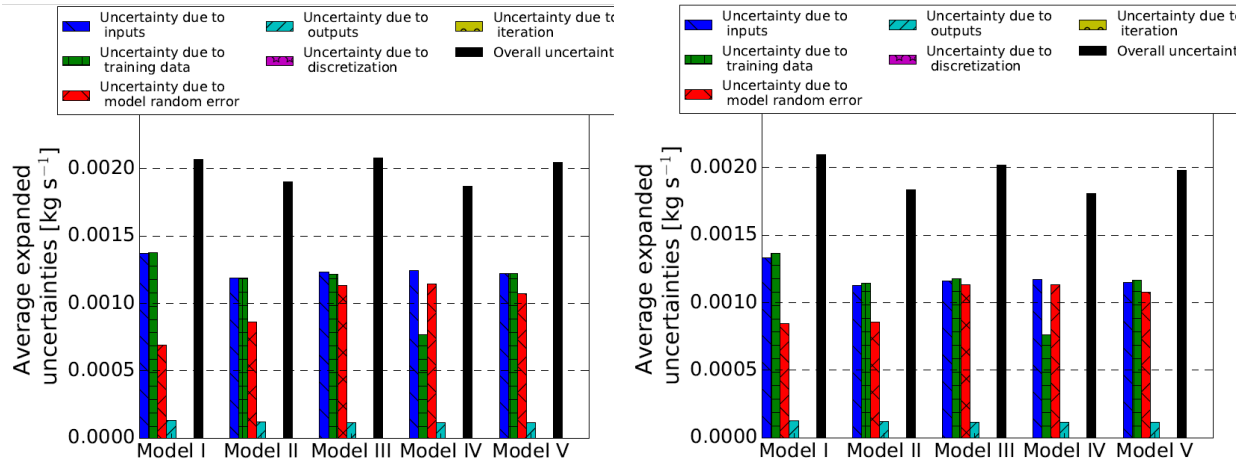


Figure 6 Average expanded uncertainties and their components under (a) training data points only and (b) all operating conditions of Compressor 2 after re-training with training data in Figure 5

Figure 6 shows that the differences of uncertainties of model random error between the training data points and all data points after re-training are not as significant as that before re-training. This shows that the new data points mitigate the potential extrapolation issues in both models.

The most reliable model according to Figure 3 is Model II for both compressors. Although it does not contain as many empirical coefficients as Model I or as many physical principles as Model III, IV and V that are advocated in current inverse modeling research, it has the least redundancy in the model structure and is complex enough to model the dependence of the mass flow rate with the independent variables. This conclusion is reinforced by the results in Table 5 which Model II is the most accurate model at all operating conditions of both compressors. Hence Model II is the model that is least affected by the randomness in measurement and extrapolation issues for these two compressors, and the calculation of the uncertainty at all operating conditions can help to approximate which model has the best accuracy at all operating conditions with measurement at limited operating conditions only.

7. Conclusions and recommendations

To conclude, the accuracy and reliability of various compressor mass flow rate models under different number of empirical coefficients and physical principles are evaluated to understand how the number of empirical coefficients and physical principles used in the models affect the accuracy and reliability of the models. It is found that redundant model structure, regardless of its empirical or physical origin, reduces the reliability of the models and may affect the accuracy of the model over its operating range. Model accuracy and reliability are the best when the models are constructed with a balance of the number of empirical coefficients and physical principles. Both empirical and semi-empirical models are subjected to extrapolation as far as the data are insufficient to explain the phenomenon reliably, and different model forms or extra data can help to reduce the effect of insufficient data. Based on the quantitative analysis, the following recommendations can be made:

a. Regression coefficients that do not help improving the model accuracy should be removed to maintain model reliability, regardless of their physical or empirical origins;
b. Appropriate choices of training data relative to the applicable range of the model can reduce the extrapolation effect on model reliability;
c. Models created by the same data set are subjected to extrapolation issues differently due to their different model structure, and the effect of extrapolation to the reliability of the models is independent of how many physical principles and empirical coefficient the models contain;

d. Comparison of uncertainty of multiple models at all operating conditions can help to approximate which model has the best accuracy over the entire operating range, and no experimental measurement in addition to the training data is needed by using this method.

8. References

[1] B. D. Rasmussen and A. Jakobsen, "Review of compressor models and performance characterizing variables," presented at the International Compressor Engineering Conference, West Lafayette, IN, 2000.

[2] G. Ding, "Recent developments in simulation techniques for vapour-compression refrigeration systems," *Int. J. Refrig.*, vol. 30, no. 7, pp. 1119–1133, Nov. 2007.

[3] P. Li, H. Qiao, Y. Li, J. E. Seem, J. Winkler, and X. Li, "Recent advances in dynamic modeling of HVAC equipment. Part 1: Equipment modeling," *HVACR Res.*, vol. 20, no. 1, pp. 136–149, Jan. 2014.

[4] X. Lin, H. Lee, Y. Hwang, and R. Radermacher, "A review of recent development in variable refrigerant flow systems," *Sci. Technol. Built Environ.*, vol. 21, no. 7, pp. 917–933, Oct. 2015.

[5] AHRI, *ANSI/AHRI Standard 540: 2004 Standard For Performance Rating Of Positive Displacement Refrigerant Compressors And Compressor Units*. Arlington, VA: Air-Conditioning, Heating and Refrigeration Institute, 2004.

[6] S. Shao, W. Shi, X. Li, and H. Chen, "Performance representation of variable-speed compressor for inverter air conditioners based on experimental data," *Int. J. Refrig.*, vol. 27, no. 8, pp. 805–815, Dec. 2004.

[7] B. Shen, "Improvement and validation of unitary air conditioner and heat pump simulation models at off-design conditions," PhD Thesis, Purdue University, 2006.

[8] L. Yang, L.-X. Zhao, C.-L. Zhang, and B. Gu, "Loss-efficiency model of single and variable-speed compressors using neural networks," *Int. J. Refrig.*, vol. 32, no. 6, pp. 1423–1432, Sep. 2009.

[9] D. I. Jähnig, D. T. Reindl, and S. A. Klein, “A semi-empirical method for representing domestic refrigerator/freezer compressor calorimeter test data,” *ASHRAE Trans.*, vol. 106, p. 122, 2000.

[10] M.-H. Kim and C. W. Bullard, “Thermal Performance Analysis of Small Hermetic Refrigeration and Air-Conditioning Compressors.,” *JSME Int. J. Ser. B*, vol. 45, no. 4, pp. 857–864, 2002.

[11] E. Winandy, C. Saavedra O, and J. Lebrun, “Experimental analysis and simplified modelling of a hermetic scroll refrigeration compressor,” *Appl. Therm. Eng.*, vol. 22, no. 2, pp. 107–120, Feb. 2002.

[12] E. Navarro, E. Granryd, J. F. Urchueguía, and J. M. Corberán, “A phenomenological model for analyzing reciprocating compressors,” *Int. J. Refrig.*, vol. 30, no. 7, pp. 1254–1265, Nov. 2007.

[13] M.-E. Duprez, E. Dumont, and M. Frère, “Modelling of reciprocating and scroll compressors,” *Int. J. Refrig.*, vol. 30, no. 5, pp. 873–886, Aug. 2007.

[14] C. P. Arora, *Refrigeration and Air Conditioning*, 3rd edition. Boston, Mass.: McGraw-Hill, 2009.

[15] T. Zakula, N. T. Gayeski, P. R. Armstrong, and L. K. Norford, “Variable-speed heat pump model for a wide range of cooling conditions and loads,” *HVACR Res.*, vol. 17, no. 5, pp. 670–691, Oct. 2011.

[16] H. Cheung and J. E. Braun, “Simulation of fault impacts for vapor compression systems by inverse modeling. Part II: System modeling and validation,” *HVACR Res.*, vol. 19, no. 7, pp. 907–921, Oct. 2013.

[17] C. O. R. Negrão, R. H. Erthal, D. E. V. Andrade, and L. W. da Silva, “A semi-empirical model for the unsteady-state simulation of reciprocating compressors for household refrigeration applications,” *Appl. Therm. Eng.*, vol. 31, no. 6–7, pp. 1114–1124, May 2011.

[18] W. Li, “Simplified steady-state modeling for hermetic compressors with focus on extrapolation,” *Int. J. Refrig.*, vol. 35, no. 6, pp. 1722–1733, Sep. 2012.

[19] V. Aute, C. Martin, and R. Radermacher, “A Study of Methods to Represent Compressor Performance Data over an Operating Envelope Based on a Finite Set of Test Data,” Air-Conditioning, Heating, and Refrigeration Institute, College Park, MD, 8013, 2015.

[20] V. Aute and C. Martin, “A Comprehensive Evaluation of Regression Uncertainty and the Effect of Sample Size on the AHRI-540 Method of Compressor Performance Representation,” in *23 rd International Compressor Engineering Conference at Purdue*, 2016.

[21] H. Cheung, S. Omer, and C. K. Bach, “A Method to Calculate Uncertainty of Empirical Compressor Maps with the Consideration of Extrapolation Effect and Choice of Training Data,” *Sci. Technol. Built Environ.*, 2017.

[22] K. K. Brown, H. W. Coleman, and W. G. Steele, “A Methodology for Determining Experimental Uncertainties in Regressions,” *J. Fluids Eng.*, vol. 120, no. 3, pp. 445–456, Sep. 1998.

[23] D. C. Montgomery, *Design and Analysis of Experiments*, 6 edition. Hoboken, NJ: Wiley, 2004.

[24] ASME, *ASME PTC 19.1-2013 Test Uncertainty Performance Test Codes*. New York: The American Society of Mechanical Engineers, 2013.

[25] S. J. Kline and F. A. McClintock, “Describing uncertainties in single-sample experiments,” *Mech. Eng.*, vol. 75, no. 1, pp. 3–8, 1953.

[26] J. Nocedal and S. Wright, *Numerical Optimization*. Springer Science & Business Media, 2006.

[27] W. L. Oberkampf and C. J. Roy, *Verification and Validation in Scientific Computing*. Cambridge University Press, 2010.

[28] A. E. Dabiri and C. K. Rice, “A Compressor Simulation Model with Corrections for the Level of Suction Gas Superheat,” in *ASHRAE Transactions*, 1981, vol. 87, p. 771.

[29] D. Kraft and I. München, “Algorithm 733: TOMP - Fortran modules for optimal control calculations,” *ACM Trans Math Soft*, pp. 262–281, 1994.

[30] S. Shrestha, I. Mahderekal, V. Sharma, and O. Abdelaziz, “Test Report #11: Compressor Calorimeter Test of R-410A Alternatives R-32, DR-5, and L-41a,” Air-Conditioning, Heating, and Refrigeration Institute, 2013.

[31] S. Shrestha, V. Sharma, and O. Abdelaziz, “Test Report #21: Compressor Calorimeter Test of R-404A Alternatives ARM-31a, D2Y-65, L-40, and R-32/R-134a (50/50),” Air-Conditioning, Heating, and Refrigeration Institute, 2013.

[32] ASHRAE, *ANSI/ASHRAE Standard 23.1-2010 Methods of Testing for Rating the Performance of Positive Displacement Refrigerant Compressors and Condensing Units that Operate at Subcritical Temperatures of the Refrigerant*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc., 2010.

[33] R. E. Walpole, R. H. Myers, S. L. Myers, and K. E. Ye, *Probability and Statistics for Engineers and Scientists*, 9 edition. Boston: Pearson, 2011.

[34] JCGM, *JCGM 100:2008 Evaluation of measurement data — Guide to the expression of uncertainty in measurement*. Joint Committee for Guides in Metrology, 2008.

[35] R. T. St. Laurent and R. D. Cook, “Leverage, Local Influence and Curvature in Nonlinear Regression,” *Biometrika*, vol. 80, no. 1, pp. 99–106, 1993.

[36] Y. A. Çengel and M. A. Boles, *Thermodynamics: An Engineering Approach*, 5 edition. Boston: McGraw-Hill Science/Engineering/Math, 2005.

Appendix Uncertainty calculation methods of different compressor mass flow rate models

This appendix describes the modification made to the uncertainty calculation method in [21] in order to calculate the uncertainty components of the estimated mass flow rate in various models. These uncertainty components are summed together to calculate the standard uncertainty of the predicted mass flow rate, which is multiplied with its Student's *t* value according to the Student's *t*-distribution to calculate its expanded uncertainty.

A.1 Uncertainty calculation method of Model I

Although Cheung et al. [21] generally described the uncertainty calculation method of Model I, it did not describe the uncertainty due to superheat correction method in [28] that is often used with Model I. With the consideration of superheat, the uncertainty components of the predicted mass flow rate can be written as Equations (16), (17), (18) and (19).

$$\Delta\hat{\dot{m}}_{comp,I,input} = \sqrt{\begin{aligned}&\left[\Delta\hat{\dot{m}}_{comp,rat,input}\left(1+\hat{\beta}_{10}\theta_1(T_{suc},P_{suc})\right)\right]^2+\left[\hat{\dot{m}}_{comp,rat}\hat{\beta}_{10}\frac{\partial\theta_1}{\partial T_{suc}}\Delta T_{suc}\right]^2\\&+\left[\hat{\dot{m}}_{comp,rat}\hat{\beta}_{10}\frac{\partial\theta_1}{\partial P_{suc}}\Delta P_{suc}\right]^2\\&+\left[\hat{\dot{m}}_{comp,rat}\hat{\beta}_{10}\left(\frac{\Delta\rho_{suc,EOS}}{\rho_{suc,rat}}-\frac{\rho_{suc}}{\rho_{suc,rat}^2}\Delta\rho_{suc,rat,EOS}\right)\right]^2\end{aligned}} \quad \textbf{(16)}$$

$$\Delta\hat{\dot{m}}_{comp,I,output} = \Delta\hat{\dot{m}}_{comp,rat,output}\frac{\hat{\dot{m}}_{comp}}{\hat{\dot{m}}_{comp,rat}} \quad \textbf{(17)}$$

$$\Delta\hat{\dot{m}}_{comp,I,train} = \Delta\hat{\dot{m}}_{comp,rat,train}\frac{\hat{\dot{m}}_{comp}}{\hat{\dot{m}}_{comp,rat}} \quad \textbf{(18)}$$

$$\Delta\hat{\dot{m}}_{comp,I,model} = \sqrt{\left(\Delta\hat{\dot{m}}_{comp,rat,model}\frac{\hat{\dot{m}}_{comp,I}}{\hat{\dot{m}}_{comp,rat}}\right)^2+\left(\hat{\dot{m}}_{comp,rat}\theta_1(T_{suc},P_{suc})\Delta\hat{\beta}_{10}\right)^2} \quad \textbf{(19)}$$

where the formula to calculate the uncertainty component of the rated mass flow rate and the uncertainty due to equation of state are described in [21]. Since the method does not involve numerical methods, its uncertainty due to numerical method is zero.

A.2 Uncertainty calculation method of Model II

The calculation method of uncertainty components in Model II is similar to that of Model I, and their expressions are shown in Equations (20), (21), (22) and (23).

$$\Delta\hat{\dot{m}}_{comp,II,input} = \sqrt{\begin{matrix}\left[\rho_{suc}(T_{suc},P_{suc})\Delta\theta_{2,input}\right]^2 + \left[\theta_2\frac{\partial\rho_{suc}}{\partial T_{suc}}\Delta T_{suc}\right]^2 \\ + \left[\theta_2\frac{\partial\rho_{suc}}{\partial P_{suc}}\Delta P_{suc}\right]^2 + \left[\Delta\rho_{suc,EOS}\theta_2\right]^2\end{matrix}} \quad \textbf{(20)}$$

$$\Delta\hat{\dot{m}}_{comp,II,output} = \rho_{suc}(T_{suc},P_{suc})\Delta\theta_{2,output} \quad \textbf{(21)}$$

$$\Delta\hat{\dot{m}}_{comp,II,train} = \rho_{suc}(T_{suc},P_{suc})\Delta\theta_{2,train} \quad \textbf{(22)}$$

$$\Delta\hat{\dot{m}}_{comp,II,model} = \rho_{suc}(T_{suc},P_{suc})\Delta\theta_{2,model} \quad \textbf{(23)}$$

where the calculation of uncertainty components of the arbitrary function θ_2 follows the method in [21].

A.3 Uncertainty calculation method of Model III

The uncertainty calculation of Model III differs from [21] that the mathematical model is a nonlinear model with regression coefficients estimated by an implicit numerical method. The formula to calculate the uncertainty components are derived as Equations (24), (25), (26), (27), (28), (29), (30), (31), (32), (33), (34) and (35).

$$\Delta\hat{\dot{m}}_{comp,III,input} = \sqrt{\sum_{\psi=T_{suc},P_{suc},P_{dis}}\left[\frac{\partial\hat{\dot{m}}_{comp,III}}{\partial\psi}\Delta\psi\right]^2 + \sum_{\psi=\rho_{suc},c_v,c_p}\left[\frac{\partial\hat{\dot{m}}_{comp,III}}{\partial\psi}\Delta\psi_{EOS}\right]^2} \quad \textbf{(24)}$$

$$\Delta\hat{\dot{m}}_{comp,III,output} = \frac{\hat{\dot{m}}_{comp,III}}{n}\sum_{i=1}^{n_{train}}\frac{\Delta\dot{m}_{comp,train,i}}{\dot{m}_{comp,train}} \quad \textbf{(25)}$$

$$\Delta\hat{\dot{m}}_{comp,III,train}$$

$$= \sqrt{\begin{aligned} &\sum_{\psi=T_{suc},P_{suc},P_{dis},\dot{m}_{comp}} \sum_{j=1}^{n_{train}} \sum_{k=1}^{n_{train}} \left(\sum_{i=0}^{m-1} \left(\frac{\partial \hat{\dot{m}}_{comp,III}}{\partial \hat{\beta}_i} \frac{\partial \hat{\beta}_i}{\partial \psi_{train,j}} \right) \Delta\psi_{train,j} \cdot \left(\frac{\partial \hat{\dot{m}}_{comp,III}}{\partial \hat{\beta}_i} \frac{\partial \hat{\beta}_i}{\partial \psi_{train,k}} \right) \Delta\psi_{train,k} \right) \\ &+ \sum_{\psi=\rho_{suc},c_p,c_v} \sum_{j=1}^{n_{train}} \sum_{k=1}^{n_{train}} \left(\sum_{i=0}^{m-1} \left(\frac{\partial \hat{\dot{m}}_{comp,III}}{\partial \hat{\beta}_i} \frac{\partial \hat{\beta}_i}{\partial \psi_{train,j}} \right) \Delta\psi_{EOS,train,j} \cdot \left(\frac{\partial \hat{\dot{m}}_{comp,III}}{\partial \hat{\beta}_i} \frac{\partial \hat{\beta}_i}{\partial \psi_{train,k}} \right) \Delta\psi_{EOS,train,k} \right) \end{aligned}} \quad (26)$$

$$\Delta\hat{\dot{m}}_{comp,III,model} = \frac{\sum_{i=1}^{n} \left(\hat{\dot{m}}_{comp,i} - \dot{m}_{comp,i} \right)^2}{n_{train} - q} \sqrt{1 + max(j_{lev}, 0)} \quad (27)$$

$$j_{lev} = \left(\frac{\partial \hat{\dot{m}}_{comp,III}}{\partial \vec{\hat{\beta}}} \right)^T \left(\boldsymbol{J}_{train}^T \boldsymbol{J}_{train} - \left[\vec{\dot{m}}_{comp,III,train} - \vec{\hat{\dot{m}}}_{comp,III,train} \right] [\boldsymbol{H}_{train}] \right)^{-1} \left(\frac{\partial \hat{\dot{m}}_{comp,III}}{\partial \vec{\hat{\beta}}} \right) \quad (28)$$

$$\boldsymbol{J}_{trai} = \begin{bmatrix} \frac{\partial \hat{\dot{m}}_{comp,III,train,1}}{\partial \hat{\beta}_0} & \frac{\partial \hat{\dot{m}}_{comp,III,train,1}}{\partial \hat{\beta}_1} & \frac{\partial \hat{\dot{m}}_{comp,III,train,1}}{\partial \hat{\beta}_2} \\ \vdots & \vdots & \vdots \\ \frac{\partial \hat{\dot{m}}_{comp,III,train,n}}{\partial \hat{\beta}_0} & \frac{\partial \hat{\dot{m}}_{comp,III,train,n}}{\partial \hat{\beta}_1} & \frac{\partial \hat{\dot{m}}_{comp,III,train,n}}{\partial \hat{\beta}_2} \end{bmatrix} \quad (29)$$

$$\boldsymbol{H}_{train,i} = \begin{bmatrix} \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_0^2} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_1 \partial \hat{\beta}_0} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_2 \partial \hat{\beta}_0} \\ \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_0 \partial \hat{\beta}_1} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_1^2} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_2 \partial \hat{\beta}_1} \\ \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_0 \partial \hat{\beta}_2} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_1 \partial \hat{\beta}_2} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_2^2} \end{bmatrix} \quad (30)$$

$$\left[\vec{\dot{m}}_{comp,train} - \vec{\hat{\dot{m}}}_{comp,train} \right] [\boldsymbol{H}_{train}] = \sum_i (\dot{m}_{comp,train,i} - \hat{\dot{m}}_{comp,train,i}) \boldsymbol{H}_{train,i} \quad (31)$$

$$\Delta\hat{\dot{m}}_{comp,III,num} = \sqrt{\hat{\dot{m}}_{comp,III,disc}^2 + \hat{\dot{m}}_{comp,III,it}^2} \quad (32)$$

$$\Delta\hat{\dot{m}}_{comp,III,disc} = \sqrt{\sum_{i=0}^{q-1}\left(\frac{\partial\hat{\dot{m}}_{comp,III}}{\partial\hat{\beta}_i}\frac{\hat{\beta}_i(h)-\hat{\beta}_i(\Gamma h)}{\Gamma^{\lambda}-1}\right)^2} \quad \textbf{(33)}$$

$$\Delta\hat{\dot{m}}_{comp,III,it} = \sqrt{\sum_{i=0}^{q-1}\left(\frac{\partial\hat{\dot{m}}_{comp,III}}{\partial\hat{\beta}_i}\frac{\hat{\beta}_i-\hat{\beta}_i[\text{with one more iteration}]}{\Upsilon_i-1}\right)^2} \quad \textbf{(34)}$$

$$\Upsilon_i = \left|\frac{\hat{\beta}_i-\hat{\beta}_i[\text{with one more iteration}]}{\hat{\beta}_i-\hat{\beta}_i[\text{with one less iteration}]}\right| \quad \textbf{(35)}$$

The uncertainty due to training data in Equation (26) considers the measurement uncertainty of the training data, the uncertainty due to equation of state in the training data and the correlation of uncertainties between data as the same uncertainty source according to JCGM Guide of Measurement 100-2008 [34]. While its derivatives of the estimated mass flow rate with respect to the regression coefficients can be calculated analytically, the derivatives of the regression coefficients with respect to the training data points are approximated by finite difference method.

The uncertainty due to model random error in Equation (27) considers Jacobian leverage instead of tangential leverage [35] to calculate the uncertainty due to model random error in [21]. This allows the nonlinearity of the mass flow rate model to be accounted for more accurately. However, since the original Jacobian leverage is not restricted to be positive and it is impossible for the uncertainty due to the incomprehensiveness of training data to offset the uncertainty due to inaccuracy of the model, the Jacobian leverage is limited to be positive in Equation (27).

Equation (32) calculates the uncertainty due to numerical method from two sources: uncertainty due to discretization in Equation (33) and uncertainty due to iteration in Equation (34).

Uncertainty due to discretization calculates the uncertainty as a result of using thresholds to approximate derivatives numerically in Equation (10). It is calculated by Richarson Extrapolation [27] and involves regression coefficients calculated by using a larger threshold and the order of accuracy of the numerical method using threshold. In this study, the multiplier of the threshold Γ is set to 2 and the order of accuracy λ of the forward difference method to approximate derivatives is found to be 1 [26].

Uncertainty due to iteration calculates the uncertainty as a result of using iterative methods. Using the Eigenvalue method [27], regression coefficients calculated by fewer and more iterations are used to calculate the uncertainty due to iteration as shown in Equations (34) and (35).

A.4 Uncertainty calculation method of Model IV

The formula of the uncertainty calculation of Model IV is slightly different from that of Model III due to the use of the forth coefficients in Equation (16). The difference lies in the calculation of the Jacobian and Hessian matrices in Equations (29) and (30) to include the forth coefficients. The new Jacobian and Hessian matrices for Model IV are shown in Equations (36) and (37).

$$\boldsymbol{J}_{train} = \begin{bmatrix} \frac{\partial \hat{\dot{m}}_{comp,IV,train,1}}{\partial \hat{\beta}_0} & \cdots & \frac{\partial \hat{\dot{m}}_{comp,IV,train,1}}{\partial \hat{\beta}_3} \\ \vdots & \vdots & \vdots \\ \frac{\partial \hat{\dot{m}}_{comp,IV,train,n}}{\partial \hat{\beta}_0} & \cdots & \frac{\partial \hat{\dot{m}}_{comp,IV,train,n}}{\partial \hat{\beta}_3} \end{bmatrix} \quad \textbf{(36)}$$

$$\boldsymbol{H}_{train,i} = \begin{bmatrix} \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_0^2} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_1 \partial \hat{\beta}_0} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_2 \partial \hat{\beta}_0} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_3 \partial \hat{\beta}_0} \\ \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_0 \partial \hat{\beta}_1} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_1^2} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_2 \partial \hat{\beta}_1} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_3 \partial \hat{\beta}_1} \\ \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_0 \partial \hat{\beta}_2} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_2 \partial \hat{\beta}_1} & \frac{\partial^2 \hat{\dot{m}}_{comp,III,train,i}}{\partial \hat{\beta}_2^2} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_3 \partial \hat{\beta}_2} \\ \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_0 \partial \hat{\beta}_3} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_1 \partial \hat{\beta}_3} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_2 \partial \hat{\beta}_3} & \frac{\partial^2 \hat{\dot{m}}_{comp,IV,train,i}}{\partial \hat{\beta}_3^2} \end{bmatrix} \quad \textbf{(37)}$$

A.4 Uncertainty calculation method of Model V

Uncertainty of Model V is calculated similarly as the uncertainty of Model III with exceptions due to the use of the polytropic coefficient in Equations (12), (13) and (14). The difference lies in the calculation of its uncertainty due to inputs and training data. Its uncertainty due to inputs, including the uncertainty due to the use of the specific entropy, is calculated by Equation (38).

$$\Delta\hat{\dot{m}}_{comp,IV,input} = \sqrt{\sum_{\psi=T_{suc},P_{suc},P_{dis}} \left[\frac{\partial\hat{\dot{m}}_{comp,III}}{\partial\psi}\Delta\psi\right]^2 + \sum_{\psi=\rho_{suc},\rho_{dis,s},s_{suc}} \left[\frac{\partial\hat{\dot{m}}_{comp,III}}{\partial\psi}\Delta\psi_{EOS}\right]^2} \quad \textbf{(38)}$$

The relative uncertainty of specific entropy is the same as that of specific heat capacity according to their relationship derived from the first and second law of thermodynamics in Equation (39) [36].

$$ds = c_v\frac{dT}{T} + \frac{Pdv}{T} \quad \textbf{(39)}$$

Its uncertainty due to training data is also modified from Equation (26) to include the uncertainty of the equation of state of the specific entropy, and the resultant formula is Equation (40).

$$\Delta\hat{\dot{m}}_{comp,IV,train} = \sqrt{\begin{array}{l}\sum_{\psi=T_{suc},P_{suc},P_{dis},\dot{m}_{comp}} \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\begin{array}{c}\sum_{i=0}^{m-1}\left(\frac{\partial\hat{\dot{m}}_{comp,IV}}{\partial\hat{\beta}_i}\frac{\partial\hat{\beta}_i}{\partial\psi_{train,j}}\right)\Delta\psi_{train,j}\cdot \\ \left(\frac{\partial\hat{\dot{m}}_{comp,IV}}{\partial\hat{\beta}_i}\frac{\partial\hat{\beta}_i}{\partial\psi_{train,k}}\right)\Delta\psi_{train,k}\end{array}\right) \\ + \sum_{\psi=\rho_{suc},\rho_{dis,s},s_{suc}} \sum_{j=1}^{n}\sum_{k=1}^{n}\left(\begin{array}{c}\sum_{i=0}^{m-1}\left(\frac{\partial\hat{\dot{m}}_{comp,IV}}{\partial\hat{\beta}_i}\frac{\partial\hat{\beta}_i}{\partial\psi_{train,j}}\right)\Delta\psi_{EOS,train,j}\cdot \\ \left(\frac{\partial\hat{\dot{m}}_{comp,IV}}{\partial\hat{\beta}_i}\frac{\partial\hat{\beta}_i}{\partial\psi_{train,k}}\right)\Delta\psi_{EOS,train,k}\end{array}\right)\end{array}} \quad \textbf{(40)}$$

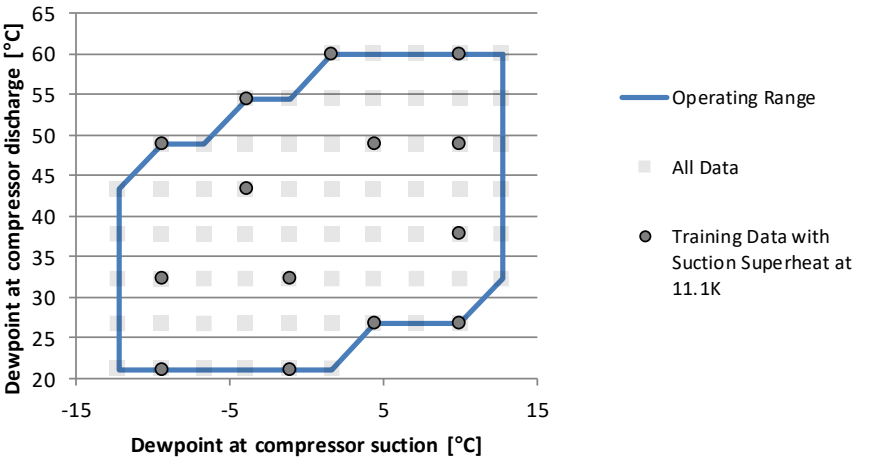
Supplementary materials for the paper "A comparison of the effect of empirical and physical modeling approaches to extrapolation capability of compressor models by uncertainty calculation: a case study with common semi-empirical compressor mass flow rate models"

Howard Cheung, Shengwei Wang*
Department of Building Services Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong
* Corresponding author: beswwang@polyu.edu.hk

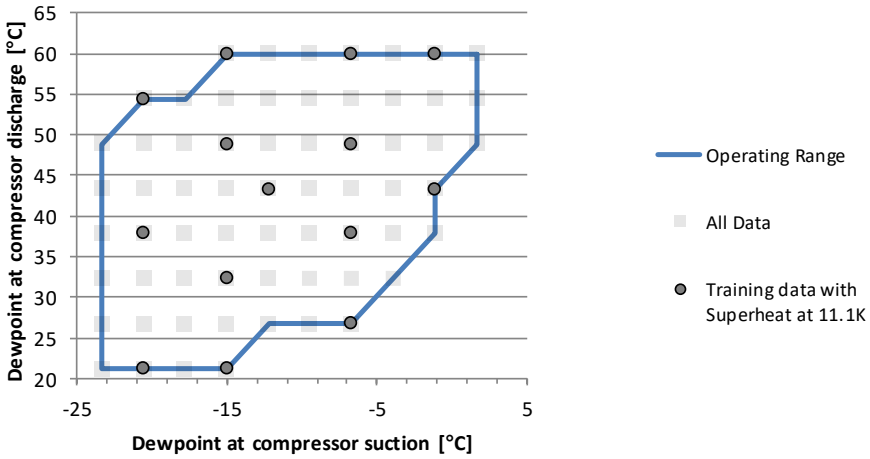
Effect of extrapolation on the accuracy of the models

While Cheung et al. provides detailed information on how the accuracy of the empirical compressor map from ANSI/AHRI Standard 540-2004 [1] is reduced by extrapolation [2] and it is generally understood that extrapolation leads to inaccuracy, the effects of extrapolation to the accuracy models studied in the paper for the two compressors are not well quantified in the paper. This section in the supplementary materials discusses the effect of extrapolation of the models of the two compressors.

Extrapolation of a regression model is the use of the model at operating conditions deviated from that of the training data. To understand how a compressor model of Compressor 1 and 2 in the paper extrapolates, the position of the training data points of the compressor mass flow rate models within the range of available data is shown in Figure 1.



(a)



(b)

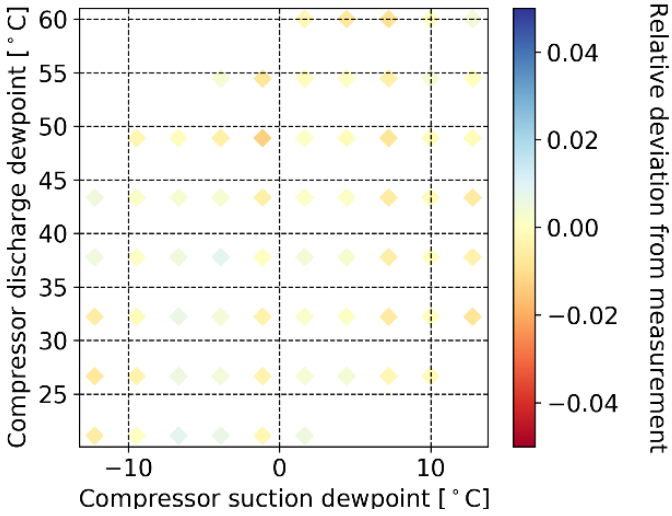
Figure 1 Illustration of compressor calorimeter operating conditions of (a) Compressor 1 and (b) Compressor 2 (identical to Figure 2 in the paper)

If extrapolation reduces the accuracy of a model, the accuracy of the models under the conditions on the blue solid line in Figure 1 should be less than the accuracy of the models under the conditions around the center of Figure 1(a) and Figure 1(b). The accuracy of the models at different data points is quantified by Equation (1).

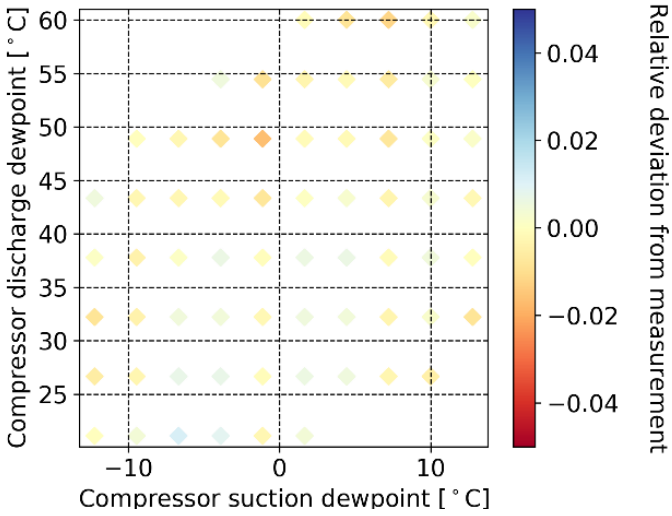
$$\text{Relative deviation} = \frac{\hat{\dot{m}}_{comp} - \dot{m}_{comp}}{\dot{m}_{comp}} \quad \textbf{(1)}$$

Equation (1) shows the calculation of relative deviation at a data point. The larger the relative deviation, the less accurate a model is at the specified condition of the data point.

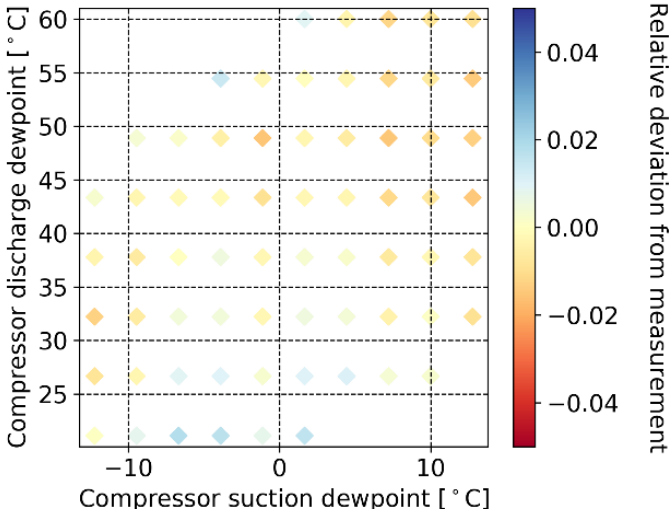
To illustrate the effect of extrapolation on the accuracy of different models, the relative deviations of different maps at different data points with compressor suction superheat at 11.1K are plotted in Figure 2.



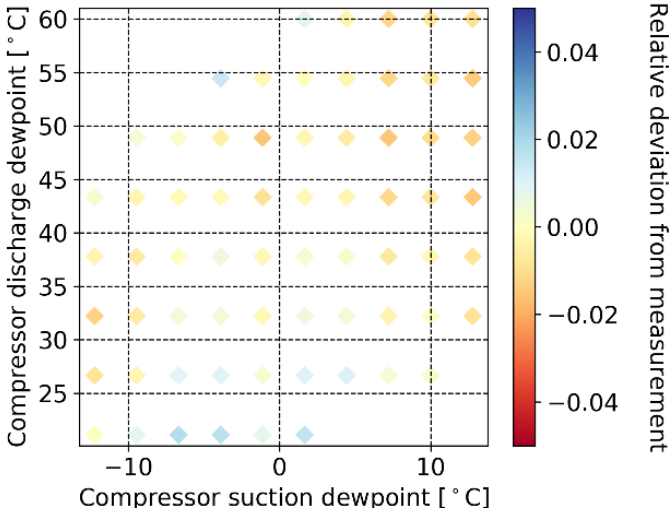
(a) Model I



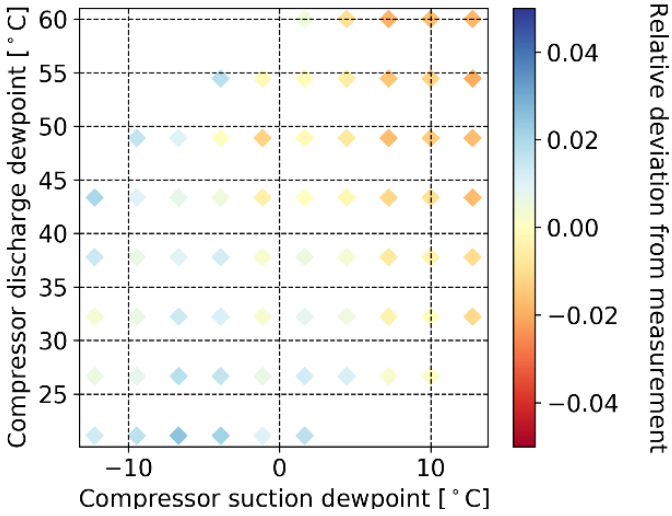
(b) Model II



(c) Model III

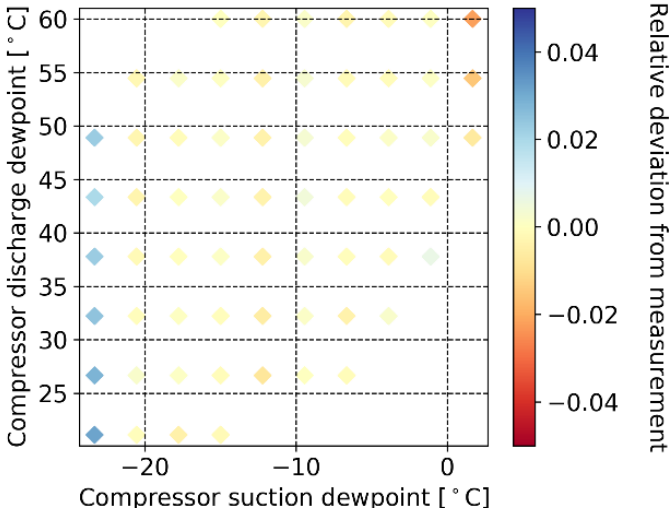


(d) Model IV

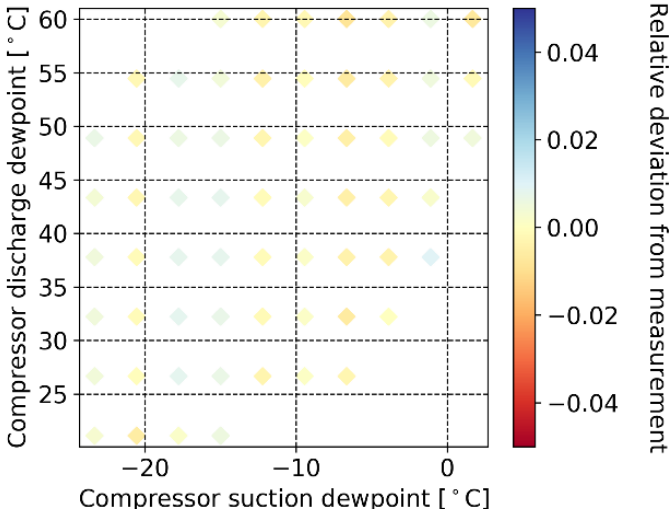


(e) Model V

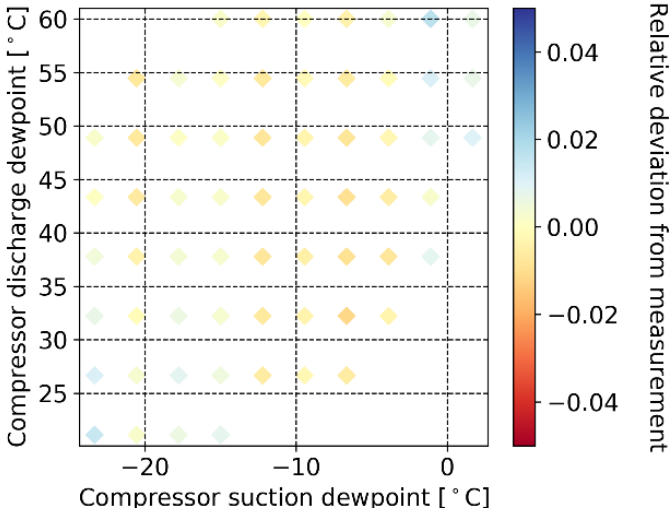
Figure 2 Relative deviations of estimation of different models under suction superheat at 11.1K and different compressor suction and discharge dewpoint for Compressor 1



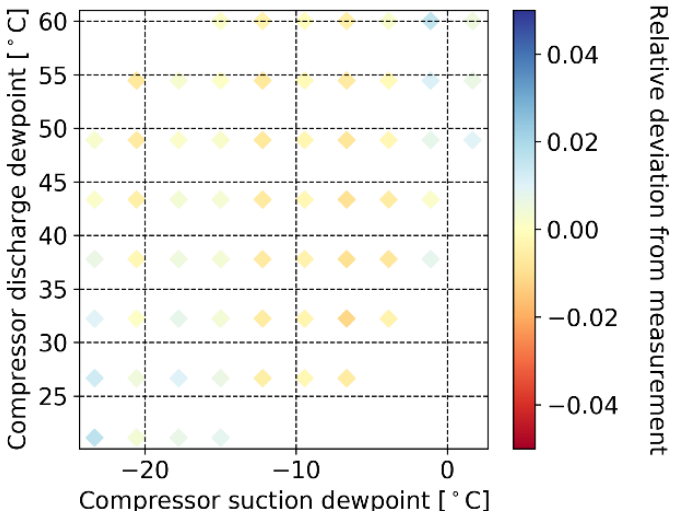
(a) Model I



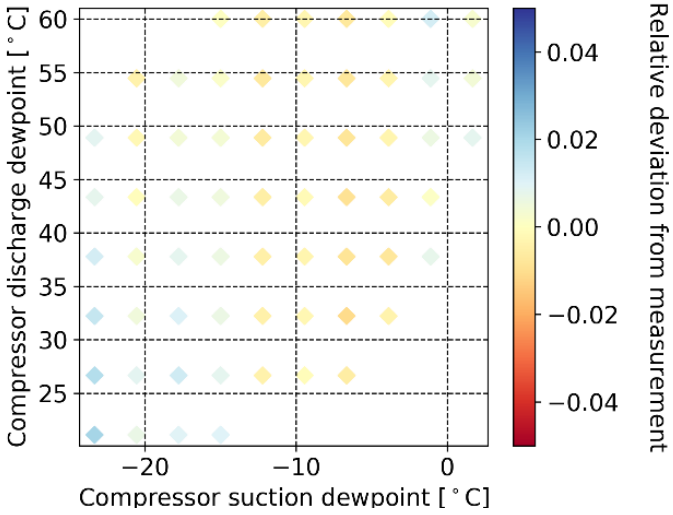
(b) Model II



(c) Model III



(d) Model IV



(e) Model V

Figure 3 Relative deviations of estimation of different models under suction superheat at 11.1K and different compressor suction and discharge dewpoint for Compressor 2

At first glance, not much significant deviation can be found in models of Compressor 2 in Figure 3 and only large relative deviations can be found at the edges of Figure 2. However, a close observation to Figure 3 reveals that large errors always occur at compressor suction temperature that is absent in the training data points of models of Compressor 2 in Figure 1. This shows that

the deterioration of model accuracy due to extrapolation occurs in all models of both compressors.

References

[1] AHRI, ANSI/AHRI Standard 540: 2004 Standard For Performance Rating Of Positive Displacement Refrigerant Compressors And Compressor Units. Arlington, VA: Air-Conditioning, Heating and Refrigeration Institute, 2004.

[2] H. Cheung, S. Omer, and C. K. Bach, "A Method to Calculate Uncertainty of Empirical Compressor Maps with the Consideration of Extrapolation Effect and Choice of Training Data," Sci. Technol. Built Environ., 2017.