Prediction of uncertainty of 10-coefficient compressor maps for extreme operating conditions

Howard Cheung

Postdoctoral Research Fellow, Ray W. Herrick Laboratories, School of Mechanical Engineering, Purdue University, 177 S. Russell St., West Lafayette, IN 47907-2031, US

E-mail: `cheung@purdue.edu`

Christian K. Bach

Assistant Professor, Mechanical and Aerospace Engineering, Oklahoma State University, 218 Engineering North, Stillwater, OK 74078-5016

E-mail: `cbach@okstate.edu`

Abstract. Empirical compressor maps are a simple and reliable approach for heating and cooling system designers to estimate compressor refrigerant mass flow rate and power consumption quickly. These maps were used for a long time since most compressor manufacturers build the maps with extensive test matrices, leading to good accuracy. However, the situation changes when engineers extrapolate the maps to investigate the compressors performance under extreme operating conditions such as for cold climate heat pump applications or under conditions with system faults. Engineers are not confident on the exact uncertainty of the extrapolation, and often claim that the inaccuracy of their studies is a result of high extrapolation uncertainty. This paper presents a method to estimate the extrapolation uncertainty due to the structure of the test matrix that trains the manufacturer maps and helps the investigators to understand if the extrapolation is the main cause of their inaccuracy. To verify that the method can estimate the uncertainty due to extrapolation, the study builds 10-coefficient compressor maps trained by different test matrices of the same size and different operating points. The maps are used to estimate the compressor performance under different operating points and their estimation uncertainties are compared. The results show that the component of the uncertainty that depends on the structure of the test matrix is small at operating conditions within the test matrix but the uncertainty grows significantly as the estimation becomes further away from the operating conditions within the test matrices.

1. Sources of uncertainty

Uncertainty of the compressor map output is the range where the true value of the map output may be relative to the map output. It consists of multiple components and can be grouped as follows:

- Uncertainty due to inputs
- Uncertainty due to training data
- Uncertainty due ot model random error
- Uncertainty due to outputs

1.1. Uncertainty due to inputs

Uncertainty due to inputs is the uncertainty propagated to the map output due to the uncertainty in the inputs to the maps. The inputs to the map (evaporating and condenser temperature) are usually obtained by converting pressure measurement to saturation temperature with refrigerant equations of state, and the estimated saturation temperature contains uncertainty from both the equation of state and the pressure measurement. The equation of state of R22 estimates saturation pressure at an uncertainty of 0.2% [1]. When the equation of state estimates an saturation temperature at a given pressure, this uncertainty is transformed into a component of the uncertainty of the saturation temperature as shown in Eqns. (1) and (2).

$$\frac{\Delta P_{sat,EOS}}{P_{sat}(T)} = 0.002 \tag{1}$$

$$\Delta T_{sat,EOS} = |\frac{\partial T_{sat}(P)}{\partial P}|\Delta P_{sat,EOS} \tag{2}$$

where $\Delta P_{sat,EOS}$ is the uncertainty of saturation pressure as a result of uncertainty of the equation of state, $P_{sat}(T)$ is the saturation pressure from temperature T, $\Delta T_{sat,EOS}$ is the uncertainty of saturation temperature as a result of uncertainty of the equation of state and $T_{sat}(P)$ is the saturation temperature at pressure P.

The component of the uncertainty due to pressure measurement in saturation temperature values is calculated by Eqn. (3).

$$\Delta T_{sat,mea} = |\frac{\partial T_{sat}(P)}{\partial P}|\Delta P_{sat,mea} \tag{3}$$

where $\Delta T_{sat,mea}$ is the uncertainty of saturation temperature as a result of pressure measurement and $\Delta P_{sat,mea}$ is the uncertainty of pressure measurement.

The overall uncertainty of the saturation temperature is given by Eqn. (4).

$$\Delta T_{sat} = \sqrt{(\Delta T_{sat,EOS})^2 + (\Delta T_{sat,mea})^2} \tag{4}$$

The uncertainty of the map output propagated from the inputs of condensing temperature and evaporating temperature can be given by Eqn. (5).

$$\Delta \hat{\dot{W}}_{input} = \sqrt{(\frac{\partial \hat{\dot{W}}}{\partial T_{evap}}\Delta T_{evap})^2 + (\frac{\partial \hat{\dot{W}}}{\partial T_{cond}}\Delta T_{cond})^2} \tag{5}$$

where $\Delta \hat{\dot{W}}_{input}$ is the uncertainty due to inputs at the map output

1.2. Uncertainty due to training data

Uncertainty due to training data is the uncertainty propagated to the map output from the training data through the map coefficients. This can be understood by considering the estimation of the map coefficients as a function of the training data as Eqn. (6).

$$\hat{\vec{\beta}} = g(T_{evap,train,1}, ..., T_{evap,train,n}, T_{cond,train,1}, ..., T_{cond,train,n}, \dot{W}_{train,1}, ..., \dot{W}_{train,n}) \tag{6}$$

The uncertainty propagated to the map output through function g and $\hat{\vec{\beta}}$ in Eqn. (6) is calculated by Eqn. (7).

$$\Delta\hat{\dot{W}}_{train} = \sqrt{\begin{aligned}&\Sigma_{j=1}^{n}(\Sigma_{i=1}^{m}(\frac{\hat{\partial}W}{\partial\beta_i}\frac{\partial\beta_i}{\partial T_{evap,train,j}})\Delta T_{evap,train,j})^2\\&+\Sigma_{j=1}^{n}(\Sigma_{i=1}^{m}(\frac{\hat{\partial}W}{\partial\beta_i}\frac{\partial\beta_i}{\partial T_{cond,train,j}})\Delta T_{cond,train,j})^2\\&+\Sigma_{j=1}^{n}(\Sigma_{i=1}^{m}(\frac{\hat{\partial}W}{\partial\beta_i}\frac{\partial\beta_i}{\partial\dot{W}_{,train,j}})\Delta\dot{W}_{train,j})^2\end{aligned}} \qquad (7)$$

where $\Delta\hat{\dot{W}}_{train}$ is the uncertainty due to training data at the map output.

1.3. Uncertainty due to model random error

In linear regression, the random error ε in Eqn. (EQUATION NUMBER TO BE ADDED) is assumed to be normally distributed around zero with some finite variance. This variance becomes part of the uncertainty of the uncertainty of the map output and can be presented in the form of confidence intervals. Statistic textbooks [2, 3] illusrated that the confidence interval of the variance can be calculated as Eqn. (8).

$$\Delta\hat{\dot{W}}_{model} = t_{n-m,1-\alpha/2}\sigma\sqrt{1 + \vec{x}^T(\mathbf{X}_{train}^T\mathbf{X}_{train})^{-1}\vec{x}} \qquad (8)$$

where $\Delta\hat{\dot{W}}_{train}$ is the uncertainty due to model random error at the map output.

One significant term in Eqn. (8) is $\vec{x}^T(\mathbf{X}_{train}^T\mathbf{X}_{train})^{-1}\vec{x}$ which is the leverage of a regression model [4]. It estimates how deviated the current input vector to the map is relative to the training data of the regression model, and its magnitude grows with the deviation. This helps to understand if the estimation is related to the training data and if the model is applicable at the current situation described by the input vector.

1.4. Uncertainty due to output

Uncertainty of the map output should represent the probable range where the true value lies relative to the map output. However, since the map is built from data obtained from measured power consumption but not the true values of the power consumption, the uncertainty components of the map output propagated from other sources only estimate the uncertainty of the estimation with the measured power consumption. Another component of uncertainty must be introduced so that the map output uncertainty is the uncertainty to the true map output. This uncertainty component can be approximated with the uncertainty of the measured power consumption with its true values as Eqn. (9).

$$\Delta\hat{\dot{W}}_{output} = \frac{1}{n}\Sigma_{i=1}^{n}\frac{\Delta\dot{W}_{train,i}}{\dot{W}_{train,i}} \qquad (9)$$

1.5. Overall uncertainty

The overall uncertainty of the map output is given by the square of the sum of all uncertainty components as Eqn. (10).

$$\Delta\hat{\dot{W}} = \sqrt{(\Delta\hat{\dot{W}}_{input})^2 + (\Delta\hat{\dot{W}}_{train})^2 + (\Delta\hat{\dot{W}}_{model})^2 + (\Delta\hat{\dot{W}}_{output})^2} \qquad (10)$$

1.6. Effect of range of training data on map output

To study how the range of training data affects map output accuracy and uncertainty, the difference between the map output and approximated measurement and the map output uncertainty for all available data points are plotted in Figure 1, and their uncertainty components are plotted in Figure 2.
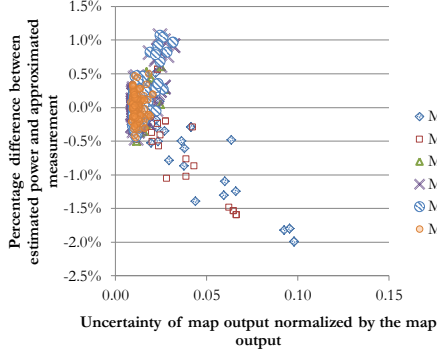

Figure 1. Change of accuracy of maps with output uncertainty in different maps.
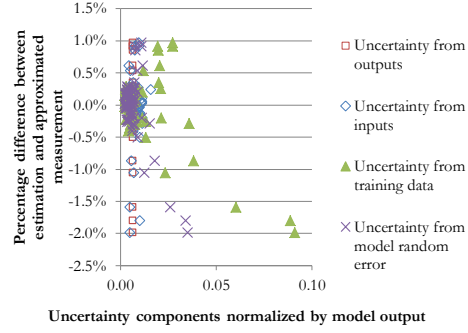
Figure 2. Change of accuracy of maps with different uncertainty components.

Figure 1 shows that inaccurate map outputs are associated with higher uncertainty and the uncertainty calculation method is a good indicator of the accuracy of the map output. Figure 2 shows that the high uncertainty of the inaccurate data points in Figure 1 are results of high uncertainty from model random error and training data.

The increase of uncertainty from model random error with inaccuracy can be explained by the leverage term $\vec{x}^T(\mathbf{X}_{train}^T\mathbf{X}_{train})^{-1}\vec{x}$ in Eqn. (8). When the inputs to the map become more different from the training data points o fthe map, the applicability and the accuracy of the map decrease. This is associated with an increase of leverage and the uncertainty from model random error. Hence the uncertainty of model random error increases as the applicability and the accuracy of the map decreases.

To explain the increase of uncertainty from training data with a decrease of map accuracy in Figure 2, the squared terms in Eqn. (7) are labeled as uncertainty from training data per measurement, and their values from two map outputs of Map 1 are plotted as histograms as shown in Figures 3 and 4.

Figure 3 shows a histogram that is more left-skewed than Figure 4. This is because Figure 3 is obtained from a data point inside the training data range of Map 1 in Figure (FIGURE NUMBER TO BE ADDED FROM TEST CASE), and the map only needs information from a few data points around $T_{evap} = -1.1°C$ and $T_{cond} = 60.0°C$ to estimate its map output. Hence the estimation does not depend on most of the data points, and their training data uncertainties do not propagate to the map output as shown by Figure 3.

However, the map output at the condition in Figure 4 is conducted with the condition at the lower left handed corner in Figure (FIGURE NUMBER HERE), and is far away from most of the training data points of Map 1. To estimate the map output at this condition, the estimation is done by extrapolation and is significantly by multiple data points in the training data. If any of these significant training data points change, the estimation result at this condition will be changed significantly. Hence much more training data points propagate their uncertainty to the map output in Figure 4 than that in Figure 3. This also explains that the uncertainty of training data is a significant component in the extrapolation uncertainty of the map output and it increases as the map applicability and accuracy decrease.
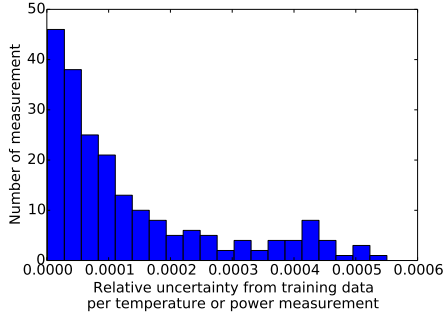
Figure 3. Histogram of uncertainty terms in Eqn. (7) for $T_{evap} = -1.1°C$ and $T_{cond} = 60.0°C$.

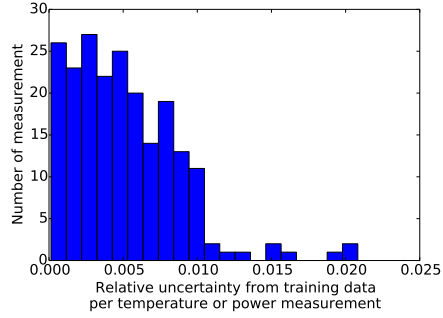**Figure 4.** Histogram of uncertainty terms in Eqn. (7) for $T_{evap} = -28.9°C$ and $T_{cond} = 26.7°C$.

1.7. Effect of number of training data points on map output

To understand how the number of training data points affect the map output, another map (Map 7) was created with the same range of training data as Map 1 and 22 data points only, and its data points are listed in the appendix. The accuracy of Map 7 is compared with that of Map 1 in Table 1.

Table 1. Coefficients of variation of Maps 1 and 7.

Map	Coefficient of variation based on training data poinst only (cov_{train})	Coefficient of variation based on all available data points (cov_{all})
Map 1	0.18%	0.31%
Map 7	0.17%	1.15%

Table 1 shows that the accuracy of Map 7 is similar to that of Map 1 at the training data points, but the analysis with all available data points shows that Map 7 is less accurate than that of Map 1. This shows that despite similar range of training data, fewer training data points reduces map accuracy. But to know if the reduction of accuracy happens at extrapolation, the accuracy of the map outputs is compared with the uncertainty from model random error and the uncertainty from training data as shown in Figures 5 and 6.

Figures 5 and 6 show that both the accuracy of Maps 1 and 7 decreases as their uncertainty from model random error and the uncertainty from training data increase, and the reduction of Map 1 accuracy in both figures are more rapid than Map 7. Since the uncertainty components in both fiures are indicators of the map applicability, Map 7, despite having the same range of training data as Map 1, is less applicable and accurate as Map 1 because it is constructed with fewer training data points. This shows that map accuracy and applicability is not only affected by the range of training data but also the number of trianing data points.

2. Conclusions

To conclude, this paper demonstrates how uncertainty of 10-coefficient compressor maps can be calculated based on the uncertainty due to inputs, the uncertainty due to model random error, the uncertainty due to training data and the uncertainty from outputs. It shows that a method
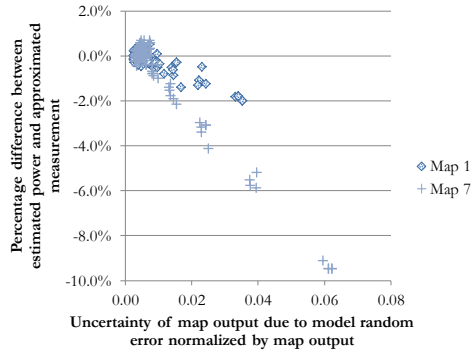
Figure 5. Change of accuracy of maps with uncertainty from model random error in Maps 1 and 7.
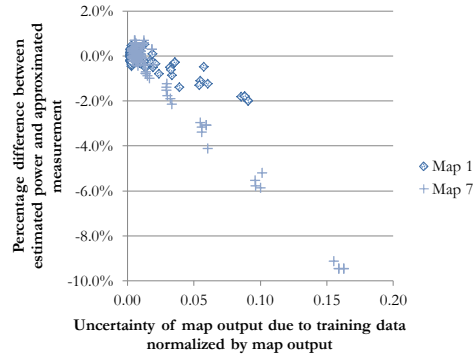

Figure 6. Change of accuracy of maps with uncertainty from training data in Maps 1 and 7.

to assess the acuracy and applicability of the map output without measured data outside the training data. It also shows that the range of training data and the number of training data points affect the accuracy and applicability of the map.

References
[1] N T B and E K C 1995 *NIST Technical Note* (Gaithersburg: National Institute of Standards and Technology)
[2] Montgomery D C 2005 *Design and Analysis of Experiments* 6th ed (Jon Wiley & Sons)
[3] Graybill F A and Iyer H K 1994 *Regression Analysis: Concepts and Applications* (Duxbury Pr)
[4] Atkinson A C 1987 *Transformation and Regression* (Oxford Statistical Science Series)