An Evaluation of Classical Steady-State Off-Line Linear Parameter Estimation Methods Applied to Chiller Performance Data

T. Agami Reddy, Ph.D., P.E.
Member ASHRAE

Klaus K. Andersen, Ph.D.

The objective of this paper is to evaluate different inverse methods with application to off-line model parameter estimation using data from a field-operated chiller. In HVAC&R data analysis, there is sometimes a need to evaluate and use estimation techniques that are more subtle than the ordinary least square (OLS) method. One example is in fault detection and diagnosis of HVAC&R equipment and systems using performance data obtained from field monitoring. By identifying a better performance model, the fault detection process is more likely to be refined and accurate. In this paper a number of exploratory, diagnostic, and classical estimation methods are reviewed to determine the circumstances in which they are likely to be superior to the OLS method. These methods are then evaluated using monitored data from a field-operated chiller. This study provides a reference on parameter estimation methods for the HVAC&R community.

INTRODUCTION

There are two broad approaches to modeling equipment and systems in engineering: *forward modeling*, in which predicting the system performance under prespecified driving conditions is the desired objective, and *inverse modeling*, in which identifying the system structure is the intended purpose. The former is primarily used for design, whereas the latter is used for control and prediction of the system response to external stimuli. Further, inverse modeling essentially consists of two distinct aspects: *model formulation* (or system identification) and *parameter estimation*. In the HVAC&R field, a number of physical models have been developed during the last decade (Bourdouxhe et al. 1999). The adoption of proper parameter estimation techniques has been largely ignored and its statistical foundations underappreciated. There is a need for HVAC&R professionals to better understand and recognize the scope and applicability of the various regression methods that have been proposed in literature, and how the estimation of the parameters in the framework of the associated identification approach has a direct impact in terms of the predictive and diagnostic ability of the identified models.

One of the several application areas of inverse methods is for fault detection and diagnosis (FDD) for HVAC&R equipment in general and for chillers in particular (Gordon and Ng 2000; Grimmelius et al. 1995; Stylianou and Nikanpour 1996; Rossi and Braun 1997). The physical understanding of the system is important to determine whether the system is working properly and to diagnose faults. Because fault detection is most often done by comparing a model under fault-free condition with monitored data, inverse methods for building models and analyzing data need to be understood, whether the FDD model is derived from physical reasoning, based purely on an empirical model, or both.

This paper presents and evaluates pertinent methods that can be applied for the parameter estimation and used to calculate prediction errors and confidence limits. The difference between

T. Agami Reddy is an associate professor with the Department of Civil and Architectural Engineering at Drexel University, and **Klaus K. Andersen** is a research associate in the Department of Mathematical Modeling at the Technical University of Denmark.

off-line (or batch or block) and online (or adaptive or real-time) identifications lies in whether the parameters are incrementally adjusted as newer data come in without having to reuse the old data. Each of these can be further divided into *recursive* methods, where all data are given the same weight in readjusting the parameter estimates, and *forgetting* methods, where more weight is given to newer data.

Many methods of estimation and diagnostic techniques are available (Cook and Weisburg 1982). The analyst needs to be generally aware of the various available parameter estimation techniques. Performance data from HVAC&R equipment and systems are usually ill-conditioned. In such cases, parameter estimation based on multivariate models requires special care.

The objective of this paper is to present and evaluate the important parameter estimation techniques useful to HVAC&R professionals. A number of exploratory, diagnostic, and estimation methods are reviewed, and the circumstances under which they are superior to the classical ordinary least squares method are discussed. Monitored data from a field-operated chiller are used to illustrate the use of the methods applied to linear steady-state models. The focus of this paper is on multivariate single-equation steady-state linear models.

BASIC CONCEPTS

The historic approach has been to use black-box (empirical or statistical) models to characterize the performance of equipment (Stoecker 1989). The most basic classification of a model is to determine whether it is *empirical* (black-box) or *physical* (gray-box). There are different degrees of "grayness," starting with completely mechanistic on one end, to completely black-box at the other end, depending on the amount of physical knowledge used in the model-building process and the types of assumptions and simplifications made.

Formulation of the model structure for gray-box models may initially require a certain amount of effort and insight. However, once the model structure has been established, identifying the model parameters is a relatively simple process. On the other hand, though proper black-box modeling requires that several model structures be evaluated before deciding on the best model, the general approach is to assume high-order polynomial models for the model structure, and estimate the corresponding parameters. Black-box models often require less time and effort to identify from monitored data and have had widespread appeal, although they do not contain physical information. Judgment and the status of available knowledge dictate which model approach to adopt for the specific case.

A second manner of classifying models is by their *order*. Zeroth models (also called cross-section models) are algebraic equations representative of the steady-state performance of the equipment or process. Higher-order models incorporate either differential equations or time series (either as a response factor series or as a transfer function representation).

Regression models may be classified by their *structure*. Single-variate or multivariate model classification depends on whether there is one or several regressor variables, whereas single-equation or multi-equation classification depends on whether there is one or several response variables. Models also are classified as linear or nonlinear; the latter must be further divided between inherently nonlinear models and those that can be made linear by suitable variable transformation.

Data may be obtained from measurements made in a laboratory setting or in the field. Laboratory equipment can be controlled so that a proper design of experiment procedure can be adopted. Laboratory data can cover the entire range of equipment operation, and the potential for errors, both bias and random, is generally much less because of the use of higher-quality sensors and proper care in placement and calibration. Performance data collected from the field are inherently less accurate.

It is advantageous when dealing with multiple regression of data to use matrix algebra because of the compactness and ease of manipulation it offers without loss in clarity (Draper and Smith 1981). For a data set of n readings that include k regressor variables x and one response variable y, the model can be expressed as follows:

$$\boldsymbol{Y}_{(n,1)} = \boldsymbol{X}_{(n,p)}\boldsymbol{\beta}_{(p,1)} + \boldsymbol{\varepsilon}_{(n,1)} \tag{1}$$

where $\boldsymbol{\beta}$ is the vector of parameters, $\boldsymbol{\varepsilon}$ is the vector of errors, p is the number of parameters in the model = $k + 1$, $\boldsymbol{Y}' = (y_1 y_2 \ldots y_n)$, $\boldsymbol{\beta}' = (\beta_0 \beta_1 \ldots \beta_k)$, $\boldsymbol{\varepsilon}' = (\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n)$, and $\boldsymbol{X}$ is the matrix of regressor variables.

Properties of Estimators

For proper parameter estimation, it is important to distinguish between an estimate, which is a specific number, and an estimator, which is a random variable. Because the search for estimators is the crux of the parameter estimation process, some fundamental notions regarding desirable properties of an estimator are presented below (Pindyck and Rubenfeld 1981). A desirable property is the *lack of bias*, in which the distribution of the estimator has the parameter value as its mean value. *Efficiency* is a measure of the level of dispersion, and lack of bias does not indicate the dispersion. The greater the efficiency associated with an estimation process, the stronger the statistical or inferential statements one can make about the estimated parameters.

In many circumstances there is a tradeoff between bias and variance of estimators. When the goal of a model is to maximize the precision of predictions, an estimator with very low variance and some bias may be more desirable than an unbiased estimator with high variance. One criterion that is useful in this regard is the goal of minimizing the *total mean square error* (TMSE), defined as

$$\text{TMSE}\ (b) = E(b - \beta)^2 = [\text{Bias}\ (b)]^2 + \text{var}(b) \tag{2}$$

When b is unbiased, the mean square error and variance of the estimator b are equal. Thus TMSE may be regarded as a generalization of the variance concept and leads to the generalized definition of the relative efficiency of two estimators, whether biased or unbiased.

An estimator is *consistent* if it converges to its true value as the sample size increases. However, to determine the best estimator, the efficiency is a more powerful criterion. Consistency is more important than lack of bias as an estimation criterion. A biased yet consistent estimator may not equal the true parameter on average, but will approximate the true parameter as the sample information grows larger.

Measurement and Model Errors

A crucial issue during parameter identification is the explicit recognition of the type of measurement inaccuracy present. This has a direct influence on the estimation method to be used. Measurement errors (ε) can manifest as additive errors [i.e., $\varepsilon_i \neq f(y_i, x_i)$], or as multiplicative errors [i.e., $\varepsilon_i = f(y_i, x_i)$], or both. Multiplicative errors occur when instrument measurement uncertainty is a function of the actual reading and not of the maximum reading, as is usually the case.

During estimation, it is important to identify which of the following statistical assumptions apply (Beck and Arnold 1977):

- $E(\varepsilon_i) = 0$ (i.e., zero mean errors without bias)
- var $(y_i) = \sigma^2$ {i.e., constant variance [note that $E(\varepsilon_i^2) = \sigma^2$ if previous assumption is satisfied]}
- $E(\varepsilon_i \cdot \varepsilon_j) = 0$ if first assumption is satisfied (i.e., no serial autocorrelation)
- var $(x_i) = 0$ (i.e., no measurement error in the regressor variables)
- ε_i has a normal probability distribution

- Errors are additive
- No prior information on model parameters

Another important source of error during model identification is *model misspecification error*, which is independent of measurement errors (Reddy et al. 1998). This arises when the functional form of the model is not appropriate. This can occur due to (1) inclusion of irrelevant variables, (2) exclusion of an important variable, (3) assumption of a linear model when the relationship is not, or (4) incorrect model order. Katipamula et al. (1994) illustrate the concept of model misspecification as applied to building thermal cooling load modeling. Other than (1), the causes can be detected by residual analysis, though the exact cause can only be identified through several attempts at model reformulations. Even if the physics of the phenomenon or of the system is well understood and can be cast in mathematical terms, experimental or parameter identifiability constraints may require that a simplified or macroscopic model be used rather than the detailed model, causing model misspecification. This is likely to cause both bias and random noise in the parameter estimation process.

Off-Line Estimation Techniques

The interrelation and applicability of the three types of estimation methods discussed in this paper are shown in the flow chart of Figure 1. Estimation by least squares (also known as the classical estimation method) is the most commonly used approach. Because it is based on minimizing the squared deviation, it is also referred to as the method of moments estimation (MME). Maximum likelihood estimation (MLE) is often superior to MME, especially when the probability distribution of the errors is not normal. MLE allows generation of estimators of unknown parameters that are generally more efficient and consistent than all other unbiased estimates. If the errors are normally distributed, MLE and MME give identical results, and in that regard MME can be viewed as a special (but important) case of MLE.

The Bayesian estimation is more subtle in that it allows prior information about the value of the estimate to be included with the experimental or numerical results. For larger samples, the Bayesian estimates and the classical estimates are practically the same. When samples are small, Bayesian estimation is particularly useful.

The classical ordinary least squares approach (OLS method) is the best method to use when there are no errors in $\boldsymbol{X}$, when there is no collinearity between the regressors, when model residuals have constant variance and are not patterned, and when the errors are normally distributed. In this case, the parameter set β is determined such that the sum of error squares function is minimized. This leads to the system of normal equations, provided matrix $\boldsymbol{X}$ is not singular (Draper and Smith 1981), of

$$\boldsymbol{X'Xb} = \boldsymbol{X'Y} \qquad \text{or} \qquad \boldsymbol{b}_{OLS} = (\boldsymbol{X'X})^{-1}\boldsymbol{X'Y} \tag{3}$$

where $\boldsymbol{b}$ is the least square estimator matrix of β.

There are no statistical assumptions regarding the distribution of errors in OLS parameter estimation. If the first four assumptions related to the measurement errors are satisfied, then the Gauss-Markov assumptions are said to be satisfied. Under these conditions, **b** is an unbiased estimator of β with the variance-covariance matrix var($\boldsymbol{b}$) given by

$$\text{var}(\boldsymbol{b})_{OLS} = \sigma^2(\boldsymbol{X'X})^{-1} \tag{4}$$

where σ^2 = mean square error (MSE) of the model error terms.

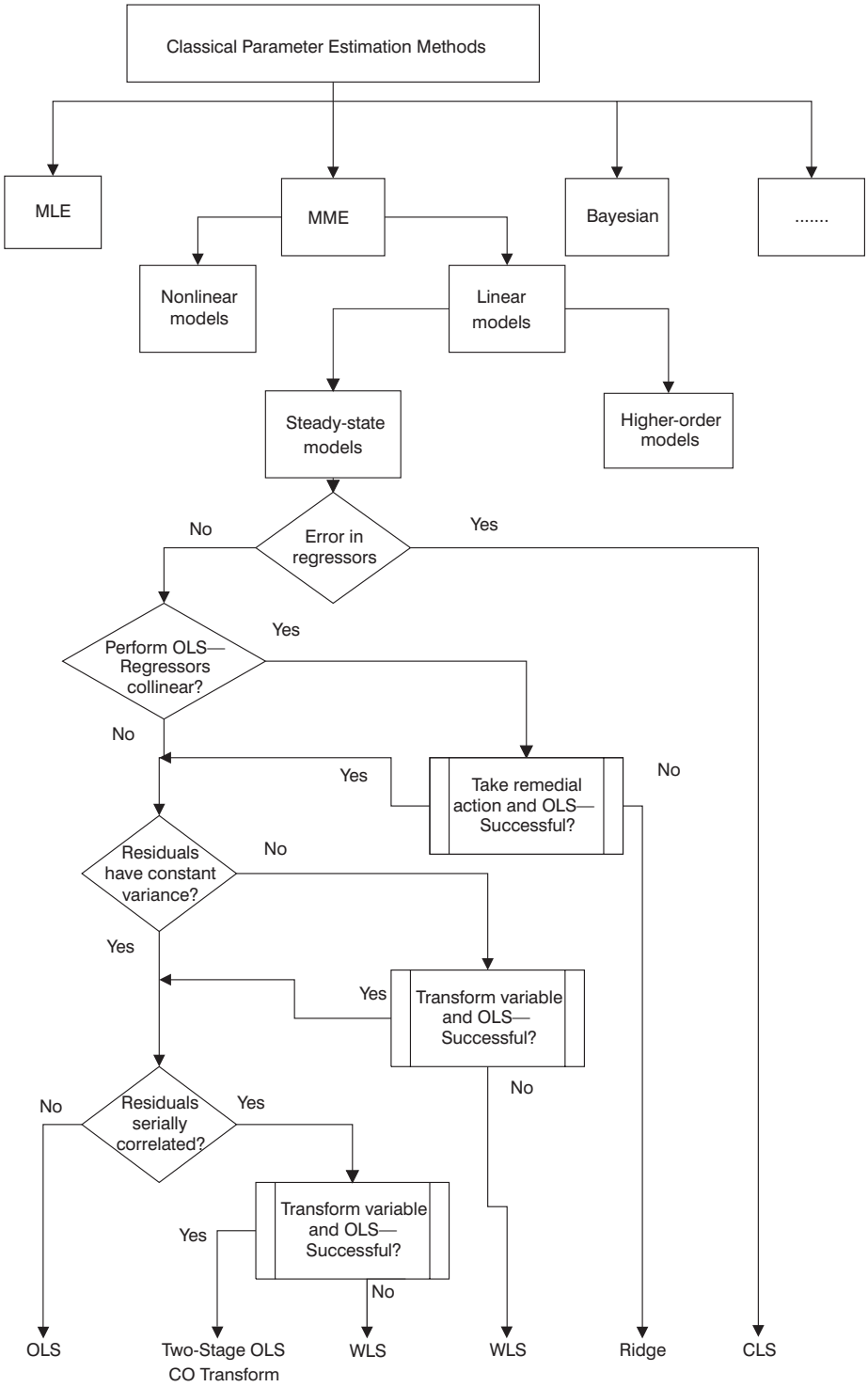


Figure 1. Guide to Selecting Appropriate Classical Parameter Estimation Methods for Linear Steady-State Models

The Gauss-Markov theorem states that of all the unbiased estimators of β, the OLS estimator has minimum variance. An unbiased estimator of σ^2 is s^2, where

$$s^2 = \frac{\varepsilon'\varepsilon}{n-p} = \frac{y'y - b'x'y}{n-p} = \frac{\text{SSE}}{n-p} \tag{5}$$

where SSE = sum of square errors.

For predictions within the range of variation of the original data, the mean and individual response values are normally distributed with the equations for variance given by the following:

For the mean response at a specific set of $\boldsymbol{x_0}$ values (called the confidence level)

$$\text{var}(\langle \hat{y}_0 \rangle)_{OLS} = s^2[\boldsymbol{X}_0 (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'_0] = \boldsymbol{X}'_0 \boldsymbol{b} \tag{6}$$

For an individual prediction (called the prediction level)

$$\text{var}(\hat{y}_0)_{OLS} = s^2[\mathbf{1} + \boldsymbol{X}_0 (\boldsymbol{X}'\boldsymbol{X})^{-1} \boldsymbol{X}'_0] \tag{7}$$

where **1** is a column vector of unity.

Model Fitting

Problems associated with model underfitting and overfitting are usually the result of a failure to identify the nonrandom pattern in time series data. Underfitting does not capture enough of the variation in the response variable for the corresponding set of regressor variables to provide an explanation. Overfitting implies including randomness in the model (i.e., attempting to fit the noise in the data). An extreme example is an attempt to fit a model with six parameters to six data points that have some inherent experimental error. The model has zero degrees of freedom and the set of six equations can be solved without error. This is clearly unphysical because the model parameters have also "explained" the random noise in the observations in a deterministic manner.

Underfitting and overfitting can be detected by performing certain statistical tests on the residuals where the residual is defined as $e_i = y_i - \hat{y}_i$. The most commonly used test involving model residuals is the Durbin-Watson (DW) statistic (Draper and Smith 1981), defined by

$$\text{DW} = \sum_{i=1}^{n} (e_i - e_{i-1})^2 / \sum_{i=1}^{n} e_i^2 \tag{8}$$

where e_i is the residual at time interval i.

The range of DW is 0 to 4. When there is no serial or autocorrelation present, the expected value of DW is 2. If the model underfits, DW would be less than 2; it would be greater than 2 for an overfitted model. Tables are available for appropriate significance tests with different numbers of regressor variables and numbers of data points. The DW statistic is only sensitive to correlated errors in adjacent observations (i.e., when only first-order autocorrelation is present). For example, if the time series has seasonal patterns, then higher autocorrelations may be present that the DW statistic will be unable to detect.

When a black-box model containing several regressors is used, stepwise regression improves the robustness of the model by reducing the number of regressors in the model and the adverse effects of multicollinearity between the remaining regressors. There are two variants of stepwise regression (Chatterjee and Price 1991):

- **Forward selection,** which starts with an equation containing the constant term, and adds into the equation the regression variable that has the highest correlation with *y*. If, based on a *t*-test, the estimated parameter is significantly different from zero, it is retained and the search for a second variable is performed. This process is continued until no other variable is statistically significant.
- **Backward elimination,** which starts with the full equation and successively drops one variable at a time based on their contribution to the reduction of error sum of squares (i.e., the one having the smallest *t*-ratio).

Many software packages use the *F*-test indicative of the overall model instead of the *t*-test on individual parameters to perform the stepwise regression. A value of $F = 4$ is often chosen, representative of the 99% statistical significance.

Correlated Regressors

Data are said to be ill-conditioned when the regressor variables of a model are correlated with each other, which results in $\boldsymbol{C} = \boldsymbol{X}'\boldsymbol{X}$ matrix becoming close to singular. This may affect the estimates of the model coefficients (unstable with large variance) and the general applicability of the estimated model. There are three commonly used diagnostic measures dependent on the matrix $\boldsymbol{C}$ that can be used to evaluate the magnitude of ill-conditioning (Belsley et al. 1980):

- The **correlation matrix,** where the diagonal elements of matrix $\boldsymbol{C}$ are scaled by unity. This allows investigation of the correlation between pairs of regressors in a qualitative manner, but may be of limited use in assessing the magnitude of overall multicollinearity of the regressor set.
- **Variance inflation factors,** which provide a better quantitative measure of the extent to which the variance of the parameter is increased or inflated. The diagonal elements of the **C** matrix are

$$C_{jj} = \frac{1}{(1 - R_j^2)} \qquad j = 1, 2, \ldots, k \tag{9}$$

  where R_j^2 is the coefficient of multiple determination resulting from regressing x_j on the other $k - 1$ variables. The stronger the linear dependency of x_j on the remaining regressors, the larger the value of R_j^2. The variance of b_j is inflated by the quantity $(1 - R_j^2)$. Thus, the variance inflation factors (VIF) $(b_j) = C_{jj}$. The VIF allow the evaluation of the joint relationship among a specified regressor and all other regressors. Many texts suggest that VIF > 10 indicate strong ill-conditioning, while $5 <$ VIF < 10 indicates a moderate problem.
- **Condition number,** which allows evaluation of the joint relationship among regressors. This measure, adopted from numerical analysis, provides useful summary information. The condition number for the matrix $\boldsymbol{C}$ is calculated as the square root of the ratio of the largest singular value or eigenvalue to the smallest. Evidence of collinearity is suggested for condition numbers >15, and corrective action is warranted when the value exceeds 30 or so (Chatterjee and Price 1991).

If OLS estimation is used with ill-conditioned data, the parameters, though unbiased, are not efficient in the sense they no longer have minimum variance. More importantly, OLS formulas understate both the standard errors in $\boldsymbol{b}$ and the model's prediction uncertainty.

With ill-conditioned data, either more data under different operating conditions should be collected or additional variables should be included in the model. Methods such as *ridge regression*, *singular value decomposition* (SVD) or *principal component analysis* (PCA) (Draper and Smith 1981) can also be used. PCA creates a set of artificial variables from the original regressor

set via a linear transform in such a way that the artificial vectors are orthogonal to each other. Subsequently, retaining those vectors that explain the largest proportion of the variation in the standardized data set allows for stability in the parameter estimation and robust model identification. If the principal components can be interpreted physically, then it can be an extremely valuable tool. Though it has been shown to be useful in fields such as the social sciences as a way of finding effective combinations of variables, PCA has had limited success in the physical and engineering sciences, and even in the HVAC&R area (Reddy and Claridge 1994).

Ridge regression for ill-conditioned data (Chatterjee and Price 1991) allows more stable estimates to be obtained. There are several alternative ways of defining and computing ridge estimates. The ridge trace is a graphical approach that allows exploratory analysis to be performed simultaneously. Because ($\boldsymbol{X'X}$) is close to singular, the approach involves introducing a known amount of "noise" via a variable k, leading to it becoming less sensitive to multicollinearity. With this approach, the parameter vector given by Equation (3) for OLS becomes:

$$\boldsymbol{b}_{Ridge} = (\boldsymbol{X'X} + k.\boldsymbol{I})^{-1}\boldsymbol{X'Y} \tag{10}$$

Values of k from 0 (the OLS case) to 1.0 are tested to determine the optimum value that yields the least model MSE (usually k is in the range of 0 to 0.2). The ridge estimators are biased but tend to be stable and have smaller variance than OLS estimators. Forecasts of the response variable tend to be more accurate and the uncertainty bands more realistic.

Parameter variance:
$$\text{var}(\boldsymbol{b})_{Ridge} = s^2(\boldsymbol{X'X} + k.\boldsymbol{I})^{-1}\boldsymbol{X'X}(\boldsymbol{X'X} + \text{k}.\boldsymbol{I})^{-1} \tag{11}$$

Prediction bands:
$$\text{var}(\hat{y}_0)_{Ridge} = s^2\{\mathbf{1} + X_0'\,[(\boldsymbol{X'X} + k.\boldsymbol{I})^{-1}\boldsymbol{X'X}(\boldsymbol{X'X} + k.\boldsymbol{I})]^{-1}X_0\} \tag{12}$$

Ridge regression should be performed with standardized variables (i.e., the individual observations subtracted by the mean and divided by the standard deviation) to remove large differences in the numerical values of the different regressors.

Residual Analysis

Once a preliminary model has been identified, the model residuals can be analyzed (either empirically/visually or by using formal statistical decision rules) in an effort to improve the model. Estimated regression parameters are usually very sensitive to outliers, and the residual behavior may invalidate the use of the OLS-based equations for parameter variance and model prediction uncertainty bands. Several books treat this aspect in detail (Belsley et al. 1980, Chatterjee and Price 1991).

Model residual analysis involves two aspects:

- **Detection of ill-conditioned model residual behavior.** The residuals can exhibit nonconstant variance and/or serial correlation, both of which require weighted least squares estimation procedures as described later. Though nonconstant variance is easy to detect visually, the cause is difficult to identify. Some common causes are model misspecification or instrument measurement errors being multiplicative rather than additive. Serial correlation (or autocorrelation) is often present in time series data collected from HVAC&R systems. Autocorrelation is present if residuals show a trend or a pattern of clusters above or below the zero value that can be discerned visually. Correlations can suggest either that additional variables have been left out of the model (model misspecification error), or that the data are autocorrelated.
- **Identification of leverage and influence data points.** The objective is to systematically identify subsets of data that have an unusual or disproportionate influence on the estimated

model, and to identify the model parameters that are most affected. A data point is said to be influential if its deletion, singly or in combination with a relatively few others, causes statistically significant changes in the fitted model coefficients. The leverage of a point quantifies the extent to which that point is *isolated* in the x-space. A matrix (called the hat matrix) is defined as

$$\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}' = p_{ij} \tag{13}$$

The diagonal element p_{ii} is defined as the leverage of the i^{th} data point. Because the diagonal elements have values between 0 and 1, their average value is equal to (p/n). Points with $p_{ii} > 3p/n$ are regarded as points with high leverage (the threshold is sometimes taken as $2p/n$). A point that has high leverage is necessarily influential. The R-Student residuals can be used to detect influence points. Large residuals are traditionally used to highlight suspect data points or data points unduly affecting the regression model. It is meaningful to explore a normalized or scaled value of the residual, namely the standardized residuals or R-Student residuals, where

$$R\text{-Student} = \frac{\varepsilon_i}{s.(1 - p_{ii})^{1/2}} \tag{14}$$

Points with $|R\text{-Student}| > 3$ can be said to be influence points that corresponds to a significance level of 99%. Sometimes a less conservative value of 2 is used, corresponding to the 95% significance level. Note that data points that satisfy both criteria (i.e., are influential and have high leverage) are the ones worthy of closer scrutiny.

Weighted least squares (WLS) can be used if a visual inspection or an analysis of model residuals suggests strong departure from OLS assumptions (nonconstant variance or serially correlated residuals). The standard errors in the parameter vector **b** and the model prediction uncertainty bands will be understated as in the previous case. The two important instances for data ordered in time are (1) when the models residuals or errors have nonuniform variance, and (2) when the errors are serially correlated though they have uniform variance.

Three possible remedies for such deficiencies, as shown in Figure 1, are to

- Transform the variables (log, Poisson, exponential, etc.) (Chatterjee and Price 1991)
- Introduce additional variables into the model and collect new data
- Use weighted least squares (*WLS*).

Weighted least squares can be adopted when some observations are deemed more important than others (i.e., outlier points are given less weight than those closer to the trend line). WLS involves weighting the model residuals differently; unlike OLS, under WLS, the parameter vector given by Equation (3) for OLS is given by

$$\boldsymbol{b}_{WLS} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y} \tag{15}$$

with an associated variance of $\text{var}(\boldsymbol{b})_{WLS} = s^2\,(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}$ and $\boldsymbol{\Sigma}$ is the weighting vector.

The individual prediction variance is

$$\text{var}(\hat{y}_0)_{WLS} = s^2[\mathbf{1} + \boldsymbol{X}_0'\,(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}_0] \tag{16}$$

In practice, specific information about Σ may not be available at the onset. A common approach is to use OLS (i.e., assume $\Sigma = \boldsymbol{I}$) and then deduce the form of $\boldsymbol{\Sigma}$ from the residuals of the regression analysis. More specifically, $\boldsymbol{\Sigma}$ is determined from the following:

$$\boldsymbol{W}'\boldsymbol{W} = \boldsymbol{\Sigma} \tag{17}$$

where the transformation matrix $\boldsymbol{W}$ is computed as described in several texts, say Pindyck and Rubinfeld (1981) and is the basis of the generalized least squares (GLS) approach.

Simpler treatments are

- The case of model residuals with nonconstant variance can be determined from a plot of the model residuals against the response variable. The range in the response variable is divided into as many regions as seem to have different variances and the standard deviation of the residuals for each of these regions is calculated. A specific value of σ_m, depending on the magnitude in the response variable to a particular observation $m(m = 1, 2, \ldots, n)$ is assigned to construct the $\boldsymbol{W}$ matrix.
- The case of serial correlation ρ in the model residuals is a result of pure autocorrelation or model misspecification. For the latter, the regressor to be included should be identified. For the former, either the $\boldsymbol{W}$ matrix approach presented above or a simpler approach involving variable transformation can be used. There are several techniques of doing so, but the two-stage Cochrane-Orcutt (CO) procedure (Chatterjee and Price 1991; Pindyck and Rubenfeld 1981) is widely used (see Figure 1).

  The CO approach can be used during parameter estimation of MLR models provided only one of the regressor variables is the cause of the pseudocorrelation. Also, a more sophisticated version of the CO method has been suggested by Hildreth and Lu (Chatterjee and Price 1991) that involves only one estimation process where the optimal value of ρ is determined along with the parameters. This, however, requires nonlinear estimation methods.

All the previous estimation methods are strictly applicable to the case when there is little or no error in the regressor variables. This is not often the case, so a method is needed that can find efficient estimates under such instances. If the variances of the measurement errors are known, the bias of the OLS estimator can be removed and a consistent estimator, called *corrected least squares* (CLS), can be derived (Fuller 1987; Draper and Smith 1981). The estimator is determined from

$$\boldsymbol{b}_{CLS} = (\boldsymbol{X}'\boldsymbol{X} - \boldsymbol{S}_{xx}^2)^{-1}(\boldsymbol{X}'\boldsymbol{Y} - \boldsymbol{S}_{xy}^2) \tag{18}$$

where $\boldsymbol{S}_{xx}^2$ is a $p \times p$ matrix with the covariance of the measurement errors and $\boldsymbol{S}_{xy}^2$ is a $p \times 1$ vector with the covariance between the regressor variables and the dependent variable.

A simple conceptual explanation is that Equation (18) performs on the estimator matrix an effect essentially the opposite of what ridge regression does. Whereas ridge regression "jiggles" or randomly enhances the dispersion in the numerical values of the X variables in order to reduce the adverse effect of multicollinearity on the estimated parameter bias, CLS tightens the variation in an attempt to reduce the effect of random error on the X variables. A more complete description of the method and its advantages over OLS is provided by Andersen and Reddy (2002) in the framework of performance data of a chiller.

Model Identification

When the model errors are not distributed normally, the model prediction uncertainty bands cannot be determined efficiently from the expression of variance. In practice, this occurs frequently, and so techniques that deal with this situation are needed. One approach is to evaluate the model errors, determine the likely shape qualitatively, and select a distribution (other than the normal one) that is likely to fit this pattern (Draper and Smith 1981). Another approach is to use MLE to develop a function that gives higher weight to some residuals (usually the smaller ones) and lower weights to others (usually the larger ones and outliers). Clearly there are several possible functions that can be used, and which one is optimal is usually difficult to ascertain. Blind use of an estimator, assuming it to be the most robust, is inadvisable (Draper and Smith 1981).

Efron and Tibshirani (1982) have argued that, given the available computing power, one should not be constrained by traditional parametric theory with its over-reliance on a small set of standard models for which theoretical solutions are available. This parallels the manner in which numerical methods have in large part replaced closed-forms solution techniques in almost all fields of engineering. Versatile numerical techniques overcome such problems as the lack of knowledge of the probability distribution of the errors and the determination of sampling distributions of such quantities as the median or of the interquartile range. There are several approximate techniques for variance estimation with no or partially known distribution functions. The two most widely used are the Monte Carlo techniques and the resampling or randomized techniques.

Monte Carlo techniques (or stochastic simulation) allow synthetic data to be generated from observations that have been corrupted with noise, usually with additive or multiplicative preselected magnitude (with or without bias) and probability distribution. Different sets of data sequences of preselected sample size can be generated, from which sampling distributions of the parameter estimates can be deduced and their sensitivity to the various assumptions evaluated. However, the synthetic generation has to be done so that serial or autocorrelation structure among the regressors is maintained.

Resampling techniques (of which the jackknife and bootstrap methods are widely used) involve drawing samples either with or without replacement from the observed data set, and performing analyses on them from which population related estimators and prediction variance can be determined (Efron and Tibshirani 1982). The jackknife method, introduced by Quenouille in 1949 and later extended by Tukey in 1958, is a technique of universal applicability that allows confidence intervals to be determined of an estimate calculated from a sample while reducing bias of the estimator.

There are several numerical schemes for implementing the jackknife scheme. One version is to divide the random sample of observations into groups of equal size, omit one group at a time, and determine what are called pseudo-estimates, the actual confidence intervals of the parameters. The bootstrap method (popularized by Efron in 1979) is similar but differs in that no groups are formed. Random samples are simply created by sampling with replacement from the observational data set (Davison and Hinkley 1997). Individual estimators deduced from such samples permit estimates and confidence intervals to be determined.

STEADY-STATE CHILLER MODELS AND THEIR APPLICATION

Inverse modeling basically consists of two distinct aspects: model formulation and parameter estimation. The model formulation can essentially be divided into gray-box and black-box models. The model structure is different under each formulation, which has direct bearing on the robustness of parameter estimation. We propose to study the effect on parameter variance and the related model prediction bands when different estimation methods are used in the framework of the different types of models of steady-state chiller performance. The first

model type is the gray-box approach, proposed by Gordon and Ng (2000). The model formulation is derived from the physics, and the unknown quantities are estimated using statistical methods. The second model is characterized as a pure black-box model, with the model structure based on empirical data and statistical methods only. The main difference between the two types of models is that the gray-box model and parameters may be interpreted physically, but the black-box model and parameters may not.

The universal thermodynamic model proposed by Gordon and Ng (2000) (GN) is a simple, universal model for chiller performance based on the first principles of thermodynamics and linearized heat losses. The model correlates the dependent chiller COP (defined as the ratio of chiller thermal cooling capacity to the electrical power consumed by the compressor) with independent parameters such as the fluid inlet temperature to the condenser, fluid temperature entering the evaporator, and the thermal cooling capacity of the evaporator. The GN model is a three-parameter model that takes the following form:

$$\left(\frac{1}{\mathrm{COP}}+1\right)\frac{T_{chi}}{T_{cdi}}-1 \;=\; a_1\frac{T_{chi}}{Q_{ch}}+a_2\frac{(T_{cdi}-T_{chi})}{T_{cdi}Q_{ch}}+a_3\frac{(1/\mathrm{COP}+1)Q_{ch}}{T_{cdi}} \tag{19}$$

where the temperatures are in absolute units.

The following terms are defined:

$$x_1=\frac{T_{chi}}{Q_{ch}},\; x_2=\frac{(T_{cdi}-T_{chi})}{T_{cdi}Q_{ch}},\; x_3=\frac{(1/\mathrm{COP}+1)Q_{ch}}{T_{cdi}},\;\text{and}\; y=\left(\frac{1}{\mathrm{COP}}+1\right)\frac{T_{chi}}{T_{cdi}}-1 \tag{20}$$

allowing Equation (19) to be rewritten as

$$y \;=\; b_0+b_1x_1+b_2x_2+b_3x_3 \tag{21}$$

where b_0 is the intercept term.

The reformulation provides a model that is linear in the four parameters. As discussed by Gordon and Ng (2000) and also by Andersen and Reddy (2002), the intercept term in this model must be zero if the model parameters are to be interpreted in terms of physical quantities. However, if the chiller model is to be used for predictive purposes only, then the model given by Equation (21) is more appropriate from a statistical point of view because of the presence of measurement errors.

Two broad categories of black-box models are *classical* and *artificial neural networks* (Haykin 1999). The black-box model is typically characterized as having little physical information incorporated in the model structure, and only describes an empirical relationship between input and output. Using the same variable as those for the physical model, a second-order linear polynomial model called a multivariate polynomial model (MP) is appropriate (DOE 1981):

$$\begin{aligned}\mathrm{COP} \;=\; &\alpha+\beta_1T_{cdi}+\beta_2T_{chi}+\beta_3Q_{ch}+\beta_4T_{cdi}^2+\beta_5T_{chi}^2\\ &+\beta_6Q_{ch}^2+\beta_7T_{cdi}T_{chi}+\beta_8T_{cdi}Q_{ch}+\beta_9T_{chi}Q_{ch}\end{aligned} \tag{22}$$

The above model has nine coefficients that need to be identified from monitored data. These coefficients, unlike the three parameters appearing in the GN model, have no physical meaning

and their magnitude cannot be interpreted in physical terms. One should only retain parameters in the model that are statistically significant, and this is determined by step-wise regression

The MP modeling approach has the advantage that it can be used in a routine manner. However, collinearity in regressors and ill-behaved residual behavior often justifies another empirical approach in which the variables are transformed so that a simpler (preferably linear) model is obtained that may overcome some of the problems of poor statistical estimation.

The monitoring of model prediction errors from either model allows fault detection when used in conjunction with control charts. The prediction errors are the deviation between the predicted and measured outputs. This deviation is then compared with the uncertainty of the model at the 3σ level. An estimate for the prediction error variance (or squared standard deviation) is given by Equation (5).

With an estimate of the variance of the prediction error, monitoring can then be done using a control chart. The width of the uncertainty bands may be determined from well-known formulas for prediction intervals (Box and Luceno 1997) or approximated by 3σ limits, so that when the system is in fault-free condition then approximately 99% of the errors are expected to fall inside the confidence limits. A fault is indicated when the prediction errors start to increase and an increasing fraction falls outside the confidence limits.

The model parameters themselves could also be monitored. A detection method based on changes in model parameters may be superior to one based on model prediction errors (Pau 1981). Furthermore, if a physical model such as the GN model is used, the parameters estimates may directly provide diagnostic ability. However, great care is required during the parameter estimation process because this advantage is lost if the OLS assumptions are violated and/or if the regressors have a certain amount of measurement error (Anderson and Reddy 2002). Fault detection based on model parameters also suffers from the fact that the number of parameters to be monitored equals those in the model. As the number of parameters increases, the monitoring becomes more difficult compared to evaluating a single measure such as the prediction errors. Furthermore, the determination of uncertainty bands becomes more complex as the number of parameters increases, especially if the parameter estimates are correlated.

CASE STUDY ANALYSIS

Hourly data from a centrifugal chiller located in Toronto, Canada, were available during an operating period from June until October 1997, and were used for the case study illustration. The analysis was performed using the following four variables: cooling capacity, compressor power, supply chilled water temperature, and cooling water temperature. The data set contained 810 observations. The data set was divided into a training set, which was used to compare how different model formulations fit the data (referred to as model training or parameter estimation), and a testing (or validating) data set, used to single out the most suitable model in terms of its predictive accuracy. The primary purposes of having training and testing data sets were to avoid overfitting and to obtain a more accurate indication of the extrapolation errors. The training set contained the first 550 data points and the testing set contained the second 260 data points. The training data set was selected such that the range of variation of the individual variables was larger than that of the same variables in the testing data set. The same types of cross-correlations among variables were present in both data sets.

Pertinent descriptive statistics for both data sets are given in Table 1 and time series plots of the four variables are shown in Figure 2. There is relatively little variation in the two temperature variables, while the load and power experience significant variations. The correlations and serial correlations among the variables are presented in Table 2. The correlations between (P, T_{chi}) and (Q_{ch}, T_{chi}) are somewhat different for the training and testing data sets, showing an increase from around 0.55 to about 0.9. For the training period, the correla-

Table 1. Descriptive Statistics for Chiller Data

	Training Data Set (550 Data Points)				Testing Data Set (260 Data Points)			
	P	Q_{ch}	T_{chi}	T_{cdi}	P	Q_{ch}	T_{chi}	T_{cdi}
Mean	222	1108	285	302	202	1011	288	302
Standard deviation	37.8	282	2.43	0.73	14.7	149	2.97	0.77
Maximum	154	517	282	299	163	630	283	297
Minimum	340	1771	292	304	230	1241	293	303

Table 2. Correlation Matrices and First-Order Serial Correlation Coefficients for Chiller Data

	Training Data Set (550 Data Points)				Training Data Set	Testing Data Set
	P	Q_{ch}	T_{chi}	T_{cdi}	One Lag	One Lag
P	—	0.98	**0.52**	0.57	0.937	0.911
Q_{ch}	0.99	—	0.62	0.59	0.936	0.933
T_{chi}	**0.86**	0.91	—	0.37	0.943	0.930
T_{cdi}	0.54	0.61	0.54	—	0.829	0.585
Testing Data Set (260 Data Points)					—	—

Table 3. Correlation Coefficient Matrix for Regressors in GN Model During Training

	Training Data Set (550 Data Points)				Training Data Set
	Y	X_1	X_2	X_3	One Lag
Y	1.0	0.93	0.82	–0.79	0.92
X_a		1.0	0.96	–0.95	0.93
X_2			1.0	–0.92	0.93
X_3				1.0	0.94

tion coefficient between (P, Q_{ch}) is 0.98 while the other correlations coefficients are less than 0.6—not very significant. A scatter plot of the thermal load against the COP is shown in Figure 3.

Except for T_{cdi}, the strengths of the first-order serial correlations are very close. This suggests that any serial correlations in the model residuals during model identification were probably due to pseudoeffects, provided the model fit was very good and the residuals were well-behaved. The serial correlation between training and testing data sets is close.

The descriptive statistics for the regressors in the physical model provided an immediate indication as to whether the data set was ill-conditioned. Because the magnitude of the three variables differed by several orders of magnitude, the ridge regression used standardized variables. The estimated correlation matrix for the regressors in the GN model is given in Table 3. There is evidence of strong multicollinearity between the regressors.

From the statistical analysis it was found that the physical model fit the field-monitored data well except at the high end (see Figure 4A). The Adj-R^2 is 99.1%, CV = 1.45%. An analysis of variance showed that there was no statistical evidence to reduce the model order. The model residuals have constant variance, as indicated by the studentized residual plots versus time and by regressor variable (Figure 4B and 4C). Most of the data are contained within bounds of 2. However, the DW statistic of 0.90 indicates that serial correlation is present. The first-order serial correlation of the model residuals is equal to 0.39, which is not very significant. Given the

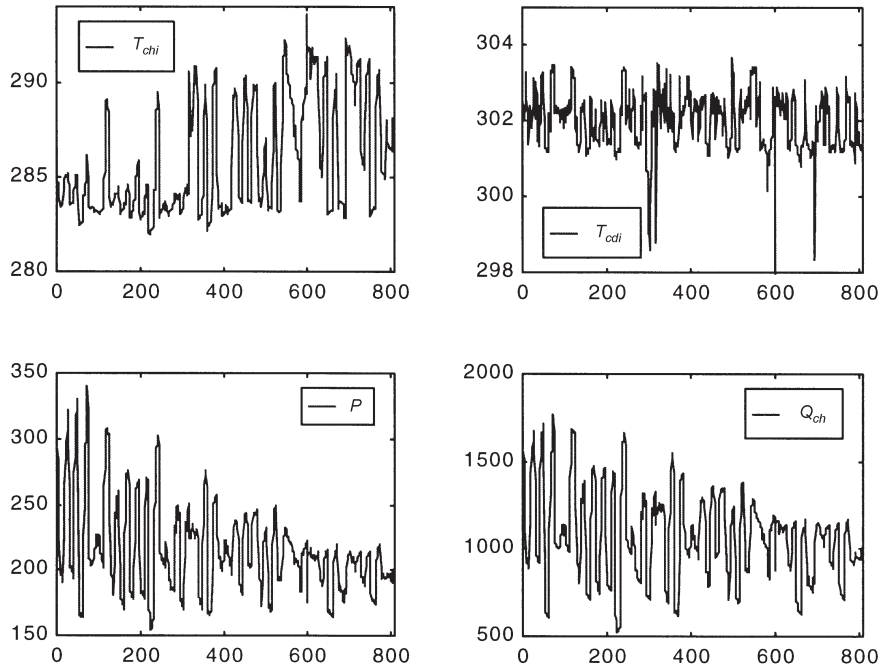


Figure 2. Time Series Data of Four Measured Variables of Centrifugal Chiller

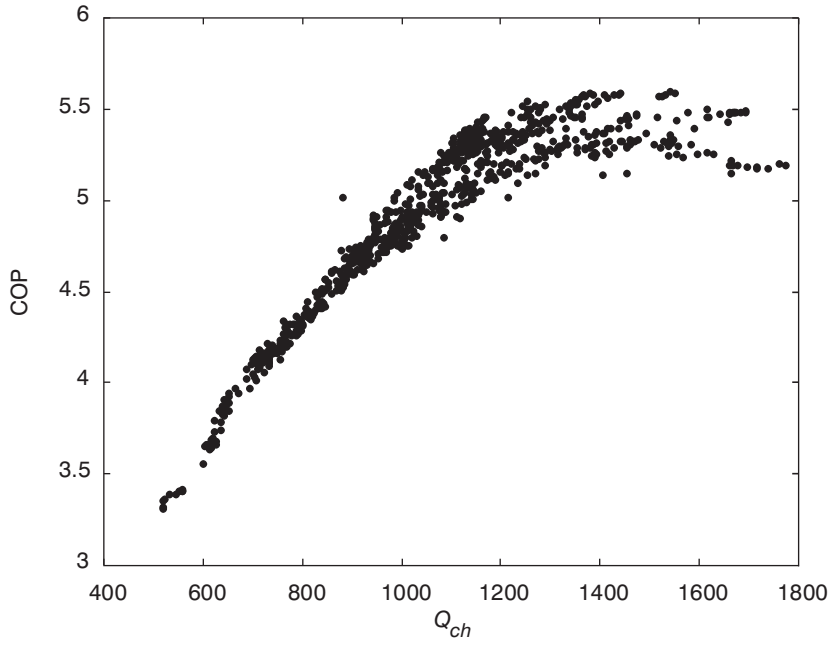


Figure 3. Scatter Plot of COP Against Thermal Cooling Load (in kW)

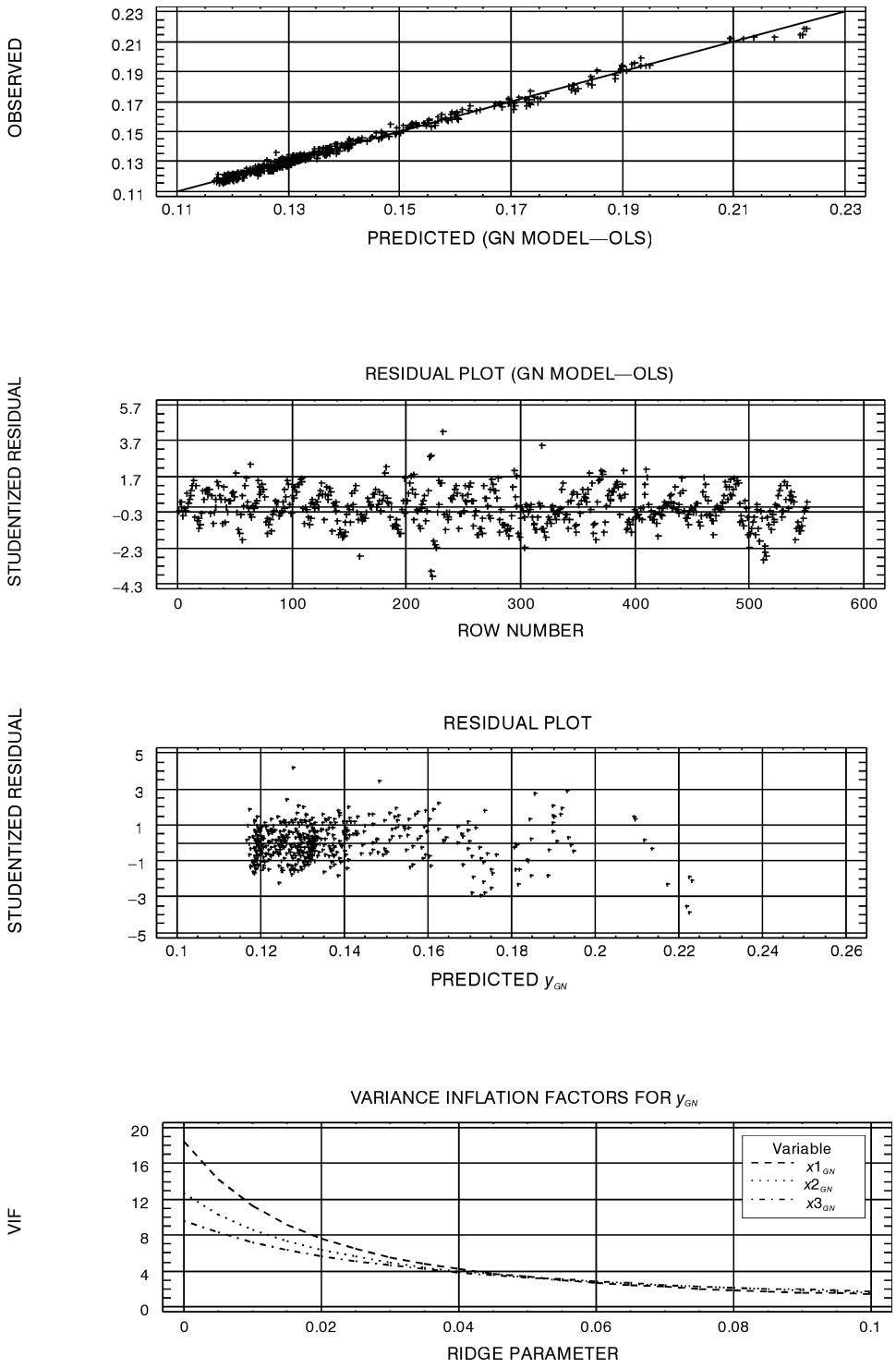


Figure 4. Analysis of Chiller Data Using GN (Gray-Box) Model

Table 4. Results of Ridge Regression Applied to GN Model

	Training Data Set					Testing Data Set	
	Adj-R^2	CV(%)	VIF(X_1)	VIF(X_2)	VIF(X_3)	CV(%)	NMBE(%)
$k = 0.0$*	99.1	1.45	18.43	12.69	9.61	1.63	–1.01
$k = 0.02$	89.8	3.02	7.67	6.39	5.71	2.86	1.91
$k = 0.04$	85.1	4.16	4.22	3.99	3.87	4.35	3.18
$k = 0.06$	82.3	4.87	2.69	2.78	2.82	5.20	3.86
$k = 0.08$	80.2	5.32	1.88	2.07	2.17	5.73	4.27

*Equivalent to OLS estimation

excellent model fit and well-behaved residuals, the first-order serial correlation present in the data is probably pseudo, and a CO transform would be appropriate in this case.

The parameters are strongly correlated and indicate ill-conditioned data according to the VIF values. The condition number of the matrix (square root of the ratio of the largest to the smallest eigenvalue) is $(2.88283/0.03554)^{1/2} = 9.0$. This does not suggest ill conditioning, because this value is lower than the threshold value of 30. Ridge analysis with the ridge factor k varied from 0 to 0.2 was performed. The ridge trace is shown in Figure 4D and the VIF factors are shown in Table 4.

The OLS estimates will be unbiased; ridge estimation will be biased but the estimation is likely to be more efficient. Table 4 shows that for the training set CV values increased as k was increased, but the VIF values of the parameters were decreased. Hence, inferences about whether ridge regression is better than OLS and which value of ridge parameter k is optimal would not be made. Using the rule-of-thumb lower bound of 5 for the VIF values indicated that $k = 0.02$ or 0.04 would be reasonable choices.

The CV and NMBE values of the models for the testing data set are also shown in Table 4. For OLS ($k = 0$), these values are 1.63% and 1.01%, indicating that the identified OLS model can provide extremely good predictions. Both indices increased as k was increased, indicating poorer predictive ability both in variability or precision and in bias. To identify a chiller model for predictions only, the analysis suggests that OLS is the best even though the data are ill-conditioned.

For the MP model, some or many of the variables may be statistically insignificant. A stepwise OLS regression was performed. Both forward selection and backward elimination techniques were evaluated using the F-ratio of 4 as the cutoff (representative of 99% statistical significance). Whereas the backward elimination retained seven terms (excluding the constant), forward selection retained only three. The Adj-R^2 and RMSE statistics were almost identical, so the forward selection model was retained for parsimony.

The final model contained the three following variables: [Q_{ch}^2, $T_{cdi}Q_{chi}$, $T_{chi}Q_{ch}$]. As shown in Figure 5, the fit is again excellent. The Adj-R^2 is 99.2% and CV is 0.95%, which are slightly better than those of the GN model. An analysis of variance also showed that there was no statistical evidence to reduce the model order. The model residuals have constant variance, as indicated by the Studentized residual plots versus time and by regressor variable (Figure 5B and 5C). Most data are contained within bounds of 2. However, the DW statistic of 0.92 indicates that serial correlation is present. Given the excellent model fit and well-behaved residuals, the first-order serial correlation present in the data is probably pseudo. The parameters are very strongly correlated (the correlation coefficients for all three variables are 0.99), indicating ill-conditioned data. The condition number of the matrix = $(2.9858/0.00005177)^{1/2} = 75.9$ suggests larger ill-conditioning than in the GN model. The ridge trace is shown in Figure 5D and the VIF factors are shown in Table 5.

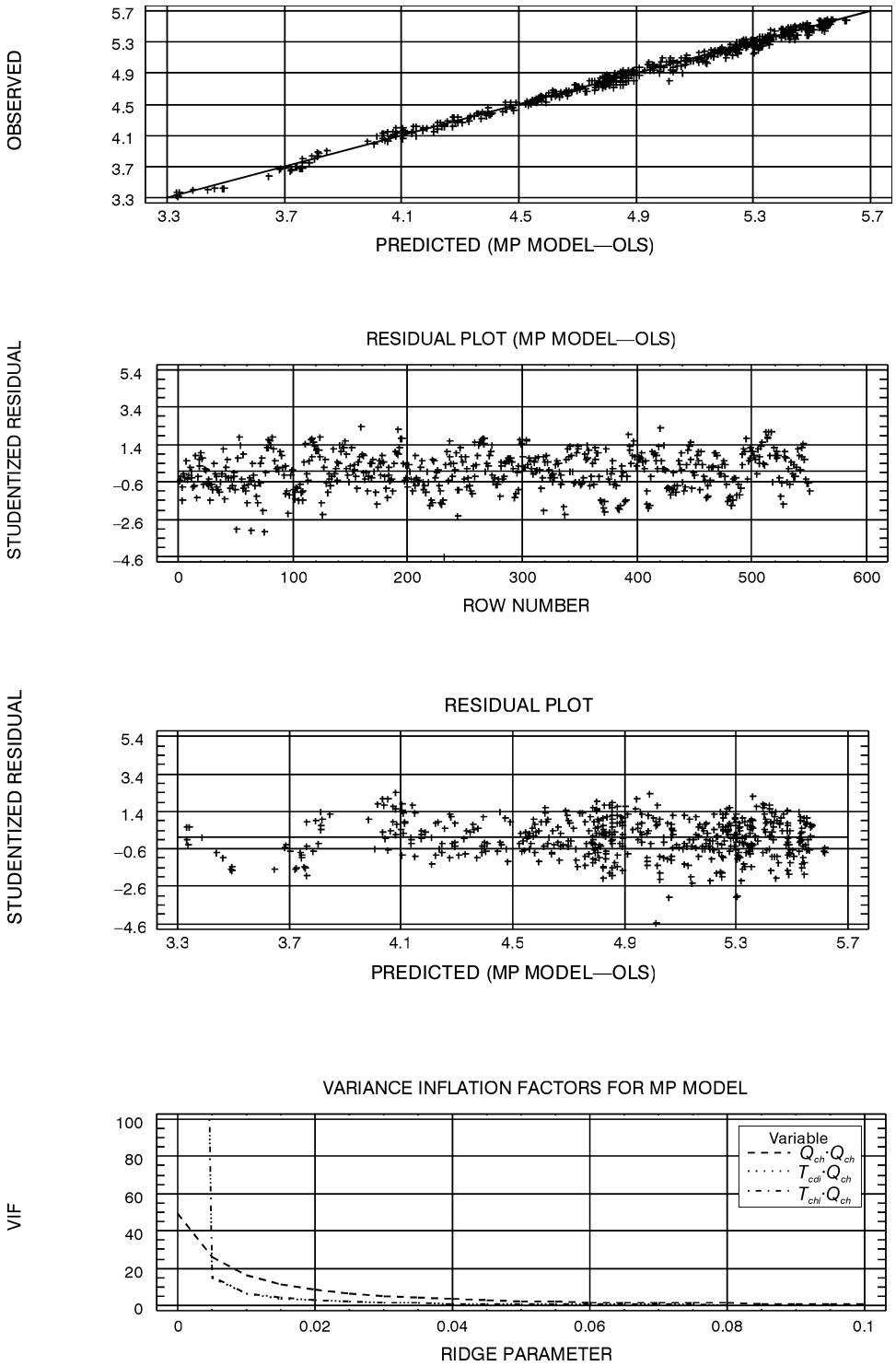


Figure 5. Analysis of Chiller Data Using MP (Black-Box) Model

Table 5. Results of Ridge Regression Applied to the MP Mode

	Training Data Set					Testing Data Set	
	Adj-R^2	CV(%)	VIF(X_1)	VIF(X_2)	VIF(X_3)	CV(%)	NMBE(%)
k = 0.0*	99.2	0.95	49.0	984.8	971.3	1.13	0.62
k = 0.005	93.1	1.81	26.3	15.0	15.2	2.86	2.38
k = 0.01	89.9	2.39	16.4	6.43	6.57	3.52	2.89
k = 0.015	87.8	2.80	11.2	3.89	4.00	3.95	3.21
k = 0.02	86.3	3.10	8.17	2.69	2.75	4.25	3.43

*Equivalent to OLS estimation

The CV values for the training set increased as k was increased from 0 to 0.025, as shown in Table 5 (along with the VIF of the regressors). The CV and NMBE values of the models for the testing data set are also shown in Table 5. For OLS (k = 0), these values are 1.13% and 0.62%, indicating that the identified OLS model can provide extremely good predictions. Further, these indices suggest the MP model to be slightly superior to the GN model in this regard.

The empirical variable transform (VT) model approach to predict COP is specific to the data set at hand. The scatter plot shown in Figure 3 indicates that the variance is not constant (e.g., there are larger variations in the thermal load when the COP is high). Exploratory data analysis identified the following black-box empirical model function exhibiting constant variance (as shown in Figure 6):

$$\frac{Q_{ch}}{\log(P)} = \beta_0 + \beta_1 Q_{ch} + \beta_2 Q_{ch}^2 \tag{23}$$

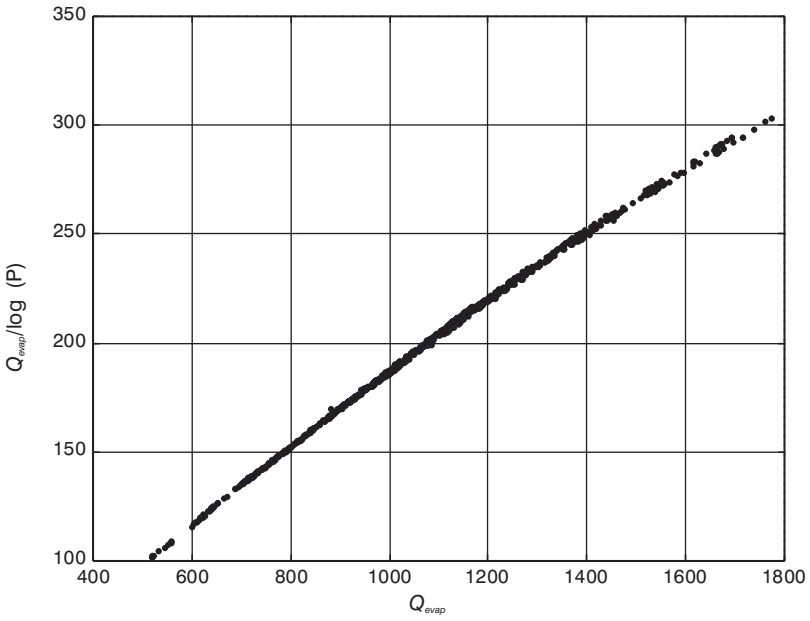


Figure 6. Scatter Plot of VT Model Showing New *y*-Variable Against Thermal Cooling Load (in kW)

The variability is almost constant, so a second-order polynomial was suggested. The model is linear in the parameters, and is likely to be more robust because the temperature variables have been dropped. This model, Equation (23), is based on empirical data from one site only and might not be adequate for other chillers/sites.

The OLS estimation used the 510 data points. The model fit was again excellent, with R^2 equal to 99.97 and CV = 2.5% (which is poorer than in the previous two models). A *t*-test indicated that there was no statistical evidence that the model order needed to be reduced. Figure 7A provides a visual indication of the fit to the model. The model residuals against time shown in Figure 7B also seem to be generally well-behaved, though not as good as in the two previous models. The DW statistic was determined to be 0.61, which indicated serial correlation in the model residuals. A plot of model residual against the *y*-value (shown in Figure 7C) indicated very strong nonuniformity in residual variance. The residual behavior strongly suggested model misspecification. In this regard the proposed VT model is definitely inferior to the previous models. Further, the VIF shown in Figure 7A suggested that regressor collinearity needed also to be considered, and this chiller model was not a viable candidate. This model approach illustrates the pitfalls of evaluating only at the R^2 and CV and not the model residual pattern. If only Figure 7B and not Figure 7C were generated, a poor model would have been identified from the data as being satisfactory.

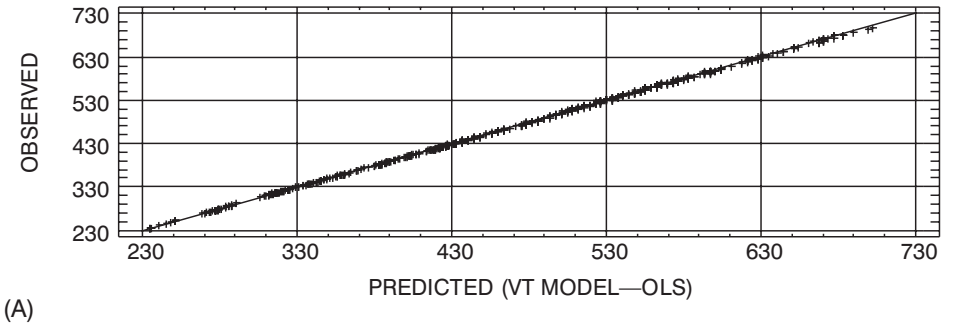
The analysis results presented were based on separating the training and testing data sets such that the effect of model extrapolation errors would be minimized. In an effort to evaluate whether the conclusions reached were sensitive to the selection, an extreme type of data split was evaluated. The data were split into two equal data sets, the training containing the lower half of the Q_{ch} data variation, and the testing containing the upper half. This would provide an evaluation of the GN and MP models under an extreme situation.

The results of the analysis are summarized in Table 6 along with the sequential data split. Both models are still very good, with the R^2 values during training being about 99% and the CV about 1%, which are very close to those found during the sequential data set split. Further, the extrapolation errors have CV values less than 2%, and the NMBE are about 1%. The parameter estimates are also consistent between both types of data splits. It can be concluded with confidence that both GN and MP models, which are linear, are not sensitive to how the training/testing data set split is selected. The fact that the model extrapolation errors are good justifies the use of both models for accurate prediction of fault-free chiller behavior.

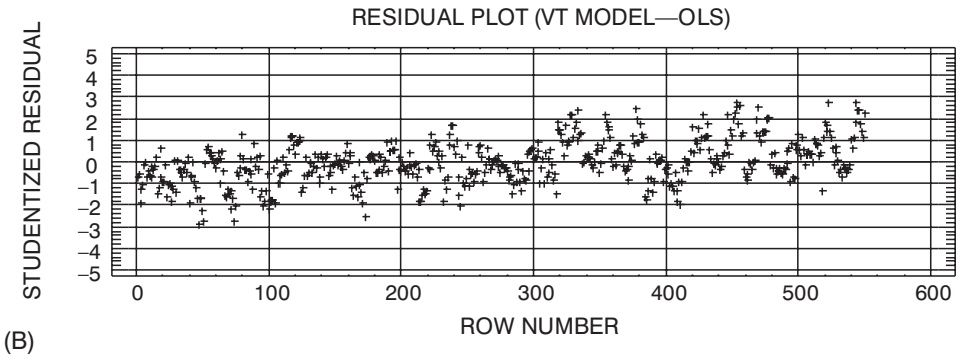
Table 6 assembles the parameter estimates of the GN and the MP models, along with the respective relative errors as percentage values. A value of 0.97% (for b_1 of the MP model) implies that the standard error of the coefficient is about 1% of its numerical magnitude, which is very good. The percentage errors are generally small, with GN holding a slight edge because it seems to give more efficient parameters estimates. The largest error is that of b_2 = 11.2% for the MP model. These standard errors can be used in tracking individual model parameters.

Table 6. OLS Model Parameters and Their Relative Errors for GN and MP Models Under Different Training/Testing Data Sets

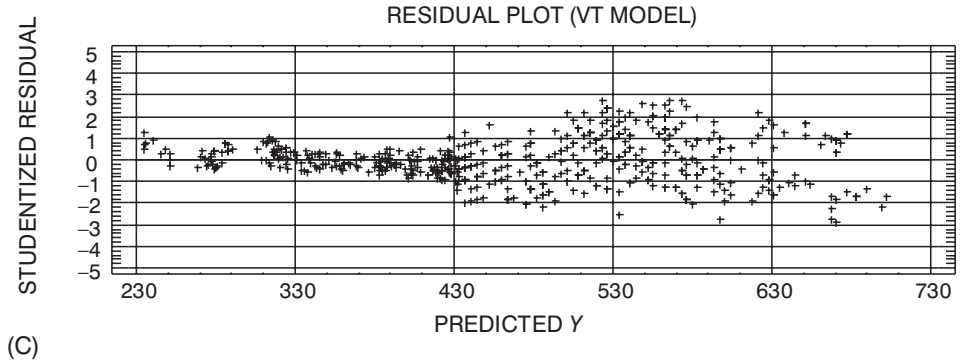
		Training		Testing		b_1		b_2		b_3	
		R^2, %	CV, %	CV, %	NMBE, %	Estimate	Standard Error/Est.	Estimate	Standard Error/Est.	Estimate	Standard Error/Est.
Sequent. data set	GN	99.1	1.45	1.63	–1.01	0.6343	0.73%	–806.4	1.83%	0.01488	1.68%
	MP	99.2	0.95	1.13	0.62	-2.27×10^{-6}	0.97%	-6.57×10^{-6}	11.2%	3.00×10^{-5}	2.53%
Extreme data set	GN	99.3	1.38	1.80	1.08	0.6418	0.86%	–842.7	2.42%	0.01478	1.80%
	MP	99.3	0.91	1.13	0.69	-2.30×10^{-6}	1.02%	-7.90×10^{-6}	12.5%	3.15×10^{-5}	3.33%



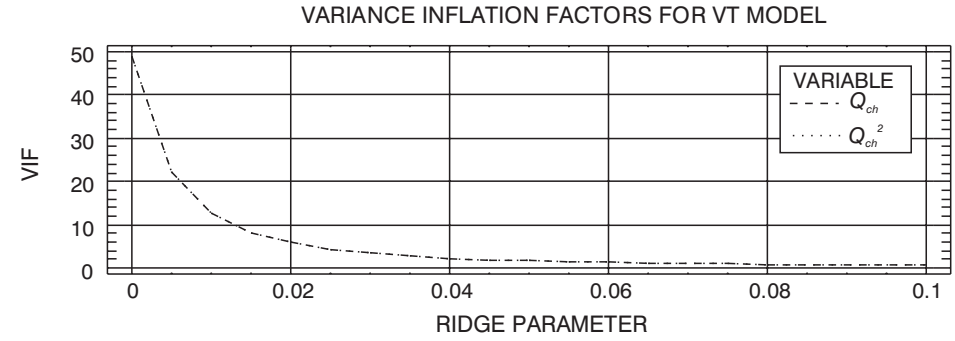
(A)



(B)



(C)



(D)

Figure 7. Analysis of Chiller Data Using VT (Black-Box) Model

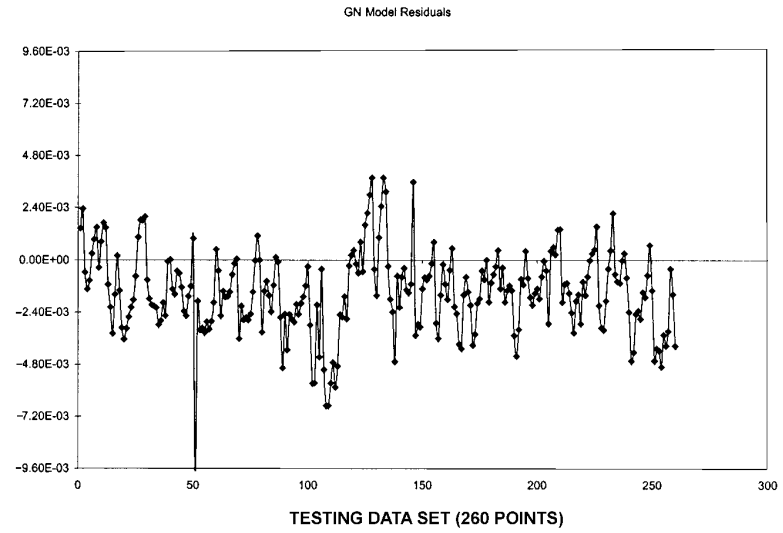


Figure 8A. Scatter Plot with Time of GN Model Residuals for Testing Data Set

The *y*-axis is scaled to ±4 times the RMSE of the model so that each division represents one standard error

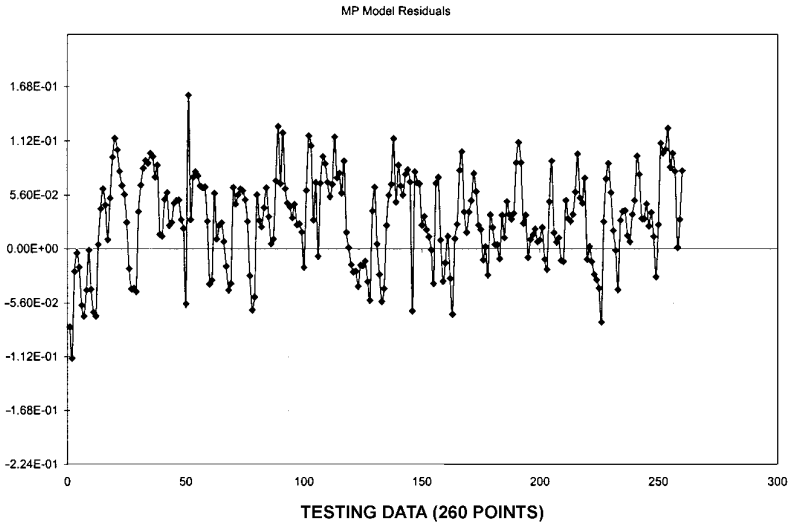


Figure 8B. Scatter Plot with Time of MP Model Residuals for Testing Data Set

The *y*-axis is scaled to ±4 times the RMSE of the model so that each division represents one standard error

An approach to automated fault detection procedure is to set up a control chart based on model prediction. Rather than use Equation (7) for predicting the prediction bands of the OLS models, a simpler and perhaps more practical method would be to use the RMSE of the model residuals determined from the training period. Control chart limits can be taken at either two or three times the model RMSE value, which are representative of the 95% and the 99% confidence level values. Figure 8 depicts such control limits for both the GN and MP models based on OLS estimation, along with the 260 testing data set points. Other than one outlier point, most of the residuals fall between two standard deviations and only a few between two and three standard deviations. For example, for the GN model only seven points fall outside the 95% bands. Because there are 260 points, about 13 points would be expected to fall outside these bands even when the process is under control.

SUMMARY AND CONCLUDING REMARKS

The objectives of this paper were to provide a brief background of inverse methods covering the two phases of model formulation and parameter estimation, and to present a case study of how monitored data from a chiller can be analyzed to identify a model to be used for fault detection. The procedure of separating the data into training and testing data sets was described, and three different steady-state models were evaluated.

It was shown that although the regressors showed strong evidence of collinearity, the ordinary least squares (OLS) model was preferred over the ridge model with this particular data set. It was shown that using only the model goodness of fits could be inappropriate for setting up control charts, because a very good model could have improper residuals. Though both the GN and MP models were equally appropriate, it was concluded that the latter was preferable due to its slightly better estimation properties even though its individual parameters had poorer efficiency. The techniques illustrated in this paper are applicable to data analysis and modeling of HVAC&R systems in general.

ACKNOWLEDGEMENTS

This research was performed as part of an ASHRAE-funded research project (RP-1139). The authors would like to acknowledge the Project Monitoring Subcommittee (PMSC) members for their constructive and helpful suggestions during this research. The authors would also like to thank Professor Jeff Haberl from Texas A&M University for kindly supplying us with the chiller data used in this paper. One of the authors (KKA) would like to acknowledge Professor Henrik Madsen of the Technical University of Denmark for financial support during his stay at Drexel University.

NOMENCLATURE

E	expected value
k	number of regression parameters in the model, ridge factor
n	number of observation data
P	electric power consumed by the chiller
p	number of model parameters (= k + 1)
Q_{ch}	thermal load on chiller
r	correlation coefficient
T_{cdi}	fluid inlet temperature to the condenser
T_{chi}	fluid inlet temperature to the chiller or evaporator
$\boldsymbol{X}$, x	regressor variable(s) or regressor matrix
$\overline{X}$	mean value of X
$\hat{X}$	model predicted value of X
$\boldsymbol{Y}$, y	response variable(s)
$\boldsymbol{b}$, β	model parameter or model parameter vector
ε	vector of model error
s^2, σ^2	mean square error, Equation (5)
ρ	autocorrelation coefficient
$\boldsymbol{\Sigma}$	weighting vector
var	model variance

Acronyms

Adj-R^2	coefficient of determination adjusted for degrees of freedom
CO	Cochrane-Orcutt procedure
COP	coefficient of performance of chiller
CLS	corrected least squares
CV-RMSE, CV	coefficient of variation of the root mean square error
DW	Durbin-Watson statistic, Equation (8)
FDD	fault detection and diagnosis
GLS	generalized least squares
GN	Gordon-Ng chiller model, Equation (19)
MBE	mean bias error
MLR	multiple linear regression
MME	method of moments estimation
MP	multivariate polynomial chiller model, Equation (22)
NMBE	normalized mean bias error
OLS	ordinary least squares
RMSE	root mean square error
SSE	sum of square errors
TMSE	total mean square error, Equation (2)
VIF	variance inflation factors, Equation (9)
VT	variable transformed chiller model, Equation (23)
WLS	weighted least squares

REFERENCES

Andersen, K.K. and T.A. Reddy. 2002. The Error in Variables (EIV) Regression Approach as a Means of Identifying Unbiased Physical Parameter Estimates: Application to Chiller Performance Data. To be published in April 2002 issue of the *International Journal of Heating, Ventilating, Air-Conditioning and Refrigeration Research.*

Beck, J.V. and K.J. Arnold. 1977. *Parameter Estimation in Engineering and Science*. New York: John Wiley & Sons.

Belsley, D.A., E. Kuh, and R.E. Welsch. 1980. *Regression Diagnostics*. New York: Wiley Science.

Bourdouxhe, J.P., M. Grodent, and J. Lebrun. 1999. *HVAC1 Toolkit—A Toolkit for Primary HVAC System Energy Calculation.* ASHRAE.

Box, G.E.P, W.G. Hunter, and J.S. Hunter. 1978. *Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building*. New York: John Wiley & Sons.

Box, G. and A. Luceno. 1997. *Statistical Control by Monitoring and Feedback Adjustment*. New York: John Wiley & Sons.

Chatterjee, S. and B. Price. 1991. *Regression Analysis by Example*, 2nd ed. New York: John Wiley & Sons.

Chen, J. and R.J. Patton. 1999. *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Boston: Kluwer Academic Publishers.

Cook, R.D. and S. Weisburg. 1982. *Residuals and Influence in Regression*. New York: Chapman and Hall.

Davison, A.C. and D.V. Hinkley. 1997. *Bootstrap Methods and Their Applications*. Cambridge, U.K.: Cambridge University Press.

DOE. 1980. *DOE-2, Reference Manual, Part 1, Version 2.1*. Lawrence Berkeley National Laboratory. Berkeley, California.

Draper, N. and H. Smith. 1981. *Applied Regression Analysis*, 2nd ed. New York: John Wiley & Sons.

Efron, B. and R.J. Tibshirani. 1982. *An Introduction to the Bootstrap.* New York: Chapman and Hall.

Fuller, W.A. 1987. *Measurement Error Models*. New York: John Wiley & Sons.

Gordon, J.M. and K.C. Ng. 2000. *Cool Thermodynamics*. Cambridge, U.K.: Cambridge International Science Publishing.

Grimmelius, H.T., J.K. Woud, and G. Been. 1995. On-Line Failure Diagnosis for Compression Refrigeration Plants. *International Journal of Refrigeration* 18(1):31-41.

Haykin, S. 1999. *Neural Networks*, 2nd ed. New Jersey: Prentice Hall.

Katipamula, S., T.A. Reddy, and D.E. Claridge. 1994. Development and Application of Regression Models to Predict Cooling Energy Consumption in Large Commercial Buildings. *ASME/JSME/JSES International Solar Energy Conference Proceedings*, p. 307, San Francisco, March.

Pau, L.F. 1981. *Failure Diagnosis and Performance Monitoring*. Marcel Dekker: New York.

Pindyck, R.S. and D.L. Rubenfeld. 1981. *Economic Models and Economic Forecasts*, 2nd ed. New York: McGraw-Hill.

Reddy, T.A. and D.E. Claridge. 1994. Using Synthetic Data to Evaluate Multiple Regression and Principal Component Analysis for Statistical Models of Daily Energy Consumption. *Energy and Buildings* 21:35-44.

Reddy, T.A., J.K. Kissock, and D.K. Ruch. 1998. Uncertainty in Baseline Regression Modeling and in Determination of Retrofit Savings. *Journal of Solar Energy Eng*. 120:185.

Rossi, T.M. and J.E. Braun. 1997. A Statistical Rule-Based Fault Detection and Diagnostic Method for Vapor Compression Air Conditioners. *International Journal of Heating, Ventilating, Air-Conditioning and Refrigeration Research* 3(1):19-37.

Stoecker, W.F. 1989. *Design of Thermal Systems*, 3rd ed. New York: McGraw-Hill.

Stylianou, M. and D. Nikanpour. 1996. Performance Monitoring, Fault Detection, and Diagnosis of Reciprocating Chillers. *ASHRAE Transactions* 102(1):615-627.