Development of a Toolkit for Calculating Linear, Change-point Linear and Multiple-Linear Inverse Building Energy Analysis Models

ASHRAE Research Project 1050-RP

Final Report

J. Kelly Kissock, Ph.D., P.E.
Jeff S. Haberl, Ph.D., P.E.
David E. Claridge, Ph.D., P.E.

November 1, 2002

UNIVERSITY of DAYTON

Department of Mechanical
and Aerospace Engineering

ENERGY SYSTEMS LABORATORY
Texas Engineering Experiment Station
Texas A&M University System

Final Report for
ASHRAE Research Project 1050-RP

Development of a Toolkit for Calculating Linear, Change-point Linear and Multiple-Linear Inverse Building Energy Analysis Models

Submitted To:

American Society of Heating, Refrigeration and Air-Conditioning Engineers
1791 Tullie Circle, NE
Atlanta, GA 30329

Prepared By:

J. Kelly Kissock
Department of Mechanical and Aerospace Engineering
University of Dayton, Dayton OH 45469-0210
937-229-2852 kkissock@udayton.edu

Jeff S. Haberl
Energy Systems Laboratory
Department of Architecture
Texas A&M University, College Station, TX 77843-3581
979-845-6065 jhaberl@esl.tamu.edu

David E. Claridge
Energy Systems Laboratory
Department of Mechanical Engineering
Texas A&M University, College Station, TX 77843-3123
979-845-1280 claridge@esl.tamu.edu

November 1, 2002

Executive Summary

ASHRAE Guideline14P specifies guidelines for measuring energy savings from building energy-conservation retrofits. A primary method for measuring retrofit savings recommended by ASHRAE Guideline 14P involves identifying a regression model of baseline, or pre-retrofit, energy use as a function of influential variables, such as weather or occupancy, which may change between the pre and post retrofit periods. The regression model is then used to predict how much energy the building would have consumed in the post-retrofit period if it had not been retrofitted. Energy savings are calculated as the difference between the model's prediction of baseline energy use and measured energy use in the post-retrofit period, or as the difference between baseline and post-retrofit models applied to the same weather year.

This report summarizes the results of ASHRAE Research Project 1050: Development of a Toolkit for Calculating Linear, Change-Point Linear and Multiple Linear Inverse Building Energy Analysis Models. The Inverse Modeling Toolkit (IMT) is a FORTRAN 90 application for developing regression models of building energy use. IMT can identify single and multi-variable least-squares regression models. It can also identify variable-base degree-day and single and multi-variable change-point models, which have been shown to be especially useful for modeling building energy use. This report includes background information about IMT and the models, instructions for its installation and operation, and the results of accuracy and robustness testing.

The report comes with an IMT CD-ROM. The IMT CD-ROM contains two folders: 'Detailed Test Results' and 'IMT Software'. The results of extensive bounds and accuracy tests, including the data files and IMT output, are in the 'Detailed Test Results' folder. IMT source and executable files, along with sample data and instruction files, are in the 'IMT Software' folder. To install IMT:

1. Copy the 'IMT Software' folder from the IMT CD-ROM to your computer.
2. Open 'Windows Explorer' and select all files in the 'IMT Software' folder on your computer.
3. Select the menu items: 'File', 'Properties'
4. In the dialog box that appears, remove the 'Read Only' attribute and click the 'Apply' button.

Acknowledgements

This work was sponsored by ASHRAE research project 1050-RP under the guidance of Technical Committee 4.7 – Energy Calculations. Atch Sreshthaputra, John Oh, Yoon Jong, Jon Tommiller, Kazim Yadullahi and Chris Schmidt provided assistance with the coding, testing and deployment of the IMT software. Numerous data sets for testing the software were provided by the Texas LoanSTAR Program. Many of the algorithms in the IMT were adapted from the EModel software developed with funding from the Texas State Energy Conservation Office. The ASHRAE 1050-RP project monitoring subcommittee of Jan F. Kreider (Chair), Moncef Krarti, Robert Sonderegger, and Agami Reddy provided valuable guidance throughout the project.

Table of Contents

1.0 Introduction

1.1 Motivations for Measuring Savings

Energy conservation retrofits are typically initiated based on predictions of how much energy and money a retrofit will save. However, predicted savings often differ substantially from savings determined by measuring energy consumption before and after a retrofit. Nadel and Keating (1991) showed that measured savings in several major residential energy conservation programs were often less than half of predicted savings (Table 1). Similar discrepancies were found in commercial building retrofit programs. In a study of over 1,700 commercial building energy retrofits, fewer than one in six came within 20% of measured results (Greely et al., 1990). Jamieson and Qualmann (1990) found that the mean deviation between predicted and measured savings in a commercial building retrofit program was 165%, even after excluding buildings with known changes in operations. Discrepancies such as these led Jamieson and Qualmann to conclude that "utility concern regarding the reliability of model predictions for the purchase of energy savings is well-founded". And Hirst et al. (1986) concluded that "large discrepancies between predicted and actual energy use ... discourage efficiency investments".

Table 1. Measured and predicted savings in major residential energy conservation programs (Nadel and Keating, 1991).

Program	Program Description	Measured/Predicted (%)
CMP Energy Manag. Assit.	Low-income grants	40
CMP Pay As You Save	Util. Grant and loan	47
CMP Energy Manag. Rebates	Rebates	15
CMP Packaged Weatherization	Standard weatherization	36
CMP Weatherization	Low-income grants	22
GPU RECAP	Performance contracting	22-44
NU Performance Contracting	Performance contracting	22
BPA Residential Weatherization	Comprehensive weatherization	40-58
Hood River Cons. Project	Comprehensive weatherization	43
SCL HELP (multiplex)	Comprehensive weatherization	117
NEES Partners - Residential	Comprehensive weatherization	107

The large discrepancies between predicted and measured savings in early energy-conservation programs highlighted the need to accurately measure energy savings. As the size and expense of energy conservation programs grew throughout the 1980s, so did the emphasis on evaluation, which became an important part of program management. Measured savings were used to verify the success of retrofits, guide the selection of future retrofits, and, in some cases, to identify and correct operational and maintenance problems, which resulted in even greater savings (Claridge et al., 1994). The importance of measured savings increased further in the late 1980s when state regulatory agencies began granting shareholder incentives based on measured Demand Side Management (DSM) program results (Fels and Keating, 1993).

In the 1990s, the move toward utility deregulation diminished the size and number of utility DSM programs. However, a new type of retrofit funding mechanism, called performance contracting, emerged in which payment for the retrofit is based on measured savings. The growing popularity of performance contracting created new incentives for developing protocols and standards for measuring savings. In response to this need, the National Association of Energy Service Contractors developed protocols for the measurement of retrofit savings in 1992. In 1994, ASHRAE began development of a guideline for measuring retrofit savings (GPC-14P). In 1994, the US Department of Energy also initiated an effort that resulted in publication of the North American Energy Measurement and Verification Protocols (USDOE, 1996a) and, later, the International Performance, Measurement and Verification Protocols (USDOE, 1997: 2001). In addition, the U.S. Federal Energy Management Program developed their own set of Measurement and Verification Guidelines for Federal Energy Projects (USDOE, 1996b).

1.2 Overview of Methods for Measuring Savings

The most straightforward way to measure energy savings is to simply compare pre and post-retrofit energy use. This method implicitly assumes that the change in energy consumption between the pre-retrofit and post-retrofit periods is caused solely by the retrofit. However, energy consumption in commercial buildings is also influenced by other factors including weather conditions, occupancy, internal loads and building operating procedures -- all of which may change between the pre and post-retrofit periods. If these changes are not accounted for, savings determined by this simple method will be erroneous. Thus, more sophisticated methods for measuring savings generally seek to adjust the pre-retrofit data, the post-retrofit data, or both, to account for these changes.

The most common adjustment discussed in the literature is for changing weather conditions between the baseline and post-retrofit periods. Weather adjustments are critical in two situations: 1) when less than a full year of energy data are available for defining the baseline behavior of a building, and 2) when the energy data being analyzed vary from year to year depending on the weather.

If a full year of baseline weather data are available, the importance of weather adjustment depends on how much annual energy use varies with weather. Several factors affect the weather-sensitivity of building energy use. One of the most important is the relative amount of weather-dependent and weather-independent energy use in the measured data. For example, sub-metered air-conditioning electricity use shows more weather variability than whole-building electricity use. In addition, *relative* weather sensitivity, measured as a fraction of average energy use, is generally greater when the *average* energy use is smaller. For example, relative heating energy use typically varies more in Miami, where average heating energy use is small, than in Boston, where the average heating energy use is more substantial. Finally, simulation studies indicate that, in general, heating energy use shows more annual variability than cooling energy use and that smaller buildings are more weather sensitive than larger buildings.

For example, in a study that simulated residential energy consumption with 20 years of measured weather data from four U.S. cities, Kissock et al. (1999) reported that annual heating gas use in northern US cities varied by up 29% while whole-building gas use varied by up to 22%. In southern US cities, annual heating gas use in northern US cities varied by up 93% and whole-building gas use varied by up to 12%. Annual residential air-conditioning electricity use varied by up 23% while whole-building electricity use varied by up to 8%. Because these variations are often of the same magnitude of the expected savings from an energy-conservation retrofit, the need for weather adjustment in residential buildings is clear.

Annual weather sensitivity is less pronounced in larger buildings, but may still be important. For example, in a simulation study of commercial building energy use in five U.S. cites, Eto (1988) reported that simulated gas consumption during cold years was 7.2% to 28.6% higher than gas use during average weather years; and during warm years gas use ranged from 2.5% to 26.4% less than during average weather years. Variations in annual sub-metered commercial-building cooling energy use may also be significant, but become almost invisible if weather-independent energy use is included in the measured data. For example, in the same study sited above, Kissock et al. reported that simulated annual whole building electricity use in large commercial buildings varied by only about 1% during the same 20 year period.

In general, two types of measured savings, actual and normalized, can be determined. Actual savings (Kissock, 1993; Cowan and Schiller, 1997, etc.) are calculated as the difference between the energy use predicted by the baseline model and measured post-retrofit energy use. Cowan and Schiller (1997) describe the steps involved in measuring actual savings in buildings as:

1) measure energy use and influential variables during baseline period
2) create a mathematical model of baseline energy use as function of influential variables
3) measure energy use and influential variables during post-retrofit period
4) apply influential variables from the post-retrofit period to the baseline model to estimate what energy use would have been without the retrofit
5) subtract predicted baseline energy use from measured post-retrofit energy use to estimate savings
6) adjust the baseline model as needed to account for changing conditions

Actual savings appear to be the most common type of savings used in energy performance contracting; however, under some circumstances, actual savings can give results that deviate substantially from predicted savings even if the retrofit equipment performs exactly as predicted. Consider, for example, the case where savings are predicted for a chiller retrofit by simulating chiller performance using typical weather data. If the weather during the post-retrofit period is substantially cooler than normal, the expected savings may never materialize even though the new chiller performed exactly as predicted.

Normalized savings (Fels, 1986, Ruch and Claridge, 1992) estimate how much energy would be saved during a 'normal' year. Calculating normalized savings requires developing a statistical model of energy use as a function of influential variables for both the pre and post-retrofit periods and then driving each model with "normal" conditions to calculate the normalized annual consumption during each period. This method typically provides the best comparison between measured and predicted savings when predicted savings are generated from simulation models using 'normal' or expected conditions.

Both actual and normalized savings require that measured energy use be characterized by one or more mathematical or statistical models. In the building energy community, models derived from measured energy use are called "inverse" models. The term "inverse" differentiates them from "forward" models in which building energy use is predicted from engineering principles. Claridge (1998) summarized the most common methods for developing inverse models of measured energy use. The primary methods include variable-base degree-day (VBDD) models, multivariate regression (MVR) models, change-point (CP) regression models, combination CP/VBDD/MVR regression models, calibrated simulation models and artificial neural network models. These methods are summarized in the following sections. Algorithms used in the ASHRAE Inverse Modeling Toolkit (IMT) are noted.

1.3 Variable-Base Degree Day Models

During the 1980s, Fels (1986) adapted the VBDD method for use in measuring savings as the PRInceton Scorekeeping Method (PRISM). The algorithm finds the base-temperature that gives the best statistical fit between energy consumption and the number of variable-base degree-days in each energy use period. PRISM was one of the first methods to include an estimate of the standard error for all regression parameters (Goldberg, 1982). The method found widespread use, especially in evaluation of residential energy conservation programs. Subsequently, PRISM was found to provide adequate fits with commercial building billing data (Eto, 1988; Haberl and Vajda, 1988; Haberl and Komer, 1990; Kissock and Fels, 1995); however, the physical interpretation of the variable-base degree-day method does not apply to commercial buildings with simultaneous heating and cooling (Rabl et al., 1992; Kissock, 1993).

The FASER (OmniComp, 1984) and Metrix Utility (Silicon Energy Corp., 2000) data analysis programs have also adapted the VBDD method for baseline modeling. Both programs use a manual search procedure to identify the balance-point temperature. Sonderegger (1998) notes that, in his experience, the optimum is rather flat and that a fairly wide range of degree-day base temperatures produce similar results.

For the IMT VBDD model, a search procedure was developed to automatically identify the balance-point temperature that produces the best-fit to the data.

1.4 Change-Point Models

In general, heating and cooling energy consumption in multi-zone commercial buildings tends to vary with ambient temperature throughout the entire range of ambient temperatures encountered. Thus, the VBDD method, which specifies a constant base energy usage below (or above) the balance-point temperature, is not appropriate. In addition, linear two-parameter regression models fail to capture the non-linear relationship between heating and cooling energy use and ambient temperature caused by system effects, such as VAV control, or latent loads (Kissock et al., 1998).

Change-point models, however, succeed at capturing both effects, and, as a consequence, have found widespread use as baseline models for measuring energy savings (Haberl et al. 1994; USDOE- IPMVP, 1997). In the statistical literature, these models are known as piece-wise linear regression models or spline fits. In these models, the data are divided into intervals and line segments fit to the data in each interval with the constraint that the line segments meet at a common point between each interval (Hudson, 1966). Algorithms for piece-wise linear regression have been developed for cases in which the change point between linear sections is known in advance (Neter et al., 1989). When the change-point is not known in advance, it is sometimes estimated by inspection (Maidment et al., 1985; Schrock and Claridge, 1989); however this method does not guarantee a "best fit".

The literature review identified three algorithms that may be applicable for best-fit change-point models. The first algorithm was published by Crawford, Dykowski and Czajkowski (1991). The procedure begins by dividing the data into n bins along the x axis. Developing simple linear regression models for each bin would result in discontinuities between the linear segments. To overcome this problem, the bin widths are varied until the lines intersect at the bin boundaries. Identifying bin boundaries that meet the constraint of continuity between line segments requires an iterative solution of two matrix equations. Testing of the procedure indicated that the initial bin boundaries can affect whether convergence will occur and the values at which convergence will occur. Although an algorithm was developed which assures convergence, the number of change-points cannot be determined in advance. Because of the uncertainty of obtaining convergence, the inability to specify the number of change points, and the reliance of the final result on the initial conditions, this method was not recommended for the Inverse Modeling Toolkit.

The second method was published by Ruch and Claridge (1992). This method develops a four-parameter change-point model of energy consumption, typically as a function of dry-bulb temperature, along with accompanying error diagnostics for the model's parameters. The algorithm finds the optimal change-point by searching within an interval known to contain the change-point. The first step is to split the data into two temperature regimes, fit ordinary least-squared lines in each regime, and calculate the intersection of the lines. This is repeated for numerous temperature regions. In the second stage, the change point is assumed and the model is fit using linear regression. From the collection of fits in the two stages, the algorithm chooses the one with the best least-squares fit. The reliability of the parameter estimated is then calculated. The algorithm was coded into a computer program called 4P in the early stages of the Texas LoanSTAR program. Unfortunately, the method did not prove to be robust when used on actual measured energy data. In addition, the prescription of defining an acceptable region for the change-point: 1) required that the data be pre-inspected and 2) created the possibility that the true best-fit change point might lie outside of that region. For these reasons, this algorithm was not recommended for the Inverse Modeling Toolkit.

The third set of algorithms identified were first described by Kissock et al. (1994) and implemented in the EModel software. These algorithms use a two-stage grid search to identify the best change point. In this method, the minimum x value is selected as the initial change point in a standard piece-wise linear regression equation. The change-point is then incremented and the regression is repeated across the range of x-values. The change point that results in the lowest RMSE is selected as the best-fit change-point temperature. This method is then repeated with a finer grid centered about the initial best-fit change point. The uncertainty with which the change-point temperature is known can be approximated as the width of the finest grid. The method is easily adaptable to three-parameter heating, three-parameter cooling and four-parameter models. A similar algorithm for five-parameter models was developed by Kissock (1996). These models have been used extensively with building energy data and have proven to be extremely robust (Haberl et al., 1998). Selected results from the regression engine have been compared to results from SAS and were found to agree to within several significant figures of precision (Kissock et al., 1994). Because of the simplicity, robustness and accuracy of these algorithms, they were chosen for use in the Inverse Modeling Toolkit.

1.5 Multivariate Regression Models

Multivariate regression can incorporate more than one independent variable into a model of building energy consumption, and as such can be a powerful tool (Forrester and Wepfer, 1984; Leslie et al., 1986; Austin, 1997; Katipamula et al., 1998). Proper care must be taken, however, when using MVR models to predict energy consumption. In general, the addition of independent variables to the model will always increase the strength of the correlation; however, the relative uncertainty (standard error) of each regression coefficient, and hence its predictive value, will decrease. In addition, multicollinearity between independent variables increases the uncertainty with which the values of the regression coefficients are known. Singular Value Decomposition (Anderson, 1990) and Principle Component Analysis (Reddy and Claridge, 1994; Ruch et al., 1993) have been shown to reduce the effects of multicollinearity.

The Inverse Modeling Toolkit uses standard multi-variable regression algorithms.

1.6 Combination CP-MVR and VBDD-MVR Models

CP and VBDD models have been shown to provide good fits between building energy use and ambient temperature. However, other variables also influence building energy use. Combination CP-MVR and VBDD-MVR models attempt to retain this ability to describe energy use as a function of ambient temperature while including the effects of additional independent variables. One approach reported in the literature (Rabl, 1992; Ruch et al. 1993; Sonderegger, 1997; Sonderegger, 1998) is to sequentially identify the change-point or base temperature and then use this result in a MVR model. An alternative approach to is to use indicator variables to produce separate CP or VBDD models for each operating or occupational mode (Austin, 1997; Kissock et al., 1998).

To develop CP-MVR models for Inverse Modeling Toolkit, the change-point algorithms developed by Kissock (1994; 1996) were extended to include multiple independent variables. Using this approach, CP-MVR models can be identified in a single step, rather than sequentially, and without breaking up the data according to operational modes. The Inverse Modeling Toolkit can also produce VBDD-MVR models by first running the VBDD model and then running the MVR model on the VBDD residual file.

1.7 Calibrated Simulation Models

In some cases, complex interaction affects, a lack of specific end-use or pre-retrofit data, or other reasons make it impractical or impossible to rely on a comparison of pre and post-retrofit data to estimate savings. In these cases, simulation models can be calibrated to available data, then adjusted to predict energy savings (Katapamula and Claridge, 1993; Wilson, 1998). The Inverse Modeling Toolkit does not include calibrated simulation capabilities.

1.8 Artificial Neural Network Models

Artificial neural networks (ANN) attempt to mimic parts of the architecture of the brain. The distributed parallel processing structure of the brain is simulated by arranging nodes in layers such that each node is connected to all of the nodes in the adjacent layers. Each node sums the inputs it receives and transmits an output signal to the other nodes to which it is connected. The output signal of each node is multiplied by a weight that is varied during the learning process. A learning algorithm trains an ANN to recognize patterns between input and output variables.

ANNs have been shown to effectively model building energy use (Anstett and Kreider, 1993; MacKay, 1994; Kissock, 1994; Feuston and Thurtell, 1994; Kreider et al., 1995) and improve control of HVAC systems (Curtiss et al., 1996). They have also been proposed as baseline models for measuring savings (Krarti et al., 1998). The Inverse Modeling Toolkit does not include ANN models.

1.9 Uncertainty of Savings

Goldberg (1982) estimated the uncertainty of VBDD parameters in the PRISM method. Cowan and Schiller (1997), among others, discuss the uncertainty of the estimated savings in terms of the money, time and equipment required to reduce the uncertainty. Kissock et al. (1993) and Katapamula et al. (1995) investigate how the length and timing and data time-periods of baseline periods affect the prediction accuracy of the baseline regression models. Kissock et al.

(1998) discuss the error in retrofit savings calculations due to varying indoor air temperature or internal gains. A complicated algorithm for estimating error associated with linear models was described by Ruch et al. (1999). The algorithm was translated into a computer code by Ruch and Kissock and tested in development versions of EModel (1994). Unfortunately, the uncertainty routines were sometimes unstable.

A simplified method of estimating the uncertainty associated with linear regression models and the determination of savings have been described by Reddy et al. (1998) and Kissock et al. (1998). In this method, the uncertainty ε_{pd} associated with predicting E as function of an independent variable T in a baseline model is:

$$\varepsilon_{pd} = t(1-\alpha/2, n-2)\ \text{RMSE} \left[1 + \frac{1}{n} + \frac{\left(T_d - \bar{T}\right)^2}{\sum_{d=1}^{n} \left(T_d - \bar{T}\right)^2} \right]^{\frac{1}{2}} \tag{1.1}$$

where:

$$\text{RMSE} = \text{Root Mean Square Error} = \sqrt{\frac{\sum_{d=1}^{n} (E_d - \hat{E})^2}{(n-2)}} \tag{1.2}$$

The t-statistic, $t(1-\alpha/2, n-p)$, is a function of the level of significance (α), the number of days in the pre-retrofit period (n), and the number of parameters in the model (p). The level of significance (α) indicates the fraction of predictions that are likely to fall outside of the prediction uncertainty bands. In practice, the value of the t-statistic is close to 1.96 for a reasonable number of pre-retrofit data points and a 5% significance (95% confidence) level. In addition, the value of the parenthetic term is usually very close to unity. Thus, ε_{pd} can be closely approximated as:

$$\varepsilon_{pd} \approx 1.96\ \text{RMSE}\ (1 + 2/n)^{1/2} \tag{1.3}$$

2.0 Installing and Running IMT

2.1 Installing IMT

IMT is a FORTRAN 90 application compiled to run on personal computers using the Microsoft Windows operating systems. To install IMT:

1. Copy the 'IMT Software' folder from the IMT CD-ROM to your computer.
2. Open 'Windows Explorer' and select all files in the 'IMT Software' folder on your computer.
3. Select the menu items: 'File', 'Properties'
4. In the dialog box that appears, remove the 'Read Only' attribute and click the 'Apply' button.

The IMT Software folder will contain the following files.

Executable version of toolkit: IMT.EXE
Source code version of the toolkit: IMT.F90
Example data files: DAILYDAT.TXT, VBDDDAT.TXT
Example instruction files: DAILYINS.TXT, VBDDINS.TXT
Required DLL files: SALFLIBC.DLL, FTN90.DLL

2.2 Running IMT

To run IMT, you can either 1) click on the IMT.EXE icon, or 2) open a DOS window and run IMT from within the DOS window. When running IMT by clicking on the IMT.EXE icon, the application will automatically open in a DOS window within the Microsoft Windows operating system. When IMT finishes executing, the DOS window will close, returning you to the Microsoft Windows operating system. When running IMT using this method, the only way to access IMT output is to edit the output files IMT.OUT and IMT.RES.

Alternately, you can first open a DOS window and then run IMT from within the window. This method of running IMT will keep the DOS window open between IMT executions and allow you to see IMT output on the screen. To run IMT using this method, first locate the DOS window application called CMD.COM or CMD.EXE. Either CMD.COM or CMD.EXE is generally located in the C:\WINDOWS\SYSTEM or C:\WINNT\SYSTEM32 folders. A shortcut icon to one of these applications may also be located in the Start, Programs menu. Open the DOS window by double clicking on the CMD.COM or CMD.EXE icon. In the DOS window, use DOS commands to operate IMT. First move to the directory in which you installed IMT. For example, if you installed IMT in the folder C:\IMT, type:

CD C:\IMT

Before running IMT, you must have a properly formatted input data file. Input data files are described in more detail in Chapter 2. To get you started, IMT comes with two sample data files. The first is a uniform time-scale data file called DAILYDAT.TXT. DAILYDAT.TXT contains daily ambient temperatures and energy consumption data from a commercial building. Use DAILYDAT.TXT to run mean, 2P, 3P, 4P, 5P and MVR models. The second input data file is a nonuniform-timescale file called VBDDDAT.TXT. VBDDDAT.TXT contains monthly energy use and occupancy data, and daily ambient temperatures. Use VBDDDAT.TXT to run VBDD models. These models are described more completely in Chapter 5.

You must give IMT instructions to locate the input data file, find the desired fields and records in the input data file, and select the proper regression model. There are two ways to give IMT operating instructions. The first is to direct it to an instruction file. To get you started, we have included two sample instruction files, DAILYINS.TXT and VBDDINS.TXT. To run IMT using the DAILYINS.TXT instruction file, type:

IMT DAILYINS.TXT

at the command prompt. You can also type:

IMT

at the command prompt, and then type:

DAILYINS.TXT

when prompted for the name of the instruction file. If the instruction file is in a different folder, type the complete path and filename of the instruction file.

The second way to give IMT operating instruction is through the keyboard. To use this method, type:

IMT

at the command prompt. When prompted whether you would like to enter instructions from an instruction file or through the keyboard, type:

0

To continue to enter instructions via the keyboard, respond to each prompt.

2.3 Running the Sample Input Data File: DAILYDAT.TXT

DAILYDAT.TXT is a uniform time-scale data file containing daily ambient temperatures and energy consumption data from a commercial building. The fields are:

1: Site number
2: Month
3: Day
4: Year
5: Group field (1 for pre-retrofit period and 2 for post-retrofit period)
6: Cooling energy use (MBtu/day)
7: Heating energy use (MBtu/day)
8: Whole building electricity use (kWh/day)
9: Average ambient temperature (F)

The first five records in DAILYDAT.TXT are shown in Figure 2.1. Data files are described more completely in Chapter 3.

114	10	16	90	1	61.8	27.23	-99	76
114	10	17	90	1	65.2	25.68	-99	79
114	10	18	90	1	44.2	35.21	-99	64
114	10	19	90	1	42.6	38.66	-99	62
114	10	20	90	1	52	32.76	-99	70

Figure 2.1. First five records from DAILYDAT.TXT input data file. The –99 values in the second to last field are "no-data" flags that indicate that no data were available on these days.

DAILYINS.TXT is an IMT instruction file that instructs IMT to generate a multivariable regression (MVR) model of cooling energy use as a function of building electricity use and ambient temperature. DAILYINS.TXT is shown in Figure 2.2. Instruction files are described more completely in Chapter 4.

```
Path and name of input data file = DAILYDAT.TXT
Value of no data flag = -99
Column number of group field = 5
Value of valid group field = 1
Residual file needed (1 yes, 0 no) = 1
Model (1:Mean,2:2p,3:3pc,4:3ph,5:4p,6:5p,7:MVR,8:HDD,9:CDD) = 7
Column number of dependent variable = 6
Number of Y1 independent variables data file (0 to 6) = 2
Column number of X1 independent variable = 8
Column number of X2 independent variable = 9
Column number of X3 independent variable = 0
Column number of X4 independent variable = 0
Column number of X5 independent variable = 0
Column number of X6 independent variable = 0
```

Figure 2.2. DAILYINS.TXT instruction file to generate a multivariable regression (MVR) model of cooling energy use as a function of building electricity use and ambient temperature.

To run IMT using the DAILYINS.TXT instruction file and DAILYDAT.TXT data input file, simply type:

IMT DAILYINS.TXT

at the command line in the DOS window. Another way to load start IMT and load the instruction file is to click on the IMT.EXE icon in the Microsoft Windows operating system, and then type:

DAILYINS.TXT

at the prompt asking for the instruction file. Either of these methods will cause IMT to produce a multivariable regression model of pre-retrofit cooling energy use as a function of building electricity use and ambient temperature in the DAILYDAT.TXT data input file. IMT output is printed to the computer screen and also to ASCII data output files IMT.OUT and IMT.RES. IMT.OUT is shown below.

```
*********************************************
  ASHRAE INVERSE MODELING TOOLKIT (1.9)
*********************************************
    Output file name = IMT.Out
*********************************************
    Input data file name =  DAILYDAT.TXT
    Model type =           MVR
    Grouping column No =    5
    Value for grouping =    1
    Residual mode =         1
    # of X(Indep.) Var =    2
    Y1 column number =      6
    X1 column number =   8
    X2 column number =   9
    X3 column number =   0 (unused)
    X4 column number =   0 (unused)
    X5 column number =   0 (unused)
    X6 column number =   0 (unused)
*********************************************
    Regression Results
  ---------------------------------------
           N =    167
  ---------------------------------------
          R2 =     0.845
  ---------------------------------------
       AdjR2 =     0.845
  ---------------------------------------
        RMSE =      6.4314
  ---------------------------------------
     CV-RMSE =    11.328%
  ---------------------------------------
           p =     0.627
  ---------------------------------------
          DW =     0.740 (p>0)
  ---------------------------------------
           a =    -50.6026 (      5.8082)
  ---------------------------------------
          X1 =      0.0035 (      0.0007)
  ---------------------------------------
          X2 =      1.2576 (      0.0421)
  ---------------------------------------
```

DAILYINS.TXT can be modified using any text editor to run different models. To run a two-parameter model of cooling energy use as a function of ambient temperature, modify DAILYINS.TXT so that the model type is 2 (2P), the dependent Y variable is field 6, the number of independent X variables is 1, and the independent X variable is field 9. Then type IMT DAILYINS.TXT at the DOS command prompt to run IMT.

To run a multivariable regression model of cooling energy use as a function of ambient temperature and whole-building electricity use, modify DAILYINS.TXT so that the model type is 7 (MVR), the dependent Y variable is field 6, the number of independent X variables is 2, the independent X1 variable is field 9, and the independent X2 variable is field 8. Then type IMT DAILYINS.TXT at the DOS command prompt to run IMT.

To run a four-parameter model of cooling energy use as a function of ambient temperature, modify DAILYINS.TXT so that the model type is 5 (4P), the dependent Y variable is field 6, the number of independent X variables is 1, and the independent X1 variable is field 9. Then type IMT DAILYINS.TXT at the DOS command prompt to run IMT.

To run a four-parameter model of cooling energy use as a function of ambient temperature, with whole-building electricity use as an additional independent variable, modify DAILYINS.TXT so that the model type is 5 (4P), the dependent Y variable is field 6, the number of independent X variables is 2, the independent X1 variable is field 9, and the independent X2 variable is field 8. Then type IMT DAILYINS.TXT at the DOS command prompt to run IMT.

In DAILYDAT.TXT, records associated with the pre-retrofit period have a value of 1 in the Group Field, and records from the post-retrofit period have a value of 2 in the Group Field. The value of the Grouping Variable in DAILYINS.TXT is 1; thus, DAILYINS.TXT instructs IMT to model only the pre-retrofit data. To model post-retrofit data instead of the pre-retrofit data, modify DAILYINS.TXT so that the value of the Grouping Variable is 2. To model all data instead of the pre-retrofit data, modify DAILYINS.TXT so that the field number of the grouping variable is 0. You can also use this feature to generate separate models for occupied and unoccupied periods, weekdays and weekends, etc., or to exclude questionable data from the model.

2.4 Running the Sample Input Data File: VBDDDAT.TXT

IMT creates VBDD models from nonuniform-timescale data files. VBDDDAT.TXT is a nonuniform-timescale data file and contains *monthly* energy use and occupancy data, and *daily* ambient temperatures. It is called a nonuniform-timescale input data file because the independent and dependent variables have different time scales. The fields in VBDDDAT.TXT are:

1: Month
2: Day
3: Year
4: Cooling energy use (units/month)
5: Group field (1 for pre-retrofit period and 2 for post-retrofit period)
6: Independent variable 1
7: Independent variable 2
8: Independent variable 3
9: Average daily ambient temperature (F)

The first 36 records in VBDDDAT.TXT are shown in Figure 2.3

12	30	1996	-99	-99	-99	-99	-99	-99
12	31	1996	-99	-99	-99	-99	-99	44
1	1	1997	215	1	30	20	5	41
1	2	1997	-99	-99	-99	-99	-99	52
1	3	1997	-99	-99	-99	-99	-99	57
1	4	1997	-99	-99	-99	-99	-99	60
1	5	1997	-99	-99	-99	-99	-99	47
1	6	1997	-99	-99	-99	-99	-99	25
1	7	1997	-99	-99	-99	-99	-99	23
1	8	1997	-99	-99	-99	-99	-99	20
1	9	1997	-99	-99	-99	-99	-99	20
1	10	1997	-99	-99	-99	-99	-99	21
1	11	1997	-99	-99	-99	-99	-99	2
1	12	1997	-99	-99	-99	-99	-99	2
1	13	1997	-99	-99	-99	-99	-99	2
1	14	1997	-99	-99	-99	-99	-99	10
1	15	1997	-99	-99	-99	-99	-99	25
1	16	1997	-99	-99	-99	-99	-99	19
1	17	1997	-99	-99	-99	-99	-99	2
1	18	1997	-99	-99	-99	-99	-99	5
1	19	1997	-99	-99	-99	-99	-99	13
1	20	1997	-99	-99	-99	-99	-99	31
1	21	1997	-99	-99	-99	-99	-99	37
1	22	1997	-99	-99	-99	-99	-99	47
1	23	1997	-99	-99	-99	-99	-99	30
1	24	1997	-99	-99	-99	-99	-99	32
1	25	1997	-99	-99	-99	-99	-99	28
1	26	1997	-99	-99	-99	-99	-99	18
1	27	1997	-99	-99	-99	-99	-99	32
1	28	1997	-99	-99	-99	-99	-99	22
1	29	1997	-99	-99	-99	-99	-99	19
1	30	1997	-99	-99	-99	-99	-99	28
1	31	1997	-99	-99	-99	-99	-99	39
2	1	1997	268	1	35	30	3	39
2	2	1997	-99	-99	-99	-99	-99	39
2	3	1997	-99	-99	-99	-99	-99	-99

Figure 2.3. First 36 records of the nonuniform-timescale input data file VBDDDAT.TXT.

VBDDINS.TXT is an IMT instruction file that instructs IMT to generate a variable-base cooling degree-day model of the whole-building electricity use in VBDDDAT.TXT. VBDDINS.TXT is shown in Figure 2.4. Instruction files are described more completely in Chapter 4.

```
Path and name of input data file = VBDDDAT.TXT
Value of no-data flag = -99
Column number of group field = 5
Value of valid group field = 1
Residual file needed (1 yes, 0 no) = 1
Model (1:Mean,2:2p,3:3pc,4:3ph,5:4p,6:5p,7:MVR,8:HDD,9:CDD) = 9
Column number of dependent variable Y = 4
Number of independent variables (0 to 6)  = 1
Column number of independent variable X1 = 9
Column number of independent variable X2 = 0
Column number of independent variable X3 = 0
Column number of independent variable X4 = 0
Column number of independent variable X5 = 0
Column number of independent variable X6 = 0
```

Figure 2.4. VBDDINS.TXT instruction file to generate a variable-base cooling degree-day model of the whole-building electricity use in VBDDDAT.TXT.

To run a variable-base cooling degree-day model of building electricity use, simply type:

IMT VBDDINS.TXT

at the command line in the DOS window. Another way to start IMT and load VBDDINS.TXT is to click on the IMT.EXE icon in the Microsoft Windows operating system, and type:

VBDDINS.TXT

at the prompt asking for the instruction file. Two IMT output files, IMT.OUT and IMT.RES will be generated in the same directory as IMT.EXE.

2.4.1 Running Combination VBDD/MVR Models

IMT does not directly support combination variable-base degree-day, multiple-variable regression (VBDD/MVR) models. However, combination VBDD/MVR models can be constructed using two steps. First, determine the best-fit VBDD model using the procedure described above. Next, use the residual file from the VBDD model as input to a MVR model. This will facilitate VBDD/MVR models since residual files from VBDD models always include the degree-days in each energy data period computed to the best-fit base temperature.

For example, to model monthly cooling energy use as a function of daily ambient temperatures and another independent variable, run IMT using VBDDINS.TXT as the instruction file. Next, modify VBDDINS.TXT so that the data input file is IMT.OUT, the model type is 7 (MVR), the number of independent X variables is 2, the independent X1 variable is field 10, and the independent X2 variable is field 6. Save it as NONUNIMVR.TXT. Then run IMT again using NONUNIMVR.TXT as the instruction file. This procedure is demonstrated using a real data file in Chapter 10.

2.5 IMT Output

Model coefficients and goodness of fit parameters are reported in the ASCII output file IMT.OUT (Chapters 6 and 7). IMT.OUT is automatically created in the same directory as IMT.EXE each time you run IMT; thus IMT.OUT will be overwritten each run. If you wish to save IMT.OUT, it must be renamed. IMT.OUT can be viewed or printed using any text editor or word processor. You may find it easiest to use the DOS Editor. To do so, type:

```
EDIT  IMT.OUT
```

at the command prompt.

If instructed to, IMT will also create a file called IMT.RES that includes all input data, predicted values of the dependent variable, and the difference between predicted and measured values of the dependent variable. IMT.RES is created in the same directory as IMT.EXE; thus IMT.RES will be overwritten each time you run IMT and instruct it to generate a residual file. If you wish to save IMT.RES, it must be renamed.

2.6 Quitting IMT and Getting Help

To get help, type "?" at the prompt. To quit, type "Q" or "Quit" at the prompt.

3.0 Input Data Files

3.1 Input Data File Format

IMT reads input data files in standard ASCII format. The data files should contain only numeric data, and contain an equal number of fields (columns) in each record (row). Data fields should be separated by one or more blank spaces.

3.2 Types of Input Data Files

IMT can read two types of input data files: uniform and nonuniform timescale. Uniform-timescale data files are composed of records in which all fields are measured over the same timescale. For example, a uniform-timescale data file would be one in which each record includes the amount of energy consumed in an hour, as the dependent-variable field, and the average occupancy and temperature over that hour, as independent-variable fields. IMT can read uniform-timescale data files of any timescale: hourly, daily, weekly, monthly, yearly, etc.

An example of a uniform-timescale data-input file, DAILYDAT.TXT, is shown in Figure 3.1. The fields are:

1: Site number
2: Month
3: Day
4: Year
5: Group field (1 for pre-retrofit period and 2 for post-retrofit period)
6: Cooling energy use (MBtu/day)
7: Heating energy use (MBtu/day)
8: Whole building electricity use (kWh/day)
9: Average ambient temperature (F)

```
114    10    16    90    1    61.8   27.23    -99    76
114    10    17    90    1    65.2   25.68    -99    79
114    10    18    90    1    44.2   35.21    -99    64
114    10    19    90    1    42.6   38.66    -99    62
114    10    20    90    1      52   32.76    -99    70
```

Figure 3.1. First five records from uniform-timescale DAILYDAT.TXT input data file. The –99 values in the second to last field are "no-data" flags that indicate that no data were available on these days.

To facilitate the use of variable-base degree-day (VBDD) models, IMT can read nonuniform-timescale data files in which the dependent variable is energy use, measured over roughly a monthly timescale, and the independent variable is ambient temperature measured on a daily timescale. An example nonuniform-timescale data-input file, VBDDDAT.TXT, is shown in Figure 3.2. The fields are:

1: Month
2: Day
3: Year
4: Monthly electricity use
5: Group field (1 for pre-retrofit period and 2 for post-retrofit period)
6: Dummy independent variable 1
7: Dummy independent variable 2
8: Dummy independent variable 3
9: Average daily temperature (F)

12	30	1996	-99	-99	-99	-99	-99	-99
12	31	1996	-99	-99	-99	-99	-99	44
1	1	1997	215	1	30	20	5	41
1	2	1997	-99	-99	-99	-99	-99	52
1	3	1997	-99	-99	-99	-99	-99	57
1	4	1997	-99	-99	-99	-99	-99	60
1	5	1997	-99	-99	-99	-99	-99	47
1	6	1997	-99	-99	-99	-99	-99	25
1	7	1997	-99	-99	-99	-99	-99	23
1	8	1997	-99	-99	-99	-99	-99	20
1	9	1997	-99	-99	-99	-99	-99	20
1	10	1997	-99	-99	-99	-99	-99	21
1	11	1997	-99	-99	-99	-99	-99	2
1	12	1997	-99	-99	-99	-99	-99	2
1	13	1997	-99	-99	-99	-99	-99	2
1	14	1997	-99	-99	-99	-99	-99	10
1	15	1997	-99	-99	-99	-99	-99	25
1	16	1997	-99	-99	-99	-99	-99	19
1	17	1997	-99	-99	-99	-99	-99	2
1	18	1997	-99	-99	-99	-99	-99	5
1	19	1997	-99	-99	-99	-99	-99	13
1	20	1997	-99	-99	-99	-99	-99	31
1	21	1997	-99	-99	-99	-99	-99	37
1	22	1997	-99	-99	-99	-99	-99	47
1	23	1997	-99	-99	-99	-99	-99	30
1	24	1997	-99	-99	-99	-99	-99	32
1	25	1997	-99	-99	-99	-99	-99	28
1	26	1997	-99	-99	-99	-99	-99	18
1	27	1997	-99	-99	-99	-99	-99	32
1	28	1997	-99	-99	-99	-99	-99	22
1	29	1997	-99	-99	-99	-99	-99	19
1	30	1997	-99	-99	-99	-99	-99	28
1	31	1997	-99	-99	-99	-99	-99	39
2	1	1997	268	1	35	30	3	39
2	2	1997	-99	-99	-99	-99	-99	39
2	3	1997	-99	-99	-99	-99	-99	-99

Figure 3.2. First 36 records of the nonuniform-timescale input data file VBDDDAT.TXT.

When creating VBDD models of nonuniform-timescale data sets, the independent variable, temperature, is assumed to be on a daily time scale. The dependent variable, energy use, is assumed to represent the total energy use from the current record to the previous energy use record. Thus, in VBDDDAT.TXT, the energy use value of 268 on 2/1/1997 represents the energy use between 1/2/1997 and 2/1/1997. The energy use value of 215 on 1/1/1997 is ignored because IMT cannot recognize the beginning of the data period that it represents. Nonuniform-timescale data files can contain any number of independent variables in addition to temperature if they are listed on the same record as the energy value. Like energy, additional independent variables are assumed to represent the total from the current record to the previous energy use record.

3.3 Number of Input Data Files

All data used in a regression model must be included in a single data file.

3.4 Size of the Input Data Files

Data from the input data file are manipulated in arrays stored in the computer's Random Access Memory (RAM). Thus, the size of the input data file is limited only by the amount of RAM available to the computer.

3.5 Data Grouping

The input data file may contain records that the user does not want to include in the model. IMT will operate on only those records that the user specifies. If you wish to use a subset of the data in the model, add a field (column) to the data input file that will indicate which records are to be used in the model. This field is called the 'Grouping Variable'. Place the same numeric value in each record that you wish to be included in the model. You will then specify this field as the "Grouping Variable" when giving IMT operating instructions. In the operating instructions, you must also specify the value of the number in the grouping field that indicates that a record is to be included in the model. Grouping fields can be used for weekday/weekend, pre-retrofit/post-retrofit and other groups.

3.6 No-Data Flag

If there exists no valid values for one or more fields in a data record, a "no-data" flag must be placed in the appropriate field to indicate that this field should not be included in a regression model. The user may select any numeric value for the no-data flag. The toolkit will ignore any record that has a no-data flag in a field on which the model is to operate. In the example input data files DAILYDAT.TXT and VBDDDAT.TXT (Figures 3.1 and 3.2), -99 is used at the no-data flag.

4.0 Operating Instructions

The user must enter instructions to the executable version of IMT. These instructions include the path and filename of the data input file, the type of regression model, and the records and fields in the data input file on which to operate. This chapter describes how to enter these instructions.

4.1 Methods of Entering Operating Instructions

The toolkit accepts operating instructions by: 1) reading an instruction file, or 2) from the keyboard as the user responds to prompts displayed on the computer screen.

4.2 Instruction File Format and Content

The instruction file must be a standard ASCII text file with 14 records. The information required on each line of the instruction file, or when entering instructions by responding to screen prompts through the keyboard, is shown in Table 4.1.

Table 4.1 Information required in the Instruction File or through keyboard entry.

Line	Information Required
1	Path and name of input data file
2	Value of no data flag
3	Column number of group field
4	Value of valid group field
5	Residual file needed (1 yes, 0 no)
6	Model (1:Mean,2:2p,3:3pc,4:3ph,5:4p,6:5p,7:MVR,8:HDD,9:CDD)
7	Column number of dependent variable
8	Number of Y1 independent variables data file (0 to 6)
9	Column number of X1 independent variable
10	Column number of X2 independent variable
11	Column number of X3 independent variable
12	Column number of X4 independent variable
13	Column number of X5 independent variable
14	Column number of X6 independent variable

On Line 1, enter the path and name of the data-input file. If the data-input file is in the same folder as IMT.EXE, the path need not be entered.

On Line 2, enter the value of the marker, called the 'no-data flag', used to denote missing data. A typical value for the no-data flag is '–99'. IMT will not use data from the data-input file in regression models if the data has the value of the no-data flag specified on Line 2.

On Line 3, enter the column number of the grouping field in the data-input file. The grouping field is an optional column in the data-input file for indicating which records should be included in the IMT regression model. If the data-input file does not have a grouping field, enter '0'. Entering '0' on Line 3 causes IMT to use all of the data in the input-data file in the regression model.

On Line 4, enter the value of the data in the grouping field that indicates that this record should be included in the regression model. For example, if the data-input file includes data from both the pre-retrofit and post-retrofit periods, a grouping field could be added to the data-input file with a value of '1' for each pre-retrofit record and a value of '2' for each post-retrofit record. The column number of this field in the data-input file should be entered in the instruction file on Line 3. To develop a regression model of pre-retrofit data, enter '1' on Line 4 of the instruction file. To develop a regression model of post-retrofit data, enter '2' on Line 4 of the instruction file.

On Line 5, enter '1' if a residual output file, IMT.RES, is desired and "0" if no residual output file is desired. Residual output files are described in more detail in Chapter 7.

On Line 6, enter the number 1 through 9 corresponding to the desired regression model. IMT regression models are described in more detail in Chapter 5.

On Line 7, enter the column number in the data-input file of the dependent variable. For example, to create a model of chiller energy use as a function of outdoor air temperature, chiller energy use would be the dependent variable and outdoor-air temperature would be the independent variable.

On Line 8, enter the number of independent variables to be used in the model. For example, to create a model of chiller energy use as a function of outdoor-air temperature and outdoor-air specific humidity, the number of independent variables would be '2'.

On Lines 9-14, enter the column number(s) in the data-input file of the independent variable(s). Enter '0' for all unused independent variables. For example, to create a model of chiller energy use as a function of outdoor-air temperature and outdoor-air specific humidity, enter '2' on Line 8 to indicate two independent variables. Then enter the column number in the data-input file of outdoor-air temperature on Line 9 and the column number of outdoor-air specific humidity on Line 10. Enter '0' on Lines 11 – 14.

An example instruction file, DAILYINS.TXT, is shown in Figure 4.1. To create different instruction files, we recommend modifying an existing instruction file by changing the values to the right of the equal signs and saving the instruction file with a different name. This will retain the field description information in the instruction file and make the instruction file easier to understand. However, IMT only reads data to the right of the equal sign on each record. Thus, the field description information to the left of, and including, the equal sign is optional.

```
Path and name of input data file = DAILYDAT.TXT
Value of no data flag = -99
Column number of group field = 5
Value of valid group field = 1
Residual file needed (1 yes, 0 no) = 1
Model (1:Mean,2:2p,3:3pc,4:3ph,5:4p,6:5p,7:MVR,8:HDD,9:CDD) = 7
Column number of dependent variable = 6
Number of Y1 independent variables data file (0 to 6) = 2
Column number of X1 independent variable = 8
Column number of X2 independent variable = 9
Column number of X3 independent variable = 0
Column number of X4 independent variable = 0
Column number of X5 independent variable = 0
Column number of X6 independent variable = 0
```

Figure 4.1. DAILYINS.TXT instruction file to generate a multivariable regression (MVR) model of cooling energy use as a function of building electricity use and ambient temperature using the DAILYDAT.TXT input data file shown in Figure 3.1.

The toolkit determines the type of input data file by the type of regression model. HDD and CDD models require nonuniform-timescale data files. All other models require uniform-timescale data files.

4.3 Entering Operating Instructions by Typing Responses to Screen Prompts

The toolkit also accepts operating instructions entered through the keyboard. The toolkit provides the user with prompts asking for the appropriate information. The information requested by the screen prompts is the same as the information in the instruction file. The user can quit the program at any time by typing “Quit”, “Q”, “quit” or “q”.

5.0 Regression Model Types

IMT supports several types of regression models since no single model type is appropriate for all types of buildings or patterns of energy use. The types of models supported by the toolkit are described below. The two-parameter and change-point models supported by the toolkit are shown in Figure 5.1. The type of model is identified by the number of regression coefficients β.

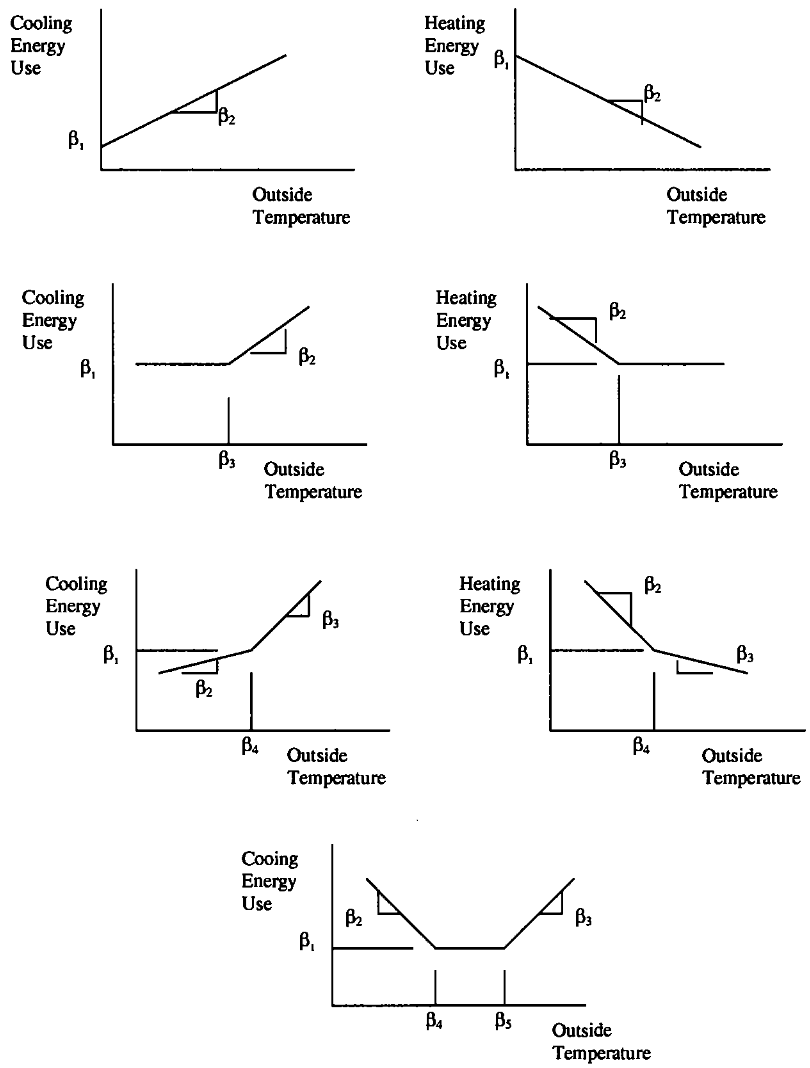

Figure 5.1 IMT change-point models. Top row: 2P cooling and heating models. Second row from top: 3P cooling and heating models. Third row from top: 4P cooling and heating models. Bottom row: 5P heating and cooling model.

5.1 Mean Model

IMT can calculate the arithmetic mean of the dependent variable. IMT mean models are appropriate for modeling building energy use that does not vary in relation to other independent variables.

5.2 Two-Parameter Model

IMT can find a simple linear regression model (2P) of type:

$$Y = \beta_1 + \beta_2 X_1 \tag{5.1}$$

where β_1 and β_2 are regression coefficients, X_1 is the independent variable and Y is the dependent variable.

2P models are appropriate for modeling building energy use that varies linearly with another single independent variable. For example, in some buildings, heating and cooling energy use varies linearly with outdoor air temperature. In IMT output, β_1 is reported as "a", and β_2 is reported as "X1".

5.3 Three-Parameter Cooling and Heating Models

IMT can find best-fit three-parameter (3P) change-point models of the type described by Kissock et al. (1994):

$$Y_c = \beta_1 + \beta_2 (X_1 - \beta_3)^+ \tag{5.2}$$

$$Y_h = \beta_1 + \beta_2 (X_1 - \beta_3)^- \tag{5.3}$$

where β_1 is the constant term, β_2 is the slope term, and β_3 is the change-point,. The ( )+ and ( )$^-$ notations indicate that the values of the parenthetic term shall be set to zero when they are negative and positive respectively.

3P models are appropriate for modeling building energy use that is varies linearly with an independent variable over part of the range of the independent variable and remains constant over the other part. For example, 3PC models, using outside air temperature as the independent variable, are often appropriate for modeling whole-building electricity use in residences electric air conditioning. Similarly, 3PH models, using outside air temperature as the independent variable, are often appropriate for modeling heating energy use in residences with gas or oil heating.

IMT can also find combination three-parameter multi-variable regression models (3P-MVR), with up to four independent variables, of the type:

$$Y_c = \beta_1 + \beta_2 (X_1 - \beta_3)^+ + \beta_4 X_2 + \beta_5 X_3 + \beta_6 X_4 \quad (5.4)$$

$$Y_h = \beta_1 + \beta_2 (X_1 - \beta_3)^- + \beta_4 X_2 + \beta_5 X_3 + \beta_6 X_4 \quad (5.5)$$

where X_1 is typically temperature, and X_2, X_3, and X_4 are optional independent variables.

In IMT output for 3PC models, β_1, the Y change point coefficient, is reported as "Ycp". β_2, the right slope coefficient, is reported as "RS". β_3, the X change point coefficient, is reported as "Xcp". IMT output for 3PH models is the same as for 3PC models except that β_2, the left slope coefficient, is reported as "LS". In IMT output for 3P-MVR models, the additional independent variable coefficients, β_4, β_5, and β_6, are reported as the number of the independent X variable, "X_2", "X_3" and "X_4"respectively.

5.4 Four-Parameter Model

IMT can find best-fit four-parameter (4P) change-point models of the type described by Kissock et al. (1994):

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_4)^+ \quad (5.6)$$

where β_1 is the constant term, β_2 is the left slope, β_3 is the right slope and β_4 is the change point. IMT can also find combination four-parameter multi-variable regression models (4P-MVR), with up to three independent variables, of the type:

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_4)^+ + \beta_5 X_2 + \beta_6 X_3 \qquad (5.7)$$

where X_1 is typically temperature, and X_2 and X_3 are optional independent variables.

Four-parameter models using outdoor air temperature as the independent variable are appropriate for modeling heating and cooling energy use in variable-air-volume systems and/or in buildings with high latent loads. In addition, these models are sometimes appropriate for describing non-linear heating and cooling consumption associated with hot-deck reset schedules and economizer cycles (Kissock, 1993).

In IMT output for 4P models, β_1, the Y change point coefficient, is reported as "Ycp". β_2 , the left slope coefficient, is reported as "LS". β_3 , the right slope coefficient, is reported as "RS". β_4 , the X change point coefficient, is reported as "Xcp". In IMT output for 4P-MVR models, the additional independent variable coefficients, β_5, and β_6, are reported as the number of the independent X variable, "X_2" and "X_3".

5.5 Five-Parameter Model

IMT can find best-fit five-parameter (5P) change-point models of the type described by Kissock (1996):

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_5)^+ \qquad (5.8)$$

where β_1 is the constant term, β_2 is the left slope, β_3 is the right slope, β_4 is the left change point, and β_5 is the right change point.

IMT can also find combination five-parameter multi-variable regression models (5P-MVR), with up to two independent variables, of the type:

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_5)^+ + \beta_6 X_2 \qquad (5.9)$$

where X_1 is typically temperature and X_2 is an optional independent variable.

Five-parameter models using outdoor air temperature as the independent variable are appropriate for modeling energy consumption data that includes both heating and cooling, such as whole-building electricity data from buildings with electric heat-pumps or both electric chillers and electric resistance heating. They are also appropriate for modeling fan electricity consumption in variable-air-volume systems.

In IMT output for 5P models, β_1, the Y change point coefficient, is reported as "Ycp". β_2, the left slope coefficient, is reported as "LS". β_3, the right slope coefficient, is reported as "RS". β_4, the left X change-point coefficient, is reported as "Xcp1". β_5, the right X change-point coefficient, is reported as "Xcp2". In IMT output for 5P-MVR models, the additional independent variable coefficient β_6, is reported as the number of the independent X variable, "X_2".

5.6 Multiple-Variable Regression Model

IMT can find multiple-variable linear regression (MVR) models, with up to six independent variables, of type:

$$Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 X_3 + \beta_5 X_4 + \beta_6 X_5 + \beta_7 X_6 \quad (5.10)$$

where β_1 through β_7 are regression coefficients, and X_1 through X_6 are independent variables. IMT does not test for multicollinearity.

In IMT output, β_1 is reported as "a", the independent variable coefficients β_2 to β_7, are reported as the number of the independent X variables, "X_1" to "X_6".

5.7 Variable-Base Heating and Cooling Degree-Day Models

IMT can find best-fit variable-base degree-day models of type:

$$Y = \beta_1 + \beta_2 \, HDD(\beta_3) \quad (5.11)$$

$$Y = \beta_1 + \beta_2 \, CDD(\beta_3) \quad (5.12)$$

where β_1 is the constant term, β_2 is the slope term, and $HDD(\beta_3)$ and $CDD(\beta_3)$ are the number of heating and cooling degree-days, respectively, in each energy data period calculated with base temperature β_3. The number of heating and cooling degree-days in each energy data period of n days is:

$$\mathrm{HDD}(\beta_3) = \sum_{i=1}^{n} (\beta 3 - \mathrm{T}_i)^+ \qquad (5.13)$$

$$\mathrm{CDD}(\beta_3) = \sum_{i=1}^{n} (\mathrm{T}_i - \beta 3)^+ \qquad (5.14)$$

where T_i is the average daily temperature.

In IMT output, β_1 is reported as "A" and β_2 is reported as "X_1". β_3, the base temperature for computing the degree days, is reported as "DD Base".

6.0 Model Uncertainty Parameters

Guideline-14P Working Draft 99.2, June 7, 1999 specifies that modeling uncertainty be estimated using three indices:

1) Coefficient of Variation of the Standard Deviation (CVSTD)

$$\text{CVSTD} = 100 \text{ x } [\sum (y_i - \bar{y})^2 /(n-1)]^{1/2} / \bar{y} \tag{6.1}$$

2) Coefficient of Variation of the Root Mean Square Error (CVRMSE)

$$\text{CVRMSE} = 100 \text{ x } [\sum (y_i - \hat{y}_i)^2 /(n-p)]^{1/2} / \bar{y} \tag{6.2}$$

3) Normalized Mean Bias Error (MBE)

$$\text{NMBE} = \frac{\sum^{n} (y_i - \hat{y}_i)}{(n - p) * \bar{y}} * 100 \tag{6.3}$$

Where:

- y dependent variable of some function of the independent variable(s)
- $\bar{y}$ arithmetic mean of the sample of n observations
- $\hat{y}$ regression model's predicted value of y
- n number of data points or periods in the baseline period
- p number of parameters or terms in the baseline model, as developed by a mathematical analysis of the baseline data.

CVSTD (Equation 6.1) is a special case of CVRMSE (Equation 5.2) for mean models with one parameter. Thus, to comply with Guideline-14P, the toolkit reports:

- CVSTD for mean models
- CVRMSE for 2 - 5 parameter and MVR models
- NMBE for all models

IMT reports the following uncertainty statistics:

- Standard Deviation, STD, for mean models:

$$\text{STD} = [\sum (y_i - \bar{y})^2 / (n-1)]^{1/2} \tag{6.4}$$

STD is a measure of the spread of data from the mean.

- Root Mean Square Error, RMSE, for all regression models:

$$\text{RMSE} = [\sum (y_i - \hat{y}_i)^2 / (n-p)]^{1/2} \tag{6.5}$$

RMSE is a measure of the spread of data from the model.

- Coefficient of multiple determination, R^2, for all regression models:

$$\text{R}^2 = 1 - \frac{\sum_n (y_i - \hat{y}_i)^2}{\sum_n (y_i - \bar{y}_i)^2} \tag{6.6}$$

R2 can be interpreted as the fraction of variation explained by the model.

- Adjusted R^2, AdjR^2, for all MVR models:

$$\text{Adj}\,\text{R}^2 = 1 - \frac{(N-1)}{(N-p)}(1-R^2) \tag{6.7}$$

In MVR models, the addition of an independent variable will always result in an increase in the model's R^2. Adjusted R^2 divides each sum of squares in R^2 by the associated degrees of freedom, and is thus a measure of the actual improvement in predictive ability from adding independent variables.

- Auto-correlation coefficient of residuals, p, for all regression models:

$$\text{p} = \frac{\sum_{t=2}^{n} e_{t-1} e_t}{\sum_{t=2}^{n} (e_{t-1})^2} \tag{6.8}$$

Least-squares regression assumes that p is approximately zero. As p gets closer to one, this assumption becomes suspect and the RMSE may underestimate the true uncertainty of the model.

- Durbin Watson statistic, DW, for all regression models:

$$\mathrm{DW} = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n}(e_t)^2} \tag{6.9}$$

DW is used to test the hypothesis that p = 0. Low DW values indicate that p > 0 and high DW values indicate that p = 0.

- Standard error of each regression coefficient. The standard error of a regression coefficient indicates the variance with which the coefficient is known. The standard error is defined such that with a probability of 1-α, the true parameter will fall within the bounds:

$$\beta_{true} = \beta_{estimated} + t(1\text{-}\alpha/2,\ n\text{-}p)\ s(\beta) \tag{6.10}$$

where t is the t distribution and $s(\beta)$ is the standard error of each regression coefficient:

$$s(\beta) = [\mathrm{MSE}\ (\mathbf{X'X})^{-1}]^{.5} \tag{6.11}$$

- The uncertainty of each change-point coefficients as given by the width of the final search interval.

Together, these measures of uncertainty allow the user to assess the fit of the model to the data, select appropriate independent variables, and to calculate the overall uncertainty of savings using the methods described in Guideline 14P.

7.0 Output Data Files

IMT reports model results by generating an output file, IMT.OUT, and an optional residual file, IMT.RES, for each model run. Both output files are standard ASCII text files.

The output file IMT.OUT includes the information entered in the operating instructions, model coefficients and goodness-of-fit parameters. A new copy of IMT.OUT is generated after each model run.

The residual file IMT.RES includes the data from the data-input file, predicted values of the dependent variable and model residuals. The residual file is generated only when requested in the instruction file or through keyboard prompts.

7.1 Path and Filename of Output File IMT.OUT

The output file IMT.OUT is placed in the same directory as the instruction file.

7.2 Content of Output File

The output file IMT.OUT includes:

- the output file name
- all information entered as operating instructions
- the regression results (defined in Chapter 6):
    - N = number of observations used in the model
    - R2 = coefficient of multiple determination
    - AdjR2 = adjusted coefficient of multiple determination
    - RMSE = root mean square error
    - CV-RMSE = coefficient of variation of root mean square error
    - p = auto-correlation coefficient
    - DW = Durbin Watson coefficient
    - model coefficients with standard errors

Model coefficients are specified using the convention:

DD Base = base temperature for degree-day calculation (F)

A = y intercept term

XN = regression coefficient for the Nth independent variable

The value in parenthesis following the model coefficient is the standard error of the coefficient.

A sample output file is shown in Figure 7.1.

```
********************************************
                   ASHRAE INVERSE MODELING TOOLKIT (1.9)
********************************************
   Output file name = IMT.Out
********************************************
   Input data file name =  VBDDDAT.TXT
   Model type =           CDD
   Grouping column No =    5
   Value for grouping =    1
   Residual mode =         1
   # of X(Indep.) Var =    1
   Y1 column number =      4
   X1 column number =   9
   X2 column number =   0 (unused)
   X3 column number =   0 (unused)
   X4 column number =   0 (unused)
   X5 column number =   0 (unused)
   X6 column number =   0 (unused)
********************************************
   Regression Results
  --------------------------------------
          N =     12
  --------------------------------------
         R2 =     0.810
  --------------------------------------
      AdjR2 =     0.810
  --------------------------------------
       RMSE =    34.454
  --------------------------------------
    CV-RMSE =    10.313%
  --------------------------------------
          p =     0.493
  --------------------------------------
         DW =     0.854 (p>0)
  --------------------------------------
   DD Base =      41
  --------------------------------------
         A =   258.0816 (15.3174)
  --------------------------------------
         X1 =    0.1993 ( 0.0305)
  --------------------------------------
```

Figure 7.1 Sample output file.

7.3 Path and Filename of Residual File

The residual file, IMT.RES, is placed in the same directory as the instruction file.

7.4 Contents of Residual File

The difference between the observed, y, and predicted, $\hat{y}$, values of the dependent variable is called the residual and is defined as:

$$\text{Residual} = y - \hat{y} \tag{7.1}$$

7.4.1 Residual File from a Uniform-Timescale Data-Input File

Each record of the residual file for uniform-timescale data-input files includes:

- all fields from the data-input file
- the predicted value of the dependent variable
- the residual

Thus, residual files for uniform-timescale data-input files have the same number of records as the data-input file and two more fields than the data-input file.

An example residual file from a uniform-timescale data-input file is shown in Figure 7.2. The first 9 columns are from the data-input file. The dependent variable is in the 6th column. The predicted value of the dependent variable is in the 10th column and the residual is in the 11th column.

114	10	16	1990	3	61.80	27.23	-99	76	45.65	16.15
114	10	17	1990	4	65.20	25.68	-99	79	48.19	17.01
114	10	18	1990	5	44.20	35.21	-99	64	35.50	8.70
114	10	19	1990	6	42.60	38.66	-99	62	33.81	8.79
114	10	20	1990	7	52.00	32.76	-99	70	40.57	11.43
114	10	21	1990	1	44.80	41.29	-99	63	34.65	10.15
114	10	22	1990	2	36.80	44.20	-99	57	29.57	7.23
114	10	23	1990	3	-99	-99	-99	58	-99	-99
114	10	24	1990	4	41.00	39.66	-99	63	34.65	6.35
114	10	25	1990	5	41.80	37.66	-99	64	35.50	6.30
114	10	26	1990	6	43.20	37.39	-99	62	33.81	9.39
114	10	27	1990	7	45.20	33.49	-99	65	36.34	8.86
114	10	28	1990	1	46.80	32.49	-99	68	38.88	7.92
114	10	29	1990	2	48.40	34.21	-99	68	38.88	9.52
114	10	30	1990	3	52.80	33.85	-99	67	38.04	14.76
114	10	31	1990	4	55.60	33.31	-99	68	38.88	16.72
114	11	1	1990	5	53.20	32.13	-99	68	38.88	14.32
114	11	2	1990	6	57.20	31.67	-99	70	40.57	16.63
114	11	3	1990	7	61.00	29.86	-99	75	44.80	16.20
114	11	4	1990	1	40.40	43.02	-99	57	29.57	10.83

Figure 7.2. First 20 records from a residual file from a uniform-timescale input data file.

7.4.2 Residual File from a Nonuniform-Timescale Data-Input File

Nonuniform-timescale data-input files are used with VBDD models. These data files contain observations of the dependent variable, typically energy use, which are usually measured over several days, and observations of the independent variable, typically temperature, which are usually measured on the daily timescale. Because residuals are calculated for each energy observation, the residual file only includes records corresponding to energy observations in the input data file. The residual file from a nonuniform-timescale data-input file includes:

- all fields from records in the data input file that have valid[1] energy values, except the average daily temperature field. The average daily temperature field is replaced with the average temperature over energy time-interval. This feature allows IMT to quickly and accurately calculate the average billing-period temperature for use by 2P, 3P, 4P or 5P models of monthly energy use.
- the number of degree days in the energy time-interval calculated to the best-fit reference temperature
- the residual calculated as difference between the predicted and observed values of energy use

An example residual file generated from a nonuniform-timescale data-input file is shown in Figure 7.3. The fields are month, day, year, average temperature during the energy period, energy use, heating degree days during the energy use period, predicted energy use, and the difference between observed and predicted energy use.

[1] Any value except the value of the no-data flag.

1	4	1979	37.2	2320	645.0	2,307.0	13.0
2	2	1979	31.6	2930	765.0	2,626.6	303.4
3	6	1979	27.8	2920	965.0	3,159.3	-239.3
4	4	1979	46.7	1530	336.0	1,484.0	46.0
5	4	1979	53.8	1150	162.0	1,020.5	129.5
6	5	1979	65.8	630	0.0	589.1	40.9
7	5	1979	69.9	510	0.0	589.1	-79.1
8	3	1979	79.1	600	0.0	589.1	10.9
9	4	1979	76.5	520	0.0	589.1	-69.1
10	3	1979	67.8	620	4.0	599.7	20.3
11	2	1979	55.8	950	148.0	983.2	-33.2
12	4	1979	49.8	1210	287.0	1,353.5	-143.5

Figure 7.3. Example residual file from a nonuniform-timescale data-input file.

8.0 Toolkit Design and Model Algorithms

8.1 Programming Language and Operating System

The toolkit is written in FORTRAN 90 and compiled using Numerical Algorithms Group, Inc., NAGWare FTN90 Compiler v2.1x. (Salford, 1996) The executable version of the toolkit, IMT.EXE runs in an MS DOS window of Microsoft Windows operating system. To run IMT.EXE, the dynamic link library files SALFLIBC.DLL and FTN90.DLL must be in the same directory as the IMT.EXE. The source code IMT.F90 is an ASCII text file and can be accessed using any text editor.

8.2 Toolkit Design

IMT is composed of a main module and a series of subroutines. Program flow for the main module is shown in Figure 8.1. Execution begins by calling the Process_Cmd_Line subroutine, where user instructions about the type of model and input data file are read and checked for errors. The Get_NumRowsCols subroutine reads the number of records and fields in the input data file and uses these values to define the dimensions of an array that will hold the input data. The Read_Data subroutine reads the input data file into the data array. If the data input file is contains non-uniform time series data, the data array is restructured and refilled in the FillDNonUni subroutine. In the Fill_XY subroutine, the subset of data selected by the user for modeling is filtered and placed in X and Y arrays for regression. The X and Y arrays are then passed to the appropriate subroutine for the model specified by the user. Model coefficients and goodness of fit parameters are calculated in the MeanM, MVR, ThreePMVR, FourPMVR, FivePMVR and VBDD subroutines. If the user specified a residual file, the file is created in the Create_Resid_File subroutine. The output data file is then created in the Creat_Out_File subroutine.

More detailed descriptions of program functionality are included in the IMT source code with comments documenting the functionality of each subroutine and code block.

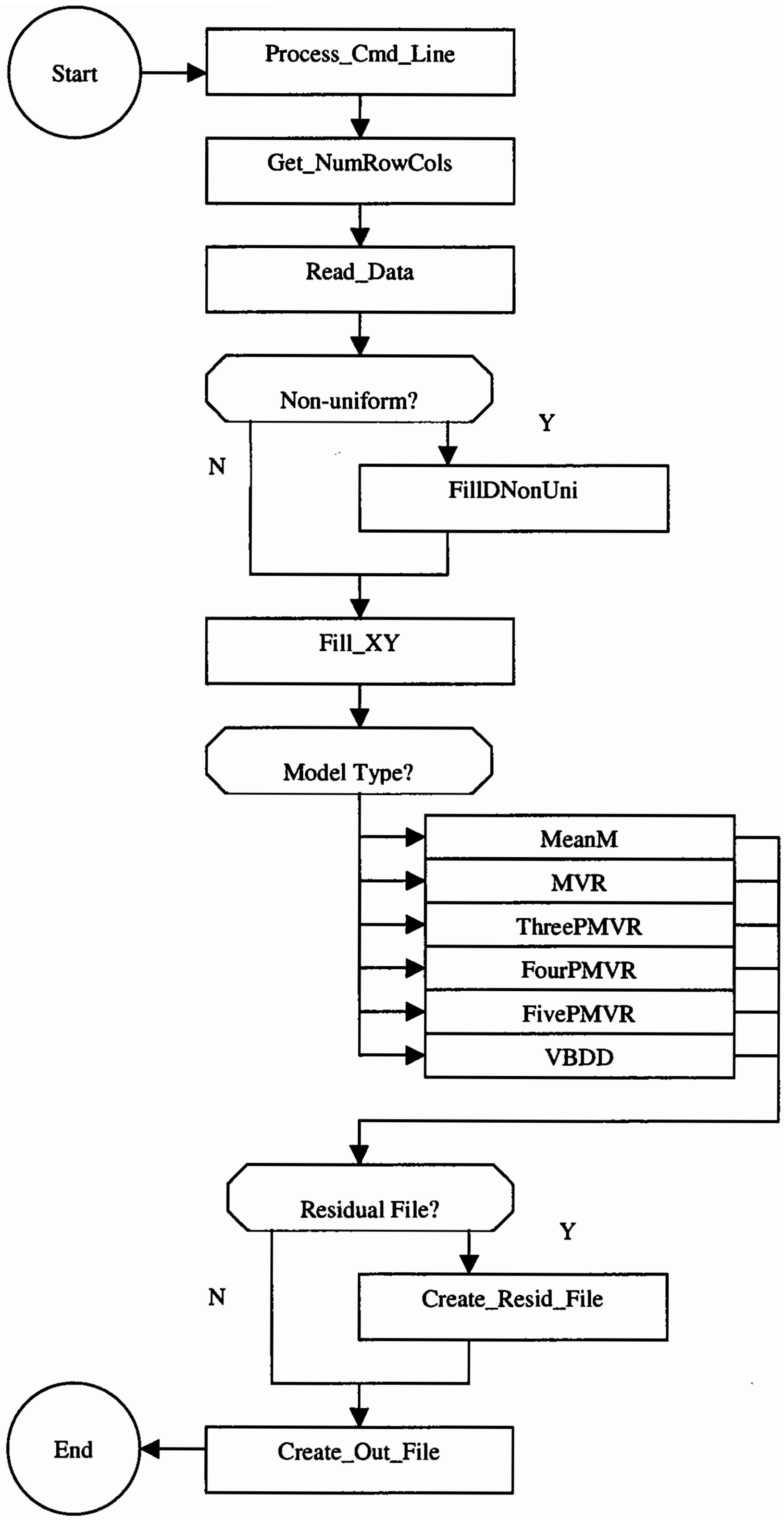

Figure 8.1 IMT main module program flow.

8.3 Least-Squares Regression Algorithm

All IMT model types except the mean model use least-squares regression to determine the model coefficients. Program flow for the IMT least-squares regression algorithm is show in Figure 8.2. Regression begins by calling the FillXY subroutine, which removes records with no-data flags and data that are not in the group specified by the grouping variable, then fills the arrays X and Y. Next, the Reg subroutine calculates least-squares regression coefficients for single or multiple independent variables. The Inf subroutine calculates inference statistics that describe the goodness-of-fit of the model. The Reg and Inf subroutines are described in the following paragraphs.

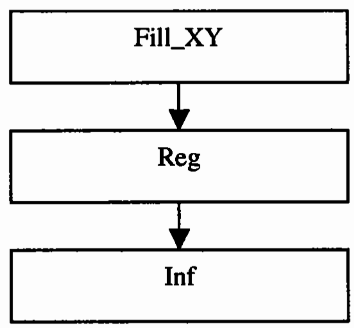

Figure 8.2 Program flow for least squares regression algorithms.

Generalized least-squares regression seeks to estimate model coefficients that minimize the sum of the squared error between predicted and actual observations. The Reg subroutine uses a matrix algebra approach to least squares regression (Neter et al., 1989). In this approach, the matrix of dependent observations, Y, is equal to the product of the matrix of independent observations, X, and the matrix of estimated regression coefficients, β, plus an error term, E.

$$Y = X\beta + E \tag{8.1}$$

Solving for β gives:

$$\beta = (X^T X)^{-1} X^T Y \tag{8.2}$$

The Reg subroutine solves Equation 8.2 by calling the Trans, Mult and Invert subroutines. The Trans, Mult and Inverse algorithms are simply computational versions of standard matrix algebra (Miller, 1981). In the Trans subroutine, the X matrix is transposed by interchanging the rows and columns. The Mult subroutine performs matrix multiplication in which the elements from one row of the first matrix are multiplied by the elements from the column of the second matrix, then summed. The Invert subroutine finds the inverse matrix of the product of X^TX.

Before operating on X, however, each X observation is normalized by the mean value of each independent variable. This normalization process provides computational stability for the Invert subroutine in cases where the values of the X observations are very large or very small.

Model residuals and inference statistics, such as R2, RMSE, CV-RMSE and the standard errors of the regression coefficients, are calculated in the Inf subroutine. To calculate the model residuals, the predicted values of the dependent variable, $\hat{Y}$, are computed from:

$$\hat{Y} = X\beta \tag{8.3}$$

The matrix of residuals, E, is then computed from:

$$\mathrm{E} = \mathrm{Y} - \hat{Y} \tag{8.4}$$

The root mean squared error, RMSE, is computed from:

$$\mathrm{RMSE} = \sqrt{\frac{\sum(Y-\hat{Y})^2}{(n-p)}} = \sqrt{\frac{Y^TY-\beta^TX^TY}{(n-p)}} \tag{8.5}$$

where n is the number of data observations and p is the number of regression coefficients. The matrix of the standard errors of the regression coefficients, S, is computed from:

$$\mathrm{S} = RMSE\sqrt{(X^TX)^{-1}} \tag{8.6}$$

The squared correlation coefficient, R^2, is computed from:

$$R^2 = 1 - \frac{\sum (Y - \hat{Y})^2}{\sum (Y - \overline{Y})^2} \tag{8.7}$$

The adjusted R^2, is computed from:

$$\text{Adjusted } R^2 = 1 - \frac{(n-1)\sum (Y - \hat{Y})^2}{(n-p-1)\sum (Y - \overline{Y})^2} \tag{8.8}$$

8.4 Change-Point Model Algorithm

IMT uses the same algorithm for finding all change-point models, including combination change-point multi-variable regression models (Kissock et al., 1994). The algorithm is demonstrated for the 3P models in the following description.

IMT can find best-fit three-parameter (3P) change-point models of type:

$$Y_c = \beta_1 + \beta_2 (X_1 - \beta_3)^+ \tag{8.9}$$

$$Y_h = \beta_1 + \beta_2 (X_1 - \beta_3)^- \tag{8.10}$$

where β_1 is the constant term, β_2 is the slope term, and β_3 is the change-point,. The ( )+ and ( )$^-$ notations indicate that the values of the parenthetic term shall be set to zero when they are negative and positive respectively. Equation 8.9 represents a 3P-Cooling model and Equation 8.10 represents a 3P-Heating model.

The best-fit change-point temperature β_3 is identified using a two-part grid-search method (Figure 8.3). In the grid-search method, the first step is to identify minimum and maximum values of X1, and to divide the interval defined by these values into ten increments of width dx. Next, the minimum value of X1 is selected as the initial value of β_3 and the model is regressed against the data to find β_1, β_2 and RMSE. The value of β_3 is then incremented by dx and the regression is repeated until β_3 has traversed the entire range of possible X values. The value of β_3 that results in the lowest RMSE is selected as the initial best-fit change-point. This method is then repeated using a finer grid of width 2 dx, centered about the initial best-fit value of β_3. The uncertainty with which the final change-point temperature is known is reported as the twice the width of the finest grid.

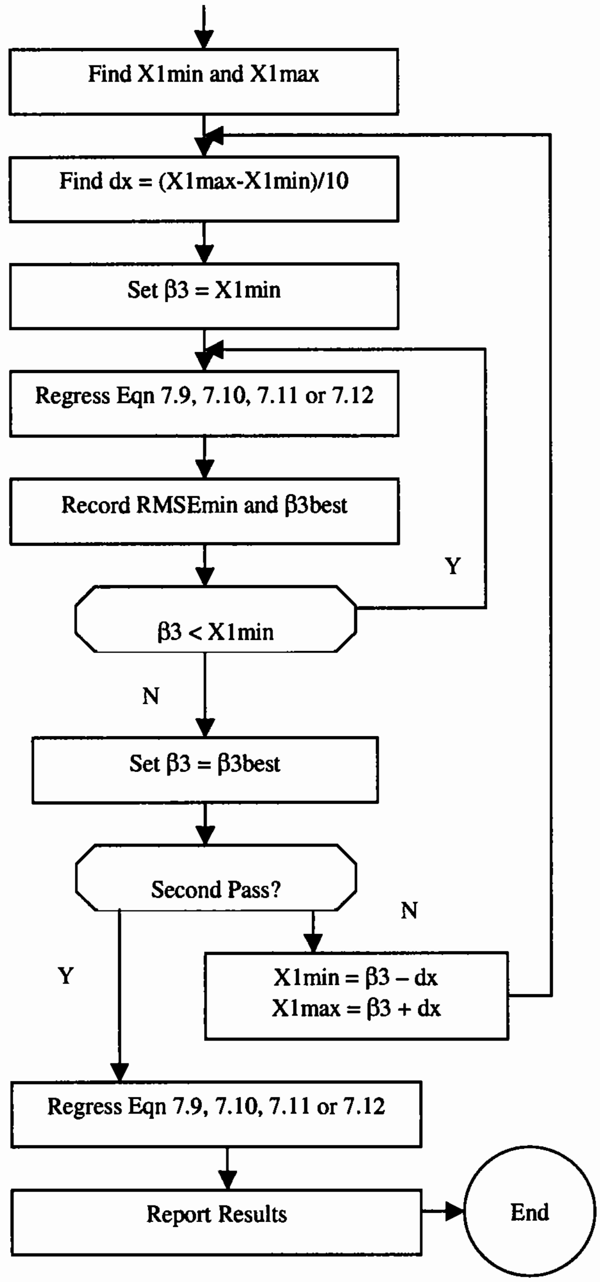

Figure 8.3 Flow diagram of algorithm for finding the best-fit change-point model.

IMT can also find combination three-parameter multi-variable regression models (3P-MVR), with up to four independent variables, of the type:

$$Y_c = \beta_1 + \beta_2 (X_1 - \beta_3)^+ + \beta_4 X_2 + \beta_5 X_3 + \beta_6 X_4 \qquad (8.11)$$

$$Y_h = \beta_1 + \beta_2 (X_1 - \beta_3)^- + \beta_4 X_2 + \beta_5 X_3 + \beta_6 X_4 \qquad (8.12)$$

where X_1 is typically temperature, and X_2, X_3, and X_4 are optional independent variables. The algorithms for finding 3P and 3P-MVR models are identical, the only difference being that the regression model is of the form of Eqns. 8.11 or 8.12 instead of Eqns. 8.9 or 8.10.

When regressing change-point models, the parenthetic + and – terms are computed with the use of an indicator variable, I. For example, in Equation 8.9, the regression equation passed to the Reg subroutine is:

$$Y = a + b\,\chi \tag{8.13}$$

Where χ represents $(X_1 - \beta_3)^+$. The numerical value of χ is computed as:

$$\chi = I\ (X_1 - \beta_3) \quad \text{where } I = 0 \text{ when } X_1 \le \beta_3 \text{ and } I = 1 \text{ when } X_1 > \beta_3 \tag{8.14}$$

IMT uses the same algorithm for finding all best-fit change-point models; the only difference is the form of the regression equation. For example, to find the best-fit 4P model, Equations 4.6 would be substituted for Eqns. 8.9, 8.10, 8.11 or 8.12 in Figure 8.3.

8.5 Variable-Base Degree-Day Model Algorithm

IMT can find best-fit variable-base degree-day (VBDD) models of the type:

$$Y = \beta_1 + \beta_2\,HDD(\beta_3) \tag{8.15}$$

$$Y = \beta_1 + \beta_2\,CDD(\beta_3) \tag{8.16}$$

where β_1 is the constant term, β_2 is the slope term, and $HDD(\beta_3)$ and $CDD(\beta_3)$ are the number of heating and cooling degree-days, respectively, in each energy data period calculated with base temperature β_3. The number of heating and cooling degree-days in each energy data period of n days is:

$$HDD(\beta_3) = \sum_{i=1}^{n} (\beta 3 - T_i)^+ \tag{8.17}$$

$$\mathrm{CDD}(\beta_3) = \sum_{i=1}^{n} (T_i - \beta 3)^+ \qquad (8.18)$$

where T_i is the average daily temperature.

To calculate VBDD models, IMT calls two subroutines: FillDNonUni and VBDD. The FillDNonUni subroutine fills and returns the arrays HDD and CDD with the heating and cooling degree days, respectively, for each energy period according to Eqns. 8.17 and 8.18. HDD(i,j) and CDD(i,j) contain the number of degree days in each energy period (i) and for base temperatures from 41 to 80 F (j). Thus, IMT's VBDD model requires that daily temperatures are reported in degrees Fahrenheit.

The best-fit VBDD model is identified using a search method (Figure 8.4) (Kissock, 1999). In this search method, Eqn. 8.15 or 8.16 is regressed using the HDDs or CDDs in each energy period for successive base temperatures, β3, from 41 F to 80 F. The base-temperature that results in the model with the highest R^2 is recorded. Eqn. 8.15 or 8.16 is regressed once more using the base-temperature that results in the model with the highest R^2, and the results are reported.

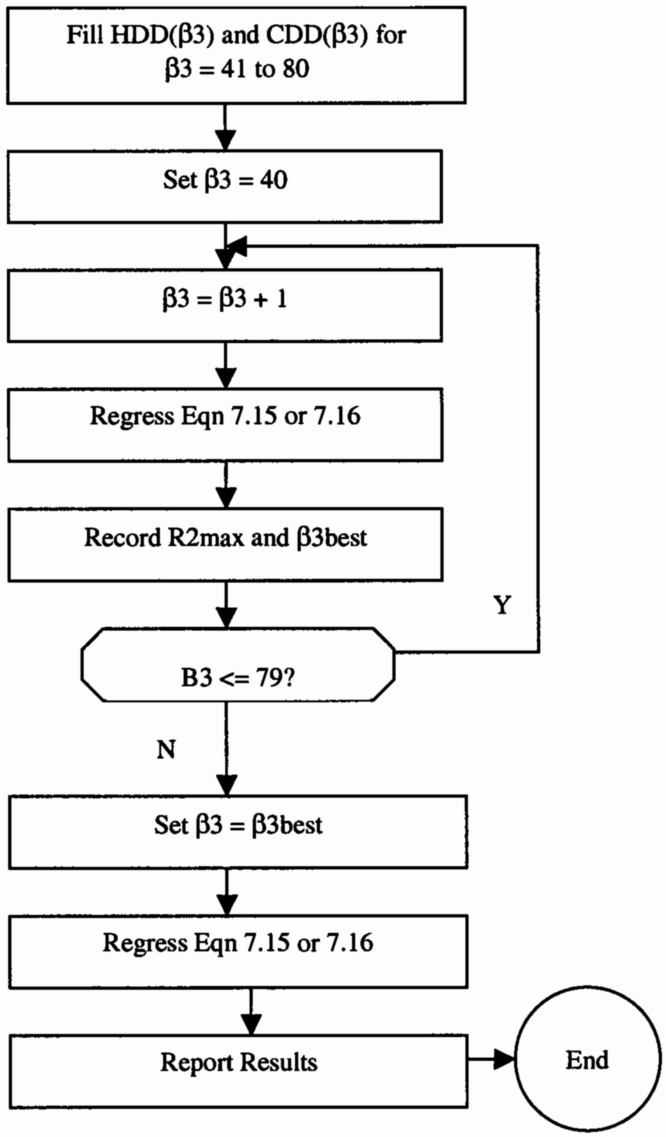

Figure 8.4 Flow diagram of algorithm for finding the best-fit variable-base degree-day model.

9.0 Toolkit Testing

Two classes of testing were performed: bounds testing and accuracy testing. Bounds tests were designed to identify the types of data sets the toolkit can reliably model. To determine these bounds, toolkit models were subjected to: 1) datasets with as few as two and as many as 9,000 observations, 2) data sets containing very large and very small numbers, 3) data sets with a variety of slopes. Section 9.1 describes these tests and summarizes the results.

To test the accuracy and determine the precision of IMT's regression algorithms, IMT mean, 2P and MVR models were benchmarked against the statistical software SAS. IMT change-point model results were compared to the data analysis software EModel and to known coefficients from synthetic data sets. Finally, toolkit HDD and CDD models were compared to PRISM HO and CO models. Section 9.2 describes these tests and summarizes the results.

Section 9.3 summarizes the robustness testing of the IMT's change-point multivariable regression (CP-MVR) and variable-base degree-day, multivariable regression (VBDD-MVR) models.

A complete description of all test results, including a CD-ROM containing the test data sets, IMT, EModel, PRISM, and SAS output is available from the Texas A&M University Energy Systems Laboratory (Sreshthaputra, et al., 2001)

9.1 Bounds Testing

Three sets of 'Bounds' testing were performed. The first set of tests was designed to insure that the toolkit could accurately model data sets containing very small and very large quantities of data. The results of these 'Quantity' tests are summarized in Section 9.1.1. The second set of tests was designed to insure that the toolkit could accurately model data sets containing very small and very large numbers. The results of these 'Magnitude' tests are summarized in Section 9.1.2. The third set of tests was designed to determine whether the toolkit could model data sets with a variety of slopes. The results of these 'Slope' tests are summarized in Section 9.1.3.

9.1.1 Summary of Quantity Testing

IMT was designed to open and read data files containing between 1 and 10,000 records (lines) of data. To test IMT's ability to open and model large data sets, data sets containing 9,000 records were created and each model was tested to see if it could accurately model these data sets. The tests showed that in all cases, valid models were generated. As an example of these tests, the data from the 9,000-record data set are plotted in Figure 9.1, and the results of a 2P model of this data set are shown in Figure 9.2.

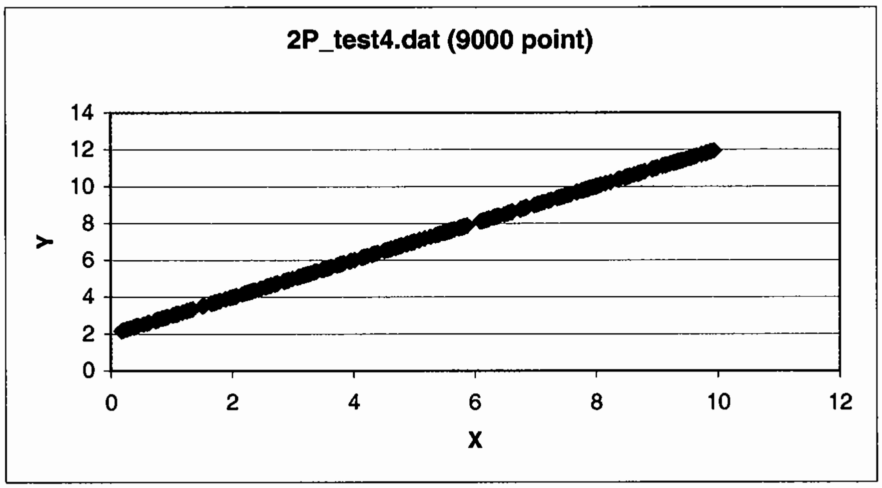

Figure 9.1. 9,000-point data set for 2P model.

```
*******************************************
ASHRAE INVERSE MODELING TOOLKIT (1.0)
*******************************************
 Output file name = IMT.Out
*******************************************
 Input data file name =  2P_test4.dat
 Model type =          2P
 Grouping column No =   0
 Value for grouping =   0
 Residual mode =        0
 # of X(Indep.) Var =   1
 Y1 column number =     2
 X1 column number =   1
 X2 column number =   0 (unused)
 X3 column number =   0 (unused)
 X4 column number =   0 (unused)
 X5 column number =   0 (unused)
 X6 column number =   0 (unused)
*******************************************
 Regression Results
 -------------------------------------
      N =  9000
 -------------------------------------
      R2 =    1.000
 -------------------------------------
    AdjR2 =    1.000
 -------------------------------------
    RMSE =    0.000
 -------------------------------------
  CV-RMSE =    0.000%
 -------------------------------------
      p =    0.876
 -------------------------------------
     DW =    0.248 (p>0)
 -------------------------------------
      a =    2.0000 ( 0.0000)
 -------------------------------------
     X1 =    1.0000 ( 0.0000)
 -------------------------------------
```

Figure 9.2. IMT 2P output for a data set with 9,000 observations.

The minimum number of valid data observations required to generate a valid model depends on the model type. The minimum number of valid data observations required for Mean, 2P and MVR models is equal to the number of regression parameters in the model, n, plus one. Thus, the minimum number of observations required for the Mean model is two, and the minimum number of observations for the 2P model is three. Similarly, the minimum number of observations required for a MVR model of the form Y = a + bX1 + cX2 is four. To verify this, the Mean, 2P and MVR regression models were tested with data sets containing n+1 data observations. The results showed that in each case, a valid model was generated. As an example of these tests, data from a three-record data set are plotted in Figure 9.3, with 2P model results in Figure 9.4.

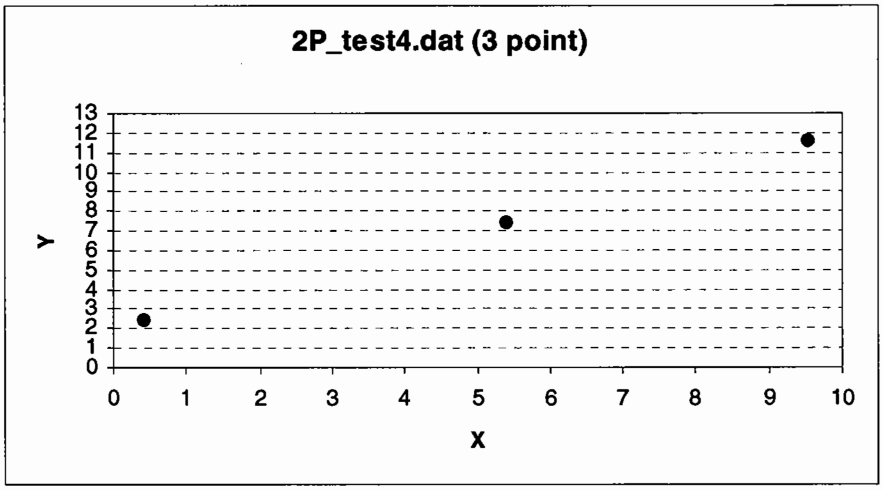

Figure 9.3 Three-point data set for 2P model.

```
*********************************************
ASHRAE INVERSE MODELING TOOLKIT (1.0)
*********************************************
 Output file name = IMT.Out
*********************************************
 Input data file name = 3ptn.txt
 Model type =          2P
 Grouping column No =   0
 Value for grouping =  **
 Residual mode =        0
 # of X(Indep.) Var =   1
 Y1 column number =     2
 X1 column number =   1
 X2 column number =   0 (unused)
 X3 column number =   0 (unused)
 X4 column number =   0 (unused)
 X5 column number =   0 (unused)
 X6 column number =   0 (unused)
*********************************************
 Regression Results
 -------------------------------------
      N =    3
 -------------------------------------
     R2 =    1.000
 -------------------------------------
   AdjR2 =    1.000
 -------------------------------------
    RMSE =    0.000
 -------------------------------------
 CV-RMSE =    0.000%
 -------------------------------------
      p =    0.118
 -------------------------------------
     DW =    0.857 (p>0)
 -------------------------------------
      a =    2.0000 ( 0.0000)
 -------------------------------------
     X1 =    1.0000 ( 0.0000)
 ---------------------------------------------
```

Figure 9.4. IMT 2P output for data seta with three observations.

The minimum number of valid data observations required for change-point models depends on the distribution of the observations. This is because IMT change-point algorithms divide the data into groups and attempt to fit 2P regression models through the data in these groups. Thus, it is possible that, under some circumstances, a change-point model can be constructed from as little as three observations. In most cases, however, the minimum number of observations required by change-point models is five. In general, multi-variable change-point models require four observations, plus one additional observation for every additional independent variable.

To verify that each change-point model could model small data sets, the 3P and 4P models were tested with five observations and the 5P model was tested with seven observations. In each case, valid models were generated. It should be noted, however, that even though a change-point regression model can be identified with a very small number of observations, the model's predictive ability may be limited.

9.1.2 Summary of Magnitude Testing

IMT models were tested with numbers with absolute values as small as 3.3×10^{-57} and as large as 1×10^{18}. In all cases, IMT ran correctly and produced output; however, IMT output fields are FORTRAN F12.4 format. Thus, only values between 9,999,999.9999 and -999999.9999 can be displayed. If a model produces a numerical result outside of this range, the IMT will display stars in the output field. In addition, the smallest numerical result IMT can display is the absolute value of 0.0001. Numerical results smaller than this are displayed as 0.0000. In cases where the numeric results cannot be properly displayed, it is recommended that the user scale the input data to smaller values. For example, the user may want to report 1,000,000 kWh as 1,000 MWh.

9.1.3 Summary of Slope Testing

IMT models were tested using data sets with slopes ranging from zero to infinite. The data sets used in slope testing are shown in Figures 9.5 – 9.7.

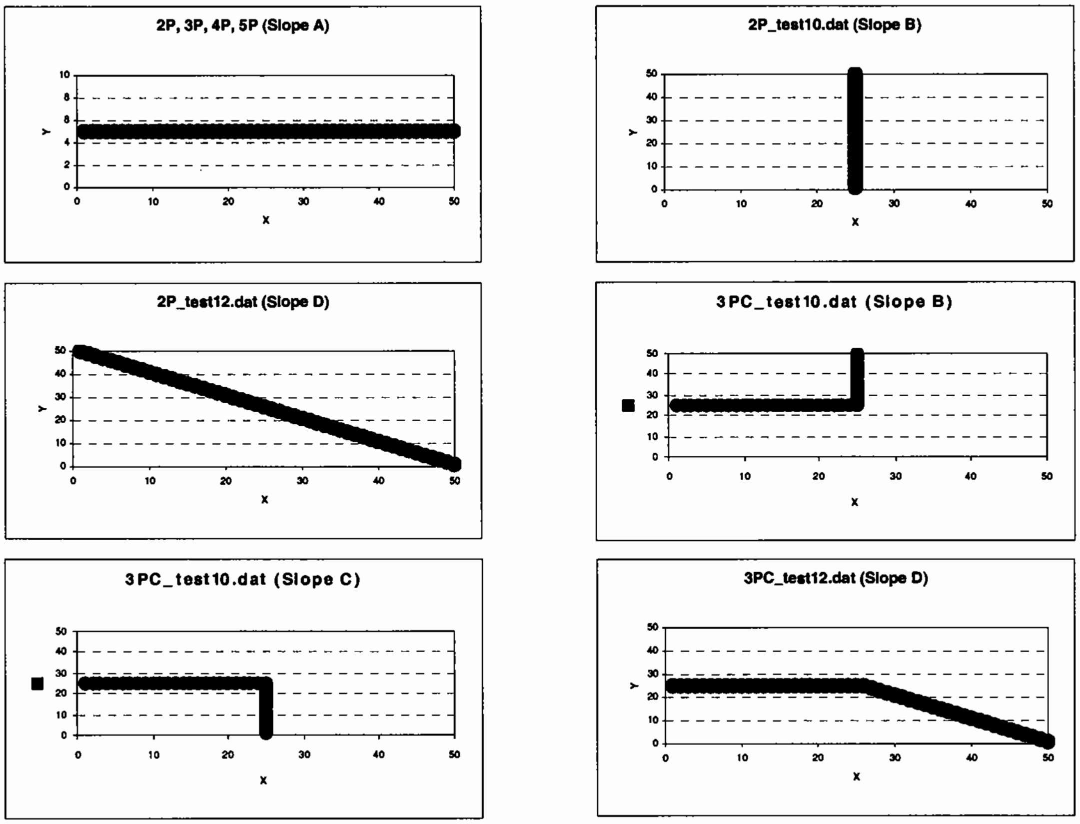

Figure 9.5. Slope testing for 2P and 3P models. IMT failed on Slope A, correctly modeled 2P-Slope D and 3P-Slope D, and could not correctly identify the vertical slopes in 2P-Slope B, 3P-Slope B and 3P-Slope C. IMT was able to correctly model data in the shapes shown in Figure 5.1.

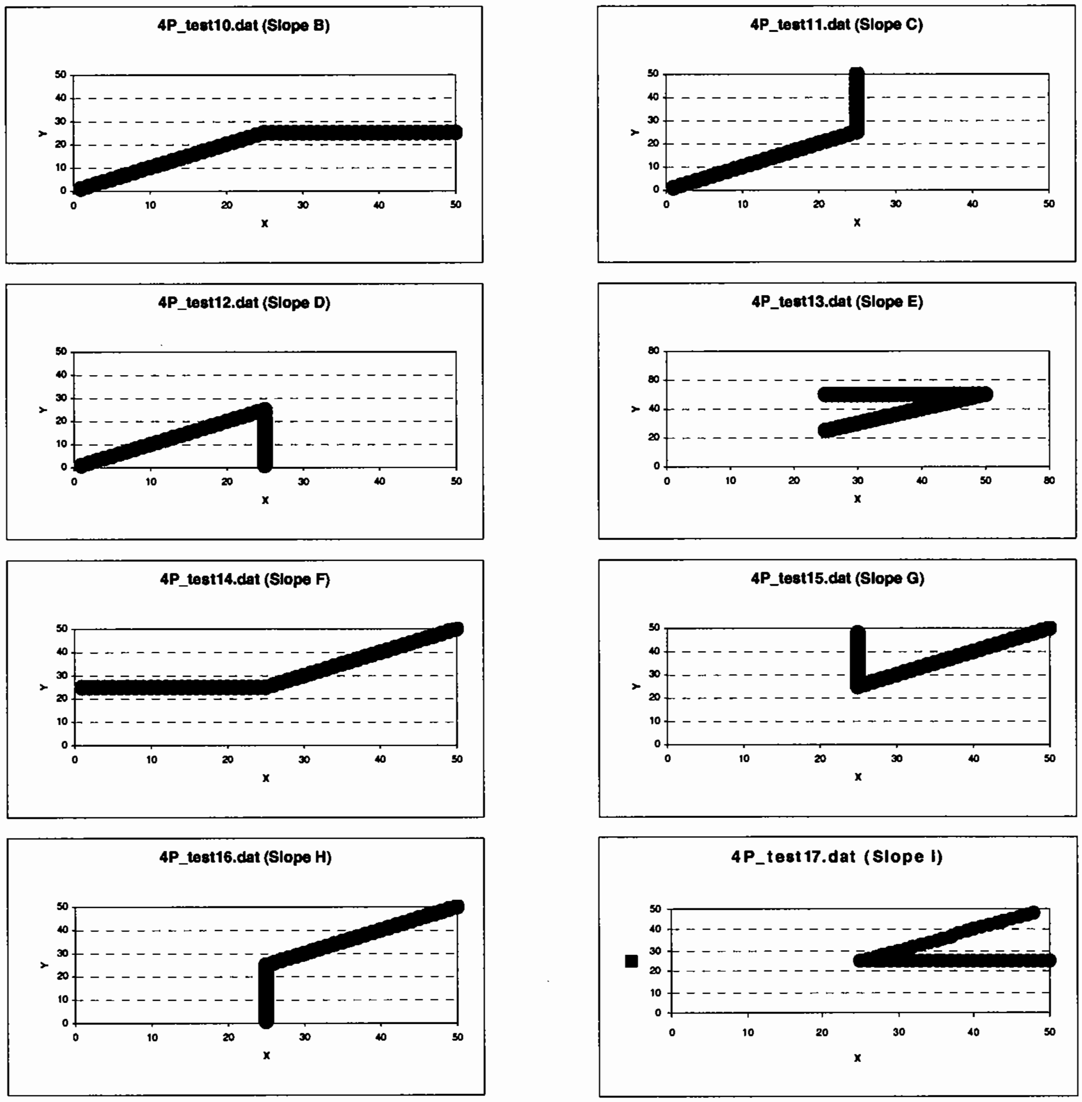

Figure 9.6. Slope testing for 4P models. IMT correctly modeled Slope B and F, could not correctly identify the vertical slopes in Slope C, D, G or H, and identified the equivalent of a 2P model through the middle of Slopes E and I.

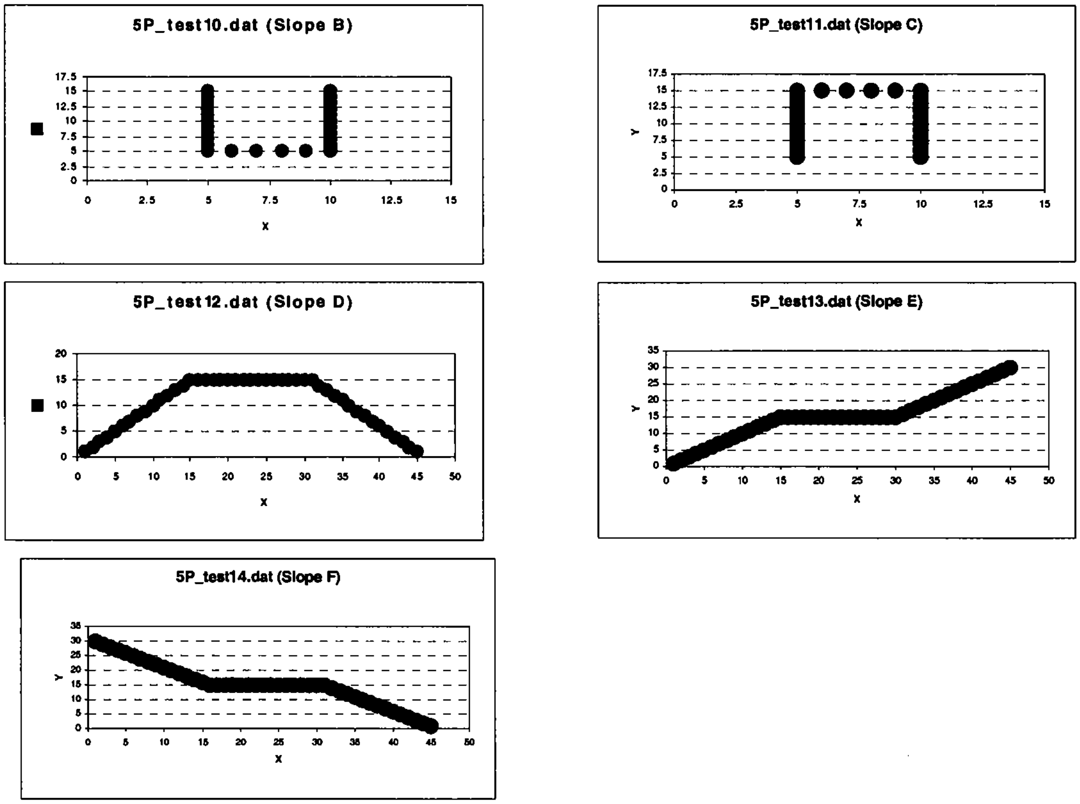

Figure 9.7. Slope testing for 5P models. IMT correctly modeled Slopes D, E and F, but could not correctly identify the infinite slopes in Slope B and C. IMT was able to correctly model data in the shapes shown in Figure 5.1.

In all cases, IMT successfully modeled data sets with non-zero and non-infinite slopes. In general, however, data sets containing data with zero and infinite slopes cannot be modeled by IMT. IMT closes without generating a useful error message when attempting to identify 2P or change-point models from data that form a horizontal line with zero slope. When attempting to identify a 2P model from data that form a perfectly vertical line (infinite slope), the reported values for the intercept and slope are unreliable. Similarly, when attempting to identify a change-point model from data that form a perfectly vertical line (infinite slope), the reported value for the change point(s) is close to the actual change point(s); however the value reported for the slope(s) is unreliable. Thus, as in all regression analysis, the user is advised to plot the data before modeling it to identify data sets that are inappropriate for the regression models.

9.2 Accuracy Testing

Four sets of accuracy and precision tests were performed. The first set of tests was designed to test the accuracy and precision of IMT's computational and regression engines by comparing IMT results with results from the widely used software SAS. The results of these tests are shown in section 9.2.1. In the second set of tests, IMT 3P, 4P and 5P change-point model results were compared to model results from EModel. The primary purpose of this set of tests test was to confirm that IMT is finding the same results as the well-tested EModel. The results of these tests are summarized in section 9.2.2. The third set of accuracy tests was designed to see how closely IMT change-point models could identify known change-points and slopes. These results are described in section 9.2.3. In the fourth set of accuracy tests, IMT variable-base heating and cooling degree-day models were compared to PRISM HO and CO models. These results are summarized in section 9.2.4.

9.2.1 Comparisons with SAS

To test the accuracy and determine the precision of IMT's regression algorithms, IMT mean, 2P and MVR models were benchmarked against the statistical software SAS (SAS, 2001). Results from these tests are shown in Tables 9.8 – 9.10. The results show good agreement between IMT and SAS to at least 4 significant figures in the regression parameters tested.

The data sets used in the mean, 2P, 3P, 4P and 5P comparison tests with SAS were from the Texas LoanSTAR database (Haberl et al., 1998). The data sets contain daily energy consumption, temperature, solar and humidity data. The data set used in the MVR comparison with SAS was also from the Texas LoanSTAR database and contains hourly energy consumption, temperature, solar and humidity data from the Texas A&M Zachry Engineering Center. The data sets used in the HDD and CDD comparisons with PRISM are from a residence in College Station, Texas and contains measured energy consumption and average daily temperature data.

Table 9.8. Mean model comparisons between IMT and SAS.

	IMT	SAS	IMT	SAS
Data Set	711	711	963	963
N	356	356	315	315
Mean	25409.281	25409.2809	1118.391	1118.39048
Std dev	2391.109	2391.1088	452.392	452.392131

Table 9.9. 2P model comparisons between IMT and SAS.

	IMT	SAS	IMT	SAS	IMT	SAS
Data Set	226	226	201	201	952	952
N	364	364	309	309	264	264
R2	0.834	0.8338	0.691	0.6906	0.728	0.728
RMSE	3082.120	3082.12038	5704.015	5704.01365	2065.280	2065.27929
A	-10227.1260	-10229	68439.5078	68439	2338.5520	2338.41956
Std (a)	791.0164	791.03445	1764.5883	1764.58622	595.1532	595.14173
X1	470.2920	470.31234	-649.0869	-649.08587	212.1381	212.13968
Std (X1)	11.0378	11.03807	24.7934	24.79339	8.0110	8.01107
	IMT	SAS	IMT	SAS		
Data Set	207-2	207-2	207-3	207-3		
N	361	361	309	309		
R2	0.861	0.8607	0.691	0.6906		
RMSE	577.486	577.48537	1932.921	1932.92099		
A	-5041.9448	-5041.74522	23192.1055	23192		
Std (a)	154.5213	154.51887	597.9664	597.96592		
X1	105.2378	105.23485	-219.9567	-219.95630		
Std (X1)	2.2346	2.23459	8.4018	8.40175		

Table 9.10. MVR model comparison between IMT and SAS.

	IMT	SAS	IMT	SAS	IMT	SAS
Data Set	MVR_1.dat	MVR_1.dat	MVR_2.dat	MVR_2.dat	MVR_3.dat	MVR_3.dat
N	8423	8423	8423	8423	8423	8423
R2	0.438	0.437626	0.873	0.873095	0.582	0.582490
ADJ R2	0.438	0.437425	0.873	0.873019	0.582	0.582242
A	473.3068	473.228982	-2354.9236	-2354.82060	1428.1320	1428.56701
Std (a)	6.8794	6.8800665	35.4836	35.482222	19.8274	19.8343238
X1	4.6356	4.637366	28.4639	28.47099	-12.9370	-12.94613
Std (X1)	0.1385	0.1384824	0.4889	0.488972	0.3020	0.3021457
X2	-1992.5603	-1995.942269	72907.6406	72901.36995	-16776.8887	-16788.49610
Std (X2)	333.7043	333.7266148	1024.9117	1024.973306	723.7756	724.0103164
X3	0.1754	0.175387	-0.0178	-0.01776	0.1297	0.12969
Std (X3)	0.0045	0.0045264	0.0149	0.014853	0.0084	0.0084356
X4			1.8636	1.86297	-0.2883	-0.28844
Std (X4)			0.0327	0.032671	0.0219	0.0219256
X5			0.1787	0.17879	0.0591	0.05928
Std (X5)			0.0188	0.018829	0.0062	0.0062429

9.2.2 Comparison of Change-Point Models Parameters with EModel

The 3P, 4P and 5P change-point model results were compared to model results from EModel (Tables 9.8-9.10). EModel was developed by the Texas A&M Energy Systems Laboratory and has been used extensively in the Texas LoanSTAR program. Because IMT and EModel use many of the same algorithms, good agreement between the results is expected. The primary purposes of this test were: 1) to confirm that IMT is finding the same results as the well-tested EModel, and 2) to understand how much the results would vary given that EModel is compiled and executed in Visual Basic and IMT is compiled and executed in FORTRAN.

In version IMT Beta 1.6, the version used for this round of testing, the output field of the standard errors of the coefficients was FORTRAN F7.4 format; hence standard errors with values greater than 99.9999 were not printed. In IMT release version 1.0, standard errors use FORTRAN F12.4 format, and standard errors between 9,999,999.999 and –9,999,999.999 can be printed.

Table 9.8. 3PC and 3PH model comparisons between IMT and EModel.

	IMT	EModel	IMT	EModel	IMT	EModel	IMT	EModel
Data Set	706	706	208	208	707	707	208-h	208-h
Model	3PC	3PC	3PC	3PC	3PH	3PH	3PH	3PH
N	358	358	358	358	365	365	357	357
R2	0.339	0.34	0.855	0.85	0.934	0.93	0.951	0.95
RMSE	870.641	870.64	4123.559	4123.56	37331.309	37331.31	1821.642	1821.64
Ycp	2417.5938	2417.5941	11145.4775	11145.4541	10248.5869	10248.4245	6001.7227	6001.7111
Std (Ycp)	55.9475	55.9475	**	332.5046	**	2671.7098	**	134.6210
LS	0.0000	0.0000	0.0000	0.0000	-8369.0127	-8369.0158	-639.4753	-639.4761
Std (LS)	0.0000	0.0000	0.0000	0.0000	**	116.5216	7.6691	7.6692
RS	87.6157	87.6158	945.5939	945.5957	0.0000	0.0000	0.0000	0.0000
Std (RS)	6.4898	6.4898	20.6595	20.6596	0.0000	0.0000	0.0000	0.0000
Xcp	56.76	56.76	59.7600	59.7600	61.6800	61.68	79.92	79.92
Std (Xcp)	1.6400	-	1.2600	-	1.6400	-	1.26000	-

Table 9.9. 4P model comparisons between IMT and EModel.

	IMT	EModel	IMT	EModel	IMT	EModel
Data Set	706	706	975	975	201	201
Model	4P	4P	4P	4P	4P	4P
N	358	358	279	279	344	344
R2	0.873	0.87	0.816	0.82	0.754	0.75
RMSE	870.641	3861.0432	263.748	263.7481	8051.226	8051.2246
Ycp	17613.2813	17613.7419	1529.8441	1529.8660	27831.1035	27831.8523
Std (Ycp)	**	1678.8280	**	193.4880	**	4807.0528
LS	343.6089	343.6469	16.8140	16.8142	562.6268	562.7605
Std (LS)	28.2882	28.2900	3.0282	3.0284	86.9673	86.9757
RS	1081.8597	1081.8386	73.5243	73.5246	1278.6936	1278.6456
Std (RS)	61.3571	61.3601	5.7500	5.7504	**	154.6463
Xcp	68.5800	68.5800	69.1200	69.12	61.6000	61.6000
Std (Xcp)	1.2600	-	1.1600	-	1.1000	-

Table 9.10. 5P model comparisons between IMT and a development version of EModel which includes a 5P model.

	IMT	EModel	IMT	EModel
Data Set	710	710	210	210
Model	5P	5P	5P	5P
N	348	348	362	362
R2	0.274	0.27	0.699	0.70
RMSE	2943.540	2943.5408	11259.003	11259.004
Xcp1	58.7007	58.7007	62.0000	62.0000
std (Xcp1)	3.0340	3.0370	2.3310	2.3333
Xcp2	61.7438	61.7438	69.0000	69.0000
std (Xcp2)	3.034	3.0370	2.3310	2.3333
Ycp	11665.7363	11665.7289	100499.6250	100499.7043
std (Ycp)	**	289.2256	**	966.8825
LS	-120.6786	-120.6790	-635.1901	-635.1859
std (LS)	11.7461	11.7462	**	102.6701
RS	47.0371	47.0376	2534.415	2543.4064
std (RS)	36.6051	36.6051	90.1323	90.1324

9.2.3 Comparison of Change-point Model Parameters with Known Parameters

In the this set of tests, IMT change-point models were derived from synthetic data sets with known slopes and change-points (Tables 9.11 – 9.13). Graphs of the synthetic data sets are shown in Figures 9.5 – 9.7. In all cases, the change-point model parameters were close to the known values, demonstrating IMT's ability to accurately model these data sets. Precise agreement is, of course, not expected because of the search and regression methods used by the toolkit.

This set of tests also points out important information about how to interpret the standard errors of the IMT change-point and slope coefficients. All IMT change-point algorithms use a two-part grid-search method, in which regression models are identified for successive change-points, until the model (and change point) that produced the lowest RMSE is identified. The standard error of the X change point(s) reported by IMT is the one-half of the width of the finest search interval. Thus, it is expected that the true value of the X change point(s) should be within the region defined by the X change point(s) plus or minus the standard error of the X change point(s). The testing reported here confirmed this expectation.

The standard errors of the other coefficients are computed using standard least-square regression methods, and can be interpreted as indicating that with 68% confidence, the true value of the coefficient is within one standard error of the reported value. However, these methods implicitly assume that the change point(s) is completely known. Thus, the standard errors on the slope and Y change point coefficients reported by IMT reflect the uncertainty of the coefficient due to the scatter of data around the regression line, but underestimate the true uncertainty with which the coefficient is known since the true value of the change-point is not exactly known.

This result is demonstrated in following tests. Consider, for example, the test results from Table 9.11 for the 3PC model. IMT reports the right slope as -1.0244 ± 0.0028, even though the true slope of -1 is outside of this interval. It is possible that the true slope may be outside of this interval even if the change-point were precisely known. However, in most cases, the reason that the standard errors of the slope and Ycp coefficients underestimate the true uncertainties is because of the uncertainty of the X change point(s).

Table 9.11. 3PC and 3PH model comparison between IMT and synthetic data sets with known slopes and change points. Data sets are identified by slope (S) and test (T). Thus, Data Set: S-D, T-12 indicates slope D, test 12 for the indicated model. A graph of Data Set: S-D, T-12 is shown in Figure 9.5.

	IMT	Known	IMT	Known
Data Set	S-D, T-12	S-D, T-12	S-D, T-12	S-D, T-12
Model	3PC	3PC	3PH	3PH
N	50	50	50	50
R2	1.000		1.000	
RMSE	0.153		0.153	
Ycp	24.9103	25	24.9103	25
std (Ycp)	.0270		.0270	
LS	0.0000	0	-1.0244	0
std (LS)	0.0000		0.0028	
RS	-1.0244	-1	0.0000	-1
std (RS)	0.0028		0.0000	
Xcp	26.7800	25	24.52	25
std (Xcp)	0.9800		0.9800	

Table 9.12. 4P model comparison between IMT and synthetic data sets with known slopes and change points. Data sets are identified by slope (S) and test (T). Thus, Data Set: S-B, T-10 indicates slope B, test 10 for the indicated model. Graphs of these data sets are shown in Figure 9.6.

	IMT	Known	IMT	Known
Data Set	S-B, T-10	S-B, T-10	S-F, T-14	S-F, T-14
Model	4P	4P	4P	4P
N	50	50	50	50
R2	1.000		1.000	
RMSE	0.124		0.124	
Ycp	24.7702	25	24.7498	25
std (Ycp)	0.9989		0.0862	
LS	1.0156	1	-0.0156	0
std (LS)	0.0028		0.0028	
RS	0.0133	0	0.9867	1
std (RS)	0.0056		0.0056	
Xcp	24.5200	25	24.5200	25
std (Xcp)	0.9800	-	0.9800	-

Table 9.13. 5P model comparison between IMT and synthetic data sets with known slopes and change points. Data sets are identified by slope (S) and test (T). Thus, Data Set: S-D, T-12 indicates slope D, test 12 for the indicated model. Graphs of these data sets are shown in Figure 9.7.

	IMT	Known	IMT	Known	IMT	Known
Data Set	S-D, T-12	S-D, T-12	S-E, T-13	S-E, T-13	S-F, T-14	S-F, T-14
Model	5P	5P	5P	5P	5P	5P
N	50	50	50	50	45	45
R2	0.998		1.000		0.999	
RMSE	0.216		0.147		.184	
Xcp1	15.667	15	15.6667	15	15.6667	16
std (Xcp1)	1.6280		1.6280		1.6280	
Xcp2	30.3333	31	30.3333	30	30.3333	31
std (Xcp2)	1.6280		1.6280		1.6280	
Ycp	15.2201	15	15.1639	15	15.1626	15
std (Ycp)	0.0457		0.0368		0.0390	
LS	0.9558	1	0.9502	1	-1.0169	-1
std (LS)	0.0078		0.0063		0.0066	
RS	-0.9582	-1	1.0159	1	-0.9501	-1
std (RS)	0.0066		0.0053		0.0066	

9.2.4 Comparisons of HDD and CDD Models with PRISM

IMT finds the best-fit models between energy use and variable-base heating or cooling degree days. Perhaps the most widely used method of correlating building energy use to variable-base degree days is the PRISM method. In this section, we compare IMT HDD and CDD model results to PRISM HO and CO model results.

Model parameters and coefficients for IMT and PRISM HDD and CDD runs are shown in Table 9.14. The energy use data are from a residence in College Station, Texas. As can be seen, all parameters and coefficients are in general agreement, considering the different algorithms used by the two methods; however, some comments are called for. First, IMT finds the base temperature to the nearest whole degree, whereas PRISM finds it to at least one decimal place. In addition, PRISM results are reported with a maximum precision of 1 decimal place, thus it is difficult to exactly compare the slopes in the HDD/HO models in Table 9.14. Finally, PRISM units for base level energy use (A in IMT, alpha in PRISM) are 'energy units per day', whereas IMT units are 'energy units per period'. Thus, to compare IMT and PRISM output, base-level energy use in IMT, A, is divided by the average days per billing period to generate a "corrected" base level energy use A,c. The values for IMT Ac and std(Ac) should and do compare to the PRISM's alpha and std(alpha).

Table 9.14. Model parameters and coefficients for IMT and PRISM runs.

HDD/HO	IMT	PRISM	IMT	PRISM
Model	HDD	HO	CDD	CO
Data Set	1308ngk	1308ngk	1308elk	1308elk
N	12	12	12	12
R2	0.969	0.977	0.897	0.901
DD Base	73	71.8	70	70.0
X1	0.1368	0.1	1.9555	1.9
Std(X1)	0.0077	0.0	0.2092	0.6
A	18.2518	0.7	476.1488	16.0
Std(A)	2.7116	0.1	53.0032	2.1
A,c	0.5951	0.7	15.6114	16.0
Std(A,c)	0.0884	0.1	1.7378	2.1
days/period	30.67		30.50	

Over twenty other comparisons between IMT and PRISM heating and cooling models were also run, with good agreement between IMT and PRISM on each run. These results are summarized in Tables 9.15 and 9.16 below.

Table 9.15. Summary of IMT HDD and PRISM HO comparisons.

Test No.		Data Type	Test	.INS file	Data File	IMT1.6	PRISM
1	VBDD_H0	GAS	HDD	VBDD_H0.ins	VBDD_H0.dat	OK	agree
2	VBDD_H1	GAS	HDD	VBDD_H1.ins	VBDD_H1.dat	OK	agree
3	VBDD_H2	GAS	HDD	VBDD_H2.ins	VBDD_H2.dat	OK	agree
4	VBDD_H3	GAS	HDD	VBDD_H3.ins	VBDD_H3.dat	OK	agree
5	VBDD_H4	GAS	HDD	VBDD_H4.ins	VBDD_H4.dat	OK	agree
6	VBDD_H5	GAS	HDD	VBDD_H5.ins	VBDD_H5.dat	OK	close
7	VBDD_H6	GAS	HDD	VBDD_H6.ins	VBDD_H6.dat	OK	agree

Table 9.16. Summary of IMT CDD and PRISM CO comparisons.

Test No.		Data Type	Test	.INS file	Data File	IMT1.6	PRISM
1	VBDD_C0	WBE	CDD	VBDD_C0.ins	VBDD_C0.dat	OK	agree
2	VBDD_C1	WBE	CDD	VBDD_C1.ins	VBDD_C1.dat	OK	agree
3	VBDD_C2	WBE	CDD	VBDD_C2.ins	VBDD_C2.dat	OK	agree
4	VBDD_C3	WBE	CDD	VBDD_C3.ins	VBDD_C3.dat	OK	agree
5	VBDD_C4	WBE	CDD	VBDD_C4.ins	VBDD_C4.dat	OK	agree
6	VBDD_C5	WBE	CDD	VBDD_C5.ins	VBDD_C5.dat	OK	agree
7	VBDD_C6	WBE	CDD	VBDD_C6.ins	VBDD_C6.dat	OK	agree
8	VBDD_C7	WBE	CDD	VBDD_C7.ins	VBDD_C7.dat	OK	agree
9	VBDD_C8	WBE	CDD	VBDD_C8.ins	VBDD_C8.dat	OK	agree
10	VBDD_C9	WBE	CDD	VBDD_C9.ins	VBDD_C9.dat	OK	agree
11	VBDD_C10	WBE	CDD	VBDD_C10.ins	VBDD_C10.dat	OK	agree
12	VBDD_C11	WBE	CDD	VBDD_C11.ins	VBDD_C11.dat	OK	agree
13	VBDD_C12	WBE	CDD	VBDD_C12.ins	VBDD_C12.dat	OK	agree
14	VBDD_C13	WBE	CDD	VBDD_C13.ins	VBDD_C13.dat	OK	agree

9.3 Testing of CP-MVR Models

IMT's change-point multi-variable regression (CP-MVR) models were tested for robustness, to identify the maximum number of independent variables that can be used with each model, and for their ability to model synthetic data sets with multiple independent variables. Robustness testing is described in section 9.3.1. Testing to determine the maximum number of independent variables is described in section 9.3.2. Testing with synthetic data sets with multiple independent variables is described in section 9.3.3.

9.3.1 Robustness Testing of CP-MVR Models

For the robustness testing, the 3PC-MVR, 3PH-MVR, 4P-MVR and 5P-MVR models were run using a data set of hourly energy consumption and weather data from the Texas A&M Zachry Engineering Center. The data set contains 8,423 records. Eleven 3P-MVR, eleven 4P-MVR and five 5P-MVR models were generated using different combinations of the independent variables. In all cases, the CP-MVR models generated error-free results.

9.3.2 Maximum Number of Independent Variables Testing of CP-MVR Models

The maximum number of independent variables allowed by each CP-MVR model is equal to the seven minus the number of change-point regression parameters.
For example, 3P-MVR models can be handle up to four independent variables; a 3P model is fit to the first independent variable and linear regression coefficients are fit to the three additional independent variables. Testing confirmed that the 3P-MVR models can handle four independent variables, the 4P-MVR model can handle three independent variables, and the 5P-MVR model can handle two independent variables.

9.3.3 Synthetic Data Testing of CP-MVR Models

Building energy use frequently varies non-linearly with outside air temperature and is affected by other variables such as occupancy and internal loads. IMT's CP-MVR models were designed to handle these cases. In this section, the 3P-MVR, 4P-MVR and 5P-MVR models were tested to determine if the CP-MVR models could indeed produce a better fit than either the

CP or MVR models alone. The testing was performed using synthetic data sets to determine how accurately the CP-MVR models could identify the true model parameters.

The first test compares 2P and MVR models to provide a point of reference for comparisons between CP, MVR and CP-MVR models. To test the 2P and MVR models, a data set was synthesized to represent a building in which chilled water energy use, CW, varies linearly with outside air temperature, T, and whole building electricity use, WBE. The specific relation used to synthesize the data was:

$$CW = -25 + 1.0\,T + 0.003\,WBE \qquad (9.1)$$

Outside air temperature T was varied from 40 to 95. WBE was set to 7,000 for five observations then 3,000 for two observations, corresponding to a weekday/weekend pattern of electricity use. A plot of the synthetic chilled water energy use CW versus outside air temperature T is shown in Figure 9.8. The weekday/weekend pattern of chilled water use is clearly evident.

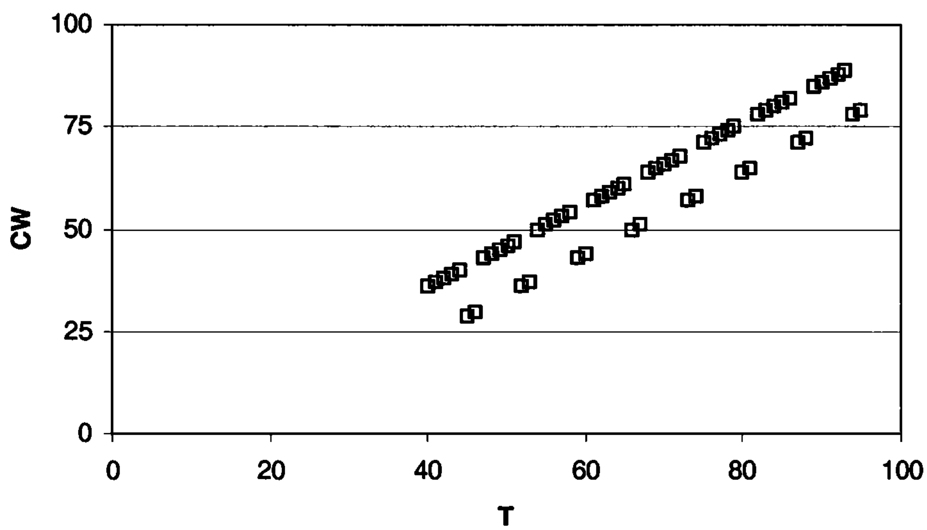

Figure 9.8 Synthesized CW versus T in a 2P-MVR pattern.

Next, a 2P model using only T, and a MVR model using both T and WBE as independent variables were run. The results demonstrate that the MVR model provided a superior fit to the data and was able to approximate the true model coefficients with good accuracy (Table 9.17).

Table 9.17. Comparison of 2P and MVR models using synthetic data.

	Synthetic Coefficients	2P Model	MVR Model
R2		0.894	1.000
CV-RMSE %		9.146	0.000
Constant	-25	-5.2137 (3.1527)	-24.9995 (0.0001)
T	1.0	0.9672 (.0454)	1.0000 (0.0000)
WBE	0.003		0.003 (0.0000)

To test the 3P-MVR models, data sets were synthesized to represent buildings in which whole-building electricity use, WBE, varies in 3P patterns with outside air temperature, T, and, in addition, also varies linearly with some other measure of occupancy, OCC. The specific relations used to synthesize the data were:

$$WBE = 50 + 1.0\ (T\text{-}67)^{+} + 0.003\ OCC \tag{9.2}$$

$$WBE = 50 + -1.0\ (T\text{-}67)^{-} + 0.003\ OCC \tag{9.3}$$

As before, outside air temperature T was varied from 40 to 95. OCC was set to 7,000 for five observations then 3,000 for two observations, corresponding to a weekday/weekend pattern of occupancy. The first data set (as generated by Eqn. 9.2) could represent a building with electric air conditioning and non-electric heating, and the second data set (as generated by Eqn. 9.3) could represent a building with electric heating and no air conditioning. Plots of the synthetic whole building electricity use WBE versus outside air temperature T are shown in Figure 9.9. The weekday/weekend patterns of energy use are clearly evident.

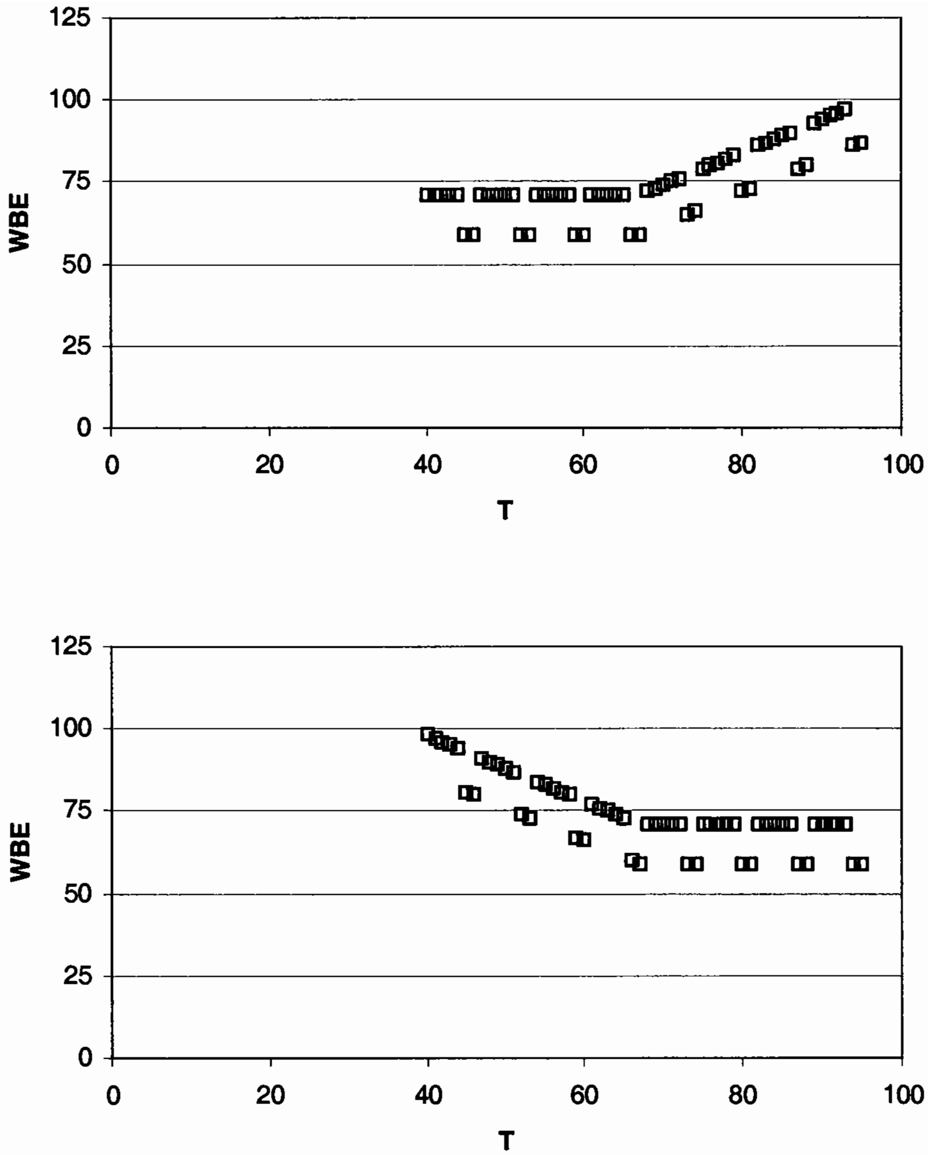

Figure 9.9. WBE versus T data in a 3PC-MVR pattern (upper figure) and a 3PH-MVR pattern (lower figure).

Next a 3P model using temperature, a MVR model using temperature and occupancy, and a combination 3P-MVR model using temperature and occupancy as independent variables were run on each data set. The results show that neither the 3P or MVR models provided as good as fit as the combination 3P-MVR models (Tables 9.18 and 9.19). In addition, the 3P-MVR models were able to approximate the true model coefficients with good accuracy.

Table 9.18. Comparison of 3PC, MVR and 3PC-MVR models using synthetic data.

	Synthetic Coefficients	3PC	MVR	3PC-MVR
R2		0.725	0.846	1.000
CV-RMSE (%)		7.352	5.551	0.216
Ycp	50	67.8155 (0.9411)	22.5180 (3.1344)	50.0435 (0.0767)
RS	1.0	0.9266 (0.0777)	5.135 (0.0345)	1.0216 (0.0024)
Xcp	67	66.4000 (1.1000)		67.5000 (1.1000)
OCC	0.003		0.0030 (0.0003)	0.0030 (0.0000)

Table 9.19. Comparison of 3PH, MVR and 3PH-MVR models using synthetic data.

	Synthetic Coefficients	3PC	MVR	3PC-MVR
R2		0.749	0.859	1.000
CV-RMSE (%)		7.397	5.589	0.217
Ycp	50	67.3199 (0.9186)	89.5176 (3.1344)	49.8460 (0.0740)
LS	-1.0	-1.0837 (0.0854)	-0.4865 (0.0345)	-0.9784 (0.0024)
Xcp	67	66.4000 (1.1000)		67.5000 (1.1000)
OCC	0.003		0.0030 (0.0003)	0.0030 (0.0000)

To test the 4P-MVR model, a data set was synthesized to represent a building in which chilled water energy use, CW, varies in a 4P pattern with outside air temperature T, and linearly with whole building electricity use WBE. The relation used to synthesize the data was:

$$CW = 50 + 0.5\,(T\text{-}67)^{-} + 1.5\,(T\text{-}67)^{+} + 0.003\ WBE \qquad (9.4)$$

Outside air temperature T was varied from 40 to 95. WBE was set to 7,000 for five observations then 3,000 for two observations, corresponding to a weekday/weekend pattern of electricity use. A plot of the synthetic chilled water energy use CW versus outside air temperature T is shown in Figure 9.10. The weekday/weekend pattern of CW use is clearly evident.

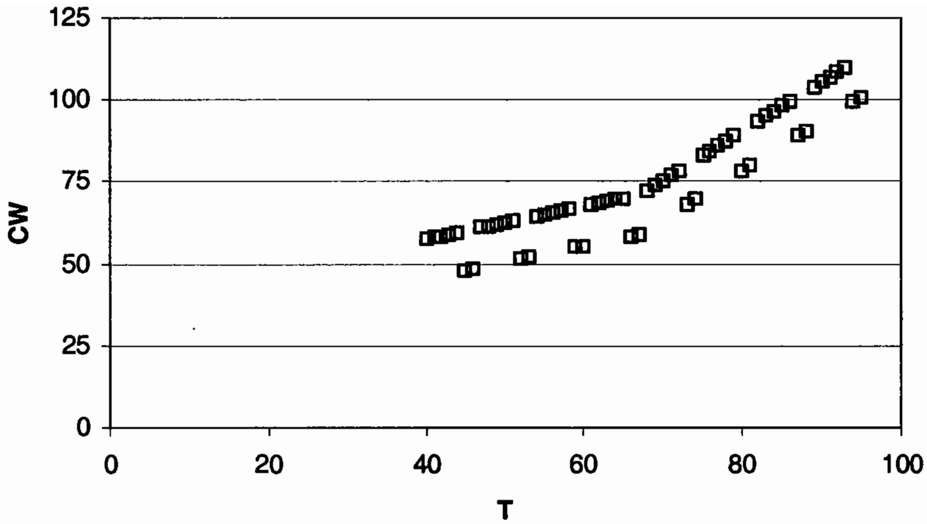

Figure 9.10 Synthesized CW versus T data in a 4P-MVR pattern.

Next a single 4P model using only T, a single MVR model using both T and WBE, and a combination 4P-MVR model using both T and WBE as independent variables were run. The results demonstrate that the 4P-MVR model provided a much superior fit to the data than either the 4P model or MVR model run separately, and was able to approximate the true model coefficients with good accuracy (Table 9.20).

Table 9.20. Comparison of 4P, MVR and 4P-MVR models using synthetic data.

	Synthetic Coefficients	4P	MVR	4P-MVR
R2		.902	0.945	1.000
CV-RMSE (%)		7.387	5.533	0.170
Ycp	50	66.9314 (9.3845)	-10.9819 (3.1344)	50.4310 (0.6055)
LS	0.5	0.4448 (0.1075)	1.0135 (0.0345)	0.5135 (0.0024)
RS	1.5	1.4568 (0.2128)	0.0030 (0.0003)	1.5135 (0.0048)
Xcp	67	66.4000 (1.1000)		67.5000 (1.1000)
WBE	0.003			0.0030 (0.0000)

To test the 5P-MVR model, a data set was synthesized to represent a building in which Whole Building Electricity Use, WBE, varies in a 5P pattern with outside air temperature, T, and linearly with some other measure of occupancy, OCC. The specific relation used to synthesize the data was:

$$WBE = 50 + -0.5\,(T\text{-}60)^{-} + 1.5\,(T\text{-}75)^{+} + 0.003\ OCC \qquad (9.5)$$

Outside air temperature T was varied from 40 to 95. OCC was set to 7000 for five observations then 3,000 for two observations, corresponding to a weekday/weekend pattern of electricity use. A plot of the Whole Building Electricity Use, WBE, versus outside air temperature, T, is shown in Figure 9.11. The weekday/weekend pattern of Whole Building Electricity Use, WBE, is clearly evident.

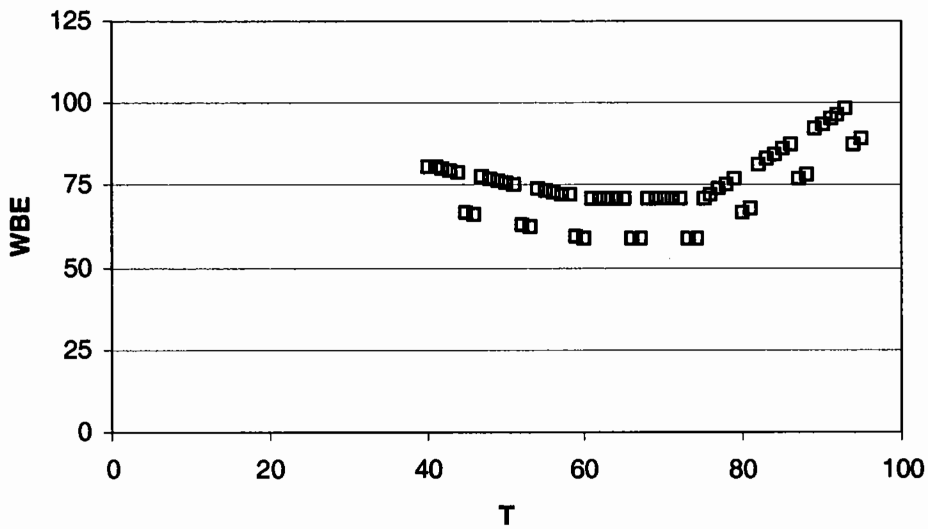

Figure 9.11 Synthesized WBE versus T in a 5P-MVR pattern.

Next a single 5P model using only T, a single MVR model using both T and WBE, and a combination 5P-MVR model using both T and WBE as independent variables were run. The results demonstrate that the 5P-MVR model provided a much superior fit to the data than either the 4P model or MVR model run separately, and was able to approximate the true model coefficients with good accuracy (Table 9.21).

Table 9.21. Comparison of 5P, MVR and 5P-MVR models using synthetic data.

	Synthetic Coefficients	5PC	MVR	5PC-MVR
R2		0.696	0.497	1.000
CV-RMSE (%)		7.355	7.1027	0.200
Ycp	50	67.4120 (1.1114)	37.4066 (5.3599)	49.8923 (0.0720)
LS	-0.5	-0.5577 (0.1283)	0.3031 (0.5900)	-0.4984 (0.0035)
RS	1.5	1.4123 (0.1283)	0.0029 (0.0005)	1.4716 (0.0035)
Xcp1	60	60.3683 (2.0350)		60.3683 (2.0350)
Xcp2	75	74.6317 (2.0350)		74.6317 (2.0350)
WBE	0.003			0.0030 (0.0000)

9.4 Testing of VBDD-MVR Models

IMT's ability to generate variable-base degree-day, multi-variable regression (VBDD-MVR) models was also tested. First, heating degree-day (HDD) and cooling degree-day (CDD) models of building energy use were generated. Next, the residual files from the HDD and CDD runs were used as input to the MVR model to generate HDD-MVR and CDD-MVR models. Five CDD-MVR and five HDD-MVR tests demonstrated the robustness of the method.

IMT's ability to generate variable-base degree-day, multi-variable regression (VBDD-MVR) models was also tested. First, heating degree-day (HDD) and cooling degree-day (CDD) models of building energy use were generated. Next, the residual files from the HDD and CDD runs were used as input to the MVR model to generate HDD-MVR and CDD-MVR models. Five CDD-MVR and five HDD-MVR tests demonstrated the robustness of this method.

10.0 Example of Model Development

Building energy use is frequently influenced by the weather and other variables. In this chapter, we describe the development of 3PC-MVR and VBDD-MVR regression models of grocery store electricity consumption as a function of outdoor air temperature and a sales indicator.

10.1 Development of a 3PC-MVR Model

One year of monthly electricity use, outdoor air temperature and sales-indicator data for a grocery store in the Cleveland, Ohio region are shown in Figure 10.1. The data fields are:

1) Month
2) Day
3) Year
4) Electricity Use (kWh/month)
5) Average Outdoor Air Temperature (F)
6) Sales indicator

1	1	2001	10335	20.7	1100
2	1	2001	9137	26.7	900
3	1	2001	7634	31.8	700
4	1	2001	9760	33.8	900
5	1	2001	9143	52.0	800
6	1	2001	11448	59.4	950
7	1	2001	10560	68.1	900
8	1	2001	9422	71.0	850
9	1	2001	12096	72.0	950
10	1	2001	10947	61.1	900
11	1	2001	8144	52.8	750
12	1	2001	9074	47.1	950

Figure 10.1. One year of monthly electricity use, outdoor air temperature and sales-indicator data for a grocery store in the Cleveland, Ohio region.

The grocery store is cooled during summer months by electric air conditioning units. No cooling is required during the winter and no electricity is used for heating. Thus, a 3PC model of electricity use versus outdoor air temperature appears to be warranted. A graph of electricity use versus outdoor air temperature, with a 3PC model, is shown in Figure 10.2. IMT output from a 3PC model is shown in Figure 10.3. As expected, the electricity use increases in the summer and the 3PC model captures that effect; however there is still significant scatter in the data, including high electricity use during the coldest month of the year December.

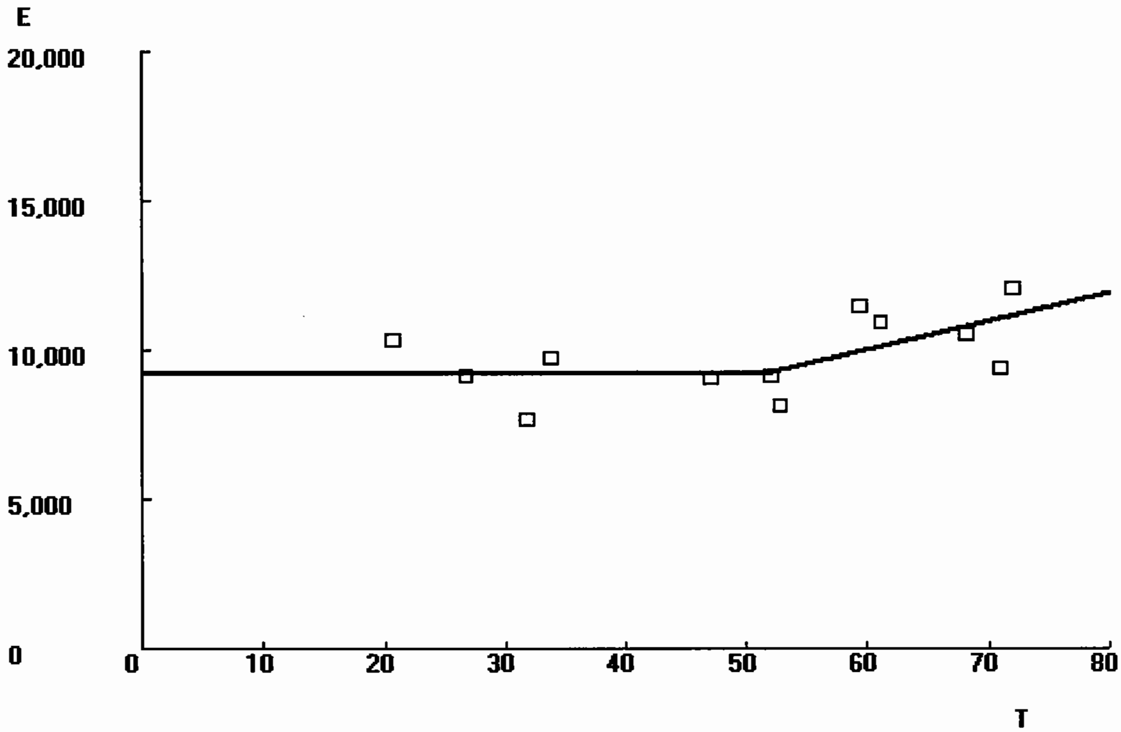

Figure 10.2. Monthly electricity use versus outdoor air temperature, with a 3PC model, for a grocery store in the Cleveland, Ohio region.

```
*********************************************
ASHRAE INVERSE MODELING TOOLKIT (1.9)
*********************************************
Output file name = IMT.Out
*********************************************
Input data file name = xx.dat
Model type =          3P Cooling
Grouping column No =   0
Value for grouping =   1
Residual mode =        0
# of X(Indep.) Var =   1
Y1 column number =     4
X1 column number =   5
X2 column number =   0 (unused)
X3 column number =   0 (unused)
X4 column number =   0 (unused)
X5 column number =   0 (unused)
X6 column number =   0 (unused)
*********************************************
Regression Results
--------------------------------------
      N =    12
--------------------------------------
     R2 =    0.372
--------------------------------------
  AdjR2 =    0.372
--------------------------------------
   RMSE =  1096.1925
--------------------------------------
CV-RMSE =    11.176%
--------------------------------------
      p =   -0.244
--------------------------------------
     DW =    2.326 (p>0)
--------------------------------------
     N1 =    5
--------------------------------------
     N2 =    7
--------------------------------------
    Ycp =  9192.1963 (   405.1517)
--------------------------------------
     LS =     0.0000 (     0.0000)
--------------------------------------
     RS =    97.2335 (    39.9271)
--------------------------------------
    Xcp =    51.4800 (     1.0260)
--------------------------------------
```

Figure 10.3. IMT output from a 3PC model of grocery store electricity use versus outdoor air temperature.

According to store management, sales increase during the holiday season in December, during the Easter holiday and in the summer. The sales volume appears to have a significant effect on store electricity use. A 3PC-MVR model that includes a sales indicator has a much improved fit ($R^2 = 0.78$) compared to the simple 3PC model ($R^2 = 0.37$) (Figure 10.4).

```
*********************************************
ASHRAE INVERSE MODELING TOOLKIT (1.9)
*********************************************
Output file name = IMT.Out
*********************************************
Input data file name = xx.dat
Model type =        3P Cooling MVR
Grouping column No =  0
Value for grouping =  1
Residual mode =      0
# of X(Indep.) Var =  2
Y1 column number =    4
X1 column number =  5
X2 column number =  6
X3 column number =  0 (unused)
X4 column number =  0 (unused)
X5 column number =  0 (unused)
X6 column number =  0 (unused)
*********************************************
Regression Results
-------------------------------------
     N =   12
-------------------------------------
    R2 =   0.783
-------------------------------------
 AdjR2 =   0.783
-------------------------------------
  RMSE =   679.4354
-------------------------------------
CV-RMSE =   6.927%
-------------------------------------
     p =   0.072
-------------------------------------
    DW =   1.707 (p>0)
-------------------------------------
    N1 =   4
-------------------------------------
    N2 =   8
-------------------------------------
   Ycp =  1835.2264 (  1749.9713)
-------------------------------------
    LS =    0.0000 (     0.0000)
-------------------------------------
    RS =   73.1916 (    20.1322)
-------------------------------------
   Xcp =   46.3500 (     1.0260)
-------------------------------------
    X2 =    8.2093 (     1.9590)
-------------------------------------
```

Figure 10.4. IMT output from a 3PC-MVR model of grocery store electricity use versus outdoor air temperature and a sales indicator.

10.2 Development of a VBDD-MVR Model

Similar modeling results can be achieved using IMT's VBDD and MVR models. To do so, the electricity use, outdoor air temperature and sales indicator data are reformatted into a non-uniform timescale data file (Figure 10.5). The data fields are:

1) Month
2) Day
3) Year

4) Average Daily Outdoor Air Temperature (F)
5) Electricity Use (kWh/month)
6) Sales indicator

12	31	2000	19.3	-99	-99
1	1	1995	40.3	10335	1100
1	2	1995	21.2	-99	-99
1	3	1995	19.4	-99	-99
1	4	1995	13.1	-99	-99
1	5	1995	7.7	-99	-99
1	6	1995	21	-99	-99
1	7	1995	29.4	-99	-99
1	8	1995	23	-99	-99
1	9	1995	21.8	-99	-99
1	10	1995	19.5	-99	-99
1	11	1995	31.4	-99	-99
1	12	1995	49.7	-99	-99
1	13	1995	58.1	-99	-99
1	14	1995	55.1	-99	-99
1	15	1995	49.9	-99	-99
1	16	1995	32.2	-99	-99
1	17	1995	29.2	-99	-99
1	18	1995	33.3	-99	-99
1	19	1995	41.4	-99	-99
1	20	1995	39.7	-99	-99
1	21	1995	29.7	-99	-99
1	22	1995	21.5	-99	-99
1	23	1995	19.3	-99	-99
1	24	1995	23.4	-99	-99
1	25	1995	20.4	-99	-99
1	26	1995	22.1	-99	-99
1	27	1995	18	-99	-99
1	28	1995	26.5	-99	-99
1	29	1995	20.6	-99	-99
1	30	1995	24.2	-99	-99
1	31	1995	20.7	-99	-99
2	1	1995	30.4	9137	900
2	2	1995	31.1	-99	-99

Figure 10.5. Part of a non-uniform timescale data file of monthly electricity use, daily outdoor air temperature and sales-indicator data for a grocery store in the Cleveland, Ohio region.

IMT output from a CDD model of these data is shown in Figure 10.6. The CDD model has a fit ($R^2 = 0.34$) and model coefficients similar to the 3PC model in Figure 10.3.

```
********************************************
  ASHRAE INVERSE MODELING TOOLKIT (1.9)
********************************************
   Output file name = IMT.Out
********************************************
   Input data file name =  xxnu.dat
   Model type =          CDD
   Grouping column No =   0
   Value for grouping =   1
   Residual mode =        1
   # of X(Indep.) Var =   1
   Y1 column number =     5
   X1 column number =   4
   X2 column number =   0 (unused)
   X3 column number =   0 (unused)
   X4 column number =   0 (unused)
   X5 column number =   0 (unused)
   X6 column number =   0 (unused)
********************************************
   Regression Results
  -------------------------------------
        N =    12
  -------------------------------------
       R2 =    0.340
  -------------------------------------
    AdjR2 =    0.340
  -------------------------------------
     RMSE =  1124.2397
  -------------------------------------
  CV-RMSE =   11.462%
  -------------------------------------
        p =   -0.274
  -------------------------------------
       DW =    2.348 (p>0)
  -------------------------------------
  DD Base =     43
  -------------------------------------
       A =  9058.4063 (   463.2720)
  -------------------------------------
      X1 =    2.0580 (     0.9072)
  -------------------------------------
```
Figure 10.6. IMT output from a CDD model of grocery store electricity use versus outdoor air temperature.

To incorporate the sales indicator into this model, the residual file from the CDD model (Figure 10.7) is used as input to the IMT's MVR model. The data fields of the residual file from the CDD model are:

1) Month
2) Day
3) Year
4) Average Daily Outdoor Air Temperature (F)
5) Actual Electricity Use (kWh/month)
6) Sales indicator
7) Cooling Degree Days (F-days/month)
8) Predicted Electricity Use (kWh/month)
9) The residual [Actual Electricity Use – Predicted Electricity Use] (kWh/month)

1.00	1.00	1995.00	20.95	10335.00	1100.00	0.00	9058.41	1276.59
2.00	1.00	1995.00	28.16	9137.00	900.00	40.80	9142.37	-5.37
3.00	1.00	1995.00	25.02	7634.00	700.00	0.00	9058.41	-1424.41
4.00	1.00	1995.00	40.49	9760.00	900.00	110.40	9285.61	474.39
5.00	1.00	1995.00	46.17	9143.00	800.00	139.00	9344.46	-201.46
6.00	1.00	1995.00	58.63	11448.00	950.00	484.40	10055.29	1392.71
7.00	1.00	1995.00	70.32	10560.00	900.00	819.70	10745.33	-185.33
8.00	1.00	1995.00	73.49	9422.00	850.00	945.20	11003.61	-1581.61
9.00	1.00	1995.00	74.01	12096.00	950.00	961.30	11036.74	1059.26
10.00	1.00	1995.00	60.35	10947.00	900.00	520.50	10129.58	817.42
11.00	1.00	1995.00	53.09	8144.00	750.00	315.60	9707.90	-1563.90
12.00	1.00	1995.00	34.99	9074.00	950.00	35.90	9132.29	-58.29

Figure 10.7. Residual file from the CDD model run.

IMT output for an MVR model of electricity use versus cooling degree-days and a sales indicator is shown in Figure 10.8. The combination CDD-MVR model has a fit (R^2 = .77) similar to the fit of the 3PC-MVR model (R^2 = .78) in Figure 10.4.

```
*********************************************
 ASHRAE INVERSE MODELING TOOLKIT (1.9)
*********************************************
  Output file name = IMT.Out
*********************************************
  Input data file name = genu.res
  Model type =          MVR
  Grouping column No =   0
  Value for grouping =   1
  Residual mode =        0
  # of X(Indep.) Var =   2
  Y1 column number =     5
  X1 column number =   7
  X2 column number =   6
  X3 column number =   0 (unused)
  X4 column number =   0 (unused)
  X5 column number =   0 (unused)
  X6 column number =   0 (unused)
*********************************************
  Regression Results
  -------------------------------------
       N =    12
  -------------------------------------
      R2 =    0.768
  -------------------------------------
   AdjR2 =    0.768
  -------------------------------------
    RMSE =   701.9841
  -------------------------------------
  CV-RMSE =    7.157%
  -------------------------------------
       p =    0.105
  -------------------------------------
      DW =    1.659 (p>0)
  -------------------------------------
       a =  1770.7170 (  1809.3533)
  -------------------------------------
      X1 =     1.9492 (     0.5671)
  -------------------------------------
      X2 =     8.2562 (     2.0234)
  -------------------------------------
```

Figure 10.8. IMT output from a MVR model of grocery store electricity use versus cooling degree-days and a sales indicator.

11.0 Glossary

Dependent variable: a variable that responds to independent variables.

Field: an individual value or piece of data from a data input file.

Grouping variable: a variable whose value indicates if the given record should be included in the model.

Independent variable: a variable used to predict the response of a dependent variable.

Instruction file: an ASCII text file containing instructions for the toolkit about the data input file and the desired type of model.

No-data flag: a numeric value inserted as a placeholder for missing or erroneous data in a data input file.

Nonuniform-timescale data file: an ASCII text file composed of records in which the dependent variable and the independent variables are measured over different timescales.

Record: one row in a data input file.

Residual: the difference between the observed value y and a model's predicted value $\hat{y}$, (Residual = y − $\hat{y}$).

Space delimited: a file in which each field in a record is separated by one or more empty spaces.

Uniform-timescale data file: an ASCII text file composed of records in which all fields are measured over the same timescale.

Weight indicator: a value used to represent the relative weight to be assigned to each observation for use in a weighted regression.

12.0 References

Anderson, D., 1990, "Electrical Usage Predictors Based on the Singular Value Decomposition Algorithm", M.S. Thesis, Civil, Environmental and Architectural Engineering Department, University of Colorado at Boulder.

Anstett, M. and J. Kreider, 1993. "Application of Neural Networking Models to Predict Energy Use". *ASHRAE Transactions*, Vol. 99, Pt. 1, pp. 505-517.

Austin, S. 1997. "Regression Analysis for Savings Verification". *ASHRAE Journal*, Vol. 39.

Claridge, D., Haberl, J., Liu, M., Houcek, J., and Aamer, A., 1994, "Can You Achieve 150% of Predicted Retrofit Savings? Is It Time for Recommissioning?", Proceedings of the ACEEE 1994 Summer Study on Energy Efficient Buildings, Pacific Grove, CA, August, pp. 5.73-5.87.

Claridge, D. 1998. "A Perspective On Methods For Analysis Of Measured Energy Data From Commercial Buildings", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 150-155.

Cowan, J. and S. Schiller. 1997. "Measuring' Energy Savings for Modernization Projects", ASHRAE Journal, August, pp. 60-62.

Crawford, R., Dykowski, R. and Czajkowski, S., 1991, "A Segmented Linear Least-Squares Modeling Procedure for Nonlinear HVAC Components", ASHRAE Transactions, Vol. 97, Pt. 2, pp. 11-18.

Curtiss, P., G. Shavit and J. Kreider. 1996. "Neural networks applied to buildings - a tutorial and case studies in prediction and adaptive control". *ASHRAE Transactions,* Vol. 102, No. 1, pp. 1141-1146.

Eto, J., 1988, "On Using Degree-days to Account for the Effects of Weather on Annual Energy Use in Office Buildings", Energy and Buildings, Vol. 12, No. 2, pp. 113 - 127.

Fels, M. 1986. "PRISM: An Introduction", *Energy and Buildings*, Vol. 9, pp. 5-18.

Fels, M. and D. Stram. 1986. "Does PRISM Distort the Energy Signature of Heat-pump Houses?", *Energy and Buildings*, Vol.9, pp. 111 - 118.

Fels, M., J. Rachlin and R. Socolow. 1986. "Seasonality of Non-heating Consumption and Its Effect on PRISM Results", *Energy and Buildings*. Vol. 9, pp.139 - 148.

Fels, M. and Keating, K., 1993, "Measurement of Energy Savings from Demand-Side Management Programs in US Electric Utilities", Annual Review of Energy and Environment, 18:57-88.

Fels, M., Kissock, J.K. and Marean, M., 1994. "Model Selection Guidelines for PRISM (Or: Now That HC PRISM Is Coming, How Will I Know When to Use It?)", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings,* Pacific Grove, CA, August.

Fels, M., Kissock, J.K., Marean, M. and Reynolds, C., 1995. "PRISM (Advanced Version 1.0) Users Guide", Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, January.

Feuston, B. and J. Thurtell. 1994. "Generalized nonlinear regression with ensemble of neural nets: the great energy predictor shootout", *ASHRAE Transactions,* Vol. 100, No. 2, pp. 1075-1080.

Forrester, J. and Wepfer, W., 1984, "Formulation of a Load Prediction Algorithm for a Large Commercial Building", *ASHRAE Transactions*, Vol. 90, Pt 1, pp. 536 - 551.

Goldberg, M., 1982. "A Geometrical Approach to Nondifferentiable Regression Models as Related to Methods for Assessing Residential Energy Conservation", Ph.D. Dissertation, Department of Statistics, Princeton University, Princeton, NJ.

Greely, K., Harris, J., and Hatcher, A., 1990, "Measured Savings and Cost-Effectiveness of Conservation Retrofits in Commercial Buildings", Lawrence Berkeley Laboratory Report - 27586, Berkeley, CA.

Haberl, Thamilseran, Reddy, Claridge, O'Neal, and Turner, 1998. "Baseline Calculations For Measurement And Verification Of Energy And Demand Savings In A Revolving Loan Program In Texas", *ASHRAE Transactions* Vol. 104, Pt. 2, pp. 841-858.

Haberl, J. and Vajda, E., 1988. "Use of Metered Data Analysis To Improve Building Operation and Maintenance: Early Results From Two Federal Complexes", *Proceedings of the ACEEE 1988 Summer Study on Energy Efficient Buildings*, Pacific Grove, CA, August, pp. 3.98 - 3.111.

Haberl, J., Komor, P. 1990. "Improving Commercial Building Energy Audits: How Daily and Hourly Data Can Help", *ASHRAE Journal,* Vol. 32, No.9, pp. 26-36, (September).

Herendeen, R., N. Hegan and L. Stiles. 1983. "Measuring Energy Savings Using Personal Trend Data", *Energy and Buildings*, Vol. 5, pp. 289-296.

Hirst, E., Clinton, J., Geller, H. and Kroner, W., 1986, *Energy Efficiency in Buildings: Progress and Promise*, American Council for an Energy Efficient Economy, Washington, D.C.

Hudson, D. 1966, 'Fitting Segmented Curves Whose Join Points Have To Be Estimated", Journal of American Statistical Association, 61, pp. 1097-1125.

Jamieson, D. and Qualmann, R., 1990, "Computer Simulation Energy Use Metering or Can We Count On Energy Savings Estimates In Designing Demand Side Programs", *Proceedings of the*

ACEEE Summer Study on Energy Efficiency in Buildings, Pacific Grove, CA, August, pp. 10.105 - 10.114.

Katipamula, S. and D. Claridge. 1993. "Use of Simplified System Models to Measure Retrofit Energy Savings", *ASME Journal of Solar Energy Engineering*, Vol. 115, No. 2, pp. 57-68.

Katipamula,S., T. Reddy and D. Claridge. 1995. Effect of Time Resolution on Statistical Modeling of Cooling Energy Use in Large Commercial Buildings, *ASHRAE Transactions*, Vol.101, Pt. 2, pp. 172-185.

Katipamula, S., T. Reddy and D. Claridge. 1998. "Multivariate Regression Modeling". *ASME Journal of Solar Energy Engineering,* Vol. 120, No. 3, pp. 177-184.

Kissock, J.K, Claridge, D.E., Haberl, J.S. and Reddy, T.A., 1992. "Measuring Retrofit Savings For the Texas LoanSTAR Program: Preliminary Methodology and Results", *Solar Engineering, 1992: Proceedings of the ASME-JSES-SSME International Solar Energy Conference*, Lahaina, HI, April.

Kissock, J., T. Agami, D. Fletcher and D. Claridge. 1993. "The Effect of Short Data Periods on the Annual Prediction Accuracy of Temperature-Dependent Regression *International Solar Engineering Conference*, pp. 455 - 463.

Kissock, J.K., 1993. "A Methodology to Measure Energy Savings in Commercial Buildings", Ph.D. Dissertation, Mechanical Engineering Department, Texas A&M University, College Station, TX, December.

Kissock, J.K, Xun,W., Sparks, R., Claridge, D., Mahoney, J. and Haberl, J., 1994. "EModel Version 1.4de", Copyright Texas A&M University, Energy Systems Laboratory, Department of Mechanical Engineering, Texas A&M University, College Station, TX, December.

Kissock, J.K., 1994. "Modeling Commercial Building Energy Use with Artificial Neural Networks", *Proceedings of the 29th Intersociety Energy Conversion Engineering Conference*, Vol. 3, pp. 1290-1295, Monterey, CA, August.

Kissock, J.K. and Fels, M., 1995. "An Assessment of PRISM's Reliability for Commercial Buildings", *Proceedings of the National Energy Program Evaluation Conference*, Chicago, IL, August.

Kissock, J.K., 1996, "Development of Analysis Tools in Support of the Texas LoanSTAR Program", University of Dayton, Department of Mechanical and Aerospace Engineering, Dayton, OH, August.

Kissock, J.K., 1997. "Tracking Energy Use and Measuring Chiller Retrofit Savings Using WWW Weather Data and New ETracker Software", *Cool Sense National Forum on Integrated Chiller Retrofits*, San Francisco, CA, June 23-24.

Kissock, Reddy and Claridge, 1998. "Ambient-Temperature Regression Analysis for Estimating Retrofit Savings in Commercial Buildings", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 168-176.

Kissock, K., H. Joseph and J. McBride. 1998. "The Effects of Varying Indoor Air Temperature and Heat Gain on the Measurement of Retrofit Savings" *ASHRAE Transactions*", Vol. 104, Pt. 2., pp. 895-900.

Kissock, K., Joseph, H., 1999. "Synthesizing Hourly Meteorological Data to Improve the Accuracy of Calibrated Simulation Models", *Proceedings of the ASME International Renewable and Advanced Energy Systems for the 21st Century Conference*, Lahaina, HI, April 11-15.

Kissock, K. 1999, "ESave", University of Dayton, Dayton, OH.

Krarti, M., J. Kreider, D. Cohen and P. Curtiss. 1998. "Estimation of Energy Savings for Building Retrofits Using Neural Networks", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 211-216.

Kreider, Claridge, Curtiss, Dodier, Haberl and Krarti, 1995. "Building Energy Use Prediction and System Identification Using Neural Networks" *ASME Journal of Solar Energy Engineering*, Vol. 117, No. 3, pp. 161-166.

Leslie, N., G. Aveta and B. Sllwinski, 1986. "Regression Based Process Energy Analysis System", *ASHRAE Transactions*, Vol. 92, Pt. 1A., pp, 93-102.

MacKay, D., 1994. "Bayesian Nonlinear Modeling for the Prediction Competition", *ASHRAE Transactions*, Vol. 100, No.2, pp. 1053-1062.

Maidment, D., Miaou, S., Nvule, D., Buchberger, S. 1985, "Analysis of Daily Water Use in Nine Cities", Center for Research in Water Resources, Bureau of Engineering Research, University of Texas, Austin, TX, CRWR 201.

Silicon Energy Corp., 2000, "Metrix Utility Accounting System", 1010 Atlantic, Alameda, CA

Miller, A., 1981, "BASIC Programs for Scientists and Engineers", Sybex, Inc., Berkeley, CA.

Nadel S. and Keating, K., 1991. "Engineering Estimates vs. Impact Evaluation Results: How Do They Compare And Why?" *Energy Program Evaluation Conference*, Chicago, pp. 24-33.

Neter, J. Wasserman, W. and Kutner, M., 1989, "Applied Linear Regression Models", Irwin Press, Boston, MA.

Rabl, A., Norford, L. and Spadaro, J., 1992. "Steady State Models for Analysis of Commercial Building Energy Data", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings,* Pacific Grove, CA, August, pp. 9.239-9.261.

Rabl, A and A, Rialhe. 1992. "Energy Signature Models for Commercial Buildings: Test With Measured Data And Interpretation", *Energy and Buildings*, Vol. 19, No. 2, pp. 143 - 154.

Rachlin, J., M. Fels and R. Socolow. 1986. "The Stability of PRISM Estimates", *Energy and Buildings*, Vol 9, pp. 149 - 157.

Reddy, T.A, Kissock, J.K. and Claridge, D.E., 1992. "Uncertainty Analysis in Estimating Building Energy Retrofit Savings in the Texas LoanSTAR Program", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings,* Pacific Grove, CA, August.

Reddy, T. and D. Claridge. 1994. "Using Synthetic Data To Evaluate Multiple Regression And Principal Component Analyses For Statistical Modeling Of Daily Building Energy Consumption", *Energy and Buildings,* Vol. 21, No.1, pp. 35-44.

Reddy, T., Saman, N., Claridge, D., Haberl, J., Turner, W. and Chalifoux, A., 1997, "Baselining methodology for facility level monthly energy use - Parts I and II", *ASHRAE Transactions*, Vol. 103, Pt. 2, pp. 336-347, 348-359.

Reddy, T., J. Kissock and D. Ruch. 1998. Uncertainty In Baseline Regression Modeling And In Determination Of Retrofit Savings. *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 185-192.

Ruch, D. and Claridge, D., 1992, "NAC for Linear and Change-Point Energy Models," *Proceedings of the 1992 ACEEE Summer Study on Energy Efficiency in Buildings,* Pacific Grove, CA, August, pp. 3.263 - 3.273.

Ruch, D. and Claridge, D., 1992. "A Four-Parameter Change-Point Model for Predicting Energy Consumption in Commercial Buildings", *ASME Journal of Solar Energy Engineering*, Vol. 114, No. 2, pp. 77 -83.

Ruch,D. and Claridge, D., 1993. "A Development and Comparison of NAC Estimates for Linear and Change-Point Energy Models for Commercial Buildings", *Energy and Buildings,* Vol. 20, No. 1, pp. 87-95.

Ruch, D, Chen, Haberl, J. and Claridge, D., 1993. "A Change-Point Principal Component Analysis (CP/PCA) Method for Predicting Energy Usage in Commercial Buildings: The PCA Model", *ASME Journal of Solar Energy Engineering*, Vol. 115, No. 2, pp. 77-84.

Ruch, D.K., Kissock, J.K. and Reddy, T.A., 1999. "Model Identification and Prediction Uncertainty of Linear Building Energy Use Models with Autocorrelated Residuals ", *ASME Journal of Solar Energy Engineering* , Vol.121, No.1, pp. 63-68.

Salford Software / Numerical Algorithms Group, 1996, NAGware FTN90 Compiler v2.1x, Oxford, UK.

SAS Instititue Inc., 2001. SAS User Manual, Metaire, LA.

Schrock, D. and Claridge, D. 1989, "Predicting Energy Usage in a Supermarket", *Proceedings of the Sixth Symposium on Improving Building Systems in Hot and Humid Climates*, Mechanical Engineering Department, Texas A&M University, Dallas, TX, October, pp. 44 – 54.

Sonderregger, R. 1997. "Energy Retrofits in Performance Contracts: Linking Modeling and Tracking". *Cool Sense National Forum on Integrated Chiller Retrofits*, San Francisco, September 23-24.

Sonderegger, R. A., 1998. "Baseline Model for Utility Bill Analysis Using Both Weather and Non-Weather-Related Variables", *ASHRAE Transactions*, Vol. 104, No. 2 , pp. 859-870.

Sreshthaputra, A., Haberl, J. and Claridge, D., 2001, "Development of a Toolkit for Calculating Linear, Change-point Linear and Mulitple-linear Inverse Building Analysis Models: Detailed Test Results", ESL-TR-01/05-01, Energy Systems Laboratory, Texas A&M University, College Station, TX.

Stram, D. and M. Fels. 1986. "The Applicability of PRISM to Electric Heating and Cooling", *Energy and Buildings*, Vol. 9, pp. 101 - 110.

United States Department of Energy, 1996a. "North American Energy Measurement and Verification Protocol", DOE/EE-0081, U.S. Department of Energy, Washington, D.C.

United States Department of Energy, 1996b. "Measurement and Verification Guidelines for Federal Energy Projects ", DOE/GO-10096-248, U.S. Department of Energy, Washington, D.C.

United States Department of Energy, 1997. "International Performance Measurement and Verification Protocol", U.S. Department of Energy, Washington, D.C.

Wilson, J., 1998. "The Significant Role of Energy Calculations in the Success of Long-Term Energy Guarantees. *ASHRAE Transactions*, Vol. 104, No.2, pp. 880-894.

APPENDIX I: SOFTWARE REQUIREMENT SPECIFICATION

Inverse Building Energy Model Toolkit
Software Requirement Specification

For
ASHRAE Research Project 1050

Authors
Kelly Kissock, University of Dayton
Jeff Haberl, Texas A&M University
Dave Claridge, Texas A&M University

Contacts
Kelly Kissock, 937-229-2852, jkissock@engr.udayton.edu
Jeff Haberl,409-845-6065, jhaberl@esl.tamu.edu
Dave Claridge, 409-845-1280, claridge@esl.tamu.edu

SRS Draft 2.1

March 17, 2000

Table of Contents

1.0 Functional Description

The objective of this research project is to develop a toolkit of well-documented FORTRAN 90 computer source code for calculating linear, change-point linear and multiple-linear regression models. The intended use of these regression models is to quantify the relationship between building energy use and one or more independent variables.

One application for the regression models would be to estimate savings from energy conservation retrofits in buildings. To do this, measured energy consumption data from the pre-retrofit, or baseline, period would be regressed against variables that influence energy consumption. The resulting "baseline" regression model could be used to estimate pre-retrofit energy consumption under post-retrofit conditions. The baseline model's estimate of pre-retrofit energy use could then be compared to measured post-retrofit energy use to determine savings from the retrofit.

To perform these functions, the toolkit shall be able to:

1) Read a data input file containing building energy use data and variables that may influence the building energy use data.
2) Read instructions from the user concerning the name and type of data input file, which data from the data input file to use, and the type of regression model to fit to the data.
3) Generate a statistical model of building energy use and find the best-fit estimates of the model parameters
4) Find the uncertainty of each model parameter, and the uncertainty of the overall model.
5) Report results to the user.

Specific requirements for each these functions are described in Chapters 2- 6.

Some users may want to incorporate parts of the toolkit into their own software applications. Other users may want a ready-to-run tool. To accommodate these different uses, the toolkit shall consist of both the source and executable code. The software requirements of the toolkit, and a proposed design for the main module are discussed in Chapter 7.

The toolkit shall be tested using data sets from the developers and from the Project Monitoring Subcommittee. Toolkit testing requirements are described in Chapter 8.

To assist in the use of the toolkit, and as a guide to understanding the models, the toolkit shall be accompanied by documentation. Toolkit documentation requirements are described in Chapter 9.

2.0 Input Data Requirements

2.1 Input Data File Format

The toolkit shall read input data files in standard ASCII format. The files shall contain numeric data only. The input data files shall be space delimited with an equal number of fields (columns) in each record (row).

2.2 Types of Input Data Files

The toolkit shall be able to read two types of input data files: uniform and nonuniform timescale. Uniform-timescale data files are composed of records in which all fields are measured over the same timescale. For example, a uniform-timescale data file would be one in which each record includes the amount of energy consumed in an hour, as the dependent-variable field, and the average occupancy and temperature over that hour, as independent-variable fields. The toolkit shall be able to read uniform-timescale data files of any timescale: hourly, daily, weekly, monthly, yearly, etc. An example of a uniform-timescale data-input file is shown in Figure 2.1

114	10	16	90	1	61.80	27.23	6036	76
114	10	17	90	1	65.20	25.68	6145	79
114	10	18	90	1	44.20	35.21	6623	64
114	10	19	90	1	42.60	38.66	6778	-99
114	10	20	90	1	52.00	32.76	6426	70
114	10	21	90	1	44.80	41.29	6651	63
114	10	22	90	1	36.80	44.20	6597	57

Figure 2.1. First seven records of a uniform-timescale data file. The fields are building id, month, day, year, grouping variable, cooling energy use, heating energy use, electricity use and ambient temperature.

Nonuniform-timescale data files are composed of records in which the dependent variable and the independent variables are measured over different timescales. Because of the widespread use of variable-base degree-day models, the toolkit shall be able to read nonuniform-timescale data files in which the dependent variable is energy use, measured over roughly a monthly timescale, and the independent variable is ambient temperature measured on a daily timescale. An example of a nonuniform-timescale data-input file is shown in Figure 2.2.

12	1	1996	233	26
12	2	1996	-99	36
12	3	1996	-99	38
12	4	1996	-99	31
12	5	1996	-99	32
12	6	1996	-99	36
12	7	1996	-99	40
12	8	1996	-99	32
12	9	1996	-99	30
12	10	1996	-99	41
12	11	1996	-99	56
12	12	1996	-99	42
12	13	1996	-99	38
12	14	1996	-99	37
12	15	1996	-99	43
12	16	1996	-99	38
12	17	1996	-99	36
12	18	1996	-99	22
12	19	1996	-99	13
12	20	1996	-99	-99
12	21	1996	-99	-99
12	22	1996	-99	-99
12	23	1996	-99	45
12	24	1996	-99	43
12	25	1996	-99	22
12	26	1996	-99	29
12	27	1996	-99	37
12	28	1996	-99	52
12	29	1996	-99	50
12	30	1996	-99	-99
12	31	1996	-99	44
1	1	1997	215	41
1	2	1997	-99	52
1	3	1997	-99	57

Figure 2.1. First 34 records of a nonuniform-timescale data file. The fields are month, day, year, monthly electricity use and daily temperature.

2.3 Number of Input Data Files

All data used in a regression model shall be included in a single data file.

2.4 Size of the Input Data Files

Data from the input data file shall be manipulated in arrays stored in the computer's Random Access Memory (RAM). Thus, the size of the input data file shall be limited only by the amount of RAM available to the computer.

2.5 Data Grouping

The input data file may contain records that the user does not want to include in the model. The toolkit shall be able to operate on only those records that the user specifies.

2.6 No-data Flag

If there exists no valid values for one or more fields in a data record, a "no-data" flag shall be placed in the appropriate field to indicate that this field shall not be included in a regression model. The user may select any numeric value for the no-data flag. The toolkit shall ignore any record that has a no-data flag in a field on which the model is to operate.

3.0 Operating Instruction Requirements

The user must enter instructions to the executable version of the toolkit. These instructions include the path and filename of the data input file, the type of regression model, and the records and fields in the data input file on which to operate.

3.1 Methods of Entering Operating Instructions

The toolkit shall accept operating instructions 1) by reading an instruction file, or 2) from the keyboard as the user responds to prompts displayed on the computer screen.

3.2 Instruction File Format and Content

The instruction file shall be a standard ASCII text file. The instruction file shall consist of 15 records of a single field each. The information required in each record shall be as shown in Table 3.1.

Table 3.1 Description of operating instruction file.

Record Number	Record Description
1	Path and name of data input file
2	Value of no-data flag
3	Field number of weight indicator
4	Field number of grouping variable
5	Value of grouping variable for record to be included in model
6	Residual output file
7	Regression model type [Mean, 2P, 3PC, 3PH, 4P, 5P, MVR, HDD, or CDD]
8	Field number of dependent (y) variable
9	Number of independent (x) variables
10	Field number of x1
11	Field number of x2
12	Field number of x3
13	Field number of x4
14	Field number of x5
15	Field number of x6

An example instruction file is shown in Figure 3.1.

```
Line 1: Name of data file (aaaaaaa.aaa) = winsave.dat
Line 2: Value of no data flag = -99
Line 3: Column number of weight indicator = 0
Line 4: Column number of group field = 5
Line 5: Value of valid group field = 1
Line 6: Residual file needed (1 yes, 0 no) = 0
Line 7: Regression model (1 to 9) = 7
Line 8: Column number of dependent variable = 6
Line 9: Number of Y1 independent variables data file (0 to 6) = 2
Line 10: Column number of X1 independent variable = 8
Line 11: Column number of X2 independent variable = 9
Line 12: Column number of X3 independent variable = 0
Line 13: Column number of X4 independent variable = 0
Line 14: Column number of X5 independent variable = 0
Line 15: Column number of X6 independent variable = 0
```

Figure 3.1. Example instruction file to generate a multivariable regression (MVR) model of cooling energy use as a function of building electricity use and ambient temperature using the sample data set shown in Figure 2.1.

The toolkit shall determine whether the type of input data file by the type of regression model. HDD and CDD models require nonuniform-timescale data files. All other models require uniform-timescale data files.

3.3 Entering Operating Instructions by Typing Responses to Screen Prompts

The toolkit shall also accept operating instructions entered through the keyboard. The toolkit shall provide the user with prompts asking for the appropriate information. The information requested by the screen prompt shall be the same as the information in the instruction file. The user shall be able to quit the program at any time by typing "q".

3.4 Error Checking

The toolkit shall check operating instructions entered by the user. If a non-valid instruction is entered from the keyboard, the toolkit shall prompt the user to enter a valid instruction. If a non-valid instruction is entered through the instruction file, the output file will report the error.

Errors in the input data file may cause errors in the execution of toolkit algorithms. The toolkit shall be designed to report useful information about the source of the error to the user whenever possible.

4.0 Model Types Supported by Toolkit

The toolkit shall include several types of regression models since no single empirical model is appropriate for all types of measured building energy use data. The types of regression models supported by the toolkit are described below.

4.1 Mean Model

The toolkit shall be able to calculate the arithmetic mean of the dependent variable.

4.2 Two-Parameter Model

The toolkit shall be able to find a simple linear regression model (2P) of type:

$$Y = \beta_1 + \beta_2 X_1 \tag{4.1}$$

where β_1 and β_2 are regression coefficients, X_1 is the independent variable and Y is the dependent variable.

4.3 Three-Parameter Cooling and Heating Models

The toolkit shall be able to find best-fit three-parameter (3P) change-point models of type:

$$Y_c = \beta_1 + \beta_2 (X_1 - \beta_3)^+ \tag{4.2}$$

$$Y_h = \beta_1 + \beta_2 (X_1 - \beta_3)^- \tag{4.3}$$

where β_1 is the constant term, β_2 is the slope term, and β_3 is the change-point,. The $( )^+$ and $( )^-$ notations indicate that the values of the parenthetic term shall be set to zero when they are negative and positive respectively.

The toolkit shall also be able to find combination three-parameter multi-variable regression models (3P-MVR), with up to four independent variables, of the type:

$$Y_c = \beta_1 + \beta_2 (X_1 - \beta_3)^+ + \beta_4 X_2 + \beta_5 X_3 + \beta_6 X_4 \quad (4.4)$$

$$Y_h = \beta_1 + \beta_2 (X_1 - \beta_3)^- + \beta_4 X_2 + \beta_5 X_3 + \beta_6 X_4 \quad (4.5)$$

where X_1 is typically temperature, and X_2, X_3, and X_4 are optional independent variables.

4.4 Four-Parameter Model

The toolkit shall be able to find best-fit four-parameter (4P) change-point models of type:

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_4)^+ \quad (4.6)$$

where β_1 is the constant term, β_2 is the left slope, β_3 is the right slope and β_4 is the change point. The toolkit shall also be able to find combination four-parameter multi-variable regression models (4P-MVR), with up to three independent variables, of the type:

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_4)^+ + \beta_5 X_2 + \beta_6 X_3 \quad (4.7)$$

where X_1 is typically temperature, and X_2 and X_3 are optional independent variables.

4.5 Five-Parameter Model

The toolkit shall be able to find best-fit five-parameter (5P) change-point models of type:

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_5)^+ \quad (4.8)$$

where β_1 is the constant term, β_2 is the left slope, β_3 is the right slope, β_4 is the left change point, and β_5 is the right change point.

The toolkit shall also be able to find combination five-parameter multi-variable regression models (5P-MVR), with up to two independent variables, of the type:

$$Y = \beta_1 + \beta_2 (X_1 - \beta_4)^- + \beta_3 (X_1 - \beta_5)^+ + \beta_6 X_2 \quad (4.9)$$

where X_1 is typically temperature and X_2 is an optional independent variable.

4.6 Multiple Variable Regression Model

The toolkit shall be able to find a multiple-variable linear regression (MVR) models, with up to six independent variables, of type:

$$Y = \beta_1 + \beta_2 X_1 + \beta_3 X_2 + \beta_4 X_3 + \beta_5 X_4 + \beta_6 X_5 + \beta_7 X_6 \qquad (4.10)$$

where β_1 through β_7 are regression coefficients, and X_1 through X_6 are independent variables.

4.7 Variable-Base Heating and Cooling Degree-Day Models

The toolkit shall be able to find best-fit variable-base degree-day models of type:

$$Y = \beta_1 + \beta_2 \, HDD(\beta_3) \qquad (4.11)$$

$$Y = \beta_1 + \beta_2 \, CDD(\beta_3) \qquad (4.12)$$

where β_1 is the constant term, β_2 is the slope term, and $HDD(\beta_3)$ and $CDD(\beta_3)$ are the number of heating and cooling degree-days, respectively, in each energy data period calculated with base temperature β_3. The number of heating and cooling degree-days in each energy data period of n days are:

$$HDD(\beta_3) = \sum_{i=1}^{n} (\beta 3 - T_i)^+ \qquad (4.13)$$

$$CDD(\beta_3) = \sum_{i=1}^{n} (T_i - \beta 3)^+ \qquad (4.14)$$

where T_i is the average daily temperature.

5.0 Model Uncertainty Parameters Reported by Toolkit

GPC-14P Working Draft 99.2, June 7, 1999 specifies that modeling uncertainty be estimated using three indices:

1) Coefficient of Variation of the Standard Deviation (CVSTD)

$$\text{CVSTD} = 100 \text{ x } [\sum (y_i - \bar{y})^2 /(n-1)]^{1/2} / \bar{y} \tag{5.1}$$

2) Coefficient of Variation of the Root Mean Square Error (CVRMSE)

$$\text{CVRMSE} = 100 \text{ x } [\sum (y_i - \hat{y}_i)^2 /(n-p)]^{1/2} / \bar{y} \tag{5.2}$$

3) Normalized Mean Bias Error (MBE)

$$\text{NMBE} = \frac{\sum^{n} (y_i - \hat{y}_i)}{(n-p) * \bar{y}} * 100 \tag{5-3}$$

Where:

y dependent variable of some function of the independent variable(s)

$\bar{y}$ arithmetic mean of the sample of n observations

$\hat{y}$ regression model's predicted value of y

n number of data points or periods in the baseline period

p number of parameters or terms in the baseline model, as developed by a mathematical analysis of the baseline data.

CVSTD (Equation 5.1) is a special case of CVRMSE (Equation 5.2) for mean models with one parameter. Thus, to comply with GPC-14P, the toolkit shall report:

- CVSTD for mean models
- CVRMSE for 2 - 5 parameter and MVR models
- NMBE for all models

In addition, the toolkit shall report the following uncertainty statistics:

- STD for mean models

$$\text{STD} = [\sum (y_i - \bar{y})^2 /(n-1)]^{1/2} \qquad (5.4)$$

- RMSE for all regression models

$$\text{RMSE} = [\sum (y_i - \hat{y}_i)^2 /(n-p)]^{1/2} \qquad (5.5)$$

- R^2 for all regression models

$$R^2 = 1 - \frac{\sum_n (y_i - \hat{y}_i)^2}{\sum_n (y_i - \bar{y}_i)^2} \qquad (5.6)$$

- Adjusted R^2 for all MVR models

$$\text{Adj}\,R^2 = 1 - \frac{(N-1)}{(N-p)}(1-R^2) \qquad (5.7)$$

  In MVR models, the addition of an independent variable will always result in an increase in the model's R^2. Adjusted R^2 divides each sum of squares in R^2 by the associated degrees of freedom, and is thus a measure of the actual improvement in predictive ability from adding independent variables.

- Standard error of each regression coefficient. The standard error of a regression coefficient indicates the variance with which the coefficient is known. The standard error is defined such that with a probability of 1-α, the true parameter will fall within the bounds:

$$\beta_{true} = \beta_{estimated} + t(1-\alpha/2, n-p)\, s(\beta) \qquad (5.8)$$

  where t is the t distribution and s(β) is the standard error of each regression coefficient:

$$s(\beta) = [\text{MSE}\, (\mathbf{X'X})^{-1}]^{.5} \qquad (5.9)$$

- The uncertainty of each change-point coefficients as given by the width of the final search interval.

Together, these measures of uncertainty will allow the user to assess the fit of the model to the data, select appropriate independent variables, and to calculate the overall uncertainty of savings using the methods described in GPC 14P.

6.0 Output Data Requirements

The toolkit shall report model results by generating an output file, and an optional residual file, for each model run. Both output files shall be standard ASCII text files.
The output file shall include the information entered in the operating instructions, model coefficients and goodness-of-fit parameters. An output file shall be generated after each model run.
The residual file shall include the data from the data input file, predicted values of the dependent variable and model residuals. A residual file shall be generated only if requested in the instruction file or through keyboard prompts.

6.1 Path and Filename of Output File

The output file shall be placed in the same directory as the instruction file. The output filename shall have the same prefix as the instruction file and a file name extension '.out'.

6.2 Content of Output File

The output file shall include:

- all information entered as operating instructions
- the number of observations used in the model
- the value and standard error of each regression coefficient
- the model goodness-of-fit parameters specified in Chapter 5.

A sample output file is shown in Figure 6.1

```
*******************************************
                  ASHRAE INVERSE METHOD TOOLKIT (1.2)
*******************************************
   Output file name = IMT.Out
*******************************************
   Input Data file name =    winsave.dat
   Regression type =   2P
   Weight column No =   0
   Grouping column No = 0
   Value for grouping = 0
   Residual mode =      0
   # of X(Indep.) Var = 1
   Y1 column number =   6
   X1 column number =  11
   X2 column number =   0 (unused)
   X3 column number =   0 (unused)
   X4 column number =   0 (unused)
   X5 column number =   0 (unused)
   X6 column number =   0 (unused)


*******************************************
   Regression Result
*******************************************
      File name = IMT.Out
*******************************************
      Model = 2P
    -------------------------------------
          N =    666
    -------------------------------------
        Yint =    -18.6454 ( 3.0841)
    -------------------------------------
          X1 =      0.8460 ( 0.0440)
    -------------------------------------
          R2 =      0.36
    -------------------------------------
       AdjR2 =      0.36
    -------------------------------------
        RMSE =     15.16
    -------------------------------------
     CV-RMSE =     38.28%
    -------------------------------------
           p =      0.92
    -------------------------------------
          DW =      0.17(p>0)
    -------------------------------------
```
Figure 6.1 Sample output file.

6.3 Path and Filename of Residual File

The residual file shall be placed in the same directory as the instruction file. The output filename shall have the same prefix as the instruction file and a file name extension '.res'.

6.4 Contents of Residual File

Each record of the residual file shall include the data from the data input file, the predicted value of the dependent variable, and the difference between the observed and predicted values of the dependent variable. The difference between the observed and predicted values of the dependent variable is called the residual and is defined as:

$$\text{Residual} = y - \hat{y} \tag{6.1}$$

6.4.1 Residual File from a Uniform-Timescale Data-Input File

An example residu20al file generated from a from a uniform-timescale data-input file is shown in Figure 6.2. The first 9 columns are from the data input file. The dependent variable is in the 6^{th} column. The predicted value of the dependent variable is in the 10^{th} column and the residual is in the 11^{th} column.

114	10	16	1990	3	61.80	27.23	-99	76	45.65	16.15
114	10	17	1990	4	65.20	25.68	-99	79	48.19	17.01
114	10	18	1990	5	44.20	35.21	-99	64	35.50	8.70
114	10	19	1990	6	42.60	38.66	-99	62	33.81	8.79
114	10	20	1990	7	52.00	32.76	-99	70	40.57	11.43
114	10	21	1990	1	44.80	41.29	-99	63	34.65	10.15
114	10	22	1990	2	36.80	44.20	-99	57	29.57	7.23
114	10	23	1990	3	-99	-99	-99	58	-99	-99
114	10	24	1990	4	41.00	39.66	-99	63	34.65	6.35
114	10	25	1990	5	41.80	37.66	-99	64	35.50	6.30
114	10	26	1990	6	43.20	37.39	-99	62	33.81	9.39
114	10	27	1990	7	45.20	33.49	-99	65	36.34	8.86
114	10	28	1990	1	46.80	32.49	-99	68	38.88	7.92
114	10	29	1990	2	48.40	34.21	-99	68	38.88	9.52
114	10	30	1990	3	52.80	33.85	-99	67	38.04	14.76
114	10	31	1990	4	55.60	33.31	-99	68	38.88	16.72
114	11	1	1990	5	53.20	32.13	-99	68	38.88	14.32
114	11	2	1990	6	57.20	31.67	-99	70	40.57	16.63
114	11	3	1990	7	61.00	29.86	-99	75	44.80	16.20
114	11	4	1990	1	40.40	43.02	-99	57	29.57	10.83

Figure 6.2. First 20 records from a residual file from a uniform-timescale input data file.

6.4.2 Residual file from a Nonuniform-Timescale Data-Input File

Nonuniform-timescale data files are used with VBDD models. These data files contain observations of the dependent variable, typically energy use, which are usually measured over several days, and observations of the independent variable, typically temperature, which are usually measured on the daily timescale. Because residuals are calculated for each energy observation, the residual file shall only include records corresponding to energy observations in the input data file. The residual file from a nonuniform-timescale data-input file shall include:

- all fields from records in the data input file which have valid[1] energy values, except the average daily temperature field. The average daily temperature field shall be replaced with the average temperature over energy time-interval
- the number of degree days in the energy time-interval calculated to the best-fit reference temperature
- the difference between the predicted and observed values of energy use

An example residual file generated from a nonuniform-timescale data-input file is shown in Figure 6.3. The fields are month, day, year, average temperature during the energy period, energy use, heating degree days during the energy use period, predicted energy use, and the difference between observed and predicted energy use.

1	4	1979	37.2	2320	645.0	2,307.0	13.0
2	2	1979	31.6	2930	765.0	2,626.6	303.4
3	6	1979	27.8	2920	965.0	3,159.3	-239.3
4	4	1979	46.7	1530	336.0	1,484.0	46.0
5	4	1979	53.8	1150	162.0	1,020.5	129.5
6	5	1979	65.8	630	0.0	589.1	40.9
7	5	1979	69.9	510	0.0	589.1	-79.1
8	3	1979	79.1	600	0.0	589.1	10.9
9	4	1979	76.5	520	0.0	589.1	-69.1
10	3	1979	67.8	620	4.0	599.7	20.3
11	2	1979	55.8	950	148.0	983.2	-33.2
12	4	1979	49.8	1210	287.0	1,353.5	-143.5

Figure 6.3. Example residual file from a nonuniform-timescale data-input file.

[1] Any value except the value of the no-data flag.

7.0 Software Requirements

7.1 Programming Language and Operating System

The toolkit shall be written in FORTRAN 90. The toolkit shall consist of both source and executable code. The executable version of the toolkit shall run in an MS DOS window of Microsoft Windows operating system.

7.2 Toolkit Design

The toolkit shall be composed of a main module and a series of subroutines. The source code shall be documented with comments describing the functionality of each subroutine and code block.

Proposed logic for the main module is shown in Figure 7.1.

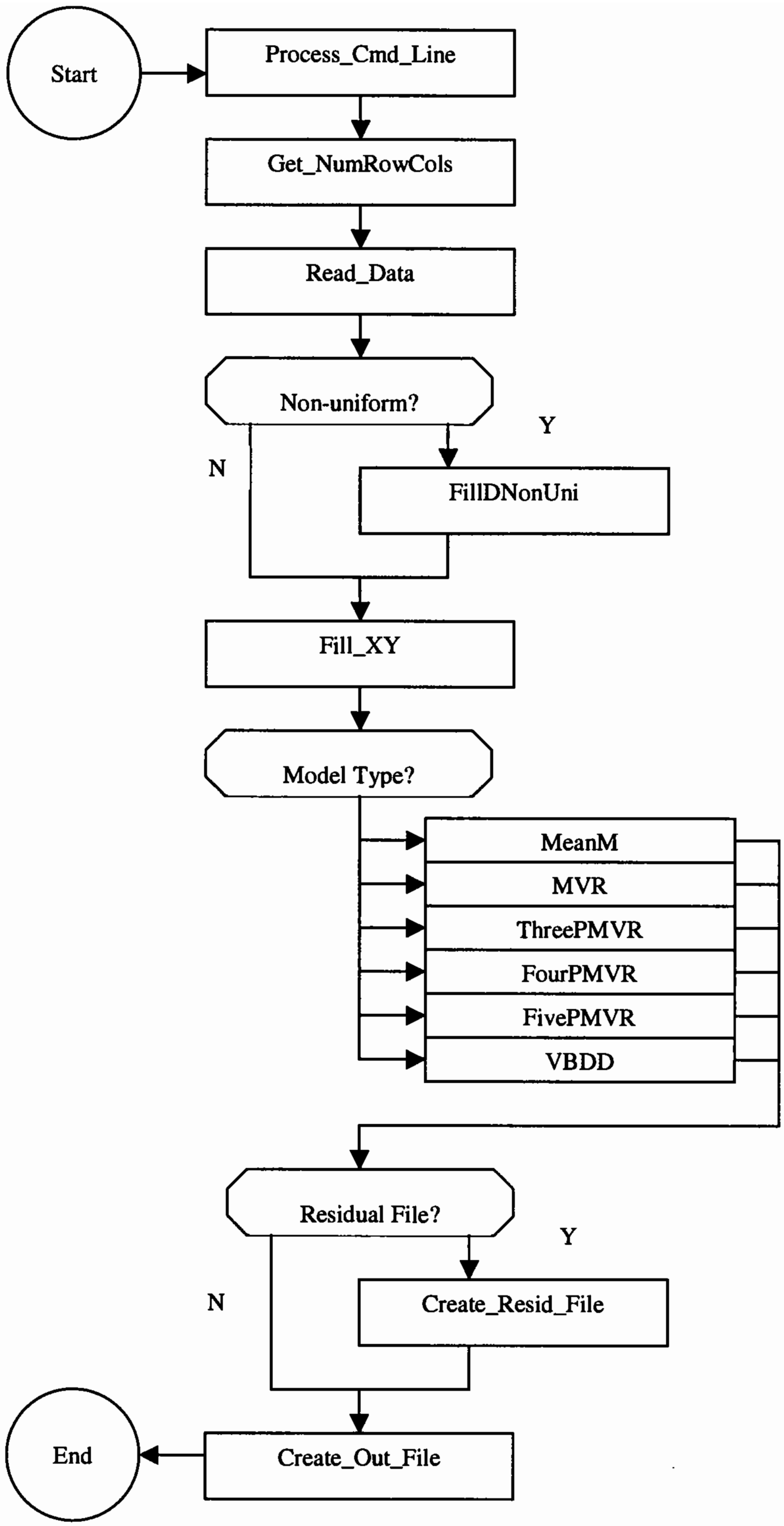
Start
Process_Cmd_Line
Get_NumRowCols
Read_Data
Non-uniform?
Y
N
FillDNonUni
Fill_XY
Model Type?
MeanM
MVR
ThreePMVR
FourPMVR
FivePMVR
VBDD
Residual File?
Y
N
Create_Resid_File
End
Create_Out_File

Figure 7.1 Flow chart of logic in toolkit's main module.

8.0 Toolkit Testing Requirements

The toolkit shall be tested for accuracy and robustness.

8.1 Accuracy of Regression Results

The accuracy of the regression routines shall be tested by comparing simple and multiple linear regression results from 20 sets of measured energy and temperature data to regression results from the widely-used statistical software SAS. The 20 test data sets shall be selected so that they include both very small and very large values, and as few as 3 and as many as 9,000 data records. The PMS may also contribute test data sets.

8.2 Accuracy of Change-Point Models

The accuracy of the change-point algorithms shall be tested using synthetic data with pre-defined change-points. The test data sets will be constructed to vary the position of the change-points relative to the range of x and y values, and to vary the spacing between data points. Ten sets of synthetic data will be generated for each type of change-point model. The change-points determined by the models will be compared to the synthetic change-points.

The values from an example synthetic data set used to test the 5P model are plotted below. The values in the data set were selected such that they fall on one of three line segments. The line segments have known slopes and are joined at change-points $x = 30$ and $x = 60$. The 5P model is tested by comparing the model's estimates of the slopes and change-points to the known slopes and change points in the synthetic data set.

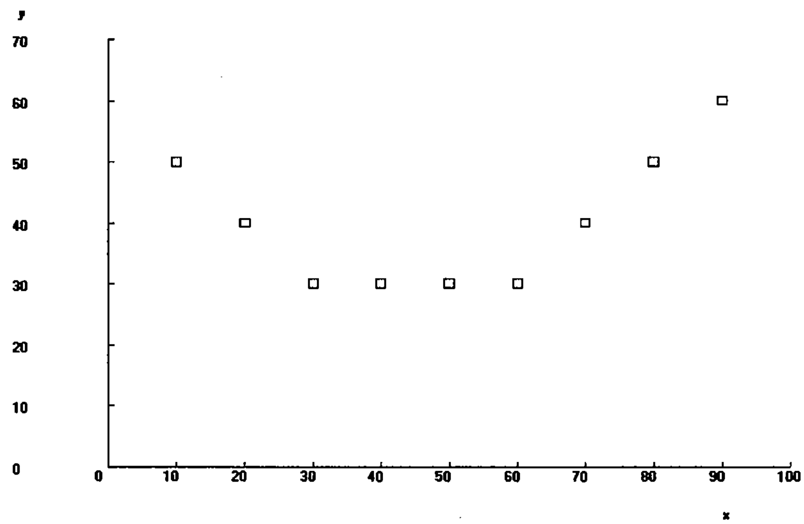

Figure 8.1 Example of synthetic data set used to test 5P model.

8.3 Robustness of Algorithms

Robustness testing shall be conducted by running 20 test sets of measured energy and temperature data through each of the regression algorithms. Problems will be noted. The PMS may contribute test data sets or test robustness as they see fit.

9.0 Toolkit Documentation Requirements

The toolkit software shall be accompanied by documentation. A proposed table of contents for the documentation is shown in Figure 9.1.

Proposed Toolkit Documentation Outline

Abstract
1.0 Introduction
- Motivation for Work
- Brief Description of Sections to Follow

2.0 Review of Published Algorithms and Selection of Algorithms for the Toolkit
3.0 Description of Input Data
- Description of Input Data Format
- Description of Error Checking Algorithm
- Use of Executable Program
- Example Application
- Source Code

4.0 Description of Least-Squares Regression Routines and Goodness-of-Fit Parameters
- Description of Method
- Description of Algorithms
- Source Code

5.0 Description of Models
- Mean model
    - Physical Basis for Model
    - Description of Algorithm
    - Use of Executable Program
    - Example Application
    - Source Code
- 2P Model
    - Physical Basis for Model
    - Description of Algorithm
    - Use of Executable Program
    - Example Application
    - Source Code
- 3P-Heating and 3P-Cooling Models
    - Physical Basis for Model
    - Description of Algorithm
    - Use of Executable Program

- Source Code
- 4P Model
  - Example Application
  - Physical Basis for Model
  - Description of Algorithm
  - Use of Executable Program
  - Example Application
  - Source Code
- 5P Model
  - Physical Basis for Model
  - Description of Algorithm
  - Use of Executable Program
  - Example Application
  - Source Code
- MVR Model
  - Physical Basis for Model
  - Description of Algorithm
  - Use of Executable Program
  - Example Application
  - Source Code
- VBDD Models
  - Physical Basis for Models
  - Description of Algorithm
  - Use of Executable Program
  - Example Application
  - Source Code

6.0 Comparison of Regression Results with Reference Statistical Software

Figure 9.1. Proposed table of contents for software documentation.

10.0 Glossary

Dependent variable: a variable that responds to independent variables.

Field: an individual value or piece of data from a data input file.

Grouping variable: a variable whose value indicates if the given record should be included in the model.

Independent variable: a variable used to predict the response of a dependent variable.

Instruction file: an ASCII text file containing instructions for the toolkit about the data input file and the desired type of model.

No-data flag: a numeric value inserted as a place holder for missing or erroneous data in a data input file.

Nonuniform-timescale data file: an ASCII text file composed of records in which the dependent variable and the independent variables are measured over different timescales.

Record: one row in a data input file.

Residual: the difference between the observed value y and a model's predicted value $\hat{y}$, (Residual = y – $\hat{y}$).

Space delimited: a file in which each field in a record is separated by one or more empty spaces.

Uniform-timescale data file: an ASCII text file composed of records in which all fields are measured over the same timescale.

Weight indicator: a value used to represent the relative weight to be assigned to each observation for use in a weighted regression.

APPENDIX II: ANNOTATED BIBLIOGRAPHY

The primary search mechanism employed was the Compendex database. Journals and proceedings searched are shown in Table 1. Each journal was searched using key words related to regression analysis and energy savings. Key words searched are shown in Table 2. Some publications from outside of these sources were also identified and included. Over one hundred papers referenced these terms. After a review of the abstracts, over 70 papers were obtained and are briefly summarized in the Annotated Bibliography.

The papers were grouped by primary subject area (Table 3). Please note, however, that there is much overlap and many papers discuss topics related to several of the groups. A summary of the findings, organized by group, is presented in the section entitled Summary of Finding.

Table 1. Journals and conference proceedings searched.

ACEEE Summer Study on Energy Efficiency in Buildings	ASME Solar Engineering Conference
ASHRAE Journal	Energy and Buildings
ASHRAE Transactions	Cool Sense National Forum on Integrated Chiller Retrofits
ASME Journal of Solar Energy Engineering	Society for Industrial and Applied Mathematics

Table 2. Keywords searched.

Regression	Energy Prediction
Energy Savings	Energy Utilization
Inverse Modeling	Energy Conservation
Modeling	Mathematical Models
Whole Building Energy Use	Numerical Analysis
Algorithms	Statistical Methods
Outdoor Air Temperature	Retrofits
Calibration	Buildings

Table 3. Subject groups for selected papers.

Motivations for Measuring Savings
Overview of Methods for Measuring Savings
Variable-Base Degree-Day Models
Change-Point Models
Multivariate Regression Models
Combination VBDD/CP/MVR Models
Calibrated Simulation Models
Artificial Neural Network Models
Advanced Regression Techniques
Uncertainty of Savings

Motivations for Measuring Savings

Hirst, E., Clinton, J., Geller, H. and Kroner, W., 1986, *Energy Efficiency in Buildings: Progress and Promise*, American Council for an Energy Efficient Economy, Washington, D.C.

Greely, K., Harris, J., and Hatcher, A., 1990, "Measured Savings and Cost-Effectiveness of Conservation Retrofits in Commercial Buildings", Lawrence Berkeley Laboratory Report - 27586, Berkeley, CA.
LBL report on BECA database of over 1,700 commercial building retrofits.

Jamieson, D. and Qualmann, R., 1990, "Computer Simulation Energy Use Metering or Can We Count On Energy Savings Estimates In Designing Demand Side Programs", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, August, pp. 10.105 - 10.114.
Account of commercial building retrofit program in which savings were predicted based on DOE2 simulations and compared with measured results.

Nadel S. and Keating, K., 1991. "Engineering Estimates vs. Impact Evaluation Results: How Do They Compare And Why?" *Energy Program Evaluation Conference*, Chicago, pp. 24-33.
Comparison of predicted and actual savings in several large DSM programs and discussion of reasons for the discrepancies.

Fels, M. and Keating, K., 1993, "Measurement of Energy Savings from Demand-Side Management Programs in US Electric Utilities", Annual Review of Energy and Environment, 18:57-88.
Describes history of DSM program evaluation. General description of methods for measuring savings including PRISM, Multiple Variable Regression, Conditional Demand Analysis and Engineering Models.

Claridge, D., Haberl, J., Liu, M., Houcek, J., and Aamer, A., 1994, "Can You Achieve 150% of Predicted Retrofit Savings? Is It Time for Recommissioning?", Proceedings of the ACEEE 1994 Summer Study on Energy Efficient Buildings, Pacific Grove, CA, August, pp. 5.73-5.87.

Describes use of measured data to identify operational and maintenance problems in the Texas LoanSTAR program, and how correction of these problems has led to savings in excess of those predicted for the retrofits.

United States Department of Energy, 1996a. "North American Energy Measurement and Verification Protocol", DOE/EE-0081, U.S. Department of Energy, Washington, D.C.
General guidelines for measurement and verification.

United States Department of Energy, 1996b. "Measurement and Verification Guidelines for Federal Energy Projects ", DOE/GO-10096-248, U.S. Department of Energy, Washington, D.C.
General guidelines for measurement and verification.

United States Department of Energy, 1997. "International Performance Measurement and Verification Protocol", U.S. Department of Energy, Washington, D.C.
General guidelines for measurement and verification.

Overview of Methods for Calculating Savings

Eto, J., 1988, "On Using Degree-days to Account for the Effects of Weather on Annual Energy Use in Office Buildings", Energy and Buildings, Vol. 12, No. 2, pp. 113 - 127.
Simulates energy consumption in a commercial building in five US cities and models it using variable-base degree-day method.

Kissock, J.K, Claridge, D.E., Haberl, J.S. and Reddy, T.A., 1992. "Measuring Retrofit Savings For the Texas LoanSTAR Program: Preliminary Methodology and Results", *Solar Engineering, 1992: Proceedings of the ASME-JSES-SSME International Solar Energy Conference*, Lahaina, HI, April.
Describes method of calculating savings from measured data using change-point models for weather adjustment.

Ruch, D. and Claridge, D., 1992, "NAC for Linear and Change-Point Energy Models," *Proceedings of the 1992 ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, August, pp. 3.263 - 3.273.
Describes calculation of Normalized Annual Consumption using linear and change-point models.

Kissock, J.K., 1993. "A Methodology to Measure Energy Savings in Commercial Buildings", Ph.D. Dissertation, Mechanical Engineering Department, Texas A&M University, College Station, TX, December.
Describes method of calculating savings from measured data using change-point models for weather adjustment. Includes the physical basis for change-point models in commercial buildings, algorithms for change-point models, and a method to estimate the uncertainty of savings.

Cowan, J. and S. Schiller. 1997. "Measuring' Energy Savings for Modernization Projects", ASHRAE Journal, August, pp. 60-62.

Describes basic method for measuring savings:

- measure energy use and influential variables during baseline period
- create mathematical model of baseline energy use as function of influential variables
- measure energy use and influential variables during post-retrofit period
- apply post-retrofit influential variables to base-line model to estimate what energy use would have been without retrofit
- subtract estimated baseline energy use for post-retrofit energy use to estimate savings

Discusses uncertainty of savings versus cost of monitoring and analysis. Recommends that cost for determining savings be no more than 5% of actual savings. Discusses baseline period, utility prices, and how to structure ESPC.

Claridge, D. 1998. "A Perspective On Methods For Analysis Of Measured Energy Data From Commercial Buildings", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 150-155.
Introduction to series of articles of measuring savings. Provides historical information on methods for measuring energy use. It summarizes the capabilities and uncertainties associated with regression modeling. Neural networks, Fourier series and spectral analysis are described. Prism technique is an adaptation of the Variable-Base Degree-Day method. Measured monthly consumption data and daily average temperature data is used to "calibrate" the model using regression analysis. Temperature dependent baselines for most commercial buildings, however, require a more general change point line. Katipamula's "Multivariate Analysis for Retrofit Savings" discusses the added complexities using multivariate analysis.

Kissock, K., Joseph, H., 1999. "Synthesizing Hourly Meteorological Data to Improve the Accuracy of Calibrated Simulation Models", *Proceedings of the ASME International Renewable and Advanced Energy Systems for the 21st Century Conference*, Lahaina, HI, April 11-15.
Describes:

- annual variation in weather
- simulation study of annual variation in energy use in a residence and large commercial building due to variation in weather
- method of synthesizing hourly meteorological data from average daily temperatures
- compares calibration error from using TMY2 data rather than real or synthetic weather data.
- concludes that calibration error from using TMY2 data is small in large commercial buildings but is large in small buildings.

Variable-Base Degree-Day Models

Goldberg, M., 1982. "A Geometrical Approach to Nondifferentiable Regression Models as Related to Methods for Assessing Residential Energy Conservation", Ph.D. Dissertation, Department of Statistics, Princeton University, Princeton, NJ.
Mathematical basis of PRISM model with uncertainty analysis.

Herendeen, R., N. Hegan and L. Stiles. 1983. "Measuring Energy Savings Using Personal Trend Data", *Energy and Buildings*, Vol. 5, pp. 289-296.
Examples of energy savings for 12 houses in central Illinois after insulation retrofit. Savings based on NAC calculations.

Fels, M. 1986. "PRISM: An Introduction", *Energy and Buildings*, Vol. 9, pp. 5-18.
A houses heating system is modeled by when the outside temperature drops below a certain value a constant amount of heating fuel is required for each drop in temperature. An iterative procedure based on Newton's method is used to find the best heating reference temperature. This occurs when heating degree-days vs. rate of energy consumption is most nearly a straight line. T is found when the mean-squared error is minimized, or equivalently the R2 value is the highest. NAC is a reliable and stable index of consumption, the other PRISM parameters provide physically meaningful indicators, whose change may not be statistically meaningful but can often suggest a reason for a change in consumption. Addresses buildings in heating-dominated climates. For cooling dominated climates and for a large solar component, more research is needed. The derivation of the physical model underlying PRISM is given along with the computation of group savings estimates and standard errors of savings estimates.

Stram, D. and M. Fels. 1986. "The Applicability of PRISM to Electric Heating and Cooling", *Energy and Buildings*, Vol. 9, pp. 101 - 110.
The use of PRISM to measure energy savings in electrically heated houses is discussed in this paper. The heating model applied to electrically heated houses without cooling performs as well as it has for gas-heated and oil-heated houses. The heating-plus-cooling model works well on houses with relatively strong cooling, while houses with erratic or weak cooling will require an alternative approach. The average cooling reference temperature estimated by the model is found to be well above the average heating reference temperature. Overall, the NAC index as the basis for savings estimated appear highly reliable for electrically heated houses. To model a house with heating and cooling both heating degree-days and cooling degree-days are used.

Fels, M. and D. Stram. 1986. "Does PRISM Distort the Energy Signature of Heat-pump Houses?", *Energy and Buildings*, Vol.9, pp. 111 - 118.
The behavior of PRISM applied to electrically heated houses with heat pumps is investigated. The nonlinear response of energy consumption to outside temperature leads to systematic distortion in the model parameters. The most important effect is on the heating reference temperature, which is reduced from the true value by about 3 degrees Celsius. The overall model performance using PRISM was found to remain high for houses using heat pumps.

Fels, M., J. Rachlin and R. Socolow. 1986. "Seasonality of Non-heating Consumption and Its Effect on PRISM Results", *Energy and Buildings*. Vol. 9, pp.139 - 148.
The effect of seasonality to non-heating energy consumption is investigated.
Non-heating fuel consumption can be modeled by a sine curve with the highest non-heating consumption in winter and the lowest in summer. The effect of non-heating energy consumption does not call for changes in the scorekeeping model.

Rachlin, J., M. Fels and R. Socolow. 1986. "The Stability of PRISM Estimates", *Energy and Buildings*, Vol 9, pp. 149 - 157.
The effect of missing energy data to PRISM models is discussed. Twelve monthly readings are optimal for the most reliable results. The NAC index is far less sensitive to missing or insufficient data than are the individual parameters. If data is available for over a one-year

period but less than another whole year, it is best to reduce the data to one full year. Energy consumption patterns change over a period of a year and including only a portion of a year will compromise the results. Missing readings are troublesome and the longer the gap the greater the problem.

Rabl, A., Norford, L. and Spadaro, J., 1992. "Steady State Models for Analysis of Commercial Building Energy Data", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, August, pp. 9.239-9.261.
Examines PRISM's suitability for modeling commercial building energy use. Cautions against standard interpretation of VBDD parameters when used to model commercial buildings.

Fels, M., Kissock, J.K. and Marean, M., 1994. "Model Selection Guidelines for PRISM (Or: Now That HC PRISM Is Coming, How Will I Know When to Use It?)", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, August.
Describes method for determining when to use HC models.

Kissock, J.K. and Fels, M., 1995. "An Assessment of PRISM's Reliability for Commercial Buildings", *Proceedings of the National Energy Program Evaluation Conference*, Chicago, IL, August.
Uses PRISM to model commercial building energy use. Finds adequate fits to monthly data.

Fels, M., Kissock, J.K., Marean, M. and Reynolds, C., 1995. "PRISM (Advanced Version 1.0) Users Guide", Center for Energy and Environmental Studies, Princeton University, Princeton, NJ, January.
Documentation for advanced version of PRISM with data graphics, Heating and Cooling model and automated model selection routine.

Change-Point Models
Neter, J. Wasserman, W. and Kutner, M., 1989, "Applied Linear Regression Models", Irwin Press, Boston, MA.
Statistical text. Describes general least-squares regression including piece-wise linear regression models.

Crawford, R., Dykowski, R. and Czajkowski, S., 1991, "A Segmented Linear Least-Squares Modeling Procedure for Nonlinear HVAC Components", ASHRAE Transactions, Vol. 97, Pt. 2, pp. 11-18.
Describes a segmented linear least-squares modeling procedure for deriving continuous, single-input, single-output models for HVAC equipment. The procedure begins by dividing the data into n bins along the x axis. Developing simple linear regression models for each bin would result in discontinuities between the linear segments. To overcome this problem, the bin widths are varied until the lines intersect at the bin boundaries. Identifying bin boundaries that meet the constraint of continuity between line segments requires an iterative solution of two matrix equations. Testing of the procedure indicated that the initial bin boundaries can affect whether converge will occur and the values at which converge will occur. Although an algorithm was

developed which assures convergence, the number of change-points cannot be determined in advance.

Ruch and Claridge, 1992. "A Four-Parameter Change-Point Model for Predicting Energy Consumption in Commercial Buildings", *ASME Journal of Solar Energy Engineering*, Vol. 114, No. 2, pp. 77 -83.
This paper develops a four-parameter change-point model of energy consumption as a function of dry-bulb temperature, along with accompanying error diagnostics for the model's parameters. The model is a generalization of the widely used three-parameter, or variable-base degree-day method. The algorithm used to fit the model to the data finds the optimal change-point temperature by searching within an interval known to contain T_{cp}. The first stage is to split the data points into two temperature regimes, fit ordinary least-squared lines in each regime, and calculate the intersection of the lines. This is repeated for numerous temperature regions. In the second stage, the T_{cp} is assumed and the model is fit using linear regression. From the collection of fits in the two stages, the algorithm chooses the one with the best least-squares fit. The reliability of the parameter estimated is then discussed. This four-parameter model is very comparable to the PRISM three-parameter model above the change-point, however, below the change-point the results vary significantly.

Ruch and Claridge, 1993. "A Development and Comparison of NAC Estimates for Linear and Change-Point Energy Models for Commercial Buildings", *Energy and Buildings*, Vol. 20, No. 1, pp. 87-95.
This paper develops the statistically rigorous methods for estimating NAC with four-parameter change-point and linear regression models. A rigorous statistical error analysis for NAC is also developed. The importance of a models goodness-of-fit and way of measuring it are discussed. The variable-base degree-day model incorporated in PRISM is basically a linear model of energy consumption as a function of temperature except that it assumes that the consumption is constant at a non-zero value to one side of a reference temperature. It is noted that NAC is linear in T rather than the average number of degree-days because the relationship between temperature and energy consumption does not break down at a reference temperature. The four-parameter change-point energy model is then given and discussed. The NAC for this model is also developed. The standard error and confidence intervals for NAC is calculated. Three models used to measure energy consumption are discussed and compared. These models include linear, PRISM, and four-parameter change-point.

Kissock, J.K, Xun,W., Sparks, R., Claridge, D., Mahoney, J. and Haberl, J., 1994. "EModel Version 1.4de", Copyright Texas A&M University, Energy Systems Laboratory, Department of Mechanical Engineering, Texas A&M University, College Station, TX, December.
Documentation for first release version of EModel with mean, two-parameter, three-parameter heating, three-parameter cooling, four-parameter, and multiple linear regression models. Includes accuracy testing against SAS.

Kissock, J.K., 1996, "Development of Analysis Tools in Support of the Texas LoanSTAR Program", University of Dayton, Department of Mechanical and Aerospace Engineering, Dayton, OH, August.

Describes new algorithm for five-parameter change-point model. Functional code is included.

Reddy, T., Saman, N., Claridge, D., Haberl, J., Turner, W. and Chalifoux, A., 1997, "Baselining methodology for facility level monthly energy use - Parts I and II", *ASHRAE Transactions*, Vol. 103, Pt. 2, pp. 336-347, 348-359.
Describes the use of change-point and variable-base degree-day models to determine weather dependent baseline energy use at large multi-building facilities.

Kissock, J.K., 1997. "Tracking Energy Use and Measuring Chiller Retrofit Savings Using WWW Weather Data and New ETracker Software", *Cool Sense National Forum on Integrated Chiller Retrofits*, San Francisco, CA, June 23-24.
Describes software that automatically selects four or five-parameter model for weather adjustment then estimates retrofit savings.

Haberl, Thamilseran, Reddy, Claridge, O'Neal, and Turner, 1998. "Baseline Calculations For Measurement And Verification Of Energy And Demand Savings In A Revolving Loan Program In Texas", *ASHRAE Transactions* Vol. 104, Pt. 2, pp. 841-858.
Measured hourly data are used to construct a baseline model. The data can then be used to predict building consumption had the retrofit not taken place. Measured post-retrofit data are compared to predicted data to determine savings. Two generic groupings of the basic modeling approach are regression models and calibrated engineering models. Regression models consist of billing and/or monitored data, utilized in one-, two-, three-, four-, or five-parameter change-point models, or MLR models. Discusses MLR, Change-point models. Discusses a large portion of one of the methods in question.

Kissock, Reddy and Claridge, 1998. "Ambient-Temperature Regression Analysis for Estimating Retrofit Savings in Commercial Buildings", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 168-176.
Describes a procedure for estimating weather-adjusted retrofit savings using ambient-temperature regression models. Selecting ambient-temperature as the sole independent regression variable is discussed. An approximate method for determining the uncertainty of savings is explained. The appropriate use of both linear and change-point models for measuring energy savings is also discussed. Ambient-Temperature is used as the single independent variable because both it eliminates problems associated with multi-collinearity problems and reduces data collection to a single easily acquired parameter. Paper discusses mathematical approach in measuring uncertainty of weather adjustment along with the best time scale to use for the data. Two- to Five-point regression models are presented to model weather-dependent energy use. The paper directly explains an algorithm used to measure energy savings. Search methods are used to model energy consumption data and the best-fit model according to statistical methods is found. This method dominates the literature found on the subject.

Multivariate Regression Models

Forrester, J. and Wepfer, W., 1984, "Formulation of a Load Prediction Algorithm for a Large Commercial Building", *ASHRAE Transactions*, Vol. 90, Pt 1, pp. 536 - 551.
Describes adaptive multivariate regression model of electrical demand for a large commercial building. Uses model to reduce peak demand.

Omnicomp, 1984, Omnicomp. Inc., State College, PA.
Energy accounting software with VBDD modeling capability.

Metrix, SRC Systems Inc., Berkeley, CA.
Energy accounting software with VBDD modeling capability.

Leslie, N., G. Aveta and B. Sllwinski, 1986. "Regression Based Process Energy Analysis System", *ASHRAE Transactions*, Vol. 92, Pt. 1A., pp, 93-102.
The results of an investigation are presented to determine which weather, production, and time-related parameters exert significant influence on energy consumption. The regression model shows that energy consumption in general depends on heating degree-days, production level, and labor force strength. Extensive gathering of production data and energy data was performed. Data gathered included production level by product class, heating degree-days, cooling degree-days, energy consumed by fuel type, labor force, direct and indirect man hours, etc. The best predictors among competing parameters were selected based on maximizing the adjusted multiple correlation coefficient. In general, heating degree-days and cooling degree-days are the most important parameter for predicting total energy consumption, with labor force strength and production level providing additional explanatory power.

Anderson, D., 1990, "Electrical Usage Predictors Based on the Singular Value Decomposition Algorithm", M.S. Thesis, Civil, Environmental and Architectural Engineering Department, University of Colorado at Boulder.
Uses singular value decomposition to reduce the effects of multicollinearity in a multivariate regression analysis of electricity use.

Ruch, Chen, Haberl and Claridge, 1993. "A Change-Point Principal Component Analysis (CP/PCA) Method for Predicting Energy Usage in Commercial Buildings: The PCA Model", *ASME Journal of Solar Energy Engineering*, Vol. 115, No. 2, pp. 77-84.
This method utilizes a Principal Component Analysis of intercorrelated influencing parameters (e.g., dry-bulb temperature, solar radiation and humidity) to predict electricity consumption in conjunction with a change-point model. This paper describes the PCA procedure and presents the results of its application in conjunction with a change-point regression, to predict whole-building electricity consumption. Comparison of the results with a traditional MLR analysis indicates that this method is a better predictor than a MLR analysis and offers more insight into the environmental and operational driving forces that influence energy consumption. The PCA method transforms the original variables into an uncorrelated set of orthogonal variables that are linear combinations of the original variables. These new variables, called principal components, retain all of the information of the original variables. Therefore, MLR can be used without compromise associated with variable correlation.

Reddy, T. and D. Claridge. 1994. "Using Synthetic Data To Evaluate Multiple Regression And Principal Component Analyses For Statistical Modeling Of Daily Building Energy Consumption", *Energy and Buildings*, Vol. 21, No.1, pp. 35-44.
Discusses multiple regression modeling and principle component analysis. MRA has been faulted as a means of predicting energy use because of the multicollinearity between the regressor variables. PCA has the potential to overcome this drawback of MRA. This paper gives a broad evaluation of each technique and some guidelines under which one approach is preferable. When using more than one climatic variable, PCA can remove the multicollinear effects in the regressor variables. Can be applied to a general discussion of measuring energy consumption in buildings.

Austin, S. 1997. "Regression Analysis for Savings Verification". *ASHRAE Journal*, Vol. 39.
Describes linear, polynomial, multiple linear regression and use of dummy variables to group data into categories. Examples include regression of chiller efficiency versus percent load and condensor water temperature, compressor output versus air temperature and washing cycles, and utility load versus time. Mentions splitting data into groups to get better fit, but not how to force lines to meet at group division (i.e. change-point model).

Katipamula, S., T. Reddy and D. Claridge. 1998. "Multivariate Regression Modeling". *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 177-184.
As a result of energy consumption in large commercial buildings being a complicated function, MLR provides better accuracy than a single variable model for modeling energy consumption. This paper also addresses the best time resolution of data to adopt to make the regression most accurate. Many independent variables have been used to perform an MLR model including, cooling-degree days, heating-degree days, wind speed and direction, humidity, refrigeration type, exhaust air, supply air, average shading in winter, average shading in summer and so on. Different buildings used different independent variables, some up to ten others as few as two. MLR models based on engineering principles are difficult to develop because they require knowledge of the HVAC system operation and how it related to other building systems. Another disadvantage of MLR is the variables should be independent of each other, which is not the case in reality.

MLR models for cooling energy consumption with DDCV systems and DDVAV systems are presented. For these models it has been determined that collinearity is not significant between T_o, T_{dp}+ and q_i at daily and hourly time scales, but is significant between T_o and q_{sol}. Five buildings in central Texas were modeled using piecewise MLR. Stepwise regression, used to show the contribution of each individual variable, is presented for the buildings that were used. The outside dry-bulb temperature is shown to account for over 87% of the cooling energy use for the DDCV model and 83% for the DDVAV model. Time scales used were monthly, daily, hourly, and HOD. Advantages and disadvantages of different time scales for modeling effort, metering and monitoring, data needed for robust modeling, applicability to savings measurements, prediction uncertainty, O&M opportunities detection and dynamic control are presented.
The MLR method is capable of measuring retrofit energy savings and identifying O&M problems.

Combination VBDD/CP/MVR Models

Rabl, A and A, Rialhe. 1992. "Energy Signature Models for Commercial Buildings: Test With Measured Data And Interpretation", *Energy and Buildings*, Vol. 19, No. 2, pp. 143 - 154.

This paper discusses the advantages of including occupancy as an additional variable to the energy signature model PRISM.

Sonderregger, R. 1997. "Energy Retrofits in Performance Contracts: Linking Modeling and Tracking". *Cool Sense National Forum on Integrated Chiller Retrofits*, San Francisco, September 23-24.

Paper begins by describing an energy accounting method for modeling utility billing data. In this method, monthly energy use from the "tuning" (baseline or pre-retrofit) period is regressed against variable base heating and cooling degree days and other independent variables. A method for predicting savings based on modifying the regression coefficients from the tuning period, and the limitations of this method, are described. The attributes and limitations of simulation models predict energy savings are described. A method to calibrate simulation results using the previous regression method is described. In this method, simulation results are regressed against the same independent variables as utility bills. The simulation results are acceptable when the ratios of coefficients from the simulation regression are similar to the ratios of coefficients from the utility regression. Savings can then be predicted from the percent change of the simulation regression coefficients.

Sonderegger, R. A., 1998. "Baseline Model for Utility Bill Analysis Using Both Weather and Non-Weather-Related Variables", *ASHRAE Transactions*, Vol. 104, No. 2 , pp. 859-870.

Paper:

- discusses how to determine tuning (or baseline) period, even in the presence of plug creep.
- proposes a general baseline regression equation using utility billing data as the dependent variable and period length, variable-base heating and cooling degree days and other influential parameters as the independent variables
- describes a method for fitting other influential data to utility billing periods.
- describes least squares regression, good-ness of fit and how to use the t statistic to find if a coefficient should be included in the model.
- describes sequential method of regression where degree-day base temperatures are selected first, then additional independent variables are added as needed.
- describes how large ranges of base-temperatures will produce the same goodness of fit.
- shows equivalence of VBDD models and mean temperature models when mean temperature models are used with daily temperatures.
- shows examples where additional non-weather variables increase the goodness-of-fit.

Kissock, J., 1999, "VBDDSave", University of Dayton, Dayton, OH.

Software to integrate utility billing and temperature data, calculate the best base temperature for a VBDD model, and use these VBDD in a MVR model.

Calibrated Simulation Models

Katipamula, S. and D. Claridge. 1993. "Use of Simplified System Models to Measure Retrofit Energy Savings", *ASME Journal of Solar Energy Engineering*, Vol. 115, No. 2, pp. 57-68.
This paper describes a method that can be used to calculate energy savings when no pre-retrofit data are available. The method is based on use of simplified calibrated system models. A VAV model was developed based on the ASHRAE TC 4.7 Simplified Energy Analysis Procedure and calibrated with post-retrofit data from a building in central Texas. Climate data, building data, and HVAC data are used to simulate the post-retrofit system. In the absence of pre-retrofit data, savings can be estimated by predicting the pre-retrofit system behavior with the use of an hourly simulation model.

Wilson, J., 1998. "The Significant Role of Energy Calculations in the Success of Long-Term Energy Guarantees. *ASHRAE Transactions*, Vol. 104, No.2, pp. 880-894.
Situations can arise that change the energy consumption behavior of a building. When this happens, changes in the pre-retrofit baseline need to be made to ensure the integrity of a performance contract. This paper discusses a calculated baseline adjustment to provide an effective method of accommodating the change while still retaining the basic tenants of the original energy guarantee. This paper presents examples of these types of situations and how to adjust the baseline calculation.

Artificial Neural Networks

Anstett, M. and J. Kreider, 1993. "Application of Neural Networking Models to Predict Energy Use". *ASHRAE Transactions*, Vol. 99, Pt. 1, pp. 505-517.
Discusses the application of an artificial neural network model to predict energy use in a complex institutional building without the need for a data acquisition system. Discussion of general building energy consumption techniques will discuss neural networks.

MacKay, D., 1994. "Bayesian Nonlinear Modeling for the Prediction Competition", *ASHRAE Transactions*, Vol. 100, No.2, pp. 1053-1062.
Winner of the 1993 energy prediction competition. Removed autocorrelation from independent variables using principle component analysis. Then used neural network method with automatic relevance determination for the regression parameters.

Feuston, B. and J. Thurtell. 1994. "Generalized Nonlinear Regression With Ensemble Of Neural Nets: The Great Energy Predictor Shootout", *ASHRAE Transactions*, Vol. 100, No. 2, pp. 1075-1080.
Discusses the use of neural networks to model whole building electric, chilled water, and hot water. Five independent parameters are used: time stamp, dry-bulb temperature, humidity ratio, solar flux, and wind speed. The technique prediction compared to actual consumption is presented. Can be applied to a general discussion of measuring energy consumption in buildings.

Kissock, J.K., 1994. "Modeling Commercial Building Energy Use with Artificial Neural Networks", *Proceedings of the 29th Intersociety Energy Conversion Engineering Conference*, Vol. 3, pp. 1290-1295, Monterey, CA, August.
General description of neural network method and comparison of simple neural networks with regression modeling.

Kreider, Claridge, Curtiss, Dodier, Haberl and Krarti, 1995. "Building Energy Use Prediction and System Identification Using Neural Networks" *ASME Journal of Solar Energy Engineering*, Vol. 117, No. 3, pp. 161-166.
This paper addresses the difficult task of predicting energy consumption well into the future without knowledge of immediately past energy consumption. Discussion of general building energy consumption techniques will discuss neural networks.

Curtiss, P., G. Shavit and J. Kreider. 1996. "Neural Networks Applied To Buildings - A Tutorial And Case Studies In Prediction And Adaptive Control". *ASHRAE Transactions*, Vol. 102, No. 1, pp. 1141-1146.
The paper discusses the use of neural networks to predict building energy use and measure retrofit savings. Produces small RMS errors. Use for broad discussion of measuring energy savings.

Krarti, M., J. Kreider, D. Cohen and P. Curtiss. 1998. "Estimation of Energy Savings for Building Retrofits Using Neural Networks", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 211-216.
Overviews the use of neural networks to estimate energy and demand savings from retrofits of commercial buildings. Also included is a brief background on neural networks along with three case studies to demonstrate how to successfully implement neural networks. Neural networks provide superior accuracy for predicting energy use in buildings. Weather data, occupancy profiles, and day types are generally considered as inputs into NN's to predict building energy use. Neural networks consist of several layers of neurons that are connected to each other via transport links. Connection strengths (weights) between the neurons are adjusted to produce the desired outcome.
Can be used to demonstrate other non-regression type methods of predicting building energy use.

Advanced Regression Techniques

Ionides, G., 1984. "Effect of Statistical Measuring Errors on the Goodness of Fit in Linear Regression". Tappi Journal. Vol. 67, No. 11, pp. 114-115.
The degree of dependence of the dependent variable on one or more independent variables is influenced both by the extent to which variables are physically related and by the precision with which the variables are measured. A technique is described in which the influence on R of statistical measuring errors in the dependent and independent variables can be separated out.

Efron, B., 1988. "Computer-Intensive Methods in Statistical Regression". *SIAM Review*. Vol. 30, No. 3, pp. 421-449.
This is a survey of modern developments in statistical regression. Topics discussed include robust regression, bootstrap measures of variability, local smoothing and cross-validation,

projection pursuit, Mallows' C_p criterion, Stein estimation, generalized regression for Poisson data, and regression methods for censored data.

Neri, F., 1989. "An Accurate And Straightforward Approach To Line Regression Analysis Of Error-Affected Experimental Data". Journal of Physics. Vol 22, pp. 215-217.
Regression technique using the minimization of the shortest distance between each experimental point and the theoretical line.

Itakura, H., 1993. "A Solution to Multiple Linear Regression Problems With Ordered Attributes". *Computers Mathematical Applications*. Vol. 25, No. 2, pp. 47-57.
A class of multiple linear regression techniques is discussed, in which the order of magnitude is constrained among regression coefficients. The problem to be solved is reduced to a quadratic programming problem in which the objective function is the residual sum of the squares in regression, and the constraints are linear ones imposed on the regression coefficients.

Nievergelt, Y., 1994. "Total Least Squared: State-of-the-Art Regression in Numerical Analysis". *SIAM Review*. V36, n2, pp. 258 - 263.
Classroom notes for regression analysis. Includes elementary algorithm for total least squares fits in numerical and applied analysis.

Iijima, M., K. Takagi, R. Takeuchi and T. Matsumoto. 1994. "Piecewise-Linear Regression On The ASHRAE Time-Series Data", *ASHRAE Transactions*, Vol. 100, Pt. 2, pp. 1088-1095.
Using piecewise-linear regression in the ASHRAE Building Energy Prediction Shootout.
Can be applied to a general discussion of measuring energy consumption in buildings.
There is a distinct difference between energy consumption on workdays and that on holidays or weekends.

Dhar,A., Reddy, T.A., Claridge, D., 1988. "Modeling Hourly Energy Use in Commercial Buildings with Fourier Series Functional Forms", *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 217-223.

Uncertainty of Savings

Reddy, T.A, Kissock, J.K. and Claridge, D.E., 1992. "Uncertainty Analysis in Estimating Building Energy Retrofit Savings in the Texas LoanSTAR Program", *Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings*, Pacific Grove, CA, August.
Describes methods to estimate the uncertainty of savings from linear regression models.

Kissock, J., T. Agami, D. Fletcher and D. Claridge. 1993. "The Effect of Short Data Periods on the Annual Prediction Accuracy of Temperature-Dependent Regression *International Solar Engineering Conference*, pp. 455 - 463.
Ideally, a full year or more of energy use and weather data should be used for empirical energy consumption models. Sometimes a full year of data is not available and one is constrained to develop a model using the ideal full year of data. This paper examines how temperature

dependent regression models of energy use based on periods of less than one-year compare to models developed using a full year's worth of data. Models using data sets of one, three and five months were explored.

Models based on three months of data for the case of chilled water consumption varied from 4% to 20%, however, heating energy use varied as much as 400%. These results are based off buildings located in central Texas. The degree of error is therefore climate dependent. Models based off of short data periods are shown to have the potential of being severally erroneous. Use data periods of a year or more.

Katipamula,S., T. Reddy and D. Claridge. 1995. Effect of Time Resolution on Statistical Modeling of Cooling Energy Use in Large Commercial Buildings, *ASHRAE Transactions*, Vol.101, Pt. 2, pp. 172-185.

The question arises as to the best time resolution is most accurate when hourly monitored data are available. This paper addresses this question by comparing monthly, daily, hourly and individual hourly or hour-of-day multiple linear regression models when applied to measured cooling consumption in commercial buildings. The advantages and disadvantages associated with each model are also presented. The outdoor dry-bulb and dew-point temperatures account for most of the variation in a buildings energy consumption. Monthly models had higher model R2 then daily, hourly, and HOD models. However, daily and HOD models proved more accurate in determining cooling energy consumption. Monthly and daily time scales are preferred because some operational parameters such as internal heat gain that change on an hourly basis is constant on a daily basis. They also require less effort in data collecting. Modeling a large commercial building using monthly data requires 12 months or more of data. Daily time scale models are most advantageous for retrofit savings determination. The HOD time scale best models O&M.

Reddy, T., J. Kissock and D. Ruch. 1998. Uncertainty In Baseline Regression Modeling And In Determination Of Retrofit Savings. *ASME Journal of Solar Energy Engineering*, Vol. 120, No. 3, pp. 185-192.

The various sources of uncertainty inherent in the estimation of measuring energy savings from baseline models and the statistics involved in determining the uncertainty is presented. Improper model residuals along with how model predictions are effected by incomplete data periods not allowing for the entire range of variation of climatic conditions are addressed.

Kissock, K., H. Joseph and J. McBride. 1998. "The Effects of Varying Indoor Air Temperature and Heat Gain on the Measurement of Retrofit Savings" *ASHRAE Transactions*", Vol. 104, Pt. 2., pp. 895-900.

Many methods used to measure energy savings between a pre- and post-retrofit period use weather dependent variables for the model. These implicitly assume that the indoor air set-point temperature and internal gains are the same before and after the retrofit. This paper develops expressions that suggest that retrofit savings are highly sensitive to minor indoor air temperature changes and internal heat gains. Many baseline energy models use only outside air temperature as an indicator of weather conditions because of the relative magnitudes of the conduction and sensible air-conditioning loads and because of the high correlation between outside air temperature and other environmental variables. In simple buildings, the accuracy of estimated savings could be significantly improved by adding indoor air-temperature in the baseline model. This paper explains how varying inside air temperature and internal heat gain affect estimated savings.

Ruch, D.K., Kissock, J.K. and Reddy, T.A., 1999. "Model Identification and Prediction Uncertainty of Linear Building Energy Use Models with Autocorrelated Residuals ", *ASME Journal of Solar Energy Engineering* , Vol.121, No.1, pp. 63-68.
Autocorrelated residuals from regression models of building energy use present problems when attempting to estimate retrofit energy savings and the uncertainty of the savings. The causes of autocorrelation in energy use models and methods of dealing with autocorrelation are discussed. To accurately predict energy use and give realistic uncertainty estimated a hybrid of ordinary least squares (OLS) and autoregressive models (AR) is used. The hybrid OLS-AR model has been proven to provide more accurate uncertainty estimated than the OLS estimate. The presence of autocorrelation causes statistical problems. Estimated prediction error bounds will be too small, leading to undue confidence being placed on the accuracy of predicted energy use. Outside temperature is an easily measured variable, whereas, humidity and internal loads are difficult and sometimes expensive to measure. Therefore, the omission of important predictor variables from the model and the consequent autocorrelation of model residuals may be unavoidable. The hybrid approach to predicting pre-retrofit energy consumption in the post-retrofit period benefits from the prediction accuracy of OLS regression coefficients but does not use the standard OLS error diagnostics that are inaccurate when autocorrelation is present. For many buildings it may not be possible to eliminate autocorrelation through model redesign, therefore, a hybrid model is used. The derivation of this model is provided along with error discussion.

APPENDIX III: DETAILED TEST RESULTS

Development of a Toolkit for Calculating Linear, Change-point Linear and Multiple-linear Inverse Building Energy Analysis Models

ASHRAE Research Project 1050-RP

Detailed Test Results

Atch Sreshthaputra

Jeff S. Haberl, Ph.D., P.E.
David E. Claridge, Ph.D., P.E.

May 2001
Updated August 2001

ENERGY SYSTEMS LABORATORY

Texas Engineering Experiment Station
Texas A&M University System
College Station, TX 77843

PREFACE

This CD-ROM contains detailed test results of the Inverse Modeling Toolkit software (IMT), which was developed for ASHRAE Research Project 1050-RP. The test files included in the CD-ROM are divided into 15 subdirectories with their names referring to the types of models that were tested (i.e., 1P, 2P, 3P_COOL, 3P_HEAT, 3P_MVR, 4P, 4P_MVR, 5P, 5P_MVR, CDD, CDD_MVR, HDD, HDD_MVR, MVR, and Site_test). Each test that was performed consists of four different IMT file types, which include: 1) IMT instruction file (.INS), 2) IMT data file (.DAT), 3) IMT output file (IMT.OUT), and 4) IMT residual file (IMT.RES). The residual files are included for those tests that needed further testing or error checking, which is the case of CDD-MVR and HDD-MVR tests.

Some tests were performed to compare the results against those calculated by other programs, which include EModel (Kissock et al., 1996), SAS (SAS Institute Inc., 2001), and PRISM (Fels, M. et al., 1986). Each EModel test contains three files, including a data file (.DAT), an instruction file (.DVN), and an output file (.DOC). IMT and EModel share the same data file (.DAT). Each SAS run contains a procedure file (.SAS) and an output file (.LST) and also shares the same data file as IMT. For each PRISM run, there are three files included: a weather file (.TPS), a data file or meter file (.MTR), and an output file (.DOC).

This report is named "summary.doc" and it is located in the main directory of the CD-ROM. The IMT program is also included in this CD in the "IMT" subdirectory.

ABSTRACT

This is the detailed test report for the ASHRAE 1050-RP project. This report presents the detailed results of the testing of IMT (Inverse Modeling Toolkit). Two kinds of testing were performed, bounds testing and accuracy testing. The bounds testing is performed in order to identify what types of data sets the IMT program can model reliably. A variety of data sets were used to test the limits of the program: 1) Data sets with as few as two and as many as 9,000 data points, 2) Data sets with very large and very small numbers, 3) Data set with a variety of slopes, and 4) Data sets with tightly packed and widely scattered observations.

In terms of accuracy test, 1P, 2P and MVR models were benchmarked against the statistical software SAS (SAS Institute Inc., 2001). The change-point model results (3P and 4P) were compared to those calculated by the data analysis software EModel (Kissock et al., 1996). Finally, the IMT's HDD and CDD models were compared to PRISM HO and CO models (Fels, M. et al., 1986).

TABLE OF CONTENTS

1. CD-ROM CONTENTS

The followings are the tests contained in this CD-ROM. X: is assumed to be the CD-ROM drive.

1.1 Mean Model, synthetic data sets (X:\1P)

1.1.1 1-Point (1P_test0)
1.1.2 2-Point (1P_test1)
1.1.3 Scattered (1P_test2)
1.1.4 Packed (1P_test3)
1.1.5 9,000-Point (1P_test4)
1.1.6 Large numbers (1P_test5)
1.1.7 Small numbers (1P_test6)

1.2 Two-Parameter Model, synthetic data sets (X:\2P)

1.2.1 2-Point (2P_test0)
1.2.2 3-Point (2P_test1)
1.2.3 Scattered (2P_test2)
1.2.4 Packed (2P_test3)
1.2.5 9,000-Point (2P_test4)
1.2.6 Large numbers (2P_test5)
1.2.7 Small numbers (2P_test6)
1.2.8 Maximum X (2P_test7)
1.2.9 Maximum Y (2P_test8)
1.2.10 Slope A (2P_test9)
1.2.11 Slope B (2P_test10)
1.2.12 Slope C (2P_test11)
1.2.13 Slope D (2P_test12)

1.3 Three-Parameter Change-Point Cooling Model, synthetic data sets (X:\3P_COOL)

1.3.1 3-Point (3PC_test0)
1.3.2 5-Point (3PC_test1)
1.3.3 Scattered (3PC_test2)
1.3.4 Packed (3PC_test3)
1.3.5 9,000-Point (3PC_test4)
1.3.6 Large numbers (3PC_test5)
1.3.7 Small numbers (3PC_test6)
1.3.8 Maximum X (3PC_test7)
1.3.9 Maximum Y (3PC_test8)
1.3.10 Slope A (3PC_test9)
1.3.11 Slope B (3PC_test10)
1.3.12 Slope C (3PC_test11)

1.3.13 Slope D (3PC_test12)

1.4 Three-Parameter Change-Point Heating Model, synthetic data sets (X:\3P_HEAT)

1.4.1 3-point (3PH_test0)
1.4.2 5-Point (3PH_test1)
1.4.3 Scattered (3PH_test1)
1.4.4 Packed (3PH_test3)
1.4.5 9,000-Point (3PH_test4)
1.4.6 Large numbers (3PH_test5)
1.4.7 Small numbers (3PH_test6)
1.4.8 Maximum X (3PH_test7)
1.4.9 Maximum Y (3PH_test8)
1.4.10 Slope A (3PH_test9)
1.4.11 Slope B (3PH_test10)
1.4.12 Slope C (3PH_test11)
1.4.13 Slope D (3PH_test12)

1.5 Three-Parameter Change-Point with Multiple Variable Regression Model (X:\3P_MVR)

1.5.1 WBCOOL VS Temperature (3PC_Mvr1)
1.5.2 WBCOOL VS Humidity Ratio (3PC_Mvr2)
1.5.3 WBCOOL VS Solar Radiation (3PC_Mvr3)
1.5.4 WBCOOL VS WBHEAT (3PC_Mvr4)
1.5.5 WBCOOL VS WBE (3PC_Mvr5)
1.5.6 WBCOOL VS Temperature, humidity ratio, solar radiation, WBH, and WBE (3PC_Mvr6)
1.5.7 WBCOOL VS Temperature, humidity ratio, solar radiation, and WBH (3PC_Mvr7)
1.5.8 WBHEAT VS Temperature (3PH_Mvr1)
1.5.9 WBHEAT VS Humidity Ratio (3PH_Mvr2)
1.5.10 WBHEAT VS Solar Radiation (3PH_Mvr3)
1.5.11 WBHEAT VS Temperature, humidity ratio, solar radiation (3PH_Mvr4)

1.6 Four-Parameter Change-Point Model, synthetic data sets (X:\4P)

1.6.1 3-Point (4P_test0)
1.6.2 5-Point (4P_test1)
1.6.3 Scattered (4P_test2)
1.6.4 Packed (4P_test3)
1.6.5 9,000-Point (4P_test4)
1.6.6 Large numbers (4P_test5)
1.6.7 Small numbers (4P_test6)
1.6.8 Maximum X (4P_test7)
1.6.9 Maximum Y (4P_test8)
1.6.10 Slope A (4P_test9)

1.6.11 Slope B (4P_test10)
1.6.12 Slope C (4P_test11)
1.6.13 Slope D (4P_test12)
1.6.14 Slope E (4P_test13)
1.6.15 Slope F (4P_test14)
1.6.16 Slope G (4P_test15)
1.6.17 Slope H (4P_test16)
1.6.18 Slope I (4P_test17)

1.7 Four-Parameter Change-Point with Multiple Variable Regression Model (X:\4P_MVR)

1.7.1 WBCOOL VS Temperature (4PC_Mvr1)
1.7.2 WBCOOL VS Humidity Ratio (4PC_Mvr2)
1.7.3 WBCOOL VS Solar Radiation (4PC_Mvr3)
1.7.4 WBCOOL VS WBHEAT (4PC_Mvr4)
1.7.5 WBCOOL VS WBE (4PC_Mvr5)
1.7.6 WBCOOL VS Temperature, humidity ratio, solar radiation, WBH, and WBE (4PC_Mvr6)
1.7.7 WBCOOL VS Temperature, humidity ratio, and solar radiation (4PC_Mvr7)
1.7.8 WBHEAT VS Temperature (4PH_Mvr1)
1.7.9 WBHEAT VS Humidity Ratio (4PH_Mvr2)
1.7.10 WBHEAT VS Solar Radiation (4PH_Mvr3)
1.7.11 WBHEAT VS Temperature, humidity ratio, solar radiation (4PH_Mvr4)

1.8 Five-Parameter Change-Point Model, synthetic data sets (X:\5P)

1.8.1 4-Point (5P_test0)
1.8.2 7-Point (5P_test1)
1.8.3 Scattered (5P_test2)
1.8.4 Packed (5P_test3)
1.8.5 9,000-Point (5P_test4)
1.8.6 Large numbers (5P_test5)
1.8.7 Small numbers (5P_test6)
1.8.8 Maximum X (5P_test7)
1.8.9 Maximum Y (5P_test8)
1.8.10 Slope A (5P_test9)
1.8.11 Slope B (5P_test10)
1.8.12 Slope C (5P_test11)
1.8.13 Slope D (5P_test12)
1.8.14 Slope E (5P_test13)
1.8.15 Slope F (5P_test14)

1.9 Five-Parameter Change-Point with Multiple Variable Regression Model (X:\5P_MVR)

1.9.1 MCC VS Temperature (5P_Mvr1)

1.9.2 MCC VS Humidity Ratio (5P_Mvr2)
1.9.3 MCC VS Solar Radiation (5P_Mvr3)
1.9.4 MCC VS Temperature, humidity ratio, and solar radiation (5P_Mvr4)
1.9.5 MCC VS Temperature and humidity ratio (5P_Mvr5)

1.10 Variable-Base Cooling Degree-Day Model (X:\CDD)

1.10.1 Whole Building Electricity Use Per Day
1.10.2 Whole Building Electricity Use Per Billing Period

1.11 Variable-Base Cooling Degree-Day with Multiple Variable Regression Model (X:\CDD_MVR)

1.11.1 WBCOOL VS Temperature (CDD_Mvr1)
1.11.2 WBCOOL VS Humidity ratio (CDD_Mvr2)
1.11.3 WBCOOL VS Solar radiation (CDD_Mvr3)
1.11.4 WBCOOL VS Temperature, humidity ratio, and solar radiation (CDD_Mvr4)
1.11.5 WBCOOL VS Temperature, humidity ratio, and solar radiation using the CDD residual file as input to the MVR model to produce CDD-MVR capabilities (CDD_Mvr5)

1.12 Variable-Base Heating Degree-Day Model (X:\HDD)

1.12.1 Whole Building Heating Energy Use Per Day
1.12.2 Whole Building Heating Energy Use Per Billing Period

1.13 Variable-Base Heating Degree-Day with Multiple Variable Regression Model (X:\HDD_MVR)

1.13.1 WBHEAT VS Temperature (HDD_Mvr1)
1.13.2 WBHEAT VS Humidity ratio (HDD_Mvr2)
1.13.3 WBHEAT VS Solar radiation (HDD_Mvr3)
1.13.4 WBHEAT VS Temperature, humidity ratio, and solar radiation (HDD_Mvr4)
1.13.5 WBHEAT VS Temperature, humidity ratio, and solar radiation using the HDD residual file as input to the MVR model to produce HDD-MVR capabilities (HDD_Mvr5)

1.14 Multiple Variable Regression Model (X:\MVR)

1.14.1 Synthetic data (MVR_0)
1.14.2 WBE VS Temperature, humidity ratio, and solar radiation (MVR_1)
1.14.3 WBCOOL VS WBE, WBHEAT, temperature, humidity ratio, and solar radiation (MVR_2)

1.14.4 WBHEAT VS WBE, WBCOOL, temperature, humidity ratio, and solar radiation (MVR_3)

1.15 LoanSTAR Data Sets (X:\Site_test)

1.15.1 Mean Model (1P_Comp)
1.15.2 Two-Parameter Model (2P_Comp)
1.15.3 Three-Parameter Cooling Model (3PC_Comp)
1.15.4 Three-Parameter Heating Model (3PH_Comp)
1.15.5 Four-Parameter Cooling Model (4P_Comp)
1.15.6 Five-Parameter Cooling Model (5P_Comp)

2. SUMMARY

The results of each test are generally presented in two tables. One table, for example Table 1.1, contains a list of files used in performing the test, including types of data sets and tests and general comments about the results.

The "Data Type" column shows types of data sets and testing. The files with .INS and .DAT extensions are the IMT instruction and data files respectively. EModel and IMT share the same input data files, .DAT files, which are formatted, space-delimited ASCII text files. The .DOC files are the EModel output files, which are MS WORD document files. The files in the "SAS" column are the SAS input (.SAS) and output files (.LST), which can be opened with any text editor program. In addition, if PRISM was used in performing the test, for each PRISM run, there are three files included: a weather file (.TPS), a data file or meter file (.MTR), and an output file (.DOC). The "Status" column summarized the results of IMT as compared to other programs used.

In Table 1.2, the detailed results from the IMT bounds testing are shown, along with the comparison testing with other programs (e.g., EModel, PRISM, and SAS).

- Table 1.1 contains a list of the files used in performing one-parameter (1P) and two-parameter (2P) tests of IMT against EModel and SAS. The input data for IMT are synthetic and generated with known values and coefficients in order to perform accuracy tests of IMT. The results indicate that the minimum number of observations for the Mean model is two data points. In terms of magnitude, IMT ran correctly and produced output of the numbers with absolute values as small as 3.3×10^{-57} and as large as 1×10^{18}. Each model was successfully tested using 9,000 observations. For the 2P model, IMT successfully modeled data sets with slopes greater than or less than zero, and slopes less than infinity (i.e., vertical). Following Table 1.1 is Table 1.2, which contains the detailed outputs from the three programs. Generally, IMT, EModel, and SAS produced outputs in good agreement with each other.

- Table 2.1 contains a list of the files used in performing three-parameter change-point cooling (3PC) and heating (3PH) model tests of IMT against EModel. The input data for IMT are synthetic and generated with known values and coefficients in order to perform accuracy tests of IMT. The results indicate that the minimum number of observations for the 3P model is five. In terms of magnitude, IMT ran correctly and produced output of the numbers with absolute values as small as 3.3×10^{-57} and as large as 1×10^{18}. Each model was successfully tested using 9,000 observations. For the 3P model, IMT successfully modeled data sets with slopes greater than or less than zero, and slopes less than infinity (i.e., vertical), but it failed to identify flat slopes (i.e., Slope A). Following Table 2.1 is Table 2.2, which contains the detailed outputs from the two programs. Generally, IMT and EModel produced outputs in good agreement with each other.

- Table 3.1 contains a list of the files used in performing four-parameter change-point (4P) models tests of IMT, also against EModel. The input data for IMT are synthetic and generated with known values and coefficients in order to perform accuracy tests of IMT. The results indicate that the minimum number of observations for the 4P model is five. In terms of magnitude, IMT ran correctly and produced output of the numbers with absolute values as small as 3.3 x 10^{-57} and as large as 1 x 10^{18}. Each model was successfully tested using 9,000 observations. For the 4P model, IMT successfully modeled data sets with slopes greater than or less than zero, and slopes less than infinity (i.e., vertical), but it failed to identify flat slopes (i.e., Slope A). Following Table 3.1 is Table 3.2, which contains the detailed outputs from the two programs. Generally, IMT and EModel produced outputs in good agreement with each other.

- Table 4.1 contains a list of the files used in performing the IMT tests of five-parameter change-point (5P) and five-parameter with multiple variable regression models (5P/MVR). For the 5P model, the input data for IMT are synthetic and generated with known values and coefficients. The results indicate that the minimum number of observations for the 5P model is seven. In terms of magnitude, IMT ran correctly and produced output of the numbers with absolute values as small as 3.3 x 10^{-57} and as large as 1 x 10^{18}. Each model was successfully tested using 9,000 observations. For the 5P model, IMT successfully modeled data sets with slopes greater than or less than zero, and slopes less than infinity, but it failed to identify flat slopes (i.e., Slope A).

  For the five-parameter change-point with multiple variable regression models (5P/MVR), the input data were obtained from the LoanSTAR database. The building used for these tests was the Zachry Engineering Center, Texas A&M University for the period of 1/1/99 to 12/31/99. The dependent variable is the energy consumption of the VAV motor control center (MCC). The independent variables include outdoor temperature, humidity ratio, and solar radiation. The results indicated that the maximum number of independent variables is two. Following Table 4.1 are Table 4.2 and Table 4.3, which contain the detailed outputs from IMT. No comparison tests were run for the 5P/MVR model.

- Table 5.1 contains a list of the files used in performing the IMT tests of three-parameter change-point with multiple variable regression models (3P/MVR). The input data were obtained from the LoanSTAR database. For the 3PC/MVR model, the dependent variable is the whole-building cooling energy consumption (WBC). The independent variables include outdoor temperature, humidity ratio, solar radiation, whole-building heating energy (WBH), and whole-building electricity consumption (WBE). For the 3PH/MVR model, the dependent variable is the whole-building heating energy consumption (WBH). The independent variables include outdoor temperature, humidity ratio, and solar radiation. The results indicated that the maximum number of independent variables is four. Following Table 5.1 is Table 5.2, which contains the detailed outputs from IMT. No comparison tests were run for the 3P/MVR model.

- Table 6.1 contains a list of the files used in performing the IMT tests of four-parameter change-point with multiple variable regression models (4P/MVR). The input data were obtained from the LoanSTAR database. For the 4PC/MVR model, the dependent variable is the whole-building cooling energy consumption (WBC). The independent variables include outdoor temperature, humidity ratio, solar radiation, whole-building heating energy (WBH), and whole-building electricity consumption (WBE). For the 4PH/MVR model, the dependent variable is the whole-building heating energy consumption (WBH). The independent variables include outdoor temperature, humidity ratio, and solar radiation. The results indicated that the maximum number of independent variables is three. Following Table 6.1 is Table 6.2, which contains the detailed outputs from IMT. No comparison tests were run for the 4P/MVR model.

- Table 7.1 contains a list of the files used in performing the IMT tests against EModel and SAS of real data using several models (e.g., 1P, 2P, 3PC, 3PH, 4P, and 5P). The input data were obtained from several LoanSTAR buildings. The "Data Type" column shows LoanSTAR building ID and data channels that were used. Generally, IMT, EModel, and SAS produced outputs in good agreement with each other. Following Table 7.1 is Table 7.2, which contains the detailed outputs from IMT, EModel, and SAS.

- Table 8.1 contains a list of the files used in performing the IMT tests of the Variable-Base Cooling Degree-Day Model (CDD) and the CDD with multiple variable regression model (CDD/MVR). For the CDD model, IMT was benchmarked against PRISM. The utility data that were used as input data for IMT were obtained from a residential building located in College Station, Texas. In order to compare with PRISM CO model, the input data were prepared for two data sets. One contains energy use per billing periods (Q) to match the slope coefficients, and the other contains energy use per day (Q/day) to match the base use coefficients. Table 8.2 contains all detailed output values from the two programs. Generally, IMT and PRISM produced outputs in good agreement with each other.

  For the CDD model with multiple variable regression model (CDD/MVR), the CDD model was run, then a residual file was used as input to the MVR model in order to produce CDD-MVR capabilities. The input data were obtained from a LoanSTAR building. The dependent variable is the whole building cooling energy consumption (WBC). The independent variables include outdoor temperature, humidity ratio, and solar radiation. Table 8.3 contains the detailed outputs from the IMT program. No comparison tests were performed for this model.

- Table 9.1 contains a list of the files used in performing the IMT tests of the Variable-Base Heating Degree-Day Model (HDD) and the HDD with multiple variable regression model (HDD/MVR). For the HDD model, IMT was benchmarked against PRISM. The utility data that were used as input data for IMT were obtained from a residential building located in College Station, Texas. In order to compare with PRISM HO model, the input data were prepared for two data sets. One contains energy use per billing periods (Q) to match the slope coefficients, and the other contains energy use per day (Q/day) to match the base use coefficients. Table 9.2 contains the detailed outputs from the two programs. Generally, IMT and PRISM produced outputs in good agreement with each other.

For the HDD model with multiple variable regression model (HDD/MVR), the HDD model was run, then a residual file was used as input to the MVR model in order to produce HDD-MVR capabilities. The input data were obtained from the LoanSTAR database. The dependent variable is the whole-building heating energy consumption (WBH). The independent variables include outdoor temperature, humidity ratio, and solar radiation. Table 9.3 contains the detailed outputs from the IMT program. No comparison tests were performed for this model.

- Table 10.1 contains a list of the files used in performing the IMT tests against EModel and SAS using the Multiple Variable Regression Model (MVR). The input data are both synthetic and real data. For real data testing, the input data were obtained from the LoanSTAR database. IMT ran and produced outputs successfully without error. EModel failed to run the MLR model with the real data sets. Following Table 10.1 is Table 10.2, which contains the detailed outputs from the three programs. Generally, IMT, EModel, and SAS produced outputs in good agreement with each other.

3. REFERENCES

Kissock, K., Wu, E., Sparks, R., and Patel, D. (1996). *EModel version 1.4 D.* College Station, TX: Texas A&M University, Energy Systems Laboratory.

Fels, M., Reynolds, C., and Stram, D. (1986). PRISMonPC. Documentation for heating-only or cooling-only estimation program: Version 4.0. *PU/CEES Report # 213A*. Princeton, NJ: The Center for Energy and Environmental Studies, The Engineering Quadrangle, Princeton University.

SAS Institute Inc. (2001). *SAS user manual.* Metairie, LA.

APPENDIX A.

SUMMARY TABLES

Table 1.1: One-parameter (1P) and two-parameter (2P) models

			IMT		EModel			SAS		Status			
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	SAS File	SAS output	IMT	EModel	SAS	Comment for IMT
1P	1P_test0	Synthetic: 1-point	1P_test0.ins	1P_test0.dat	1P_test0.dat	1P_test0.dvn	1P_test0.doc	1P_test0.sas	1P_test0.lst	Stop	OK	agree	Program cannot run. Error message shows illegal operation; Access Violation. IMT needs at least 2 points
	1P_test1	Synthetic: 2-point	1P_test1.ins	1P_test1.dat	1P_test1.dat	1P_test1.dvn	1P_test1.doc	1P_test1.sas	1P_test1.lst	OK	agree	agree	
	1P_test2	Synthetic: Scattered	1P_test2.ins	1P_test2.dat	1P_test2.dat	1P_test2.dvn	1P_test2.doc	1P_test2.sas	1P_test2.lst	OK	agree	agree	
	1P_test3	Synthetic: Packed	1P_test3.ins	1P_test3.dat	1P_test3.dat	1P_test3.dvn	1P_test3.doc	1P_test3.sas	1P_test3.lst	OK	agree	agree	
	1P_test4	Synthetic: 9,000-point	1P_test4.ins	1P_test4.dat	1P_test4.dat	1P_test4.dvn	1P_test4.doc	1P_test4.sas	1P_test4.lst	OK	agree	agree	
	1P_test5	Synthetic: Large	1P_test5.ins	1P_test5.dat	1P_test5.dat	1P_test5.dvn	1P_test5.doc	1P_test5.sas	1P_test5.lst	OK	agree	agree	IMT can run 18-digit numbers, but the output is F10.3, hence largest output is 999,999.999.
	1P_test6	Synthetic: Small	1P_test6.ins	1P_test6.dat	1P_test6.dat	1P_test6.dvn	1P_test6.doc	1P_test6.sas	1P_test6.lst	OK	agree	agree	IMT can run 57-decimal point number, output is F10.3 hence smallest output is 0.001.
2P	2P_test0	Synthetic: 2-point	2P_test0.ins	2P_test0.dat	2P_test0.dat	2P_test0.dvn	2P_test0.doc	2P_test0.sas	2P_test0.lst	OK	Overflow	agree	
	2P_test1	Synthetic: 3-point	2P_test1.ins	2P_test1.dat	2P_test1.dat	2P_test1.dvn	2P_test1.doc	2P_test1.sas	2P_test1.lst	OK	agree	agree	
	2P_test2	Synthetic: Scattered	2P_test2.ins	2P_test2.dat	2P_test2.dat	2P_test2.dvn	2P_test2.doc	2P_test2.sas	2P_test2.lst	OK	agree	agree	
	2P_test3	Synthetic: Packed	2P_test3.ins	2P_test3.dat	2P_test3.dat	2P_test3.dvn	2P_test3.doc	2P_test3.sas	2P_test3.lst	OK	agree	agree	
	2P_test4	Synthetic: 9,000-point	2P_test4.ins	2P_test4.dat	2P_test4.dat	2P_test4.dvn	2P_test4.doc	2P_test4.sas	2P_test4.lst	OK	agree	agree	
	2P_test5	Synthetic: Large	2P_test5.ins	2P_test5.dat	2P_test5.dat	2P_test5.dvn	2P_test5.doc	2P_test5.sas	2P_test5.lst	OK	disagree	disagree	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	2P_test6	Synthetic: Small	2P_test6.ins	2P_test6.dat	2P_test6.dat	2P_test6.dvn	2P_test6.doc	2P_test6.sas	2P_test6.lst	OK	agree	agree	IMT can run 57-decimal point number, but output is F12.4 hence smallest output is 0.0001.
	2P_test7	Synthetic: Max Size (X)	2P_test7.ins	2P_test7.dat	2P_test7.dat	2P_test7.dvn	2P_test7.doc	2P_test7.sas	2P_test7.lst	OK	Overflow	agree	Program can run the maximum X value of about 20,000,000.
	2P_test8	Synthetic: Max Size (Y)	2P_test8.ins	2P_test8.dat	2P_test8.dat	2P_test8.dvn	2P_test8.doc	2P_test8.sas	2P_test8.lst	OK	Overflow	agree	IMT can run 19-digit number, but the output is F12.4, hence largest output is 9,999,999.9999
	2P_test9	Synthetic: Slope A = 0	2P_test9.ins	2P_test9.dat	2P_test9.dat	2P_test9.dvn	2P_test9.doc	2P_test9.sas	2P_test9.lst	Stop	agree	agree	Program stops with no error message.
	2P_test10	Synthetic: Slope B = infinite	2P_test10.ins	2P_test10.dat	2P_test10.dat	2P_test10.dvn	2P_test10.doc	2P_test10.sas	2P_test10.lst	OK	N/A	N/A	IMT runs, but slope and intercept numbers are meaningless
	2P_test11	Synthetic: Slope C = 1	2P_test11.ins	2P_test11.dat	2P_test11.dat	2P_test11.dvn	2P_test11.doc	2P_test11.sas	2P_test11.lst	OK	agree	agree	
	2P_test12	Synthetic: Slope D = -1	2P_test12.ins	2P_test12.dat	2P_test12.dat	2P_test12.dvn	2P_test12.doc	2P_test12.sas	2P_test12.lst	OK	agree	agree	

Table 1.2: One-parameter (1P) and two-parameter (2P) models

TEST		Data Type	IMT		EModel		SAS	
1P	1P_test0	Synthetic: 1-point	N/A	N/A	Mean = 5.25, Std Dev = 0.00, CV-StDev = 0.0%	agree	Mean = 5.25, Std Dev = 0.00, CV-StDev = 0.0%	agree
	1P_test1	Synthetic: 2-point	Mean = 5.100, Std Dev = 0.141, CV-StDev = 2.773%	OK	Mean = 5.10, Std Dev = 0.14, CV-StDev = 2.8%	agree	Mean = 5.1, Std Dev = 0.14142136, CV-StDev = 2.77296777%	agree
	1P_test2	Synthetic: Scattered	Mean = 4.718, Std Dev = 2.829, CV-StDev = 59.975%	OK	Mean = 4.72, Std Dev = 2.83, CV-StDev = 60.0%	agree	Mean = 4.71769, Std Dev = 2.82943194, CV-StDev = 59.9749441%	agree
	1P_test3	Synthetic: Packed	Mean = 4.983, Std Dev = 0.279, CV-StDev = 5.601%	OK	Mean = 4.98, Std Dev = 0.28, CV-StDev = 5.6%	agree	Mean = 4.982722, Std Dev = 0.27905763, CV-StDev = 5.60050567%	agree
	1P_test4	Synthetic: 9,000-point	Mean = 5.087, Std Dev = 2.898, CV-StDev = 56.969%	OK	Mean = 5.09, Std Dev = 2.90, CV-StDev = 57.0%	agree	Mean = 5.08703221, Std Dev = 2.89804482, CV-StDev = 56.9692641%	agree
	1P_test5	Synthetic: Large	Mean = 67297.289, Std Dev = 137075.516, CV-StDev = 203.687%	OK	Mean = 67297.29, Std Dev = 137075.52, CV-StDev = 203.7%	agree	Mean = 67297.2857, Std Dev = 137075.521, CV-StDev = 203.686552%	agree
	1P_test6	Synthetic: Small	Mean = 0.001, Std Dev = 0.008, CV-StDev = 715.378%	OK	Mean = 0.00, Std Dev = 0.01, CV-StDev = 715.4%	agree	Mean = 0.00112609, Std Dev = 0.00805578, CV-StDev = 715.378489%	agree
2P	2P_test0	Synthetic: 2-point	Slope = 1.0000, Intercept = 2.0000, CV-RMSE = 0.000%, R^2 = 1.000	OK	N/A	N/A	Slope = 1.00000, Intercept = 2.00000, CV-RMSE = 0.0%, R^2 = 1.0000	agree
	2P_test1	Synthetic: 3-point	Slope = 1.0000, Intercept = 2.0000, CV-RMSE = 0.000%, R^2 = 1.000	OK	Slope = 1.0000, Intercept = 2.0000, CV-RMSE = 0.000%, R^2 = 1.000	agree	Slope = 1.00000, Intercept = 2.00000, CV-RMSE = 0.0%, R^2 = 1.0000	agree
	2P_test2	Synthetic: Scattered	Slope = 1.4975, Intercept = 35.6834, CV-RMSE = 39.298%, R^2 = 0.366	OK	Slope = 1.4975, Intercept = 35.6834, CV-RMSE = 39.3%, R^2 = 0.37	agree	Slope = 1.49750, Intercept = 35.68344, CV-RMSE = 39.298%, R^2 = 0.366	agree
	2P_test3	Synthetic: Packed	Slope = 1.0050, Intercept = 2.3368, CV-RMSE = 1.038%, R^2 = 1.000	OK	Slope = 1.0050, Intercept = 2.3368, CV-RMSE = 1.0%, R^2 = 1.00	agree	Slope = 1.00498, Intercept = 2.33684, CV-RMSE = 1.03810%, R^2 = 0.9996	agree
	2P_test4	Synthetic: 9,000-point	Slope = 1.0000, Intercept = 2.0000, CV-RMSE = 0.000%, R^2 = 1.000	OK	Slope = 1.0000, Intercept = 2.0000, CV-RMSE = 0.0%, R^2 = 1.00	agree	Slope = 1.00000, Intercept = 2.00000, CV-RMSE = 0%, R^2 = 1.0000	agree
	2P_test5	Synthetic: Large	Slope = 1.0000, Intercept = -2.8669, CV-RMSE = 0.000%, R^2 = 1.000	OK	Slope = 1.0000, Intercept = 1.0101, CV-RMSE = 0.0%, R^2 = 1.000	disagree	Slope = 1.0000, Intercept = 1.01013, CV-RMSE = 0%, R^2 = 1.0000	disagree
	2P_test6	Synthetic: Small	Slope = 1.0000, Intercept = 0.0000, CV-RMSE = 0.000%, R^2 = 1.000	OK	Slope = 1.0000, Intercept = 0.0000, CV-RMSE = 0.000%, R^2 = 1.000	agree	Slope = 1.0000, Intercept = 0.0000, CV-RMSE = 0.000%, R^2 = 1.000	agree
	2P_test7	Synthetic: Max Size (X)	N/A	N/A	N/A	N/A	Slope = 4.019652E-7, Intercept = 1.50000, CV-RMSE = 0.000%, R^2 = 1.0000	agree
	2P_test8	Synthetic: Max Size (Y)	Slope = 111111.0078, Intercept = -111110.0313, CV-RMSE = 0.000%, R^2 = 1.000	OK	N/A	N/A	Slope = 111111, Intercept = -111110, CV-RMSE = 0%, R^2 = 1.0000	agree
	2P_test9	Synthetic: Slope A = 0	N/A	N/A	Slope = 0.0000, Intercept = 5.0000, CV-RMSE = 0.0%, R^2 = 0.00	agree	Slope = 0, Intercept = 5.00000, CV-RMSE = 0.0%, R^2 = 0	agree
	2P_test10	Synthetic: Slope B = inf	Slope = 5100.0000, Intercept = 127500.0000, CV-RMSE = N/A, R^2 = N/A	OK	N/A	N/A	N/A	N/A
	2P_test11	Synthetic: Slope C = 1	Slope = 1.0000, Intercept = 0.0000, CV-RMSE = 0.000%, R^2 = 1.000	OK	Slope = 1.0000, Intercept = 0.0000, CV-RMSE = 0.000%, R^2 = 1.000	agree	Slope = 1.0000, Intercept = 0.0000, CV-RMSE = 0.000%, R^2 = 1.000	agree
	2P_test12	Synthetic: Slope D = -1	Slope = -1.0000, Intercept = 51.0000, CV-RMSE = 0.000%, R^2 = 1.000	OK	Slope = -1.0000, Intercept = 51.0000, CV-RMSE = 0.0%, R^2 = 1.00	agree	Slope = -1.00000, Intercept = 51.00000, CV-RMSE = 0%, R^2 = 1.0000	agree

Table 2.1: Three-parameter change-point cooling (3PC) and three-parameter change-point heating (3PH) models

			IMT		EModel			Status		
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	IMT	EModel	Comment for IMT
3PC	3PC_test0	Synthetic: 3-point	3PC_test0.ins	3PC_test0.dat	3PC_test0.dat	3PC_test0.dvn	3PC_test0.doc	Stop	Overflow	Unknown Floating Exception
	3PC_test1	Synthetic: 5-point	3PC_test1.ins	3PC_test1.dat	3PC_test1.dat	3PC_test1.dvn	3PC_test1.doc	OK	agree	
	3PC_test2	Synthetic: Scattered	3PC_test2.ins	3PC_test2.dat	3PC_test2.dat	3PC_test2.dvn	3PC_test2.doc	OK	agree	
	3PC_test3	Synthetic: Packed	3PC_test3.ins	3PC_test3.dat	3PC_test3.dat	3PC_test3.dvn	3PC_test3.doc	OK	agree	
	3PC_test4	Synthetic: 9,000-point	3PC_test4.ins	3PC_test4.dat	3PC_test4.dat	3PC_test4.dvn	3PC_test4.doc	OK	agree	
	3PC_test5	Synthetic: Large	3PC_test5.ins	3PC_test5.dat	3PC_test5.dat	3PC_test5.dvn	3PC_test5.doc	OK	agree	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	3PC_test6	Synthetic: Small	3PC_test6.ins	3PC_test6.dat	3PC_test6.dat	3PC_test6.dvn	3PC_test6.doc	OK	agree	IMT can run 16-decimal point numbers, but output is F12.4 hence smallest output is 0.0001
	3PC_test7	Synthetic: Max Size (X)	3PC_test7.ins	3PC_test7.dat	3PC_test7.dat	3PC_test7.dvn	3PC_test7.doc	OK	agree	IMT can run 19-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	3PC_test8	Synthetic: Max Size (Y)	3PC_test8.ins	3PC_test8.dat	3PC_test8.dat	3PC_test8.dvn	3PC_test8.doc	OK	agree	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	3PC_test9	Synthetic: Slope A = 0	3PC_test9.ins	3PC_test9.dat	3PC_test9.dat	3PC_test9.dvn	3PC_test9.doc	Stop	OK	Floating point devided by 0
	3PC_test10	Synthetic: Slope B = infinite	3PC_test10.ins	3PC_test10.dat	3PC_test10.dat	3PC_test10.dvn	3PC_test10.doc	OK	agree	xcp is close by slope is wrong
	3PC_test11	Synthetic: Slope C = infinite	3PC_test11.ins	3PC_test11.dat	3PC_test11.dat	3PC_test11.dvn	3PC_test11.doc	OK	agree	xcp is close by slope is wrong
	3PC_test12	Synthetic: Slope D = -1	3PC_test12.ins	3PC_test12.dat	3PC_test12.dat	3PC_test12.dvn	3PC_test12.doc	OK	agree	
3PH	3PH_test0	Synthetic: 3-point	3PH_test0.ins	3PH_test0.dat	3PH_test0.dat	3PH_test0.dvn	3PH_test0.doc	Stop	Overflow	Unknown Floating Point Exception
	3PH_test1	Synthetic: 5-point	3PH_test1.ins	3PH_test1.dat	3PH_test1.dat	3PH_test1.dvn	3PH_test1.doc	OK	agree	
	3PH_test2	Synthetic: Scattered	3PH_test2.ins	3PH_test2.dat	3PH_test2.dat	3PH_test2.dvn	3PH_test2.doc	OK	agree	
	3PH_test3	Synthetic: Packed	3PH_test3.ins	3PH_test3.dat	3PH_test3.dat	3PH_test3.dvn	3PH_test3.doc	OK	agree	
	3PH_test4	Synthetic: 9,000-point	3PH_test4.ins	3PH_test4.dat	3PH_test4.dat	3PH_test4.dvn	3PH_test4.doc	OK	agree	
	3PH_test5	Synthetic: Large	3PH_test5.ins	3PH_test5.dat	3PH_test5.dat	3PH_test5.dvn	3PH_test5.doc	OK	agree	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	3PH_test6	Synthetic: Small	3PH_test6.ins	3PH_test6.dat	3PH_test6.dat	3PH_test6.dvn	3PH_test6.doc	OK	agree	IMT can run 16-decimal point numbers, but output is F12.4 hence smallest output is 0.0001
	3PH_test7	Synthetic: Max Size (X)	3PH_test7.ins	3PH_test7.dat	3PH_test7.dat	3PH_test7.dvn	3PH_test7.doc	OK	agree	IMT can run 19-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	3PH_test8	Synthetic: Max Size (Y)	3PH_test8.ins	3PH_test8.dat	3PH_test8.dat	3PH_test8.dvn	3PH_test8.doc	OK	agree	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	3PH_test9	Synthetic: Slope A = 0	3PH_test9.ins	3PH_test9.dat	3PH_test9.dat	3PH_test9.dvn	3PH_test9.doc	Stop	OK	Program stops responding, but EModel runs OK
	3PH_test10	Synthetic: Slope B = infinite	3PH_test10.ins	3PH_test10.dat	3PH_test10.dat	3PH_test10.dvn	3PH_test10.doc	OK	agree	xcp is close but slope is wrong
	3PH_test11	Synthetic: Slope C = infinite	3PH_test11.ins	3PH_test11.dat	3PH_test11.dat	3PH_test11.dvn	3PH_test11.doc	OK	agree	xcp is close but slope is wrong
	3PH_test12	Synthetic: Slope D = 1	3PH_test12.ins	3PH_test12.dat	3PH_test12.dat	3PH_test12.dvn	3PH_test12.doc	OK	agree	

Table 2.2: Three-parameter change-point cooling (3PC) and three-parameter change-point heating (3PH) models

TEST		Data Type	IMT		EModel	
3PC	3PC_test0	Synthetic: 3-point	N/A	Stop	N/A	Overflow
	3PC_test1	Synthetic: 5-point	Ycp = 5.0077, LS = 0.0000, RS = 1.0136, Xcp = 10.1200, R^2 =1.000, CV-RMSE = 0.38%	OK	Ycp = 5.0077, LS = 0.0000, RS = 1.0136, Xcp = 10.1200, R^2 =1.00, CV-RMSE = 0.4%	agree
	3PC_test2	Synthetic: Scattered	Ycp = 44.0362, LS = 0.0000, RS = 3.4570, Xcp = 24.5200, R^2 = 0.547, CV-RMSE = 40.237%	OK	Ycp = 44.0362, LS = 0.0000, RS = 3.4570, Xcp = 24.52, R^2 = 0.55, CV-RMSE = 40.2%	agree
	3PC_test3	Synthetic: Packed	Ycp = 27.4776, LS = 0.0000, RS = 1.0379, Xcp = 25.5000, R^2 = 0.999, CV-RMSE = 0.854%	OK	Ycp = 27.4776, LS = 0.0000, RS = 1.0379, Xcp = 25.5000, R^2 = 1.0, CV-RMSE = 0.9%	agree
	3PC_test4	Synthetic: 9,000-point	Ycp = 4.9997, LS = 0.0000, RS = 0.9996, Xcp = 4.9984, R^2 = 1.000, CV-RMSE = 0.008%	OK	Ycp = 4.9997, LS = 0.0000, RS = 0.9996, Xcp = 4.9984, R^2 = 1.00, CV-RMSE = 0.0%	agree
	3PC_test5	Synthetic: Large	Ycp = 443143.7500, LS = 0.0000, RS = 1.0000, Xcp = 443145.4688, R^2 = 1.000, CV-RMSE = 0.000%	OK	Ycp = 443143.7328, LS = 0.0000, RS = 1.0000, Xcp = 443145.45, R^2 = 1.00, CV-RMSE = 0.0%	agree
	3PC_test6	Synthetic: Small	Ycp = 0.0005, LS = 0.0000, RS = 1.0039, Xcp = 0.0005, R^2 = 1.000, CV-RMSE = 0.084%	OK	Ycp = 0.0005, LS = 0.0000, RS = 1.0039, Xcp = 0.0005, R^2 = 1.00, CV-RMSE = 0.1%	agree
	3PC_test7	Synthetic: Max Size (X)	Ycp = 5.0091, LS = 0.0000, RS = 1.0039, Xcp = 5004678.50, R^2 = 1.000, CV-RMSE = 0.190%	OK	Ycp = 5.0091, LS = 0.0000, RS = 1.0039, Xcp = 5004678.78, R^2 = 1.000, CV-RMSE = 0.2%	agree
	3PC_test8	Synthetic: Max Size (Y)	Ycp = 4996527.00, LS = 0.0000, RS = 986543.8125, Xcp = 5.0045, R^2 = 1.000, CV-RMSE = 0.169%	OK	Ycp = 4996526.875, LS = 0.0000, RS = 986543.9533, Xcp = 5.0045, R^2 = 1.000, CV-RMSE = 0.2%	agree
	3PC_test9	Synthetic: Slope A = 0	N/A	Stop	Ycp = 5.0000, LS = 0.0000, RS = 0.0000, Xcp = 26.4800, R^2 = 0.00, CV-RMSE = 0.0%	agree
	3PC_test10	Synthetic: Slope B = inf	Ycp = 25.00, LS = 0.0000, RS = 13.0208, Xcp = 24.0400, R^2 = 0.571, CV-RMSE = 17.523%	OK	Ycp = 25.00, LS = 0.0000, RS = 13.0208, Xcp = 24.0400, R^2 = 0.57, CV-RMSE = 17.5%	agree
	3PC_test11	Synthetic: Slope C = inf	Ycp = 25.00, LS = 0.0000, RS = -12.0192, Xcp = 24.0400, R^2 = 0.536, CV-RMSE = 28.812%	OK	Ycp = 25.00, LS = 0.0000, RS = -12.0192, Xcp = 24.0400, R^2 = 0.54, CV-RMSE = 28.8%	agree
	3PC_test12	Synthetic: Slope D = -1	Ycp = 24.9103, LS = 0.0000, RS = -1.0244, Xcp = 26.48, R^2 = 1.000, CV-RMSE = 0.804%	OK	Ycp = 24.9103, LS = 0.0000, RS = -1.0244, Xcp = 26.48, R^2 = 1.00, CV-RMSE = 0.8%	agree
3PH	3PH_test0	Synthetic: 3-point	N/A	Stop	N/A	Overflow
	3PH_test1	Synthetic: 5-point	Ycp = 5.0016, LS = -1.2535, RS = 0.0000, Xcp = 8.9800, R^2 = 1.0000, CV-RMSE = 0.078%	OK	Ycp = 5.0016, LS = -1.2535, RS = 0.0000, Xcp = 8.9800, R^2 = 1.0000, CV-RMSE = 0.1%	agree
	3PH_test2	Synthetic: Scattered	Ycp = 77.4479, LS = -2.9475, RS = 0.0000, Xcp = 25.5000, R^2 = 0.407, CV-RMSE = 30.563%	OK	Ycp = 77.4479, LS = -2.9475, RS = 0.0000, Xcp = 25.5000, R^2 = 0.41, CV-RMSE = 30.6%	agree
	3PH_test3	Synthetic: Packed	Ycp = 30.0132, LS = -3.0545, RS = 0.0000, Xcp = 24.5200, R^2 = 0.9850, CV-RMSE = 6.168%	OK	Ycp = 30.0132, LS = -3.0545, RS = 0.0000, Xcp = 24.5200, R^2 = 0.99, CV-RMSE = 6.2%	agree
	3PH_test4	Synthetic: 9,000-point	Ycp = 5.0063, LS = -0.9989, RS = 0.0000, Xcp = 4.9984, R^2 = 0.969, CV-RMSE = 4.626%	OK	Ycp = 5.0063, LS = -0.9989, RS = 0.0000, Xcp = 4.9984, R^2 = 0.97, CV-RMSE = 4.6%	agree
	3PH_test5	Synthetic: Large	Ycp = 9370096.00, LS = -1.0000, RS = 0.0000, Xcp = 9370100.00, R^2 = 1.000, CV-RMSE = 0.000%	OK	Ycp = 9370096.00, LS = -1.0000, RS = 0.0000, Xcp = 9370100.00, R^2 = 1.000, CV-RMSE = 0.000%	agree
	3PH_test6	Synthetic: Small	Ycp = 0.0005, LS = -0.9871, RS = 0.0000, Xcp = 0.0005, R^2 = 1.000, CV-RMSE = 0.451%	OK	Ycp = 0.0005, LS = -0.9871, RS = 0.0000, Xcp = 0.0005, R^2 = 1.000, CV-RMSE = 0.5%	agree
	3PH_test7	Synthetic: Max Size (X)	Ycp = 5.0000, LS = 0.0000, RS = 0.0000, Xcp = 5004678.500, R^2 = 1.000, CV-RMSE = 0.003%	OK	Ycp = 5.0000, LS = 0.0000, RS = 0.0000, Xcp = 5004678.7790, R^2 = 1.000, CV-RMSE = 0.0%	agree
	3PH_test8	Synthetic: Max Size (Y)	Ycp = 5004020.0000, LS = -992540.2500, RS = 0.0000, Xcp = 5.0045, R^2 = 1.000, CV-RMSE = 0.022%	OK	Ycp = 5004018.7144, LS = -992539.8949, RS = 0.0000, Xcp = 5.0045, R^2 = 1.000, CV-RMSE = 0.0%	agree
	3PH_test9	Synthetic: Slope A = 0	N/A	Stop	Ycp = 5.0000, LS = 0.0000, RS = 0.0000, Xcp = 30.4000, R^2 = 0.00, CV-RMSE = 0.0%	agree
	3PH_test10	Synthetic: Slope B = inf	Ycp = 25.0000, LS = -26.0417, RS = 0.0000, Xcp = 25.4800, R^2 = 0.571, CV-RMSE = 17.523%	OK	Ycp = 25.0000, LS = -26.0417, RS = 0.0000, Xcp = 25.4800, R^2 = 0.56, CV-RMSE = 17.5%	agree
	3PH_test11	Synthetic: Slope C = inf	Ycp = 25.0000, LS = 24.0000, RS = 0.0000, Xcp = 25.5000, R^2 = 0.581, CV-RMSE = 27.390%	OK	Ycp = 25.0000, LS = 24.0000, RS = 0.0000, Xcp = 25.5000, R^2 = 0.58, CV-RMSE = 27.4%	agree
	3PH_test12	Synthetic: Slope D = 1	Ycp = 24.9103, LS = 1.0244, RS = 0.0000, Xcp = 24.52, R^2 = 1.000, CV-RMSE = 0.804%	OK	Ycp = 24.9103, LS = 1.0244, RS = 0.0000, Xcp = 24.52, R^2 = 1.000, CV-RMSE = 0.8%	agree

Table 3.1: Four-parameter change-point (4P) model

			IMT		EModel			Status		
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	Status	EModel	Comment for IMT
4P	4P_test0	Synthetic: 3-point	4P_test0.ins	4P_test0.dat	4P_test0.dat	4P_test0.dvn	4P_test0.doc	Wrong	Error	IMT returns wrong LS value. Emodel stops with a divided-by-zero comment
	4P_test1	Synthetic: 5-point	4P_test1.ins	4P_test1.dat	4P_test1.dat	4P_test1.dvn	4P_test1.doc	OK	OK	
	4P_test2	Synthetic: Scattered	4P_test2.ins	4P_test2.dat	4P_test2.dat	4P_test2.dvn	4P_test2.doc	OK	OK	
	4P_test3	Synthetic: Packed	4P_test3.ins	4P_test3.dat	4P_test3.dat	4P_test3.dvn	4P_test3.doc	OK	OK	
	4P_test4	Synthetic: 9,000-point	4P_test4.ins	4P_test4.dat	4P_test4.dat	4P_test4.dvn	4P_test4.doc	OK	OK	
	4P_test5	Synthetic: Large	4P_test5.ins	4P_test5.dat	4P_test5.dat	4P_test5.dvn	4P_test5.doc	OK	OK	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	4P_test6	Synthetic: Small	4P_test6.ins	4P_test6.dat	4P_test6.dat	4P_test6.dvn	4P_test6.doc	OK	OK	IMT can run 16-decimal point numbers, but output is F12.4 hence smallest output is 0.0001
	4P_test7	Synthetic: Max Size (X)	4P_test7.ins	4P_test7.dat	4P_test7.dat	4P_test7.dvn	4P_test7.doc	OK	OK	IMT can run 19-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	4P_test8	Synthetic: Max Size (Y)	4P_test8.ins	4P_test8.dat	4P_test8.dat	4P_test8.dvn	4P_test8.doc	OK	OK	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	4P_test9	Synthetic: Slope A (flat)	4P_test9.ins	4P_test9.dat	4P_test9.dat	4P_test9.dvn	4P_test9.doc	Stop	OK	IMT stops responding.
	4P_test10	Synthetic: Slope B =1	4P_test10.ins	4P_test10.dat	4P_test10.dat	4P_test10.dvn	4P_test10.doc	OK	OK	
	4P_test11	Synthetic: Slope C = infinite	4P_test11.ins	4P_test11.dat	4P_test11.dat	4P_test11.dvn	4P_test11.doc	OK	OK	xcp is close by slope is wrong
	4P_test12	Synthetic: Slope D = infinite	4P_test12.ins	4P_test12.dat	4P_test12.dat	4P_test12.dvn	4P_test12.doc	OK	OK	xcp is close by slope is wrong
	4P_test13	Synthetic: Slope E = double	4P_test13.ins	4P_test13.dat	4P_test13.dat	4P_test13.dvn	4P_test13.doc	OK	OK	IMT slope is avg of double slopes
	4P_test14	Synthetic: Slope F = 1	4P_test14.ins	4P_test14.dat	4P_test14.dat	4P_test14.dvn	4P_test14.doc	OK	OK	
	4P_test15	Synthetic: Slope G = infinite	4P_test15.ins	4P_test15.dat	4P_test15.dat	4P_test15.dvn	4P_test15.doc	OK	OK	xcp is close but slope is wrong
	4P_test16	Synthetic: Slope H = infinite	4P_test16.ins	4P_test16.dat	4P_test16.dat	4P_test16.dvn	4P_test16.doc	OK	OK	xcp is close but slope is wrong
	4P_test17	Synthetic: Slope I = double	4P_test17.ins	4P_test17.dat	4P_test17.dat	4P_test17.dvn	4P_test17.doc	OK	OK	IMT slope is avg of double slopes

Table 3.2: Four-parameter change-point (4P) model

TEST		Data Type	IMT		EModel	
4P	4P_test0	Synthetic: 3-point	Ycp = 5.1867, LS = -11.5093, RS = 0.9993, Xcp = 5.2000, R^2 =1.000, CV-RMSE = 0.000%	Wrong	N/A	Error
	4P_test1	Synthetic: 5-point	Ycp = 10.000, LS = 0.5000, RS = 1.0000, Xcp = 10.0000, R^2 =1.000, CV-RMSE = 0.000%	OK	Ycp = 10.000, LS = 0.5000, RS = 1.0000, Xcp = 10.0000, R^2 =1.00, CV-RMSE = 0.0%	OK
	4P_test2	Synthetic: Scattered	Ycp = 62.1740, LS = 0.4749, RS = 4.5700, Xcp = 23.5400, R^2 = 0.721, CV-RMSE = 29.309%	OK	Ycp = 62.1739, LS = 0.4749, RS = 4.5700, Xcp = 23.5400, R^2 = 0.72, CV-RMSE = 29.3%	OK
	4P_test3	Synthetic: Packed	Ycp = 53.9740, LS = 0.5725, RS = 2.5368, Xcp = 52.9200, R^2 = 0.962, CV-RMSE = 12.307%	OK	Ycp = 53.9740, LS = 0.5725, RS = 2.5368, Xcp = 52.9200, R^2 = 0.96, CV-RMSE = 12.3%	OK
	4P_test4	Synthetic: 9,000-point	Ycp = 5.0115, LS = 0.3544, RS = 0.9912, Xcp = 4.9984, R^2 = 0.923, CV-RMSE = 9.897%	OK	Ycp = 5.0115, LS = 0.3544, RS = 0.9912, Xcp = 4.9984, R^2 = 0.92, CV-RMSE = 9.9%	OK
	4P_test5	Synthetic: Large	Ycp = 500468.2500, LS = 0.5000, RS = 1.0000, Xcp = 500468.0000, R^2 = 1.000, CV-RMSE = 0.000%	OK	Ycp = 500468.0219, LS = 0.5000, RS = 1.0000, Xcp = 500468.0000, R^2 = 1.00, CV-RMSE = 0.0%	OK
	4P_test6	Synthetic: Small	Ycp = 0.0000, LS = 0.5020, RS = 1.0017, Xcp = 0.0000, R^2 = 1.000, CV-RMSE = 0.053%	OK	Ycp = 0.0000, LS = 0.5020, RS = 1.0017, Xcp = 0.0000, R^2 = 1.00, CV-RMSE = 0.1%	OK
	4P_test7	Synthetic: Max Size (X)	Ycp = 5.0211, LS = 0.0000, RS = 0.0000, Xcp = 5004678.5000, R^2 = 1.000, CV-RMSE = 0.162%	OK	Ycp = 5.0211, LS = 0.0000, RS = 0.0000, Xcp = 5004678.7800, R^2 = 1.00, CV-RMSE = 0.2%	OK
	4P_test8	Synthetic: Max Size (Y)	Ycp = 4986407.0000, LS = 490802.1563, RS = 989593.0000, Xcp = 5.0045, R^2 = 1.000, CV-RMSE = 0.147%	OK	Ycp = 4986413.2746, LS = 490803.3669, RS = 989591.8723, Xcp = 5.0045, R^2 = 1.00, CV-RMSE = 0.1%	OK
	4P_test9	Synthetic: Slope A (flat)	N/A	Stop	Ycp = 5.0000, LS = 0.0000, RS = 0.000, Xcp = 5.0045, R^2 = 0.00, CV-RMSE = 0.0%	OK
	4P_test10	Synthetic: Slope B =1	Ycp = 24.7702, LS = 1.0156, RS = 0.0133, Xcp = 24.5200, R^2 = 1.000, CV-RMSE = 0.650%	OK	Ycp = 24.7702, LS = 1.0156, RS = 0.0133, Xcp = 24.5200, R^2 = 1.00, CV-RMSE = 0.7%	OK
	4P_test11	Synthetic: Slope C = infinite	Ycp = 24.0401, LS = 1.0000, RS = 14.0207, Xcp = 24.0400, R^2 = 0.860, CV-RMSE = 21.876%	OK	Ycp = 24.0400, LS = 1.0000, RS = 14.0208, Xcp = 24.0400, R^2 = 0.86, CV-RMSE = 21.9%	OK
	4P_test12	Synthetic: Slope D = infinite	Ycp = 24.0401, LS = 1.0000, RS = -11.0193, Xcp = 24.0400, R^2 = 0.447, CV-RMSE = 42.556%	OK	Ycp = 24.0401, LS = 1.0000, RS = -11.0192, Xcp = 24.0400, R^2 = 0.45, CV-RMSE = 42.6%	OK
	4P_test13	Synthetic: Slope E = double	Ycp = 47.2500, LS = 0.5000, RS = 0.5000, Xcp = 44.5000, R^2 = 0.200, CV-RMSE = 17.389%	OK	Ycp = 48.2500, LS = 0.5000, RS = 0.5000, Xcp = 46.5000, R^2 = 0.20, CV-RMSE = 17.4%	OK
	4P_test14	Synthetic: Slope F = 1	Ycp = 24.7498, LS = -0.0156, RS = 0.9867, Xcp = 24.5200, R^2 = 1.000, CV-RMSE = 0.392%	OK	Ycp = 24.7498, LS = -0.0156, RS = 0.9867, Xcp = 24.5200, R^2 = 1.00, CV-RMSE = 0.4%	OK
	4P_test15	Synthetic: Slope G = infinite	Ycp = 26.0118, LS = -10.0535, RS = 1.0008, Xcp = 26.0000, R^2 = 0.514, CV-RMSE = 14.080%	OK	Ycp = 26.0000, LS = -10.0400, RS = 1.0000, Xcp = 26.0000, R^2 = 0.51, CV-RMSE = 14.1%	OK
	4P_test16	Synthetic: Slope H = infinite	Ycp = 25.9759, LS = 12.9596, RS = 1.0016, Xcp = 26.0000, R^2 = 0.875, CV-RMSE = 20.625%	OK	Ycp = 26.0000, LS = 13.0000, RS = 1.0000, Xcp = 26.0000, R^2 = 0.88, CV-RMSE = 20.6%	OK
	4P_test17	Synthetic: Slope I = double	Ycp = 30.0001, LS = 0.5000, RS = 0.5000, Xcp = 35.0000, R^2 = 0.209, CV-RMSE = 24.027%	OK	Ycp = 25.2500, LS = 0.5000, RS = 0.5000, Xcp = 25.5000, R^2 = 0.21, CV-RMSE = 24.0%	OK

Table 4.1: Five-parameter change-point (5P) and five-parameter change-point with multiple variable regression models (5P/MVR)

			IMT		Status		
TEST		Data Type	IMT file	Data File	IMT	EModel	Comment for IMT
5P	5P_test0	Synthetic: 4-point	5P_test0.ins	5P_test0.dat	Stop	N/A	Unknow Floating Point Exception
	5P_test1	Synthetic: 7-point	5P_test1.ins	5P_test1.dat	OK	N/A	
	5P_test2	Synthetic: Scattered	5P_test2.ins	5P_test2.dat	OK	N/A	
	5P_test3	Synthetic: Packed	5P_test3.ins	5P_test3.dat	OK	N/A	
	5P_test4	Synthetic: 9,000-point	5P_test4.ins	5P_test4.dat	OK	N/A	
	5P_test5	Synthetic: Large	5P_test5.ins	5P_test5.dat	OK	N/A	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	5P_test6	Synthetic: Small	5P_test6.ins	5P_test6.dat	OK	N/A	IMT can run 16-decimal point numbers, but output is F12.4 hence smallest output is 0.0001
	5P_test7	Synthetic: Max Size (X)	5P_test7.ins	5P_test7.dat	OK	N/A	IMT can run 19-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	5P_test8	Synthetic: Max Size (Y)	5P_test8.ins	5P_test8.dat	OK	N/A	IMT can run 18-digit numbers, but the output is F12.4, hence largest output is 9,999,999.9999.
	5P_test9	Synthetic: Slope A (flat)	5P_test9.ins	5P_test9.dat	Stop	N/A	Floating point divided by zero
	5P_test10	Synthetic: Slope B = 2, inf	5P_test10.ins	5P_test10.dat	OK	N/A	xcp is close but slope is wrong
	5P_test11	Synthetic: Slope C = 2, inf	5P_test11.ins	5P_test11.dat	OK	N/A	xcp is close but slope is wrong
	5P_test12	Synthetic: Slope D =1, -1	5P_test12.ins	5P_test12.dat	OK	N/A	
	5P_test13	Synthetic: Slope E = 1,1	5P_test13.ins	5P_test13.dat	OK	N/A	
	5P_test14	Synthetic: Slope F = -1,-1	5P_test14.ins	5P_test14.dat	OK	N/A	
5P/MVR	5P_Mvr1	MCC/Temp	5P_Mvr1.ins	5P_Mvr1.dat	OK	N/A	
	5P_Mvr2	MCC/Hum.ratio	5P_Mvr2.ins	5P_Mvr2.dat	OK	N/A	
	5P_Mvr3	MCC/Solar	5P_Mvr3.ins	5P_Mvr3.dat	OK	N/A	
	5P_Mvr4	MCC/(temp, hum.ratio, solar)	5P_Mvr4.ins	5P_Mvr4.dat	Stop	N/A	For 5P model, Number of X variables must be greater than 0 and less than 3.
	5P_Mvr5	MCC/(temp, hum.ratio)	5P_Mvr5.ins	5P_Mvr5.dat	OK	N/A	

Table 4.2: Five-parameter change-point model (5P)

TEST		Data Type	IMT	
5P	5P_test0	Synthetic: 4-point	N/A	Stop
	5P_test1	Synthetic: 7-point	Xcp1 = 4.0743, Xcp2 = 6.9257, Ycp = 4.8723, LS = -2.4657, RS = 2.4657, R^2 = 0.999, CV-RMSE = 1.078%	OK
	5P_test2	Synthetic: Scattered	Xcp1 = 23.8116, Xcp2 = 27.0724, Ycp = 7.1445, LS = -0.8616, RS = 1.4693, R^2 = 0.465, CV-RMSE = 45.680%	OK
	5P_test3	Synthetic: Packed	Xcp1 = 17.2947, Xcp2 = 28.7053, Ycp = 13.6478, LS = -0.8945, RS = 1.0204, R^2 = 0.939, CV-RMSE = 7.164%	OK
	5P_test4	Synthetic: 9,000-point	Xcp1 = 2.9624, Xcp2 = 5.9237, Ycp = 2.9931, LS = -1.0216, RS = 0.9710, R^2 = 0.946, CV-RMSE = 6.798%	OK
	5P_test5	Synthetic: Large	Xcp1 = 3348554.5000, Xcp2 = 6660803.0000, Ycp = 3323121.2500, LS = -0.9901, RS = 1.0063, R^2 = 1.000, CV-RMSE = 0.153%	OK
	5P_test6	Synthetic: Small	Xcp1 = 0.0000, Xcp2 = 0.0000, Ycp = 0.0000, LS = -0.9902, RS = 1.0064, R^2 = 1.000, CV-RMSE = 0.000%	OK
	5P_test7	Synthetic: Max Size (X)	Xcp1 = 3348554.2500, Xcp2 = 6660803.0000, Ycp = 4.9983, LS = 0.0000, RS = 0.0000, R^2 = 1.000, CV-RMSE = 0.153%	OK
	5P_test8	Synthetic: Max Size (Y)	Xcp1 = 2.9826, Xcp2 = 7.0274, Ycp = 4993588.5000, LS = N/A, RS = 1008166.3750, R^2 = 1.000, CV-RMSE = 0.145%	OK
	5P_test9	Synthetic: Slope A (flat)	N/A	Stop
	5P_test10	Synthetic: Slope B = 2, inf	Xcp1 = 5.5556, Xcp2 = 9.4444, Ycp = 5.0000, LS = -8.2500, RS = 7.8000, R^2 = 0.195, CV-RMSE = 38.946%	OK
	5P_test11	Synthetic: Slope C = 2, inf	Xcp1 = 5.5556, Xcp2 = 9.4444, Ycp = 15.0000, LS = 8.2500, RS = -6.6000, R^2 = 0.182, CV-RMSE = 27.866%	OK
	5P_test12	Synthetic: Slope D =1, -1	Xcp1 = 15.6667, Xcp2 = 30.3333, Ycp = 15.2201, LS = 0.9558, RS = -0.9582, R^2 = 0.998, CV-RMSE = 2.231%	OK
	5P_test13	Synthetic: Slope E = 1,1	Xcp1 = 15.6667, Xcp2 = 30.3333, Ycp = 15.1639, LS = 0.9502, RS = 1.0159, R^2 = 1.000, CV-RMSE = 1.055%	OK
	5P_test14	Synthetic: Slope F = -1,-1	Xcp1 = 15.6667, Xcp2 = 30.3333, Ycp = 15.1626, LS = -1.0169, RS = 0.9501, R^2 = 0.999, CV-RMSE = 1.199%	OK

Table 4.3: Five-parameter change-point with multiple variable regression model (5P/MVR)

TEST		Data Type	IMT	
5P/MVR	5P_MVR1	MCC/temp	Xcp1 = 44.2178, Xcp2 = 52.5467, Ycp =107.2245, LS = 0.7187, RS = 2.3912, R^2 =0.632, CV-RMSE =14.301%	OK
	5P_MVR2	MCC/Hum.Ratio	Xcp1 = 0.0167, Xcp2 = 0.0189, Ycp =174.6631 LS = 4500.1250, RS = 2631.7185, R^2 =0.332, CV-RMSE = 19.275%	OK
	5P_MVR3	MCC/Solar	Xcp1 = 39.5037, Xcp2 = 157.7781, Ycp =158.6373, LS = 0.5226, RS = 0.0287, R^2 = 0.184, CV-RMSE = 21.311%	OK
	5P_MVR4	MCC/(temp, hum.ratio, solar)	N/A	Stop

	5P_MVR5	MCC/(temp, hum.ratio)	Xcp1 = 44.2178, Xcp2 = 52.5467, Ycp =104.8277, LS = 0.5614, RS = 2.2900, X2 = 357.1251R2 =0.634, CV-RMSE =14.275%	OK

Table 5.1: Three-parameter change-point with multiple variable regression model (3P/MVR)

			IMT		EModel			Status		
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	IMT	EModel	Comment for IMT
3PC/MVR	3PC_MVR1	WBC/Temp	3PC_Mvr1.ins	3PC_Mvr1.dat	3PC_Mvr1.dat	3PC_Mvr1.dvn	3PC_Mvr1.doc	OK	agree	
	3PC_MVR2	WBC/Hum.Ratio	3PC_Mvr2.ins	3PC_Mvr2.dat	3PC_Mvr2.dat	3PC_Mvr2.dvn	3PC_Mvr2.doc	OK	agree	
	3PC_MVR3	WBC/Solar	3PC_Mvr3.ins	3PC_Mvr3.dat	3PC_Mvr3.dat	3PC_Mvr3.dvn	3PC_Mvr3.doc	OK	agree	
	3PC_MVR4	WBC/WBH	3PC_Mvr4.ins	3PC_Mvr4.dat	3PC_Mvr4.dat	3PC_Mvr4.dvn	3PC_Mvr4.doc	OK	agree	
	3PC_MVR5	WBC/WBE	3PC_Mvr5.ins	3PC_Mvr5.dat	3PC_Mvr5.dat	3PC_Mvr5.dvn	3PC_Mvr5.doc	OK	agree	
	3PC_MVR6	WBC/(temp, hum.ratio, solar, WBH, WBE)	3PC_Mvr6.ins	3PC_Mvr6.dat	N/A	N/A	N/A	Stop	N/A	For 3P model, Number of X variables must be greater than 0 and less than 5.
	3PC_MVR7	WBC/(temp, hum.ratio, solar, WBH)	3PC_Mvr7.ins	3PC_Mvr7.dat	N/A	N/A	N/A	OK	N/A	
3PH/MVR	3PH_MVR1	WBH/temp	3PH_Mvr1.ins	3PH_Mvr1.dat	3PH_Mvr1.dat	3PH_Mvr1.dvn	3PH_Mvr1.doc	OK	agree	
	3PH_MVR2	WBH/Hum.Ratio	3PH_Mvr2.ins	3PH_Mvr2.dat	3PH_Mvr2.dat	3PH_Mvr2.dvn	3PH_Mvr2.doc	OK	agree	
	3PH_MVR3	WBH/Solar	3PH_Mvr3.ins	3PH_Mvr3.dat	3PH_Mvr3.dat	3PH_Mvr3.dvn	3PH_Mvr3.doc	OK	agree	
	3PH_MVR4	WBH/(temp, hum.ratio, solar)	3PH_Mvr4.ins	3PH_Mvr4.dat	N/A	N/A	N/A	OK	N/A	

Table 5.2: Three-parameter change-point with multiple variable regression model (3P/MVR)

TEST		Data Type	IMT		EModel	
3PC/MVR	3PC_MVR1	WBC/Temp	Ycp = 795.0877, LS = 0.0000, RS = 60.9633, Xcp = 50.0480, R^2 =0.770, CV-RMSE = 19.590%	OK	Ycp = 795.0990, LS = 0.0000, RS = 60.9631, Xcp = 50.0480, R^2 =0.77, CV-RMSE = 19.6%	agree
	3PC_MVR2	WBC/Hum.Ratio	Ycp = 1017.3652, LS = 0.0000, RS = 133112.3125, Xcp = 0.0041, R^2 =0.613, CV-RMSE = 25.428%	OK	Ycp = 1017.3527, LS = 0.0000, RS = 133113.9397, Xcp = 0.0041, R^2 =0.61, CV-RMSE = 25.4%	agree
	3PC_MVR3	WBC/Solar	Ycp = 1887.0750, LS = 0.0000, RS = 0.9254, Xcp = 21.2894, R^2 =0.095, CV-RMSE = 38.889%	OK	Ycp = 1887.0672, LS = 0.0000, RS = 0.9254, Xcp = 21.2894, R^2 =0.10, CV-RMSE = 38.9%	agree
	3PC_MVR4	WBC/WBH	Ycp = 2512.1580, LS = 0.0000, RS = -2.0852, Xcp = 19.3800, R^2 =0.435, CV-RMSE = 30.727%	OK	Ycp = 2512.1302, LS = 0.0000, RS = -2.0851, Xcp = 19.3800, R^2 =0.44, CV-RMSE = 30.7%	agree
	3PC_MVR5	WBC/WBE	Ycp = 976.0828, LS = 0.0000, RS = 3.8803, Xcp = 530.6744, R^2 =0.382, CV-RMSE = 32.133%	OK	Ycp = 976.0774, LS = 0.0000, RS = 3.8804, Xcp = 530.6744, R^2 =0.38, CV-RMSE = 32.1%	agree
	3PC_MVR6	WBC/(temp, hum.ratio, solar, WBH, WBE)	N/A	Stop	N/A	N/A
	3PC_MVR7	WBC/(temp, hum.ratio, solar, WBH)	Ycp = 604.8068, LS = 0.0000, RS = 42.0486, Xcp = 56.0448, X2 = 65233.5039, X3 = 0.2642, X4 = -0.0899, R^2 = 0.838, CV_RMSE = 16.483%	OK	N/A	N/A
3PH/MVR	3PH_MVR1	WBH/temp	Ycp = 47.5498, LS = -18.356, RS = 0.0000, Xcp = 78.5328, R^2 =0.556, CV-RMSE = 76.806%	OK	Ycp = 47.5497, LS = -18.3562, RS = 0.0000, Xcp = 78.5328, R^2 =0.56, CV-RMSE = 76.8%	agree
	3PH_MVR2	WBH/Hum.Ratio	Ycp = 36.6893, LS = -40973.7813, RS = 0.0000, Xcp = 0.0161, R^2 =0.443, CV-RMSE = 86.006%	OK	Ycp = 36.6869, LS = -40974.2867, RS = 0.0000, Xcp = 0.0161, R^2 =0.44, CV-RMSE = 86.0%	agree
	3PH_MVR3	WBH/Solar	Ycp = 143.3547, LS = -0.1088, RS = 0.0000, Xcp = 1021.8912, R^2 =0.014, CV-RMSE = 114.480%	OK	Ycp = 141.0558, LS = -0.1088, RS = 0.0000, Xcp = 1043.1806, R^2 =0.01, CV-RMSE = 114.5%	agree
	3PH_MVR4	WBH/temp, hum.ratio, solar	Ycp = 157.5026, LS = -14.5659, RS = 0.0000, Xcp = 81.5312, X2 =-10100.2783, X3 = 0.0735, R^2 =0.584, CV-RMSE = 74.369%	OK	N/A	N/A

Table 6.1: Four-parameter change-point with multiple variable regression model (4P/MVR)

			IMT		EModel			Status		
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	IMT	EModel	Comment for IMT
4PC/MVR	4PC_MVR1	WBC/Temp	4PC_Mvr1.ins	4PC_Mvr1.dat	4PC_Mvr1.dat	4PC_Mvr1.dvn	4PC_Mvr1.doc	OK	agree	
	4PC_MVR2	WBC/Hum.Ratio	4PC_Mvr2.ins	4PC_Mvr2.dat	4PC_Mvr2.dat	4PC_Mvr2.dvn	4PC_Mvr2.doc	OK	agree	
	4PC_MVR3	WBC/Solar	4PC_Mvr3.ins	4PC_Mvr3.dat	4PC_Mvr3.dat	4PC_Mvr3.dvn	4PC_Mvr3.doc	OK	agree	
	4PC_MVR4	WBC/WBH	4PC_Mvr4.ins	4PC_Mvr4.dat	4PC_Mvr4.dat	4PC_Mvr4.dvn	4PC_Mvr4.doc	OK	agree	
	4PC_MVR5	WBC/WBE	4PC_Mvr5.ins	4PC_Mvr5.dat	4PC_Mvr5.dat	4PC_Mvr5.dvn	4PC_Mvr5.doc	OK	agree	
	4PC_MVR6	WBC/(temp, hum.ratio, solar, WBH, WBE)	4PC_Mvr6.ins	4PC_Mvr6.dat	N/A	N/A	N/A	Stop	N/A	For 4P model, Number of X variables must be greater than 0 and less than 4.
	4PC_MVR7	WBC/(temp, hum.ratio, solar)	4PC_Mvr7.ins	4PC_Mvr7.dat	N/A	N/A	N/A	OK	N/A	
4PH/MVR	4PH_MVR1	WBH/temp	4PH_Mvr1.ins	4PH_Mvr1.dat	4PH_Mvr1.dat	4PH_Mvr1.dvn	4PH_Mvr1.doc	OK	agree	
	4PH_MVR2	WBH/Hum.Ratio	4PH_Mvr2.ins	4PH_Mvr2.dat	4PH_Mvr2.dat	4PH_Mvr2.dvn	4PH_Mvr2.doc	OK	agree	
	4PH_MVR3	WBH/Solar	4PH_Mvr3.ins	4PH_Mvr3.dat	4PH_Mvr3.dat	4PH_Mvr3.dvn	4PH_Mvr3.doc	OK	agree	
	4PH_MVR4	WBH/(temp, hum.ratio, solar)	4PH_Mvr4.ins	4PH_Mvr4.dat	N/A	N/A	N/A	OK	N/A	

Table 6.2: Four-parameter change-point with multiple variable regression model (4P/MVR)

TEST		Data Type	IMT		EModel	
4PC/MVR	4PC_MVR1	WBC/Temp	Ycp = 1041.2838, LS = 20.1739, RS = 62.1365, Xcp = 54.5456, R^2 =0.772, CV-RMSE = 19.539%	OK	Ycp = 1041.4609, LS = 20.2269, RS = 62.1283, Xcp = 54.5456, R^2 =0.77, CV-RMSE = 19.5%	agree
	4PC_MVR2	WBC/Hum.Ratio	Ycp = 3038.7427, LS = 133298.4219, RS = 28105.3594, Xcp = 0.0191, R^2 =0.614, CV-RMSE = 25.403%	OK	Ycp = 3038.7405, LS = 133298.3360, RS = 28104.6640, Xcp = 0.0191, R^2 =0.61, CV-RMSE = 25.4%	agree
	4PC_MVR3	WBC/Solar	Ycp = 2198.5732, LS = 9.1839, RS = 0.4249, Xcp = 42.5788, R^2 =0.117, CV-RMSE = 38.427%	OK	Ycp = 2198.1833, LS = 9.1725, RS = 0.4253, Xcp = 42.5788, R^2 =0.12, CV-RMSE = 38.4%	agree
	4PC_MVR4	WBC/WBH	Ycp = 1561.4342, LS = -10.4266, RS = -0.5610, Xcp = 140.9000, R^2 =0.580, CV-RMSE = 26.485%	OK	Ycp = 1561.9606, LS = -10.4200, RS = -0.5621, Xcp = 140.9000, R^2 =0.58, CV-RMSE = 26.5%	agree
	4PC_MVR5	WBC/WBE	Ycp = 2297.9739, LS = 7.7264, RS = 1.6664, Xcp = 791.7968, R^2 =0.455, CV-RMSE = 30.171%	OK	Ycp = 2297.7176, LS = 7.7225, RS = 1.6678, Xcp = 791.7968, R^2 =0.46, CV-RMSE = 30.2%	agree
	4PC_MVR6	WBC/(temp, hum.ratio, solar, WBH, WBE)	N/A	Stop	N/A	N/A
	4PC_MVR7	WBC/(temp, hum.ratio, solar)	Ycp = 632.2288, LS = 9.2148, RS = 43.2151, Xcp = 57.5440, X2 = 65536.2734, X3 = 0.2506, R^2 = 0.838, CV-RMSE = 16.464%	OK	N/A	N/A
4PH/MVR	4PH_MVR1	WBH/temp	Ycp = 74.4480, LS = -18.3837, RS = -3.3026, Xcp = 77.0336, R^2 =0.772, CV-RMSE = 19.539%	OK	Ycp = 74.4114, LS = -18.3867, RS = -3.2995, Xcp = 77.0336, R^2 =0.56, CV-RMSE = 76.7%	agree
	4PH_MVR2	WBH/Hum.Ratio	Ycp = 341.5157, LS = -75755.0625, RS = -27462.0859, Xcp = 0.0071, R^2 =0.449, CV-RMSE = 85.563%	OK	Ycp = 341.5722, LS = -75709.9671, RS = -27469.4929, Xcp = 0.0071, R^2 =0.45, CV-RMSE = 85.6%	agree
	4PH_MVR3	WBH/Solar	Ycp = 215.8252, LS = -0.0468, RS = -0.7951, Xcp = 723.8395, R^2 =0.021, CV-RMSE = 114.085%	OK	Ycp = 215.8266, LS = -0.0468, RS = -0.7951, Xcp = 723.8396, R^2 =0.02, CV-RMSE = 114.1%	agree
	4PH_MVR4	WBH/(temp, hum.ratio, solar)	Ycp = 489.4759, LS = -22.7663, RS = -9.3594, Xcp = 56.0448, X2 = -12879.3535, X3 = 0.0078, R^2 =0.592, CV-RMSE = 73.610%	OK	N/A	N/A

Table 7.1: LoanSTAR data sets using several models (1P, 2P, 3PC, 3PH, 4P, and 5P)

			IMT		EModel			SAS		Status			
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	SAS file	SAS Output	IMT	EModel	SAS	Comment
1P	1P_comp1	711: wbe	711.ins	711.dat	711.dat	711.dvn	711.doc	711.sas	711_out.lst	OK	agree	agree	
	1P_comp2	963: wbe	963.ins	963.dat	963.dat	963.dvn	963.doc	963.sas	963_out.lst	OK	agree	agree	
	1P_comp3	208: mcc	208_2.ins	208_2.dat	208_2.dat	208_2.dvn	208_2.doc	208_2.sas	208_2_out.lst	OK	agree	agree	
	1P_comp4	210: mcc	210_1.ins	210_1.dat	210_1.dat	210_1.dvn	210_1.doc	210_1.sas	210_1_out.lst	OK	agree	agree	
2P	2P_comp1	226: wbc/temp	226.ins	226.dat	226.dat	226.dvn	226.doc	226.sas	226_out.lst	OK	agree	agree	
	2P_comp2	201: wbh/temp	201.ins	201.dat	201.dat	201.dvn	201.doc	201.sas	201_out.lst	OK	agree	agree	
	2P_comp3	952: wbe/temp	952.ins	952.dat	952.dat	952.dvn	952.doc	952.sas	952_out.lst	OK	agree	agree	
	2P_comp4	207: wbc/temp	207_1.ins	207_1.dat	207_1.dat	207_1.dvn	207_1.doc	207_1.sas	207_1_out.lst	OK	agree	agree	
	2P_comp5	207: wbh/temp	207_2.ins	207_2.dat	207_2.dat	207_2.dvn	207_2.doc	207_2.sas	207_2_out.lst	OK	agree	agree	
3PC	3PC_comp1	706: wbe/temp	706.ins	706.dat	706.dat	706.dvn	706.doc	N/A	N/A	OK	agree	-	
	3PC_comp2	208: wbc/temp	208_1.ins	208_1.dat	208_1.dat	208_1.dvn	208_1.doc	N/A	N/A	OK	agree	-	
	3PC_comp3	209: wbc/temp	209.ins	209.dat	209.dat	209.dvn	209.doc	N/A	N/A	OK	agree	-	
3PH	3PH_comp1	707: wbh/temp	707.ins	707.dat	707.dat	707.dvn	707.doc	N/A	N/A	OK	agree	-	
	3PH_comp2	208: wbh/temp	208_3.ins	208_3.dat	208_3.dat	208_3.dvn	208_3.doc	N/A	N/A	OK	agree	-	
4P	4P_comp1	208: wbc/temp	208_4.ins	208_4.dat	208_4.dat	208_4.dvn	208_4.doc	N/A	N/A	OK	agree	-	
	4P_comp2	975: wbc/temp	975.ins	975.dat	975.dat	975.dvn	975.doc	N/A	N/A	OK	agree	-	
	4P_comp3	201: wbc/temp	201.ins	201.dat	201.dat	201.dvn	201.doc	N/A	N/A	OK	agree	-	
	4P_comp4	205: wbc/temp	205.ins	205.dat	205.dat	205.dvn	205.doc	N/A	N/A	OK	agree	-	
5P	5P_comp1	210: wbc/temp	210_2.ins	210_2.dat	N/A	N/A	N/A	N/A	N/A	OK	N/A	-	
	5P_comp2	710: wbe/temp	710.ins	710.dat	N/A	N/A	N/A	N/A	N/A	OK	N/A	-	

Table 7.2: LoanSTAR data sets using several models (1P, 2P, 3PC, 3PH, 4P, and 5P)

TEST		Data Type	IMT		EModel		SAS	
1P	1P_comp1	711: wbe	Ymean = 25409.281, StdDev = 2391.109, CV-StDev = 9.410%	OK	Ymean = 25409.28, StdDev = 2391.11, CV-StDev = 9.4%	agree	Ymean = 25409.2809, StdDev = 2391.1088, CV-StDev = 9.41037572%	agree
	1P_comp2	963: wbe	Ymean = 1118.391, StdDev = 452.392, CV-StDev = 40.450%	OK	Ymean = 1118.39, StdDev = 452.39, CV-StDev = 40.5%	agree	Ymean = 1118.39048, StdDev = 452.392131, CV-StDev = 40.4502846%	agree
	1P_comp3	208: mcc	Ymean = 1253.028, StdDev = 152.335, CV-StDev = 12.157%	OK	Ymean = 1253.03, StdDev = 152.33, CV-StDev = 12.2%	agree	Ymean = 1252.50279, StdDev = 152.222448, CV-StDev = 12.1534618%	agree
	1P_comp4	210: mcc	Ymean = 2573.383, StdDev = 72.946, CV-StDev = 2.835%	OK	Ymean = 2573.38, StdDev = 72.95, CV-StDev = 2.8%	agree	Ymean = 2573.29006, StdDev = 73.0254868, CV-StDev = 2.83782571%	agree
2P	2P_comp1	226: wbc/temp	Yint = -10227.1260, Slope = 470.2920, R^2 = 0.834, CV-RMSE = 13.538%	OK	Yint = -10228.5602, Slope = 470.3123, R^2 = 0.83, CV-RMSE = 13.5%	agree	Yint = -10229, Slope = 470.31234, R^2 = 0.8338, CV-RMSE = 13.53846%	agree
	2P_comp2	201: wbh/temp	Yint = 68439.5078, Slope = -649.0869, R^2 = 0.691, CV-RMSE = 24.767%	OK	Yint = 68439.4669, Slope = -649.0859, R^2 = 0.69, CV-RMSE = 24.8%	agree	Yint = 68439, Slope = -649.08587, R^2 = 0.6906, CV-RMSE = 24.76693%	agree
	2P_comp3	952: wbe/temp	Yint = 2338.5520, Slope = 212.1381, R^2 = 0.728, CV-RMSE = 11.645%	OK	Yint = 2338.4196, Slope = 212.1397, R^2 = 0.73, CV-RMSE = 11.6%	agree	Yint = 2338.41956, Slope = 212.13968, R^2 = 0.7280, CV-RMSE = 11.64547%	agree
	2P_comp4	207: wbc/temp	Yint = -5041.9448, Slope = 105.2378, R^2 = 0.861, CV-RMSE = 27.592%	OK	Yint = -5041.7452, Slope = 105.2349, R^2 = 0.86, CV-RMSE = 27.6%	agree	Yint = -5041.74522, Slope = 105.23485, R^2 = 0.8607, CV-RMSE = 27.59200%	agree
	2P_comp5	207: wbh/temp	Yint = 23192.1055, Slope = -219.9567, R^2 = 0.691, CV-RMSE = 24.767%	OK	Yint = 23192.0949, Slope = -219.9563, R^2 = 0.69, CV-RMSE = 24.8%	agree	Yint = 23192, Slope = -219.95630, R^2 = 0.6906, CV-RMSE = 24.76704%	agree
3PC	3PC_comp1	706: wbe/temp	Ycp = 2417.5983, LS = 0.0000, RS = 87.6157, Xcp = 56.7600, R^2 = 0.339, CV-RMSE = 30.578%	OK	Ycp = 2417.5941, LS = 0.0000, RS = 87.6158, Xcp = 56.7600, R^2 = 0.34, CV-RMSE = 30.6%	agree	N/A	-
	3PC_comp2	208: wbc/temp	Ycp = 11145.4775, LS = 0.0000, RS = 945.5939, Xcp = 59.7600, R^2 = 0.855, CV-RMSE = 18.214%	OK	Ycp = 11145.4541, LS = 0.0000, RS = 945.5957, Xcp = 59.7600, R^2 = 0.85, CV-RMSE = 18.2%	agree	N/A	-
	3PC_comp3	209: wbc/temp	Ycp = 78198.4688, LS = 0.0000, RS = 2117.7173, Xcp = 52.2000, R^2 = 0.768, CV-RMSE = 12.495%	OK	Ycp = 78198.5274, LS = 0.0000, RS = 2117.7136, Xcp = 52.2000, R^2 = 0.77, CV-RMSE = 12.5%	agree	N/A	-
3PH	3PH_comp1	707: wbh/temp	Ycp = 10248.5869, LS = -8369.0127, RS = 0.0000, Xcp = 61.6800, R^2 = 0.934, CV-RMSE = 26.455%	OK	Ycp = 10248.4245, LS = -8369.0158, RS = 0.0000, Xcp = 61.6800, R^2 = 0.93, CV-RMSE = 26.5%	agree	N/A	-
	3PH_comp2	208: wbh/temp	Ycp = 6001.7227, LS = -639.4753, RS = 0.0000, Xcp = 79.9200, R^2 = 0.951, CV-RMSE = 13.166%	OK	Ycp = 6001.7111, LS = -639.4761, RS = 0.0000, Xcp = 79.9200, R^2 = 0.95, CV-RMSE = 13.2%	agree	N/A	-
4P	4P_comp1	208: wbc/temp	Ycp = 17613.2813, LS = 343.6089, RS = 1081.8597, Xcp = 68.5800, R^2 = 0.873, CV-RMSE = 17.054%	OK	Ycp = 17613.7419, LS = 343.6469, RS = 1081.8386, Xcp = 68.5800, R^2 = 0.87, CV-RMSE = 17.1%	agree	N/A	-
	4P_comp2	975: wbc/temp	Ycp = 1529.8441, LS = 16.8140, RS = 73.5243, Xcp = 69.1200, R^2 = 0.816, CV-RMSE = 13.204%	OK	Ycp = 1529.8660, LS = 16.8142, RS = 73.5246, Xcp = 69.1200, R^2 = 0.82, CV-RMSE = 13.2%	agree	N/A	-
	4P_comp3	201: wbc/temp	Ycp = 27831.1035, LS = 562.6268, RS = 1278.6936, Xcp = 61.6000, R^2 = 0.754, CV-RMSE = 20.844%	OK	Ycp = 27831.8523, LS = 562.7605, RS = 1278.6456, Xcp = 61.6000, R^2 = 0.75, CV-RMSE = 20.8%	agree	N/A	-
	4P_comp4	205: wbc/temp	Ycp = 9191.9424, LS = 61.4300, RS = 173.2227, Xcp = 68.6200, R^2 = 0.850, CV-RMSE = 6.125%	OK	Ycp = 9191.8404, LS = 61.4281, RS = 173.2247, Xcp = 68.6200, R^2 = 0.85, CV-RMSE = 6.1%	agree	N/A	-
5P	5P_comp1	210: wbc/temp	Xcp1 = 62.0000, Xcp2 = 69.0000, Ycp = 100499.6250, LS = -635.1901, RS = 2534.4150, R^2 = 0.699, CV-RMSE = 9.511%	OK	N/A	N/A	N/A	-
	5P_comp2	710: wbe/temp	Xcp1 = 58.7007, Xcp2 = 61.7438, Ycp = 11665.7363, LS = -120.6786, RS = 47.0371, R^2 = 0.274, CV-RMSE = 21.866%	OK	N/A	N/A	N/A	-

Table 8.1: Variable-Base Cooling Degree-Day (CDD) and the CDD with multiple variable regression models (CDD/MVR)

			IMT		PRISM			Status		
TEST		Data Type	IMT file	Data File	Meter File	Weather File	Output	IMT	PRISM	Comment
CDD	VBDD_C1	WBE	VBDD_C1.ins	VBDD_C1.dat	1308el_1.mtr	Bcs_1.tps	VBDD_C1.doc	OK	agree	
(Q)	VBDD_C3	WBE	VBDD_C3.ins	VBDD_C3.dat	1308el_3.mtr	Bcs_1.tps	VBDD_C3.doc	OK	agree	
	VBDD_C4	WBE	VBDD_C4.ins	VBDD_C4.dat	1308el_4.mtr	Bcs_1.tps	VBDD_C4.doc	OK	agree	
	VBDD_C5	WBE	VBDD_C5.ins	VBDD_C5.dat	1308el_5.mtr	Bcs_1.tps	VBDD_C5.doc	OK	agree	
	VBDD_C6	WBE	VBDD_C6.ins	VBDD_C6.dat	1308el_6.mtr	Bcs_1.tps	VBDD_C6.doc	OK	agree	
	VBDD_C7	WBE	VBDD_C7.ins	VBDD_C7.dat	1308el_7.mtr	Bcs_1.tps	VBDD_C7.doc	OK	agree	
	VBDD_C8	WBE	VBDD_C8.ins	VBDD_C8.dat	1308el_8.mtr	Bcs_1.tps	VBDD_C8.doc	OK	agree	
	VBDD_C9	WBE	VBDD_C9.ins	VBDD_C9.dat	1308el_9.mtr	Bcs_1.tps	VBDD_C9.doc	OK	agree	
	VBDD_C10	WBE	VBDD_C10.ins	VBDD_C10.dat	1308el_10.mtr	Bcs_1.tps	VBDD_C10.doc	OK	agree	
	VBDD_C11	WBE	VBDD_C11.ins	VBDD_C11.dat	1308el_11.mtr	Bcs_1.tps	VBDD_C11.doc	OK	agree	
	VBDD_C12	WBE	VBDD_C12.ins	VBDD_C12.dat	1308el_12.mtr	Bcs_1.tps	VBDD_C12.doc	OK	agree	
	VBDD_C13	WBE	VBDD_C13.ins	VBDD_C13.dat	1308el_13.mtr	Bcs_1.tps	VBDD_C13.doc	OK	agree	
CDD	VBDD_C1d	WBE	VBDD_C1d.ins	VBDD_C1d.dat	1308el_1.mtr	Bcs_1.tps	VBDD_C1.doc	OK	agree	
(Q/day)	VBDD_C2d	WBE	VBDD_C2d.ins	VBDD_C2d.dat	1308el_2.mtr	Bcs_1.tps	VBDD_C2.doc	OK	agree	
	VBDD_C3d	WBE	VBDD_C3d.ins	VBDD_C3d.dat	1308el_3.mtr	Bcs_1.tps	VBDD_C3.doc	OK	agree	
	VBDD_C4d	WBE	VBDD_C4d.ins	VBDD_C4d.dat	1308el_4.mtr	Bcs_1.tps	VBDD_C4.doc	OK	agree	
	VBDD_C5d	WBE	VBDD_C5d.ins	VBDD_C5d.dat	1308el_5.mtr	Bcs_1.tps	VBDD_C5.doc	OK	agree	
	VBDD_C6d	WBE	VBDD_C6d.ins	VBDD_C6d.dat	1308el_6.mtr	Bcs_1.tps	VBDD_C6.doc	OK	agree	
	VBDD_C7d	WBE	VBDD_C7d.ins	VBDD_C7d.dat	1308el_7.mtr	Bcs_1.tps	VBDD_C7.doc	OK	agree	
	VBDD_C8d	WBE	VBDD_C8d.ins	VBDD_C8d.dat	1308el_8.mtr	Bcs_1.tps	VBDD_C8.doc	OK	agree	
	VBDD_C9d	WBE	VBDD_C9d.ins	VBDD_C9d.dat	1308el_9.mtr	Bcs_1.tps	VBDD_C9.doc	OK	agree	
	VBDD_C10d	WBE	VBDD_C10d.ins	VBDD_C10d.dat	1308el_10.mtr	Bcs_1.tps	VBDD_C10.doc	OK	agree	
	VBDD_C11d	WBE	VBDD_C11d.ins	VBDD_C11d.dat	1308el_11.mtr	Bcs_1.tps	VBDD_C11.doc	OK	agree	
	VBDD_C12d	WBE	VBDD_C12d.ins	VBDD_C12d.dat	1308el_12.mtr	Bcs_1.tps	VBDD_C12.doc	OK	agree	
	VBDD_C13d	WBE	VBDD_C13d.ins	VBDD_C13d.dat	1308el_13.mtr	Bcs_1.tps	VBDD_C13.doc	OK	agree	
CDD/MVR	CDD_MVR1	WBC/Temp	CDD_Mvr1.ins	CDD_Mvr1.dat	N/A	N/A	N/A	OK	N/A	
	CDD_MVR2	WBC/Hum.Ratio	CDD_Mvr2.ins	CDD_Mvr2.dat	N/A	N/A	N/A	Stop	N/A	Error in Subroutine Invert. Matrix is singular
	CDD_MVR3	WBC/Solar	CDD_Mvr3.ins	CDD_Mvr3.dat	N/A	N/A	N/A	OK	N/A	

	CDD_MVR4	WBC/(temp, hum.ratio, solar)	CDD_Mvr4.ins	CDD_Mvr4.dat	N/A	N/A	N/A	Error	N/A	IMT ignores X2 and X3. Only X1 is used in the calculation. The .INS and output files show different parameters.
	CDD_MVR5	WBC/(temp, hum.ratio, solar)	CDD_Mvr5.ins	CDD_Mvr5.dat	N/A	N/A	N/A	OK	N/A	Use CDD residual file as input to the MVR model to mimic CDD-MVR capabilities

Table 8.2: Variable-Base Cooling Degree-Day Model (CDD)

TEST		Data Type	IMT		PRISM	
CDD (Q)	VBDD_C1	WBE	DD = 70, Base Use = 479.7767, Cooling Slope = 2.2761, R^2 = 0.971	OK	DD = 69, Base Use = 16.05/day, Cooling Slope = 2.18, R^2 = 0.9516	DD & Slope agree
	VBDD_C2	WBE	DD = 70, Base Use = 487.9663, Cooling Slope = 2.2588, R^2 = 0.973	OK	DD = 69, Base Use = 15.91/day, Cooling Slope = 2.19, R^2 = 0.9527	DD & Slope agree
	VBDD_C3	WBE	DD = 70, Base Use = 501.4227, Cooling Slope = 2.2278, R^2 = 0.972	OK	DD = 69, Base Use = 16.30/day, Cooling Slope = 2.16, R^2 = 0.9534	DD & Slope agree
	VBDD_C4	WBE	DD = 72, Base Use = 494.6992, Cooling Slope = 2.5731, R^2 = 0.964	OK	DD = 69, Base Use = 15.84/day, Cooling Slope = 2.19, R^2 = 0.9514	DD & Slope agree
	VBDD_C5	WBE	DD = 73, Base Use = 497.9115, Cooling Slope = 2.7893, R^2 = 0.964	OK	DD = 70, Base Use = 15.84/day, Cooling Slope = 2.35, R^2 = 0.9515	DD & Slope agree
	VBDD_C6	WBE	DD = 71, Base Use = 482.4804, Cooling Slope = 2.3642, R^2 = 0.956	OK	DD = 70.26, Base Use = 15.87/day, Cooling Slope = 2.31, R^2 = 0.9602	DD & Slope agree
	VBDD_C7	WBE	DD = 71, Base Use = 482.5603, Cooling Slope = 2.1845, R^2 = 0.892	OK	DD = 69.78, Base Use = 16.04/day, Cooling Slope = 1.96, R^2 = 0.9194	DD & Slope agree
	VBDD_C8	WBE	DD = 69, Base Use = 481.3739, Cooling Slope = 1.6488, R^2 = 0.849	OK	DD = 69.00, Base Use = 15.93/day, Cooling Slope = 1.80, R^2 = 0.9140	DD & Slope agree
	VBDD_C9	WBE	DD = 70, Base Use = 468.6212, Cooling Slope = 1.9651, R^2 = 0.871	OK	DD = 69.14, Base Use = 15.86/day, Cooling Slope = 1.92, R^2 = 0.8879	DD & Slope agree
	VBDD_C10	WBE	DD = 70, Base Use = 462.1427, Cooling Slope = 1.9794, R^2 = 0.868	OK	DD = 69.33, Base Use = 15.91/day, Cooling Slope = 1.87, R^2 = 0.8932	DD & Slope agree
	VBDD_C11	WBE	DD = 70, Base Use = 490.9287, Cooling Slope = 1.9138, R^2 = 0.852	OK	DD = 68.66, Base Use = 15.99/day, Cooling Slope = 1.78, R^2 = 0.8812	DD & Slope agree
	VBDD_C12	WBE	DD = 68, Base Use = 431.1980, Cooling Slope = 1.7898, R^2 = 0.877	OK	DD = 68.50, Base Use = 16.49/day, Cooling Slope = 1.72, R^2 = 0.8888	DD & Slope agree
	VBDD_C13	WBE	DD = 66, Base Use = 470.6246, Cooling Slope = 1.6957, R^2 = 0.880	OK	DD = 66.00, Base Use = 13.94/day, Cooling Slope = 1.62, R^2 = 0.8936	DD & Slope agree
CDD (Q/day)	VBDD_C1d	WBE	DD = 72, Base Use = 17.1774/day, Cooling Slope = 0.0827, R^2 = 0.914	OK	DD = 69, Base Use = 16.05/day, Cooling Slope = 2.18, R^2 = 0.9516	DD & Base Use agree
	VBDD_C2d	WBE	DD = 72, Base Use = 17.2799/day, Cooling Slope = 0.0824, R^2 = 0.914	OK	DD = 69, Base Use = 15.91/day, Cooling Slope = 2.19, R^2 = 0.9527	DD & Base Use agree
	VBDD_C3d	WBE	DD = 72, Base Use = 17.3927/day, Cooling Slope = 0.0821, R^2 = 0.914	OK	DD = 69, Base Use = 16.30/day, Cooling Slope = 2.16, R^2 = 0.9534	DD & Base Use agree
	VBDD_C4d	WBE	DD = 72, Base Use = 16.7694/day, Cooling Slope = 0.0838, R^2 = 0.905	OK	DD = 69, Base Use = 15.84/day, Cooling Slope = 2.19, R^2 = 0.9514	DD & Base Use agree
	VBDD_C5d	WBE	DD = 73, Base Use = 16.8511/day, Cooling Slope = 0.0909, R^2 =0.907	OK	DD = 70, Base Use = 15.84/day, Cooling Slope = 2.35, R^2 = 0.9515	DD & Base Use agree
	VBDD_C6d	WBE	DD = 70, Base Use = 16.0906/day, Cooling Slope = 0.0655, R^2 = 0.923	OK	DD = 70.26, Base Use = 15.87/day, Cooling Slope = 2.31, R^2 = 0.9602	DD & Base Use agree
	VBDD_C7d	WBE	DD = 70, Base Use = 16.1181/day, Cooling Slope = 0.0613, R^2 = 0.892	OK	DD = 69.78, Base Use = 16.04/day, Cooling Slope = 1.96, R^2 = 0.9194	DD & Base Use agree
	VBDD_C8d	WBE	DD = 70, Base Use = 16.2195/day, Cooling Slope = 0.0593, R^2 = 0.853	OK	DD = 69.00, Base Use = 15.93/day, Cooling Slope = 1.80, R^2 = 0.9140	DD & Base Use agree
	VBDD_C9d	WBE	DD = 71, Base Use = 15.8395/day, Cooling Slope = 0.0657, R^2 = 0.925	OK	DD = 69.14, Base Use = 15.86/day, Cooling Slope = 1.92, R^2 = 0.8879	DD & Base Use agree
	VBDD_C10d	WBE	DD = 70, Base Use = 15.5676/day, Cooling Slope = 0.0617, R^2 = 0.920	OK	DD = 69.33, Base Use = 15.91/day, Cooling Slope = 1.87, R^2 = 0.8932	DD & Base Use agree
	VBDD_C11d	WBE	DD = 70, Base Use = 16.6691/day, Cooling Slope = 0.0591, R^2 = 0.912	OK	DD = 68.66, Base Use = 15.99/day, Cooling Slope = 1.78, R^2 = 0.8812	DD & Base Use agree
	VBDD_C12d	WBE	DD = 70, Base Use = 15.7325/day, Cooling Slope = 0.0618, R^2 = 0.924	OK	DD = 68.50, Base Use = 16.49/day, Cooling Slope = 1.72, R^2 = 0.8888	DD & Base Use agree
	VBDD_C13d	WBE	DD = 66, Base Use = 12.5029/day, Cooling Slope = 0.0534, R^2 = 0.922	OK	DD = 66.00, Base Use = 13.94/day, Cooling Slope = 1.62, R^2 = 0.8936	DD & Base Use agree

Table 8.3: CDD with multiple variable regression models (CDD/MVR)

TEST		Data Type	IMT	
CDD/MVR	CDD_MVR1	WBC/Temp	DD = 55, Base Use = 21861.4785, Cooling Slope = 1727.6812, R2 = 0.90, CV-RMSE = 12.064%	OK
	CDD_MVR2	WBC/Hum.Ratio	N/A	Stop
	CDD_MVR3	WBC/Solar	DD = 80, Base Use = 30885.9727, Cooling Slope = 3.8399, R^2 = 0.145, CV-RMSE = 35.235%	OK

	CDD_MVR4	WBC/(temp, hum.ratio, solar)	N/A	Error
	CDD_MVR5	WBC/(temp, hum.ratio, solar)	a = -26437.8164, X1 = 713.8923, X2 = 1332734.1250, X3 = 1.9659, R2 = 0.771, CV-RMSE = 17.675%	OK

Table 9.1: Variable-Base Heating Degree-Day (HDD) and the HDD with multiple variable regression models (HDD/MVR)

			IMT		PRISM			Status		
TEST		**Data Type**	**IMT file**	**Data File**	**Meter File**	**Weather File**	**Output**	**IMT**	**PRISM**	**Comment**
HDD	VBDD_H1	GAS	VBDD_H1.ins	VBDD_H1.dat	1308ng_1.mtr	Bcs_1.tps	VBDD_H1.doc	OK	agree	
(Q)	VBDD_H2	GAS	VBDD_H2.ins	VBDD_H2.dat	1308ng_2.mtr	Bcs_1.tps	VBDD_H2.doc	OK	agree	
	VBDD_H3	GAS	VBDD_H3.ins	VBDD_H3.dat	1308ng_3.mtr	Bcs_1.tps	VBDD_H3.doc	OK	agree	
	VBDD_H4	GAS	VBDD_H4.ins	VBDD_H4.dat	1308ng_4.mtr	Bcs_1.tps	VBDD_H4.doc	OK	agree	
	VBDD_H5	GAS	VBDD_H5.ins	VBDD_H5.dat	1308ng_5.mtr	Bcs_1.tps	VBDD_H5.doc	OK	close	IMT has a lower DD base by 3 F.
	VBDD_H6	GAS	VBDD_H6.ins	VBDD_H6.dat	1308ng_6.mtr	Bcs_1.tps	VBDD_H6.doc	OK	agree	
HDD	VBDD_H1d	GAS	VBDD_H1d.ins	VBDD_H1d.dat	1308ng_1.mtr	Bcs_1.tps	VBDD_H1.doc	OK	agree	
(Q/day)	VBDD_H2d	GAS	VBDD_H2d.ins	VBDD_H2d.dat	1308ng_2.mtr	Bcs_1.tps	VBDD_H2.doc	OK	agree	
	VBDD_H3d	GAS	VBDD_H3d.ins	VBDD_H3d.dat	1308ng_3.mtr	Bcs_1.tps	VBDD_H3.doc	OK	agree	
	VBDD_H4d	GAS	VBDD_H4d.ins	VBDD_H4d.dat	1308ng_4.mtr	Bcs_1.tps	VBDD_H4.doc	OK	agree	
	VBDD_H5d	GAS	VBDD_H5d.ins	VBDD_H5d.dat	1308ng_5.mtr	Bcs_1.tps	VBDD_H5.doc	OK	close	
	VBDD_H6d	GAS	VBDD_H6d.ins	VBDD_H6d.dat	1308ng_6.mtr	Bcs_1.tps	VBDD_H6.doc	OK	agree	
HDD/MVR	HDD_MVR1	WBH/Temp	HDD_Mvr1.ins	HDD_Mvr1.dat	N/A	N/A	N/A	OK	N/A	
	HDD_MVR2	WBH/Hum.Ratio	HDD_Mvr2.ins	HDD_Mvr2.dat	N/A	N/A	N/A	OK	N/A	
	HDD_MVR3	WBH/Solar	HDD_Mvr3.ins	HDD_Mvr3.dat	N/A	N/A	N/A	Stop	N/A	Error in Subroutine Invert. Matrix is singular
	HDD_MVR4	WBH/(temp, hum.ratio, solar)	HDD_Mvr4.ins	HDD_Mvr4.dat	N/A	N/A	N/A	Error	N/A	IMT ignores X2 and X3. Only X1 is used in the calculation. The .INS and output files show different parameters.
	HDD_MVR5	WBH/(temp, hum.ratio, solar)	HDD_Mvr5.ins	HDD_Mvr5.dat	N/A	N/A	N/A	OK	N/A	Use HDD residual file as input to the MVR model to mimic HDD-MVR capabilities

Table 9.2: Variable-Base Heating Degree-Day Model (HDD)

TEST		Data Type	IMT		PRISM	
HDD (Q)	VBDD_H1	GAS	DD = 69, Base Use = 18.5144, Heating Slope = 0.1689	OK	DD = 69.79, Base Use = 0.72/day, Heating Slope = 0.15, R^2 = 0.9841	DD & Slope agree
	VBDD_H2	GAS	DD = 69, Base Use = 17.5255, Heating Slope = 0.1712	OK	DD = 70.30, Base Use = 0.66/day, Heating Slope = 0.15, R^2 = 0.9822	DD & Slope agree
	VBDD_H3	GAS	DD = 70, Base Use = 15.5614, Heating Slope = 0.1653	OK	DD = 70.59, Base Use = 0.62/day, Heating Slope = 0.15, R^2 = 0.9817	DD & Slope agree
	VBDD_H4	GAS	DD = 70, Base Use = 15.3872, Heating Slope = 0.1657	OK	DD = 70.00, Base Use = 0.62/day, Heating Slope = 0.16, R^2 = 0.9811	DD & Slope agree
	VBDD_H5	GAS	DD = 68, Base Use = 15.8379, Heating Slope = 0.1805	OK	DD = 70.00, Base Use = 0.60/day, Heating Slope = 0.15, R^2 = 0.9668	close
	VBDD_H6	GAS	DD = 66, Base Use = 16.3870, Heating Slope = 0.1922	OK	DD = 66.00, Base Use = 0.64/day, Heating Slope = 0.18, R^2 = 0.8757	DD & Slope agree
HDD (Q/day)	VBDD_H1d	GAS	DD = 70, Base Use = 0.6979/day, Heating Slope = 0.0050, R^2 =0.957	OK	DD = 69.79, Base Use = 0.72/day, Heating Slope = 0.15, R^2 = 0.9841	DD & Base Use agree
	VBDD_H2d	GAS	DD = 70, Base Use = 0.6458/day, Heating Slope = 0.0051, R^2 = 0.957	OK	DD = 70.30, Base Use = 0.66/day, Heating Slope = 0.15, R^2 = 0.9822	DD & Base Use agree
	VBDD_H3d	GAS	DD = 70, Base Use = 0.6245/day, Heating Slope = 0.0051, R^2 = 0.959	OK	DD = 70.59, Base Use = 0.62/day, Heating Slope = 0.15, R^2 = 0.9817	DD & Base Use agree
	VBDD_H4d	GAS	DD = 70, Base Use = 0.5954/day, Heating Slope = 0.0052, R^2 = 0.959	OK	DD = 70.00, Base Use = 0.62/day, Heating Slope = 0.16, R^2 = 0.9811	DD & Base Use agree
	VBDD_H5d	GAS	DD = 70, Base Use = 0.6037/day, Heating Slope = 0.0052, R^2 = 0.957	OK	DD = 70.00, Base Use = 0.60/day, Heating Slope = 0.15, R^2 = 0.9668	close
	VBDD_H6d	GAS	DD = 67, Base Use = 0.6496/day, Heating Slope = 0.0057, R^2 = 0.850	OK	DD = 66.00, Base Use = 0.64/day, Heating Slope = 0.18, R^2 = 0.8757	DD & Base Use agree

Table 9.3: HDD with multiple variable regression models (HDD/MVR)

TEST		Data Type	IMT	
HDD/MVR	HDD_MVR1	WBH/Temp	DD = 70, Base Use = 3130.9380, Heating Slope = 562.0746, R2 = 0.507, CV-RMSE = 75.997%	OK
	HDD_MVR2	WBH/Hum.Ratio	DD = 80, Base Use = 6735.5371, Heating Slope = -8.6924, R2 = 0.004, CV-RMSE = 107.999%	OK
	HDD_MVR3	WBH/Solar	N/A	Stop
	HDD_MVR4	WBH/(temp, hum.ratio, solar)	N/A	Error
	HDD_MVR5	WBH/(temp, hum.ratio, solar)	a = 29076.8477, X1 = -303.1802, X2 = -115652.0469, X3 = 0.0730, R^2 = 0.579, CV-RMSE = 50.796%	OK

Table 10.1: Multiple Variable Regression Model (MVR)

			IMT		EModel			SAS		Status			
TEST		Data Type	IMT file	Data File	Data File	.DVN File	Output	SAS File	SAS output	IMT	Emodel	SAS	Comment for IMT
MVR	MVR_0	Synthetic data	MVR_0.ins	MVR_0.dat	MVR_0.dat	MVR_0.dat	MVR_0.dvn	MVR_0.sas	MVR_0.lst	OK	agree	agree	
	MVR_1	WBE/(temp, hum.ratio,solar)	MVR_1.ins	MVR_1.dat	MVR_1.dat	MVR_1.dat	MVR_1.dvn	MVR_1.sas	MVR_1.lst	OK	Overflow	agree	
	MVR_2	WBC/(WBE, WBH, temp, hum.ratio, solar)	MVR_2.ins	MVR_2.dat	MVR_2.dat	MVR_2.dat	MVR_2.dvn	MVR_2.sas	MVR_2.lst	OK	Overflow	agree	
	MVR_3	WBH/(WBE, WBC, temp, hum.ratio, solar)	MVR_3.ins	MVR_3.dat	MVR_3.dat	MVR_3.dat	MVR_3.dvn	MVR_3.sas	MVR_3.lst	OK	Overflow	agree	

Table 10.2: Multiple Variable Regression Model (MVR)

TEST		Data Type	IMT		EModel		SAS	
MVR	MVR_0	Synthetic data	a = 1.0000, X1 = 1.0000, X2 = 2.0000, X3 = 3.0000, X4 = 4.0000, X5 = 5.0000, X6 = 6.0000	OK	Yint = 1.0000, X1 = 1.0000, X2 = 2.0000, X3 = 3.0000, X4 = 4.0000, X5 = 5.0000, X6 = 6.0000	agree	Yint = 1.0000, X1 = 1.0000, X2 = 2.0000, X3 = 3.0000, X4 = 4.0000, X5 = 5.0000, X6 = 6.0000	agree
	MVR_1	WBE/(temp, hum.ratio,solar)	a = 473.3068, X1 = 4.6356, X2 = -1992.5603, X3 = 0.1754	OK	N/A	N/A	Yint = 473.228982, X1 = 4.637366, X2 = -1995.942269, X3 = 0.175387	agree
	MVR_2	WBC/(WBE, WBH, temp, hum.ratio, solar)	a = -2354.9236, X1 = 28.4639, X2 = 72907.6406, X3 = -0.0178, X4 = 1.8636, X5 = 0.1787	OK	N/A	N/A	Yint = -2354.82060, X1 = 28.47099, X2 = 72901.36995, X3 = -0.01776, X4 = 1.86297, X5 = 0.17879	agree
	MVR_3	WBH/(WBE, WBC, temp, hum.ratio, solar)	a = 1428.1320, X1 = -12.9370, X2 = -16776.8887, X3 = 0.1297, X4 = -0.2883, X5 = 0.0591	OK	N/A	N/A	Yint = 1428.56701, X1 = -12.94613, X2 = -16788.49610, X3 = 0.12969, X4 = -0.28844, X5 = 0.05928	agree