ornl

OAK RIDGE NATIONAL LABORATORY

LOCKHEED MARTIN

ORNL/CON-457

Validation of the National Energy Audit (NEAT) with Data from a Gas Utility Low-Income Residential Weatherization Program in New York State

Michael B. Gettings
Linda G. Berry
Mark A. Beyer
Jonathan B. Maxwell

MANAGED AND OPERATED BY
LOCKHEED MARTIN ENERGY RESEARCH CORPORATION
FOR THE UNITED STATES
DEPARTMENT OF ENERGY

ORNL-27 (3-96)

This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information, P.O. Box 62, Oak Ridge, TN 37831; prices available from (423) 576-8401, FTS 626-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ORNL/CON-457

Energy Division

VALIDATION OF THE NATIONAL ENERGY AUDIT (NEAT) WITH DATA FROM A GAS UTILITY LOW-INCOME RESIDENTIAL WEATHERIZATION PROGRAM IN NEW YORK STATE

Michael B. Gettings and Linda G. Berry
Oak Ridge National Laboratory

Mark A. Beyer and Jonathan B. Maxwell
Aspen Systems Corporation

January 1998

Prepared by
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6205
managed by
LOCKHEED MARTIN ENERGY RESEARCH CORP.
for the
U.S. DEPARTMENT OF ENERGY
under contract number DE-AC05-96OR22464

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vii

ABSTRACT ix

ACKNOWLEDGMENTS xi

1. INTRODUCTION 1
    1.1 THE NATIONAL ENERGY AUDIT (NEAT) 1
    1.2 DATABASE FROM A NEW YORK GAS UTILITY 2
    1.3 OPPORTUNITY FOR A LOW-COST NEAT VALIDATION STUDY 3
    1.4 REPORT ORGANIZATION 4

2. RESEARCH QUESTIONS AND METHODS 5
    2.1 PREVIOUS RESEARCH ON AUDIT REALIZATION RATES 5
    2.2 METHODOLOGICAL ISSUES IN THE DETERMINATION OF REALIZATION RATES 7
    2.3 AUDIT-PREDICTED SAVINGS 11
        2.3.1 The Utility Database 11
        2.3.2 Ensuring the Accuracy of NEAT Predictions 11
        2.3.3 Controlling External Sources of Error 12
            2.3.3.1 Poor implementation practices. 12
            2.3.3.2 Inaccurate and questionable inputs. 13
            2.3.3.3 Demographic and behavioral changes. 14
        2.3.4 Identifying Potential Internal Sources of Error in NEAT 15
            2.3.4.1 Inaccurate assumptions. 16
            2.3.4.2 Inaccurate internal algorithms. 17
    2.4 SOURCES OF ERROR IN MEASURED SAVINGS ESTIMATED FROM BILLING ANALYSES 18
    2.5 RATINGS OF DATA QUALITY 18

3. RESULTS 20
    3.1 PROGRAM TARGETED HIGH-USAGE HOUSEHOLDS 20
    3.2 ESTIMATES OF PRE- AND POST-WEATHERIZATION CONSUMPTION 20
    3.3 NEAT-PREDICTED SAVINGS 24
    3.4 REALIZATION RATES 24
    3.5 ANALYSIS OF REASONS FOR VARIATIONS IN REALIZATION RATES 26

4. COMPARISONS OF NEAT VERSUS DOE-2.1E MODELING RESULTS 28
4.1 INTRODUCTION 28
4.2 SIMULATIONS OF PRE-RETROFIT CONSUMPTION 28
4.2.1 Set-Up Methodology 29
4.2.1.1 Building selection 29
4.2.1.2 Weather normalization 29
4.2.1.3 Building variables 30
4.2.1.4 Bill reconciliation 30
4.2.1.5 Output 30
4.2.2 Pre-Retrofit Consumption Results 30
4.3 METHODS OF SIMULATION FOR MEASURE-SPECIFIC AND MONTHLY ENERGY USE 34
4.3.1 NEAT Measure-Specific and Monthly Energy Use Simulation Procedure 34
4.3.2 DOE-2 Model Measure and Post-Retrofit Simulation Procedure 34
4.3.3 Measure-by-Measure Comparisons 35
4.3.3.1 Infiltration Reduction 35
4.3.3.2 Wall Insulation 36
4.3.3.3 Heating System Replacement 36
4.3.3.4 Crawl Space Insulation 37
4.3.3.5 Duct Insulation 37
4.3.3.6 Attic Insulation 37
4.3.3.7 Measure Analysis Summary 37
4.4 COMPARISONS OF NEAT-PREDICTED SAVINGS TO DOE-2 SIMULATED SAVINGS AND OF MODELED SAVINGS TO ACTUAL USE AS MEASURED BY BILLING DATA 38
4.4.1 Comparisons of Pre-Retrofit Consumption, Post-Retrofit Consumption, and Savings Estimates Obtained with the Three Methods 38
4.4.2 Influence of Balance Point Temperature Assumptions 41

5. SUMMARY AND CONCLUSIONS 43
5.1 REALIZATION RATES 43
5.2 VARIATIONS IN REALIZATION RATES BY SUBGROUPS 44
5.3 IDENTIFIED SOURCES OF ERROR 44
5.3.1 Implementation Errors 45
5.3.2 Inaccuracies in Model Assumptions and Algorithms 46
5.3.3 Behavioral Factors 47
5.4 NEAT'S VALUE AS A TOOL FOR SELECTING MEASURES 47
5.5 NEAT PREDICTIONS ARE NOT A SUBSTITUTE FOR MEASUREMENTS OF PROGRAM PERFORMANCE 48

REFERENCES 49

LIST OF TABLES

Table 2.1. Potential Sources of Error and Possible Corrective Actions in the Determination of Realization Rates 9
Table 3.1. Characteristics of the New York Gas Utility Program Dwellings as Compared to Characteristics of Houses in the New York Weatherization Assistance Program 20
Table 3.2. Comparison of PRISM NAC and NAHC Estimates to NEAT Predictions 21
Table 3.3. NEAT Predictions of Savings With and Without Billing Data Adjustment 24
Table 3.4. Realization Rates Calculated with Data Sets with Varying Levels of Data Quality 25
Table 4.1. Conduction Design Heat Loss by Component from NEAT and DOE-2 for House One and House Two 33
Table 4.2. Savings by Measure for House One 36
Table 4.3. Savings by Measure for House Two 36
Table 4.4. Comparisons of Pre-Retrofit Consumption, Post-Retrofit Consumption, and Savings Predictions Obtained from NEAT, DOE 2.1, and Billing Data 39
Table 4.5. Comparisons of Savings Predictions of NEAT and DOE 2.1 to Savings Estimated from Billing Data Using a Fixed versus a Variable Balance Point Temperature 41
Table 5.1. Realization Rates Calculated with Data Sets with Varying Levels of Data Quality 43

LIST OF FIGURES

Figure 3.1. Comparison of Pre-Weatherization NAHC Values from PRISM to Pre-Weatherization Consumption Predicted by NEAT 22
Figure 3.2. Comparison of Pre-Weatherization NAC Values from PRISM to Pre-Weatherization Consumption Predicted by NEAT 22
Figure 3.3. Comparison of Post-Weatherization NAHC Values from PRISM to Post-Weatherization Consumption Predicted by NEAT 23
Figure 3.4. Comparison of Post-Weatherization NAC Values from PRISM to Post-Weatherization Consumption Predicted by NEAT 23
Figure 3.5. Comparison of Savings Estimated in PRISM Analysis to Savings Predicted by NEAT with Uncorrected Inputs 27
Figure 3.6. Comparison of Savings Estimated in PRISM Analysis to Savings Predicted by NEAT with Corrected Inputs 27
Figure 4.1. Comparison of Monthly Energy Use Before Retrofit from Utility Bills, DOE-2, and NEAT for House One 31
Figure 4.2. Comparison of Monthly Energy Use Before Retrofit from Utility Bills, DOE-2, and NEAT for House Two 31
Figure 4.3. Comparison of Slopes of CCF/HDD Lines Before Retrofit from Billing Data, DOE-2, and NEAT for House One 32
Figure 4.4. Comparison of Slopes of CCF/HDD Lines Before Retrofit from Billing Data, DOE-2 and NEAT for House Two 32
Figure 4.5. Post-Retrofit Slopes of CCF/HDD Lines for House One 40
Figure 4.6. Post-Retrofit Slopes of CCF/HDD Lines for House Two 40

ABSTRACT

This study uses two approaches to the validation of the National Energy Audit (NEAT). The first consists of comparisons of audit-predicted savings to savings observed in a pre- and post-retrofit analysis of metered gas consumption. Here, realization rates, which are the ratio of measured savings to audit-predicted savings, are examined for 49 houses to determine how accurately NEAT predicts savings in a field setting.

The second approach involves assessing the accuracy of NEAT's internal algorithms by comparing its results to those obtained with another engineering model, DOE-2.1E, which is an industry standard. In this analysis, both engineering models are applied to two houses, using the same building description data, and measure-specific estimates of savings are compared.

Because the calculation of realization rates is very sensitive to decisions made by the analyst concerning which cases to include or exclude, and how to deal with outlying data, realization rates were calculated for several subsets of the total sample of 49 dwellings. Depending upon the subset of dwellings used, NEAT's realization rates varied from 36% to 63%. The most credible estimates fall between 50% and 70%. There was good agreement between the results obtained from NEAT and from DOE-2.1E

This NEAT validation study, as well as two previous studies, found that NEAT's typical realization rates fall between 50% and 70%. Another important finding was that realization rates vary a great deal depending upon the exact subset of houses that is examined. These findings mean that unvalidated NEAT predictions of savings should not be used to characterize program performance. Such predictions cannot be used as accurate substitutes for evaluations based on measured fuel consumption data. The high similarity of NEAT and DOE-2 results suggests that typical NEAT realization rates of less than 100% are probably not due to inaccuracies in its internal engineering algorithms. The lower savings estimated by the billing analysis, as compared to both of the engineering models, are probably due to implementation and behavioral factors. Changes in the patterns of the usage of secondary fuels was strongly associated with low realization rates.

ACKNOWLEDGMENTS

We appreciate the support for this project that was provided by Gail McKinley, Director of the Office of state and Community Programs, U.S. Department of Energy. Our technical reviewers, Terry Sharp of Oak Ridge National Laboratory and Greg Dalhoff of Dalhoff and Associates, provided many useful comments. We are grateful to both of them for their careful and insightful reviews.

VALIDATION OF THE NATIONAL ENERGY AUDIT (NEAT) WITH DATA FROM A GAS UTILITY LOW-INCOME RESIDENTIAL WEATHERIZATION PROGRAM IN NEW YORK STATE

Michael B. Gettings and Linda G. Berry
Oak Ridge National Laboratory

Mark A. Beyer and Jonathan B. Maxwell
Aspen Systems Corporation

1. INTRODUCTION

In July of 1992, nine major New York utilities initiated three-year pilot weatherization programs for their low-income customers. These programs, collectively referred to as the Utility Low-Income Energy Efficiency Program (ULIEEP), were required by the New York State Public Service Commission (PSC) Order 89-M-124. The PSC order directed the nine regulated utilities (which included gas, electric, and combination gas/electric utilities) in New York to invest $10 million annually to provide energy-efficiency services to low-income customers (Applied Energy Group, Inc. 1996).

The set of weatherization services provided under ULIEEP varied to some extent across the utilities, and according to whether clients were a utility's heating or nonheating customers. Most of the programs offered insulation and infiltration measures to space heating customers, hot water measures to water heating customers, and, in the case of electric and combination utilities, lighting measures, such as compact fluorescents.

One of the New York utilities that implemented a ULIEEP program contacted Oak Ridge National Laboratory (ORNL) early in its program planning process to inquire about the National Energy Audit (NEAT). This gas utility decided to use NEAT for their ULIEEP program, and contracted with ORNL to tailor the NEAT audit to their needs and to provide technical assistance.

1.1 THE NATIONAL ENERGY AUDIT (NEAT)

NEAT is a computer audit tool developed by ORNL for the U.S. Department of Energy's Weatherization Assistance Program (the Program). The Program provides energy-efficiency services to about 80,000 low-income households per year, and is implemented in all 50 states. Currently more than half of the states use NEAT. NEAT's engineering algorithms are based on those developed for Lawrence Berkeley Laboratory's Computerized Instrumented Residential Audit (CIRA) (Krigger, Adams and Gettings, 1997).

The National Energy Audit is designed to select the most cost-effective combination of energy-efficiency measures for installation in each weatherized home. According to the National Energy Audit Manual (Krigger, Adams, and Gettings, 1997), NEAT follows eight steps to select the most cost-effective set of measures:

- NEAT guides the auditor through the process of entering data that describe the dwelling's size and type of exterior walls, floors, windows, doors, attic areas, foundation spaces, air conditioners, and heating systems. These input files are used later along with other information to model the dwelling's energy use before and after weatherization with various possible sets of measures.
- NEAT computes heat loss, heat gain, and energy required to keep the house at a specific thermostatic set point.
- NEAT can review more then 33 possible conservation measures (including attic, wall, kneewall, sill, floor, foundation, and duct insulation; heating systems retrofits of vent dampers, intermittent ignition devices, flame retention head burners, and furnace tune-ups; heating system replacements, smart thermostats; storm and low-e windows; window films, sunscreens, and awnings; evaporative coolers, and air conditioner replacements) and check their applicability to the house as described by the auditor. The auditor may limit the review of possible measures to a subset of the total, if desired, by turning off measures in the set-up routine. For the ULIEEP program, on which this study is based, all but eight measures were turned off.
- NEAT calculates savings for each applicable conservation measure and computes its discounted savings-to-investment ratio (SIR). The SIR depends upon how many times a measure will repay the initial investment during its lifetime. Parameter files (containing materials costs, measure lifetimes, fuel prices, fuel price escalation rates, discount rates, and weather data) along with engineering algorithms, and the input files that describe the dwelling all enter into the SIR calculations.
- NEAT ranks the energy-efficiency measures in order of their SIR.
- Interacted SIRs are then determined by assuming the measures are added to the house in the order of their initial ranking. For example, the measure with the highest initial SIR is assumed to be installed then the SIR of the second ranked measure is interacted with the first. If the second ranked measure's interacted SIR still indicates cost effectiveness, it is assumed to be installed in the determination of the interacted SIRs of the measures yet to be added.
- NEAT again ranks the conservation measures by their interacted SIRs and prepares a list of essential materials.
- If actual fuel consumption data are entered, NEAT will adjust predicted measure savings to actual (as opposed to modeled) fuel consumption. To use this feature, the auditor must obtain data on actual fuel consumption from the gas or electric utility company that supplies heating/cooling fuel to the dwelling.

1.2 DATABASE FROM A NEW YORK GAS UTILITY

The New York gas utility that used NEAT in its ULIEEP program agreed to provide ORNL with a database on the audited homes. This utility database included the NEAT input files, which give a detailed description of each audited dwelling; the parameter files, which include materials costs, fuel prices, and weather data; and a listing of NEAT predictions of energy savings, taken from NEAT output files, for the weatherization measures recommended for each house. Only eight of the measures NEAT can evaluate were actually evaluated and installed by the New York utility program. Therefore, the database that the utility prepared included NEAT-predicted savings only for the following eight measures:

- attic insulation,
- wall insulation,
- basement insulation,
- crawlspace insulation,
- rim joist insulation,
- duct insulation;
- heating system replacements, and
- air infiltration reduction.

The utility also provided billing data for the pre- and post-weatherization year for most of the weatherized homes.

1.3 OPPORTUNITY FOR A LOW-COST NEAT VALIDATION STUDY

Because of a continuing interest in improving the performance of NEAT, and because several improved versions of NEAT had been developed since the first 1990 field test of the audit in North Carolina (Sharp, 1994), ORNL wanted to perform additional NEAT validation studies. The North Carolina field test of NEAT provided initial validation study results for an early version of NEAT (Sharp, 1994). Access to the New York utility information, described in Section 1.2, offered a new opportunity to access a detailed database derived from a later version of NEAT at no cost. With this ULIEEP database, it is possible to study the energy-saving predictions of NEAT for eight measures in a cold climate state, and to assess how well the predicted savings match the measured savings derived from an analysis of billing data.

This study uses two approaches to NEAT validation. The first consists of comparisons of audit-predicted savings to savings observed in a pre- and post-retrofit analysis of metered gas consumption data for all homes with complete data. Here, realization rates, which are the ratio of measured savings to audit-predicted savings, are examined for 49 houses to determine how accurately NEAT predicts savings in a field setting.

The second approach involves assessing the accuracy of NEAT's internal algorithms by comparing its results to those obtained with another engineering model, DOE-2.1E, which is an industry standard. In this analysis, both engineering models are applied to two houses, using the same building description data. The first step simulates a baseline, or pre-retrofit consumption, with each model. The baseline results obtained are then compared with each other and with measured pre-weatherization consumption data. Next, the overall post-retrofit consumption, and the effects of specific measures, are simulated with both models and the results are compared. Finally, the predicted savings obtained with both models are compared with each other and to the savings estimates obtained from billing data.

1.4 REPORT ORGANIZATION

The next section (Section 2) discusses the research questions addressed and the methods used in this study. Section 3 presents the results of the analysis of realization rates, where NEAT predictions of energy savings are compared to the savings estimates obtained from the analysis of the gas billing data for 49 houses. Section 4 compares NEAT modeling results to DOE-2.1E modeling results for two houses. Section 5 summarizes the study's conclusions and offers recommendations.

2. RESEARCH QUESTIONS AND METHODS

A variety of research questions are addressed in this study. The major questions are:

- What are the NEAT predicted gas savings for the measures that were installed in the ULIEEP dwellings?
- What are the measured gas savings derived from a PRISM[1] analysis of utility billing records for the ULIEEP dwellings?
- How closely do the predicted savings match the measured savings? What are the patterns of variation?
- To the extent that NEAT-predicted and measured savings differ, why do they differ?
- Are NEAT-predicted savings consistent with savings predictions from another engineering model, the DOE-2.1E, which is an industry standard? To the extent that they differ, why do they differ?

In short, the purposes of this study are to determine **realization rates** for NEAT and to identify reasons, if possible, for realization rates that differ from 100%. A realization rate is the ratio of measured savings to audit-predicted savings. It is usually expressed as a percentage. If the measured and predicted savings match exactly the realization rate would be equal to 100%. If savings are less than predictions, the realization rate is less than 100%, and if savings exceed predictions, the rate is more than 100%.

2.1 PREVIOUS RESEARCH ON AUDIT REALIZATION RATES

Previous research shows that the National Energy Audit (NEAT), like many other residential engineering energy audits, has typical realization rates well below 100% (Sharp, 1994; Dalhoff, 1996). This study performs an additional measurement of NEAT's realization rate, using a data set from the New York gas utility's low-income weatherization program. The realization rate is estimated by comparing NEAT predictions for 49 New York houses to the results of a billing analysis for the same 49 houses. Using this data set, we also explored the reasons for any observed discrepancies between the NEAT-predicted and measured savings (Section 3).

[1] PRISM, or the Princeton Scorekeeping Method, is a procedure (with standard software) that uses the consumption data in utility billing records along with weather data to estimate a weather-corrected index of consumption. PRISM also calculates weather-adjusted savings by comparing pre-and post-weatherization consumption, and estimates several indicators of energy use patterns including Normalized Annual Consumption (NAC), Normalized Annual Heating Consumption (NAHC), baseload energy use, and the balance point, or reference temperature, of the house. See Fels, Kissock, Marean, and Reynolds, 1995.

A second approach to NEAT validation used in this study is comparisons of NEAT predictions of energy consumption and savings with the predictions of the DOE-2.1E model, which is an industry standard. Findings from the DOE-2.1E model validation are presented in Section 4. The DOE-2.1E approach to NEAT validation was expected to help identify any deficiencies in NEAT's internal algorithms and assumptions, while the study of NEAT's realization rates was expected to reflect a variety of possible causes for the gaps between predicted and measured savings, including the influence of field implementation errors and client behaviors. Our hope was that this two-fold approach to NEAT validation would suggest ways of modifying NEAT, and of improving its implementation, that would move its realization rate closer to 100%.

In general, audit-predicted savings based on engineering algorithms tend to produce estimates of energy savings that are higher than those produced by the analysis of metered fuel consumption data (Nadel and Keating, 1991). In their review of 42 impact evaluations conducted by electric utilities on a variety of types of conservation programs, Nadel and Keating (1991) found that engineering estimates of savings were typically higher than impact evaluation estimates (based on metered data). Among the eleven residential retrofit program evaluations they examined, for example, engineering estimates were significantly higher then measured savings estimates in all but two cases. In eight of the eleven cases the realization rates were less than 50%. A more recent example of the tendency of engineering audits to estimate higher savings than impact evaluations can be found in a study prepared by Barakat and Chamberlin on the Consolidated Edison Utility Low-Income Energy-Efficiency Program (ULIEEP). In their evaluation of Consolidated Edison's ULIEEP, Barakat and Chamberlin, Inc. (1995) found realization rates of 55% to 63%.

Another study of realization rates in the Demand-Side Management programs of the four largest California utilities summarized the results of over 50 evaluation studies of programs operating between 1990 and 1992 (Brown and Mihlmester, 1994). The programs reviewed included efforts in several end-use sectors (including commercial, industrial, agricultural and residential). Some of the programs offered direct assistance (measures are installed at no cost to the customer) and some shared savings incentives (where both the utility and the customer invest in measures and share the benefits of reduced bills). Brown and Mihlmester (1994) found that the lowest realization rates were in residential direct assistance programs (with a mean of 68% and a median of 53%). The realization rates for commercial, industrial, and agricultural programs were higher than residential realization rates, and the shared incentive programs had higher realization rates than the direct assistance programs.

Although realization rates are typically less than 100%, Nadel and Keating (1991) report rates as high as 117% and as low as 15%. The mean realization rate in the eleven residential retrofit validation studies they reviewed was 44%, and the median was 40%. Previous studies of NEAT reported realization rates of 57% (Sharp, 1994) and 54% (Dalhoff, 1996).

As Nadel and Keating (1991) point out, many factors affect the relationship between engineering estimates of savings and measurements of savings in impact evaluations. In several studies that they reviewed the discrepancies between predicted and realized savings were attributed to post-retrofit decreases in the use of secondary fuels. For example, houses using electricity as a primary heating fuel and wood as a secondary fuel, began using much less wood after weatherization. Because significantly less wood was used after retrofit, the measured decreases in electricity consumption were small, much smaller than had been predicted. In other validation studies, discrepancies were attributed to inaccurate assumptions used in the engineering calculations, and to performing calculations on prototype buildings without considering how much energy the occupants of the actual buildings were using. Quality control problems in measure installations were another identified cause of low realization rates. Brown and Mihlmester (1994) indicated that the factors associated with higher realization rates in the residential sector included calibrated engineering models, relatively low predicted savings, and evaluations that fail to discount savings for free riders, takeback effects, and poor measure retention.

2.2 METHODOLOGICAL ISSUES IN THE DETERMINATION OF REALIZATION RATES

Decisions made by the analyst during the calculation of realization rates will strongly influence outcomes. The issue of how to deal with outliers, for example, can strongly influence results (especially when samples are small). The previous effort by ORNL to determine realization rates for an earlier version of NEAT in a sample of 18 North Carolina homes illustrates the strong effect the treatment of outliers can have on results. In the North Carolina study, the average audit-predicted space-heating savings were reported as 24.4 Mbtu as compared to an average savings estimate based on metered consumption data of 13.9 Mbtu. These findings produce the reported realization rate for the audit of 57% (Sharp, 1994). However, some examination of the data presented in the report shows that much of the gap between predicted and measured savings is accounted for by two or three houses. If two houses, one with very low measured savings and one with a post-retrofit increase in consumption, are dropped from the sample, the realization rate increases to 75%. With the elimination of a third house, the realization rate rises to 81%. In our analysis of realization rates, therefore, particular attention will be given to data quality and to the treatment of outliers. In addition, results will be reported and discussed for several subsets of the data, which have different levels of data quality and exclusion of outliers.

It is difficult to determine a realization rate with precision because the results are influenced by so many possible sources of error. In his review of efforts to validate several building energy analysis simulation programs (including DOE2.1C, BLAST, and SERIRES), Judkoff (1988) points out that a typical building energy simulation program contains hundreds of variables and parameters. Because of their complexity, it is impossible to test all of the possible combinations in the model. As a result, each

comparison between predicted and measured building energy performance represents only one data point in an immense multi-dimensional space. There are many possible levels of model validation, depending upon how much control over potential sources of error one implements. The simplest form of validation, Judkoff (1988) states, is the comparison of actual building energy use to model predictions with no attempt to control sources of error. Even if good agreement is obtained in such a comparison, he argues, it is difficult to interpret the results because all of the possible sources of error are operating simultaneously and may be offsetting in their effects. Judkoff (1988) identifies several types of errors that he divides into two groups as follows:

External error types include:

- differences between actual weather at the building site and weather input to the model;
- differences between actual occupant behavior and the behavioral patterns assumed by the model;
- user error in obtaining and recording the data on building characteristics, and its physical and thermal properties that are asked for in the model's input files.

Internal error types include:

- differences between the actual thermal transfer mechanisms taking place in a building and the simplified simulation of these mechanisms;
- errors or inaccuracies in the numerical solution of models;
- coding errors.

He then states that the more informative types of validation will control or eliminate various combinations of these error types. He continues with a discussion of the types of data and analysis needed to achieve control of various types of errors (Judkoff, 1988).

Table 2.1 summarizes the potential sources of error in this study's effort to determine NEAT's realization rates and lists the types of corrective action that were taken and that could be taken to deal with these errors. Table 2.1 includes the error sources identified by Nadel and Keating (1991) and by Judkoff (1988) as well as some additional types of potential error that the authors identified. In the column showing possible corrective actions in Table 2.1, the actions that were actually implemented in this study are highlighted in bold type.

While correcting obvious audit input errors (related to the building description) almost always seemed justified, it was difficult to decide how much adjustment of billing data, how much exclusion of outlying billing data points, and how much elimination of houses with extreme consumption values was justifiable. In Sections 2.3 and 2.4 the types of errors found in our databases are described in more detail and the corrective actions which we chose to deal with these error types are discussed.

Table 2.1. Potential Sources of Error and Possible Corrective Actions in the Determination of Realization Rates

POTENTIAL SOURCE OF ERROR	POSSIBLE CORRECTIVE ACTIONS[a]
AUDIT PREDICTIONS EXTERNAL	
Errors in Input Data • weather input data versus actual weather • user error in obtaining and/or recording data on building characteristics	**Check inputs for consistency and reasonableness and correct where feasible.** **Weather normalize the consumption data.** Reinspect houses to verify measurements.
Audit Is Applied to Protoype Buildings Without Considering How Much Energy the Actual Building is Using	**Apply audit only to actual buildings and adjust for the amount of consumption reported in the billing data**
Missing Input Data	**Assign average/default values.** Restrict sample to houses with complete input data.
INTERNAL	
Inaccurate assumptions about thermostat setpoints; R-value of wall, floor, and attic surfaces; percentage of dwelling that is heated, occupant behavior, internal gains, etc.	**Collect data that can be used to check the accuracy of the internal assumptions and to modify any that are not correct.**[b]
Differences between actual physical mechanisms (such as thermal transfer mechanisms) taking place in the house and the simplified model simulation of the mechanisms.	**Validation with other models (i.e., DOE2.1E)** Validation with field studies of the assumed physical processes operating in the house.
Internal algorithms not correct • errors in numerical solution of the models • errors in coding	**Validation with other models (i.e., DOE2.1E)** **Validation through tests of ability to produce consistent and reproducible results.**
IMPLEMENTATION PRACTICES	
All recommended measures not installed	**Switch off all measures not installed before running the audit.** Subtract estimated savings for the measures not installed. Restrict the sample to only those houses where recommended and installed measures match exactly.

[a] The corrective actions that were actually implemented in this study of NEAT realization rates are highlighted in bold type.
[b] Survey data were available to check some, but not all, of the assumptions in NEAT. See Section 2.3.4.1

POTENTIAL SOURCE OF ERROR	POSSIBLE CORRECTIVE ACTIONS
All installed measures not recommended	**Force audit to estimate savings for all installed measures** Add in estimated savings (from another source) for measures installed but not recommended by the audit. Restrict the sample to only those houses where recommended and installed measures match exactly.
Poor quality installations	Inspect installations and restrict sample to those houses with high quality installations
BILLING ANALYSIS	
Missing or incomplete data	**Restrict sample to houses with complete 12 months of data for both the pre- and the post-weatherization years.**
Billing data do not track weather data	**Apply PRISM diagnostic results to eliminate poorly fitted models.**
Outliers in input data points	**Use robust PRISM which assigns lower weights to outlying values.** **Adjust the outlying points so that they fall on the regression line of consumption versus heating degree days.** Restrict the sample to only those houses without extreme outlying points in the billing data.
Houses with negative savings	**Remove houses with negative savings from the sample.**
OCCUPANT BEHAVIORS THAT IMPACT CONSUMPTION	
Occupancy changes in the dwelling	**Restrict sample to houses without occupancy changes**
Changes in household's demographic composition (numbers, ages)	Conduct pre- and post-weatherization surveys to identify changes and restrict sample to houses without changes, or statistically adjust for changes
Changes in appliance stock, usage habits	Conduct pre- and post-weatherization surveys to identify changes and restrict sample to houses without changes, or statistically adjust for changes
Fuel switching (including post-weatherization use changes in use of secondary fuels)	Conduct pre- and post-weatherization surveys to identify changes. **Restrict sample to houses without fuel use changes.**

2.3 AUDIT-PREDICTED SAVINGS

2.3.1 The Utility Database

The utility database included measure-specific predicted savings for each house. These predictions were taken from NEAT output files for the eight measures (attic, wall, basement, crawlspace, rim joist, and duct insulation; heating system replacements, and air infiltration reduction) that the utility auditors evaluated with NEAT. The rest of the measures that NEAT can evaluate were eliminated, by switching them off in the set-up files, before the ULIEEP auditors ran NEAT. Using the utility database we defined total NEAT-predicted savings as the sum of the predicted savings for the set of NEAT measures that were recommended and installed in each house. For the most part, all NEAT-recommended measures for a house were installed in that house. The ULIEEP auditors were instructed to install all of the NEAT-recommended measures, and generally did so unless some physical barrier prevented an installation.

2.3.2 Ensuring the Accuracy of NEAT Predictions

A major purpose of this study was to gather information that might be useful for improving the performance of NEAT. Therefore, it was essential to obtain the most accurate predictions that NEAT is capable of producing. To obtain the most accurate possible predictions, one must understand and control all potential sources of error that are external to the NEAT engine as completely as possible. Accordingly, a thorough check of all inputs and the strictest possible control of all other external sources of error were necessary. The first author of this paper, who is the developer of NEAT as well as an experienced provider of training in its use, is especially well qualified to identify and correct potential sources of error in its implementation.

To achieve the desired level of understanding and control over external sources of error, we first attempted to replicate the NEAT results reported in the utility database by using identical input and parameter files. In this way, we hoped to understand the process used by the auditors in complete detail. Unfortunately, our efforts to replicate exactly the results shown in the utility database by rerunning NEAT on the same houses failed. Although the same measures were almost always recommended, the predicted savings for each could not always be reproduced. Because several years had passed since the ULIEEP audit runs were completed, the available information on possible reasons for the discrepancies was limited. Finally, because we were unable, even after extensive efforts, to produce identical results, the decision was made to use the results of the carefully controlled and documented ORNL NEAT runs.

In the ORNL runs, the NEAT input files (containing information on dwelling-specific measurements and characteristics) and parameter files (containing fuel prices, fuel escalation rates, measure lifetimes, materials costs, weather data, and discount rates)

were used as they were created by the utility auditors except that any obvious errors were corrected, if possible. Some examples of obvious errors that could be corrected included:

- window dimensions that were entered in the wrong units were changed to the correct units;
- billing data that were entered for the wrong house were changed to data for the correct house; and
- an efficiency rating for an existing furnace, given as 96%, was changed to an average value, because no efficiency measurement had been performed and the recorded value was unreasonably high.

A more detailed discussion of the process of correcting input errors is contained in Section 2.3.3.2.

2.3.3 Controlling External Sources of Error

In the ORNL NEAT runs, we attempted to consider, and where possible to remove, the potential sources of error that were external to the algorithms used within the NEAT computer program itself. There are many potential external sources of error including

- poor implementation practices,
- inaccurate input measurements, and
- uncontrolled behavioral changes.

Each source of external error is described below and the methods used to deal with it are discussed.

2.3.3.1 Poor implementation practices. A complete understanding of and strict control of implementation practices is necessary if one is to measure the accuracy of audit predictions with any precision. Three possible implementation-related reasons for discrepancies between audit-predicted and measured savings are:

- all of the measures with predicted savings are not actually installed;
- all of the installed measures are not included in the set of measures with predicted savings; and
- the characteristics of installations are not as is assumed by the audit model.

To reduce the first two sources of error the utility database was checked to see which measures were reported as having been installed in each house. Then, the NEAT set-up files were modified to force NEAT to estimate savings for exactly the set of measures that
were reported to have been installed in each dwelling. All other measures were switched off in the set-up files before performing a NEAT run for each dwelling. In addition, NEAT was forced to estimate savings for every measure that was reported to have been

installed by setting the required Savings-to-Investment Ratio (SIR) to zero. In practice, these modifications to the set-up procedures did not produce much change in the output

files. Switching on only the measures reported to have been installed did not usually change the set of recommended measures with predicted savings because the utility consistently tried to install all of the measures that NEAT had recommended in their audit runs. Changing the required SIR from 1.0 to zero, also did not change the pattern of recommended measures a great deal, because the preferred order of installation, from highest to lowest SIR, remained in the same sequence. A test was conducted to see if the NEAT set-up configuration described above would produce lists of recommended measures similar to those produced without such modifications to the set-up files. The test confirmed that the lists of recommended measures were almost always identical.

A final implementation issue concerns the quality of installations. Engineering estimates normally assume perfect installation of the measures, whereas reality may not reflect such perfection. For example, voids in wall insulation may occur, not simply because of worker error, but because of building construction details which make it difficult or impossible to insulate the entire envelope. Such abnormalities will produce less savings than predicted by the engineering estimates. No data were available on installation quality.

2.3.3.2 Inaccurate and questionable inputs. Simple audit procedures assume average conditions for many of the factors that affect energy use. Advanced audits, in contrast, require that measurements on many of the specific conditions present in each individual dwelling be used as inputs to a model. NEAT input screens, for example, ask for measurements of many dwelling-specific characteristics, such as the areas and types of windows, walls, doors, foundations, and attics. NEAT also asks for input information on heating systems and air conditioning equipment efficiencies. Information on air leakage can be entered into NEAT if blower door readings are available, or default values may be used. The completeness and accuracy with which input data are measured and recorded will affect the correctness of NEAT predictions.

Two inputs to NEAT that are optional, but which, when employed, should improve the accuracy of its savings predictions are actual pre-retrofit fuel consumption records and blower door test results from before and after the completion of air sealing and infiltration work. When actual pre-retrofit consumption records are included as part of the input data, NEAT will adjust predicted savings to fit with the observed consumption. If pre- and post-weatherization blower door test results are entered, they will be used to compute predicted energy savings due to air infiltration reduction more accurately. Both of these optional inputs were used in all of the ORNL NEAT runs to improve the accuracy of the NEAT predictions. A third optional input, that will improve accuracy if it is used, is the result of a heating system efficiency test. This test appears to have been performed in all except two of the sample dwellings. The efficiency test results were entered where available and default values were used for the rest of the dwellings.

Because of the time and expense, which would be involved, it was not feasible to directly check the accuracy of the input measurements for the ULIEEP dwellings by reexamining or remeasuring the dwellings. The input data were checked, however, for their reasonableness and consistency. Building descriptions that were not completely consistent, reasonable, or understandable were found in almost one-third of the cases. In a number of dwellings, which were single-family units, for example, no attic and/or no foundation space was described. The wall areas, floor areas, foundation areas, attic areas and number of floors were described in a variety of inconsistent ways in many houses. Some of the inconsistent descriptions included opposing walls of different dimensions, or attic and foundation areas with different sizes and shapes. Sometimes it was unclear what type of dwelling was being described and how many living units were in the building. Extremely high air leakage rates (i.e., greater then 10,000 CFM_{50})[2] were reported in four houses, and the infiltration rates after weatherization showed little change in these houses.

In a significant number of houses, it was apparent that, because of inconsistencies, some of the reported values had to be incorrect even though it was not possible to determine what the correct values should be. When a more correct value could be determined, it was substituted for the initial value to create a corrected input data file, and NEAT was run again on the corrected input file. Missing values for infiltration rates or furnace efficiency measurements were replaced with average values in the corrected input data files. Overall, problems related to inconsistent or missing building input data were fairly common. These building description input errors introduce an uncertain amount of error into the NEAT predictions of savings. Sometimes correction of input errors led to an increase in a dwelling's predicted savings, and sometimes it led to a decrease. Overall, predictions based on the corrected input files did not differ greatly from those based on the uncorrected input files. Predictions based on the uncorrected input files averaged 97.6 Mbtu per dwelling per year, while predictions based on the corrected input files averaged 104.7 Mbtu per dwelling. Realization rates calculated with the uncorrected and corrected input files also did not differ greatly.

2.3.3.3 Demographic and behavioral changes. Behavioral factors clearly account for much of the variation in residential energy use. Several studies have shown that consumption in physically identical residential units varies by as much as a factor of three because of the energy-related habits of the occupants (Socolow, 1978; Cramer, Miller, Craig, Hackett, Deitz, Vine, Levine, and Kowalczyk, 1985). Similarly, if the initial set of occupants moves out and is replaced by a new set of occupants during the study period, this change alone can cause dramatic changes in a dwelling's energy consumption. Most studies of the energy savings due to weatherization exclude dwellings that have occupancy changes from the sample. In this study, too, all dwellings with occupancy changes were excluded from the analysis.

---

[2] CFM_{50} is the number of cubic feet per minute of air flowing through the fan housing of a blower door when the house pressure is 50 pascals (0.2 inches of water). This figure is the usual way of comparing the airtightness of buildings that are tested with a blower door.

Many changes in energy usage occur without any program intervention. The addition, removal, or replacement of energy-consuming equipment (such as a furnace, air conditioner, water heater, or refrigerator) can significantly impact consumption. Changes in the number or ages of people in the home, in the amount of time they spend at home, and in their habits and activities can all strongly influence consumption. Structural changes, initiated by the occupants, such as adding rooms, or replacing windows, also can affect usage. Fuel price changes, or fuel shortages, may change consumption patterns as well.

Many evaluations use a control group to account for behavioral influences on energy use. With this approach, it is assumed that the control group represents what the treatment group would have done in the absence of the program intervention. Some evaluations collect survey data, before and after weatherization, on equipment holdings, demographics, and energy-related behaviors. This survey information is then incorporated into statistical models and used to adjust estimates of program-induced savings. This study has no control group. There was limited survey information that included data on reported thermostat setpoint, conditioned and unconditioned floor areas, hot water use, number of occupants, and the fuel used for cooking and hot water heating. These survey data were insufficient to control fully for behavioral effects. Nevertheless, it is certain that such effects are present in the energy consumption data that are being analyzed.

It is usual in an evaluation of energy savings to see at least a small percentage of treatment houses that increase their consumption after weatherization. Four houses in this study increased their consumption after weatherization. Three of these four had several substantial insulation measures installed. Such increases in consumption occur when behavioral factors overwhelm the effects of the energy-efficiency improvements.

In an audit validation study, it is unclear how best to deal with dwellings that show increases in consumption. On the one hand, houses that show increases in consumption could be considered outliers that should be removed because, from an engineering perspective, they are not useful for determining how well the audit predicts savings. On the other hand, in field settings, one should always expect to see some treatment houses with real consumption increases because behavioral effects are pervasive and significant. In part, how to deal with such dwellings depends upon a study's purpose. If the primary question is to determine how well audit predictions can be used to represent what the results of an impact evaluation designed to measure program-induced effects would be, then these houses must be included. If the primary question is to determine how to improve the audit, then examining such dwellings would not be useful. Our approach will be to examine realization rates both with and without houses with extreme values in the billing analyses.

2.3.4 Identifying Potential Internal Sources of Error in NEAT

In addition to the potential external sources of audit prediction error discussed above there are also several potential internal sources of error. Identifying, and correcting, any internal sources of error in NEAT were major goals of this study. The two main potential internal sources of error are inaccuracies in NEAT assumptions and inaccuracies in its engineering algorithms. These two potential sources of error and the methods used for dealing with them are discussed below.

2.3.4.1 Inaccurate assumptions. All audits must make some assumptions about average conditions. It is too burdensome, if not impossible, to measure every factor that might influence energy use. Using assumptions about average conditions, will, of course, lead to predictions that will differ widely from measurements of energy use in some specific households. Unusual occupant habits, behavior, or lifestyles may have large impacts on energy use that are not accounted for when an audit assumes average conditions, such as thermostat settings of between 65° and 75°, or an average number of occupants. The issue that must be addressed is not whether using assumptions about average conditions will sometimes over- or underestimate savings. No set of assumptions, no matter how well they represent average conditions, will eliminate such random errors. When assumptions appear to consistently produce over- or under- estimates of savings, however, this suggests that they introduce significant biases that should be corrected.

Some of the assumptions made in NEAT about average conditions include the following:

- Thermostat setpoint of 68^0 F.
- Average internal heat gain of 2900 Btu/hour representing the sum of heat output from a refrigerator, range, lighting, television, and water heater, as prescribed in the ASHRAE Handbook: Fundamentals (1993). The minimum value for the water heater was used, since this appliance is often not in the conditioned space[3]
- An average occupancy of two adults. This may have differed from the assumption used in the original runs performed by the New York utility, where heat output from additional occupants is added if the living space floor area was over 1000 ft^2.
- Infiltration levels as indicated by blower-door measurements. No additional leakage is attributed to window and door openings.
- Except for its insulation, HVAC ductwork was assumed to have been examined and repaired prior to execution of the audit.

For some of these assumptions, survey data could be used to check their accuracy. For example, thermostat settings reported by the respondents ranged from 65^0 F to 90^0 F with an average of 70.5^0 and two not specified. If the setpoint indicated in the utility's client

[3] For several houses for which the utility's database indicated substantially greater hot water, electric, and/or gas use (for other than HVAC), an additional NEAT run was performed with increased internal heat gains.

survey database indicated a setpoint that varied by more than 2^0 F from the assumed value of 68^0 F, an additional NEAT run was performed with the database value.

The average number of occupants for the houses used in the study actually averaged 3.8, compared to the assumed number of two. However, the heat given off by occupants is actually minor compared to the increase in appliance energy use that may result from their presence. No survey data specifying the average Btu/hour internal heat gain from appliances were available. However, the data do specify the average hot water use, the number of baths and showers, and the average number of clothes dryer loads in each household. These could be used as indicators for when the internal heat gain from appliances, etc., might be higher than average. In two instances, additional NEAT runs were performed where the internal gains were increased.

The assumption that the HVAC ductwork had been repaired (large leaks patched) prior to execution of the audit is likely not accurate. If any such work was performed, it was probably performed at the time of other measure installation. Note, however, that the pre-retrofit billing data would also have reflected any major duct leakage, giving consumptions greater than NEAT would predict. Comparison of pre-retrofit NEAT predicted and billed consumptions do not indicate any such significant effect. Assuming that duct leakage repair has been performed prior to application of the NEAT audit is, however, a weakness in the audit which will be addressed in future versions.

Like most engineering audits, NEAT essentially models an empty house and assumes that walls and floors are bare. In reality, homeowners often have a portion of their walls or floors covered by furniture, shelving, cupboards, drapes, decorations, storage (particularly in attics or basements), etc. Some of these items may add resistance to heat flow through these surfaces which is not accounted for in the engineering calculations. The higher the pre-retrofit R-value of a building component the less actual savings are obtained from adding more insulation to the component. This is particularly true when the initial value is low. Thus, actual savings from insulating building components can be less than those predicted based on the assumption of empty houses.

Although the empty house assumption probably has the effect of lowering NEAT realization rates to some extent, we do not consider it worthwhile to modify NEAT to reduce this effect. Incorporating the fact that household items may add resistance to heat flow would be too difficult to measure, quantify, and model. In addition, such an effort probably would produce only a small improvement in accuracy.

2.3.4.2 Inaccurate internal algorithms. Another possible internal source of error could be inaccuracies in the assumptions and calculation routines built into an advanced audit's computer codes. Given all of the possible sources of error, identifying which algorithms may be inaccurate through comparisons with billing data alone is impossible. A more effective approach to identifying possible inaccuracies in an audit's internal codes, is to use identical inputs in two or more engineering models. If the model outputs differ significantly, the analyst can then explore how their assumptions and algorithms differ,

and try to determine which of them seems to be more correct. The findings from a detailed comparison and analysis of the NEAT and DOE-2.1E model results obtained for the same two houses are presented in Section 4.

2.4 SOURCES OF ERROR IN MEASURED SAVINGS ESTIMATED FROM BILLING ANALYSES

Additional potential sources of input error, which also may affect the determination of realization rates, can be found in billing data. In this study, in order to reduce error in PRISM weather-normalized consumption estimates, only dwellings with at least twelve months of billing data for both the before and after weatherization time periods were retained.

Efforts were made to check the accuracy of the billing data and to correct anomalous values. Advanced PRISM was used to correct for estimated values and to adjust outlying data points by applying the robust PRISM routine which assigns lower weights to anomalous points (Fels, Kissock, Marean and Reynolds, 1995, pp. II-73 to II-75). In addition, individual plots of consumption versus heating degree days were examined for possible errors. Most of the billing data was reasonably well behaved. In five dwellings, however, extreme outliers among the monthly readings were seen. These outlying data points were changed by replacing them with values that fell on the regression line for consumption versus heating degree days. These points were adjusted, not because we believed the meter readings were erroneous, but because these real, but extreme, variations in consumption would not represent situations (such as a long absence from the residence) that would occur in both the pre- and post-retrofit billing periods.

2.5 RATINGS OF DATA QUALITY

The process of checking NEAT input file and billing data quality made it apparent that some houses had more consistent, reasonable, and understandable information than others. As was described in the sections above, efforts were made to improve the correctness and consistency of both data sets where possible. Nevertheless, a good deal of variation in data quality remained. Because data were available on only 49 houses, we did not wish to drop any of them from the sample. On the other hand, realization rates for individual houses varied from less than zero to more than 200%, and it seemed important to consider how realization rates would change as data quality improved or declined. Accordingly, a scheme for rating the overall quality of the data upon which a dwelling's realization rate would be calculated was developed.

Separate criteria were established to rank the confidence placed in the quality of the building input files for the NEAT runs and to rank the confidence placed in the quality of the metered consumption data that were the input to PRISM. The criteria used for the NEAT input files included whether or not the house layout could be determined

from the input data, and the number and significance of the inconsistencies or problems. Some examples of the types of data inconsistencies found include:

- NEAT building description input not totally consistent with the utility's data base information;
- the number of stories not reflected in the total floor, foundation, or attic areas specified;
- uncertainty in the number of living units being described by the NEAT building description data and how many of the units were included in the billing data;
- unexplained specification for the fraction of building heated;
- substantially different opposing wall or attic/foundation areas;
- no attic or foundation spaces specified despite only one living unit in building or all units in the building weatherized (if only one unit of a two or more unit building were weatherized, this would not be an inconsistency);
- no infiltration data specified;
- unusually high leakage rates (above 10,000 CFM_{50} with little improvement after weatherization);
- replacement furnace with minimal efficiency (63% AFUE) or existing furnace with unusually high efficiency (96%);
- unusually high thermostat setpoint indicated (90°F);
- window dimensions in wrong units.

For the billing data the quality ranking criteria were the R^2 from the PRISM run, the magnitude of difference between the estimated pre- and post- baseloads, the degree of scatter, and the amount of adjustment, if any, made to outlying data points. After the separate ratings for building data and metered data quality were completed, they were combined into a single index, which ranged from one to twelve. Using these combined ratings dwellings were ranked by overall data quality and realization rates were examined for various levels of data quality. The results are shown in Table 3.4 of Section 3.4.

3. RESULTS

3.1 PROGRAM TARGETED HIGH-USAGE HOUSEHOLDS

The New York gas utility program targeted high-usage households. As is shown in Table 3.1, these houses consumed an average of 266 Mbtu of natural gas in the pre-weatherization year. This is almost twice as much energy as the average low-income dwelling that participated in the New York Weatherization Assistance Program in 1990 consumed (New York State Energy Research and Development Authority and New York Department of State, 1993). In addition, the ULIEEP houses consumed much more energy than houses in previous NEAT validation studies. In the North Carolina study (Sharp, 1994), average pre-weatherization consumption was 50 Mbtu, reflecting the mild southern climate. In the Iowa study (Dalhoff, 1996), with a much colder climate, average pre-weatherization consumption was 131 Mbtu. The dwellings in the ULIEEP also had more heated square footage and were more energy inefficient than the households in the other groups (Table 3.1).

Table 3.1. Characteristics of the New York Gas Utility Program Dwellings as Compared to Characteristics of Houses in the New York Weatherization Assistance Program.

	New York Houses in Gas Utility Program	Average for Houses in New York Weatherization Assistance Program[a]
Average Heated Square Footage	1650	1370
Average Pre-Weatherization Consumption (in Mbtu)	266	134
Average Btu/Square Foot/ Heating Degree Day	25.2	16.5

[a]Source: New York State Energy Research and Development Authority and New York Department of State, 1993

3.2 ESTIMATES OF PRE- AND POST-WEATHERIZATION CONSUMPTION

The estimate of total natural gas consumption provided by PRISM that is adjusted to long-term weather conditions is called normalized annual consumption (NAC).

Although NAC is the key measure of fuel usage reported by PRISM, three other indicators of fuel usage also are calculated. These indicators are the normalized annual heating consumption (NAHC), the baseload energy consumption, and the balance point temperature, or reference temperature, of the house. The NAHC is the part of total gas consumption that fluctuates with changes in temperature. The baseload estimate is the part of gas consumption that stays constant regardless of temperature variations. The reference temperature estimates the outdoor temperature at which the heating system begins to operate. As a house becomes more energy-efficient its reference temperature decreases.

NAC is the most robust estimator produced by PRISM. NAHC and baseload estimators are less robust because energy consumption varies seasonally for many reasons, including changes in solar gain, groundwater temperatures, and occupant behavior. PRISM cannot distinguish among the factors causing seasonal fluctuations in usage. As a result, NAHC, baseload, and reference temperature are significantly influenced by all of the factors that vary by season, while these factors have little influence on NAC values.

Even though NAHC is a less robust estimator than NAC, it is conceptually more similar to NEAT estimates of consumption because NEAT models only heating consumption. Therefore, we will compare NEAT estimates of pre- and post- consumption to both the NAHC and the NAC values obtained from PRISM. In the calculation of realization rates, however, only NAC values are used. As Table 3.2 shows, for the New York gas utility program houses, the average pre-weatherization NAC value

Table 3.2. Comparison of PRISM NAC and NAHC Estimates to NEAT Predictions

	PRISM NAC	PRISM NAHC	NEAT Predicted Annual Consumption
Average Pre-Weatherization Consumption (in Mbtu)	266	217	254
Average Post-Weatherization Consumption (in Mbtu)	213	171	149

was 266 Mbtu and the average pre-weatherization NAHC value was 217 Mbtu, while the NEAT-estimate of average pre-weatherization consumption was 254 Mbtu. In his Iowa validation study, Dalhoff (1996) reported that NEAT tended to overpredict pre-weatherization consumption. His conclusion was based on a comparison of NAHC values to NEAT estimates. In our New York data set, NEAT exhibited a similar tendency to estimate larger pre-weatherization consumption as compared to the NAHC values

estimated by PRISM (Fig. 3.1). However, the NEAT estimates of pre-weatherization consumption tended to be slightly lower than the PRISM estimates of NAC values (Fig. 3.2).

In our New York data set, as in the Iowa study (Dalhoff, 1996), NEAT predictions of post-weatherization consumption were typically less than the NAHC estimates (Fig. 3.3). The average NAHC value for post consumption was 171 Mbtu, while the average NEAT prediction for post-weatherization consumption was 149 Mbtu. NAC estimates, at an average of 213 Mbtu, exceeded NEAT predictions (Fig. 3.4) by an even larger amount.

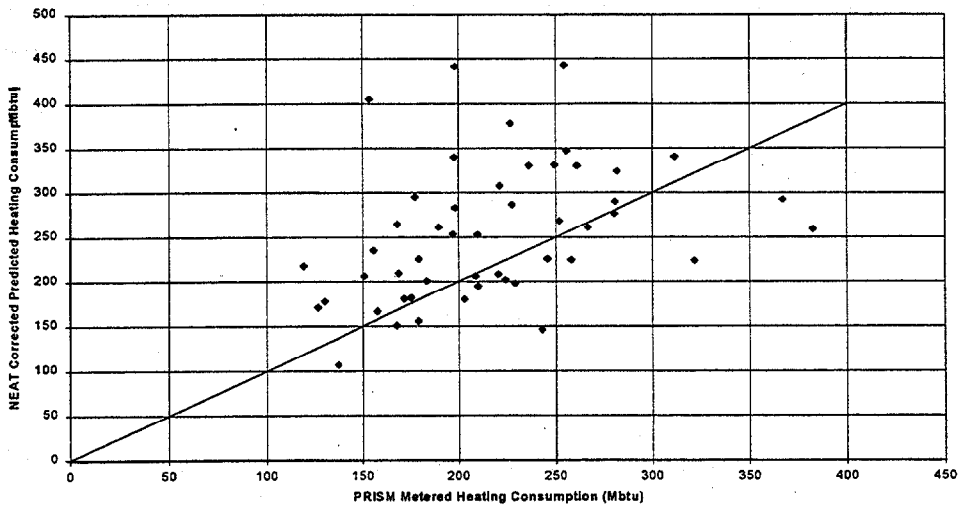

Figure 3.1. Comparison of Pre-Weatherization NAHC Values from PRISM to Pre-Weatherization Consumption Predicted by NEAT

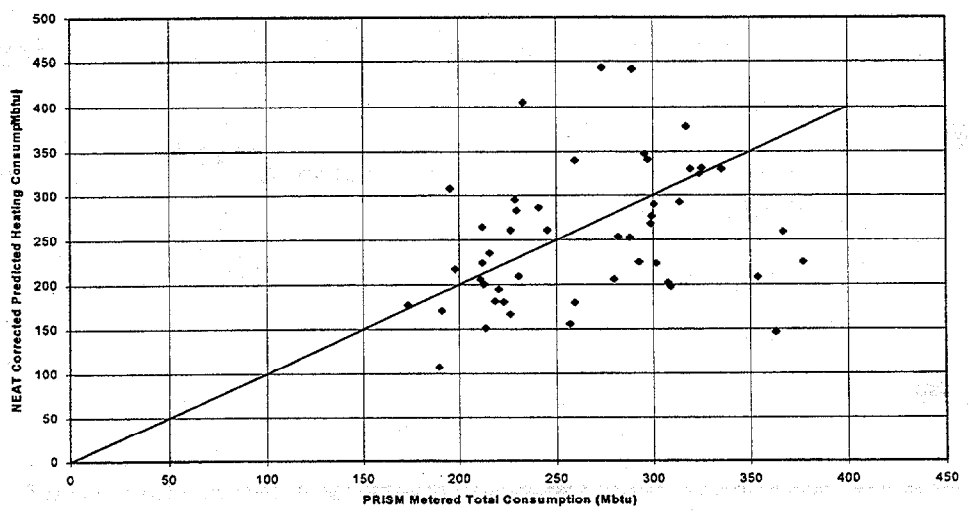

Figure 3.2. Comparison of Pre-Weatherization NAC Values from PRISM to Pre-Weatherization Consumption Predicted by NEAT

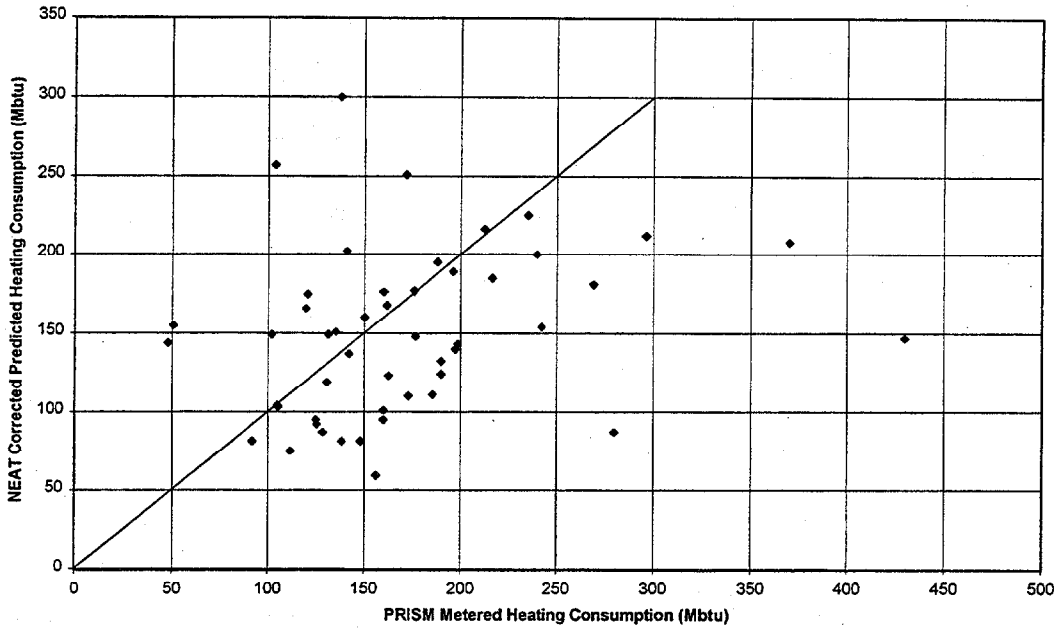

Figure 3.3. Comparison of Post-Weatherization NAHC Values from PRISM to Post-Weatherization Consumption Predicted by NEAT

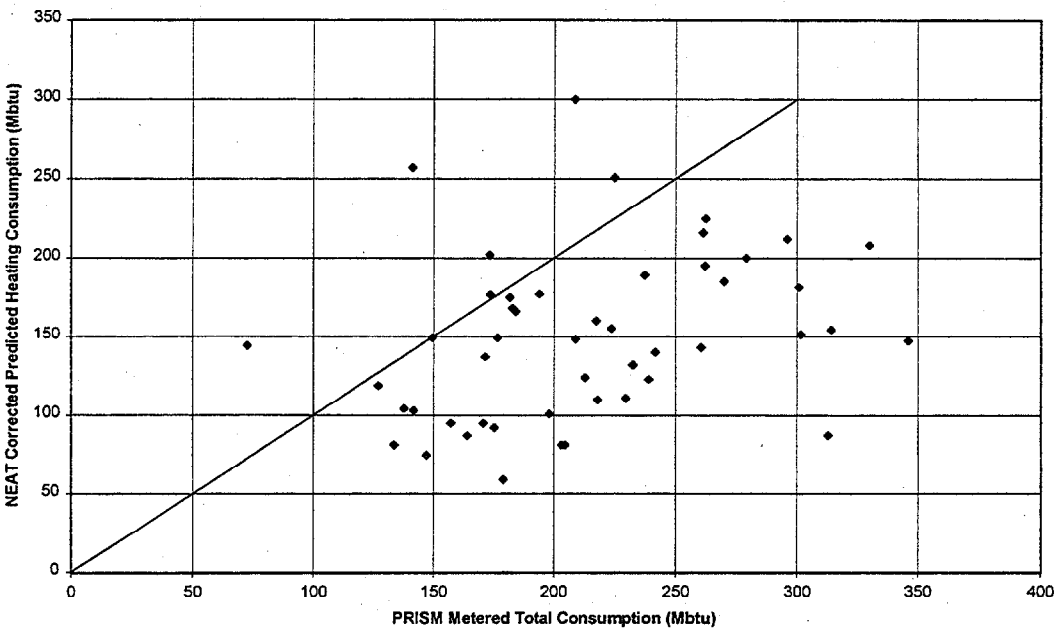

Figure 3.4. Comparison of Post-Weatherization NAC Values from PRISM to Post-Weatherization Consumption Predicted by NEAT

3.3 NEAT-PREDICTED SAVINGS

NEAT predictions of annual savings based on the corrected input files averaged 104.7 Mbtu per dwelling without any adjustment to billing data. Adjustment of NEAT predictions of savings to the amount of consumption shown in the billing data produced lower predicted savings, as is shown in Table 3.3. The realization rates both with and without billing data adjustment are shown in Table 3.4.

Table 3.3. NEAT Predictions of Savings With and Without Billing Data Adjustment

	With Billing Data Adjustment	Without Billing Data Adjustment
Average Predicted Annual Savings (in Mbtu)	88.1	104.7

3.4 REALIZATION RATES

As was explained in Section 2.5, the high variability in data quality suggested the need to develop a scheme for rating the overall quality of the data upon which a dwelling's realization rate would be calculated. Using this rating scheme, three groups of houses were defined: dwellings with high, medium, and low data quality. It was expected that the highest data quality group would have the highest realization rate. However, as Table 3.4 shows, the realization rates (with billing data adjustment) turned out to be about the same for the high data quality group (n=12) and for the total sample of 49 houses (69% and 66%, respectively). The average realization rate for the combined set of high and medium data quality houses is somewhat lower at 57%.

Because it was surprising it see little relationship between data quality and realization rates, a closer examination of the patterns of variation seemed warranted. When the standard deviations associated with the realization rates for the three groups of houses (dwellings with high, medium, and low data quality) presented in Table 3.4 were examined, they showed that there was no statistically significant relationship between data quality and realization rates. In other words, the standard deviations (which were as large as the mean values) for the three groups were so great that the differences among the group means were not significant.

The patterns of variation in the realization rates, when examined for smaller subsets of dwellings, showed that the lack of statistical association between realization rates and data quality occurred primarily because realization rates above 100% were concentrated in the houses with the poorest data quality. Among the ten houses with the lowest data quality, more than half of them had realization rates above 100%. Among the houses with high or medium data quality, in contrast, only about 15% of the houses had realization rates above 100%. Therefore, when the houses with poor data quality are

dropped, and only the medium and high data quality houses remain, the average realization rate decreases from 66% to 57%, while the standard deviations remain unchanged.

When both the poor and medium data quality houses are dropped, and only the twelve houses with the highest data quality are considered, the realization rate increases to 69%. Although the standard deviation for the twelve houses with high data quality was about the same as for the other two groups shown in Table 3.4, most of the variation occurs in the homes with lower data quality rankings within this group. In particular, among the 5 houses in ranks 8 through 12, realization rates vary from 2% to 216%. Among the seven houses with the best data quality (ranks 1 through 7), in contrast, the range is from 21% to 138%. Only one house among the top seven has a realization rate above 100%, and this group's standard deviation is comparatively low (less than 0.36).

Realization rates for several other subsets of dwellings were examined too. For example, if all of the eleven dwellings, out of the total sample of 49, that had realization rates above 100% are excluded, the average realization rate falls to 42%. If the eleven dwellings with the lowest realization rates are excluded (this subset included eight houses with realization rates of less than 25%) the group average rises to 83%. Finally, if the eleven houses from both the highest and the lowest part of the distribution are excluded, then the average realization rate becomes 54%.

Clearly, realization rates vary a great deal depending upon which houses the analyst decides to exclude from the sample. It is unclear which set of houses provides the best estimate of NEAT's realization rate. When only three groups were examined (dwellings with high, medium, and low data quality), there was no association between data quality and realization rates. A closer examination of smaller subsets of houses, suggested that higher data quality may be associated with lower standard deviations.

Table 3.4. Realization Rates Calculated with Data Sets with Varying Levels of Data Quality

Data Sets	Average Realization Rate With Billing Adjustment	Average Realization Rate Without Billing Adjustment
High Data Quality (n=12)	69%	58%
High and Medium Data Quality (n=31)	57%	51%
All houses- High, Medium and Poor Data Quality (n=49)	66%	59%

However, because so many houses with the poorest data quality had realization rates above 100%, these data do not demonstrate an association between high data quality and high realization rates. The fact that previous validation studies reported realization rates between 50% and 60%, suggests that our result of 57% for the high and medium data quality subsets of homes is most consistent with previous findings.

The overall distribution of PRISM estimates of savings (based on NAC values) versus NEAT predictions of savings is shown in Fig. 3.5. Clearly, NEAT tends to predict higher savings than are actually realized. This is the case even if the extreme values are eliminated from the plot shown in Fig. 3.5. In addition, correcting errors in NEAT inputs does little to change the pattern of overprediction (Fig. 3.6).

3.5 ANALYSIS OF REASONS FOR VARIATIONS IN REALIZATION RATES

Previous studies of NEAT realization rates found that the audit-predicted savings were more closely matched to actual savings in some types of houses than in others. Sharp (1994) concluded that for houses with pre-weatherization consumption of less than 80 Mbtu and audit-predicted savings of less than 40 Mbtu the NEAT predictions were, in most cases, very close to the measured values. On the other hand, in houses with pre-weatherization consumption of more than 80 Mbtu and audit-predicted savings of more than 40 Mbtu, NEAT-predicted consumption and savings were usually higher than measured values. Dalhoff (1996) reported that the difference between audit-predicted savings and measured savings in Iowa occurred both because of NEAT's tendency to overpredict pre-weatherization heating consumption by about 8% and to underpredict post-weatherization heating consumption by about 11%. Our New York results, however, showed only a slight tendency to overpredict pre-weatherization heating consumption combined with a strong tendency to underpredict post-weatherization heating consumption. The closer correspondence of the pre-weatherization NEAT predictions to the PRISM estimates in the New York program may be due to the targeting of high users.

In this study, several subgroups of dwellings were examined to see if certain subgroups had higher or lower realization rates than others. We found few significant relationships. Realization rates did not appear to be associated with the amount or type of insulation installed, nor were they correlated with pre-weatherization consumption levels or the pre-weatherization energy use index of Btu/square foot/heating degree day. Other factors we examined, which showed little or no relationship to realization rates, included the heated floor area, the number of occupants, and the thermostat setpoint. The one factor that appeared to have a strong relationship to the realization rates was the use of secondary heating fuels. Among the 41 houses with no use of secondary heating fuels, realization rates without billing adjustment averaged 63%. Among the eight houses that used a secondary heating fuel, the average realization rate was 36%.

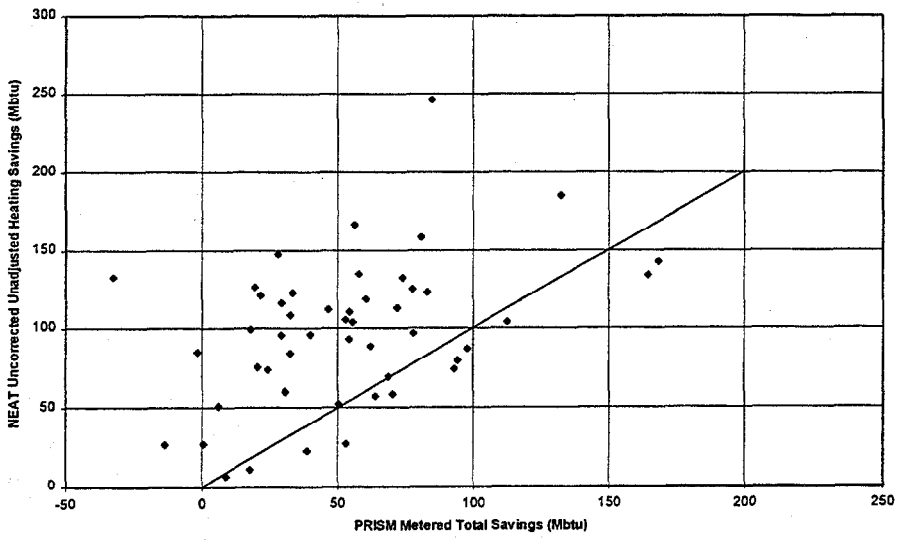

Figure 3.5. Comparison of Savings Estimated in PRISM Analysis to Savings Predicted by NEAT with Uncorrected Inputs

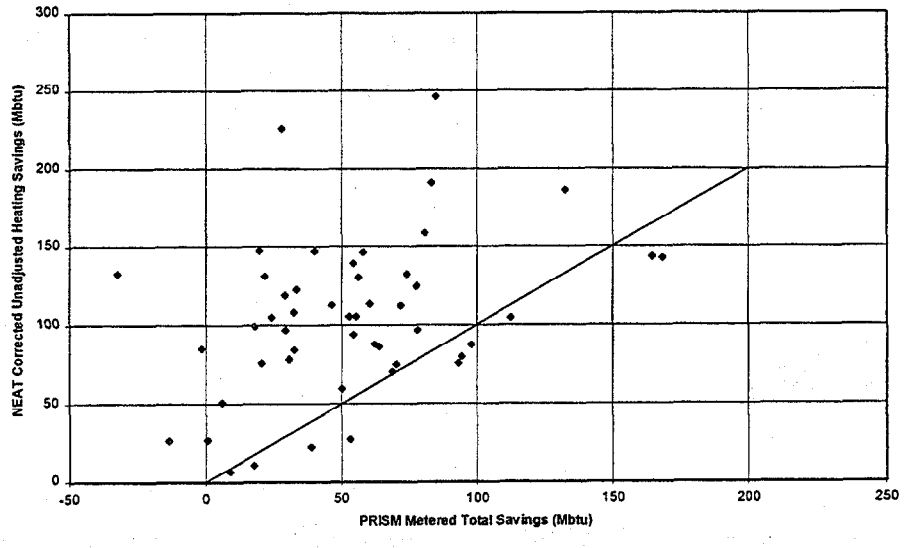

Figure 3.6. Comparison of Savings Estimated in PRISM Analysis to Savings Predicted by NEAT with Corrected Inputs

4. COMPARISONS OF NEAT VERSUS DOE-2.1E MODELING RESULTS

4.1 INTRODUCTION

To assess the accuracy of NEAT's internal algorithms its results were compared to those obtained with another model, DOE-2.1E. DOE-2.1E is an hourly energy use simulation model that is a more sophisticated and detail intensive tool than NEAT. DOE-2.1E, which is an industry standard, is designed for the skilled energy modeler/researcher rather than the residential energy auditor.

Using the same building description data, two houses were analyzed with both models. The primary goal was to assess the accuracy of NEAT's engineering algorithms by comparing their performance to the results obtained with the more detailed and sophisticated DOE-2.1E. The focus of the effort was on measures that affect heating systems and building envelopes because they are more complex, and perhaps more susceptible to error, than water-heating and other non-HVAC measures. A secondary goal was to assess how well each model's predictions matched the actual savings measured with billing data.

The three major steps involved in the assessment were as follows:

- First, the baseline, or pre-retrofit energy consumption was simulated with the two models. These results were then compared to each other, and to the pre-weatherization consumption measured with billing data (Section 4.2).
- Second, the post-retrofit consumption effects of individual measures and of all measures combined were simulated for the two houses with the two models. No billing analysis took place in this step because it is not possible to analyze the effects of individual measures with billing data (Section 4.3).
- Third, the total predicted savings for the two houses with the two models were compared to each other, and to the savings estimates obtained from an analysis of billing data (Section 4.4).

4.2 SIMULATIONS OF PRE-RETROFIT CONSUMPTION

The pre-retrofit consumptions of the two test houses were simulated using NEAT 5.2 and DOE-2.1E. The computed monthly energy use profiles obtained with the two models were compared to each other, and to the actual monthly use recorded in the billing data. Section 4.2.1 presents the methodology used and discusses the choices the analysts made in setting up the models. Section 4.2.2 presents the results of the comparisons of model results and billing data.

4.2.1 Set-Up Methodology

This section describes the set-up procedures that were used to model the pre-retrofit consumption of the two houses. The subsections below cover building selection, weather normalization, building variables, billing reconciliation, and output characteristics.

4.2.1.1 Building selection. Two houses were selected that had eight different HVAC-related measures installed. Each building had at least 18 months of pre-retrofit billing history data and seven months of post-retrofit billing data that included all winter months.

One of the houses is a 2200 ft^2, two-story frame home with masonite siding. Pre-weatherized insulation levels are listed as R-5 for the walls with 8" of fiberglass batts in the attic. Windows are mostly single-pane with wood frames. Some have existing metal storms. The leakage rate measured with a blower-door prior to weatherization is about 3500 CFM_{50}. Beneath the house is a partial basement and crawlspace, the majority below-grade. The heating system is a natural gas, forced-air furnace located in the basement.

The second house is a 1900 ft^2, two-story frame house with aluminum siding. Pre-weatherized insulation levels are listed as R-19 in the majority of the attic and 70% of the walls are insulated with blown cellulose. Windows are mostly single-pane with wood frames. Some have metal storms. The pre-weatherized leakage rate is high at 7553 CFM_{50}. This house has a below-grade basement where a forced-air, natural gas furnace is located.

4.2.1.2 Weather normalization. Each building's pre-retrofit energy use was analyzed separately based on billing data, NEAT, and DOE-2.1E. The long-run weather data used for NEAT and DOE-2.1E were slightly different from each other. The 1991-1993 weather matching the billing data differed significantly from both long run histories not only because of year-to-year variations but also because meter read dates vary from month to month. In order to normalize the results for comparison among NEAT, DOE-2.1E, and bills, energy use was expressed as a relationship between HDD_{65}/day and daily average gas use (ccf/day) for the month. An additional benefit of this approach is that billing analysis can be performed and annualized without a full year of data.

The relationship between HDD_{65}/day and ccf/day is approximately proportional for residential structures. Non-linearities are due primarily to seasonal variations in infiltration and solar gain, and variation in the waste heat of the pilot light. These non-linearities tend to make the relationship slightly concave upward rather than perfectly linear.

The slope of the least squares curve fit of ccf/HDD_{65} defines the building's characteristic heating energy use ratio. This ratio of ccf/ HDD_{65} is what we used to compare energy use between NEAT, DOE-2, and actual consumption.

4.2.1.3 Building variables. The general approach was to standardize the input parameters between NEAT and DOE-2 so that performance variations would be due to internal methodological differences rather than to input data differences. For NEAT simulation, the model was left with all set-up parameters in their original condition as installed. For DOE-2, layers were created to match NEAT's constructions, as they are described in the North Carolina Field Test Computerized Audit: Description and Operation (Gettings, 1990). Exterior walls, attic, window, door, and basement ceiling DOE-2 layers were all based on these North Carolina Field Test Report descriptions. Unlike NEAT, DOE-2 has the flexibility to model multiple zone buildings. This feature was utilized by modeling sub-spaces as separate zones from the conditioned space. Underground walls were assigned effective U-values based on data from non-NEAT sources (McQuiston and Parker, 1988, pp. 161-167; 1993 ASHRAE Handbook: Fundamentals, 1993, pp. 25.10-25.12). For other variables where corresponding NEAT data were not available, such as wall color and surface characteristics, reasonable assumptions were used based on DOE-2 reference tables and intuition. For the remaining DOE-2 variables, DOE-2 was allowed to determine it's own default value.

There was one modeling element that DOE-2 offers and NEAT does not that was disabled: heating system part load performance characteristics. The decision to disable this feature was driven by the fact that seasonal average efficiency data were provided. An additional consideration was that heating system capacity often is poorly estimated in the field due to lack of data and/or system degradation. Estimated capacity significantly affects part load calculations. Because this feature was not used, DOE-2 modeled savings should slightly overstate savings for most measures.

4.2.1.4 Bill reconciliation. Because this whole exercise tests NEAT's and DOE-2's ability to estimate energy use compared to historical use, no reconciliation to bills was performed on either model.

4.2.1.5 Output. Both NEAT (the monthly output debug version) and DOE-2 produce design condition component performance parameters and whole building estimated monthly energy use. Because billing data reflects monthly use, these data were the focus for quantitative comparisons. Comparisons of NEAT's design UA values with DOE-2's design heat loss were also performed between the base models.

4.2.2 Pre-Retrofit Consumption Results

The NEAT and DOE-2 pre-retrofit models produced similar results for both buildings based on monthly energy use, as shown in Figure 4.1 and Figure 4.2. **Note that these graphs do not reflect any weather normalization.**

Plotted against the HDD/day and ccf/day axes, the slopes of the billing, NEAT, and DOE-2 curves reflect weather-normalized sensitivity to temperature. The slopes do not vary from one another by more than 10%, as shown in Figure 4.3 and Figure 4.4. This

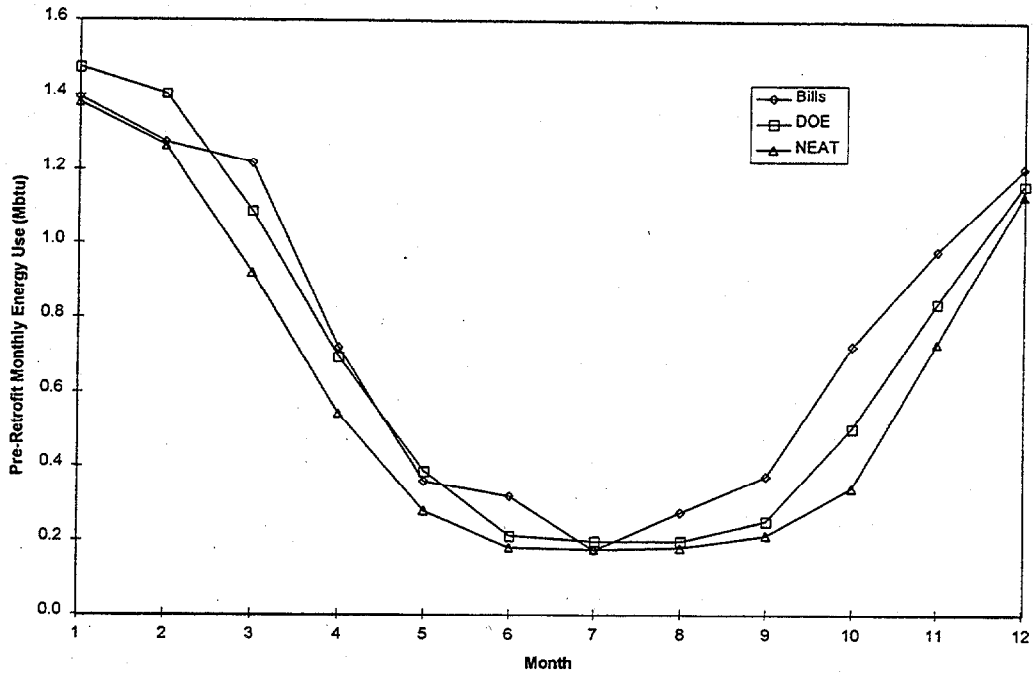

Figure 4.1. Comparison of Monthly Energy Use Before Retrofit from Utility Bills, DOE-2, and NEAT for House One.

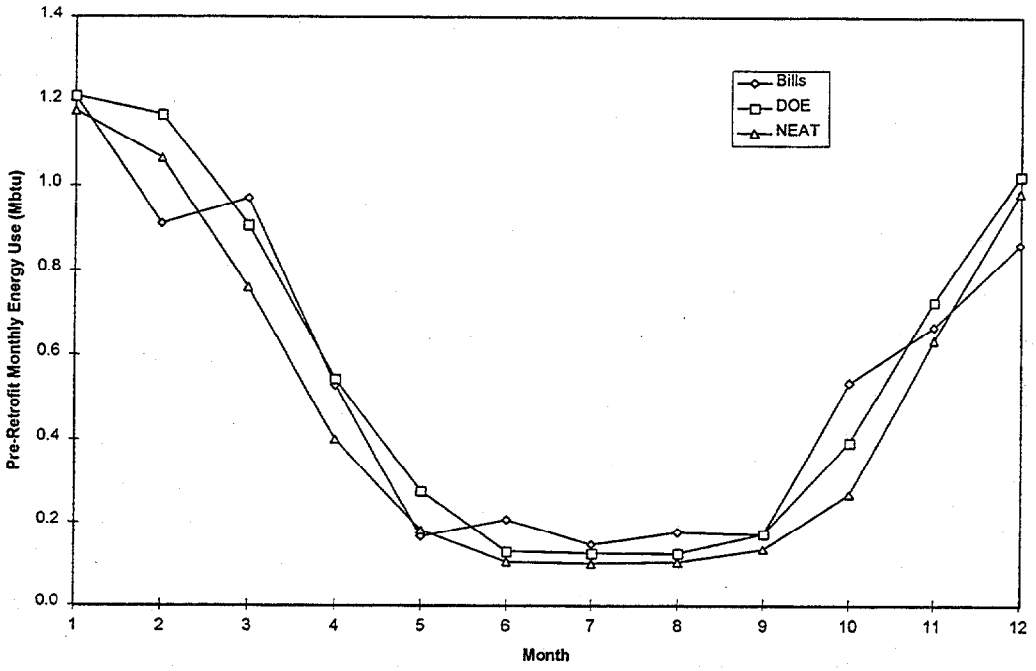

Figure 4.2. Comparison of Monthly Energy Use Before Retrofit from Utility Bills, DOE-2, and NEAT for House Two.

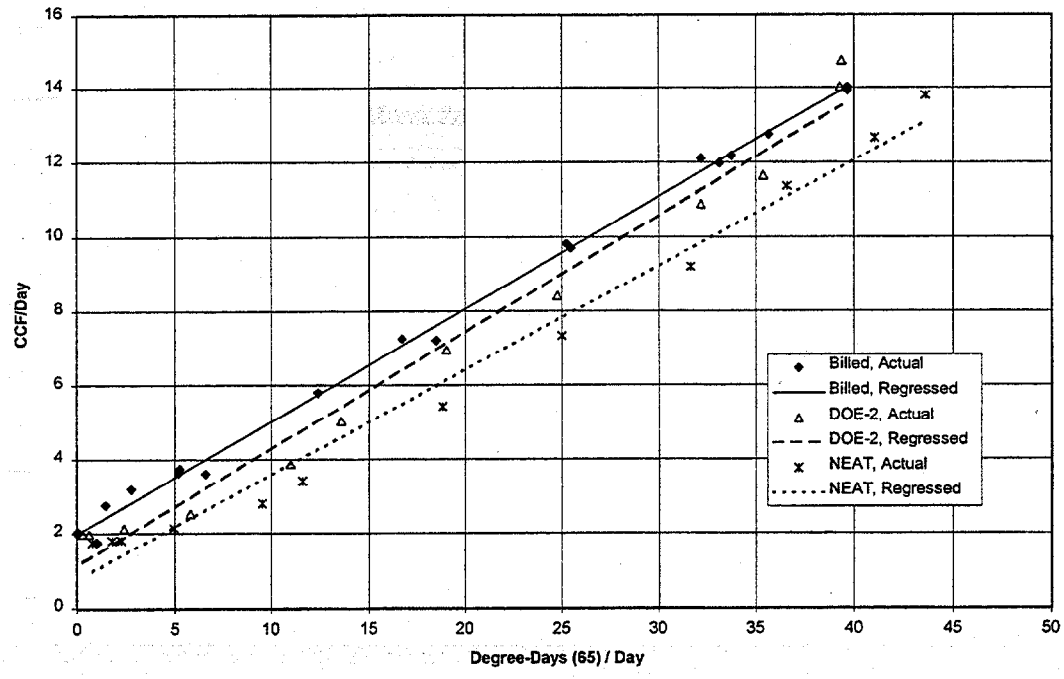

Figure 4.3. Comparison of Slopes of CCF/HDD Lines Before Retrofit from Billing Data, DOE-2, and NEAT for House One.

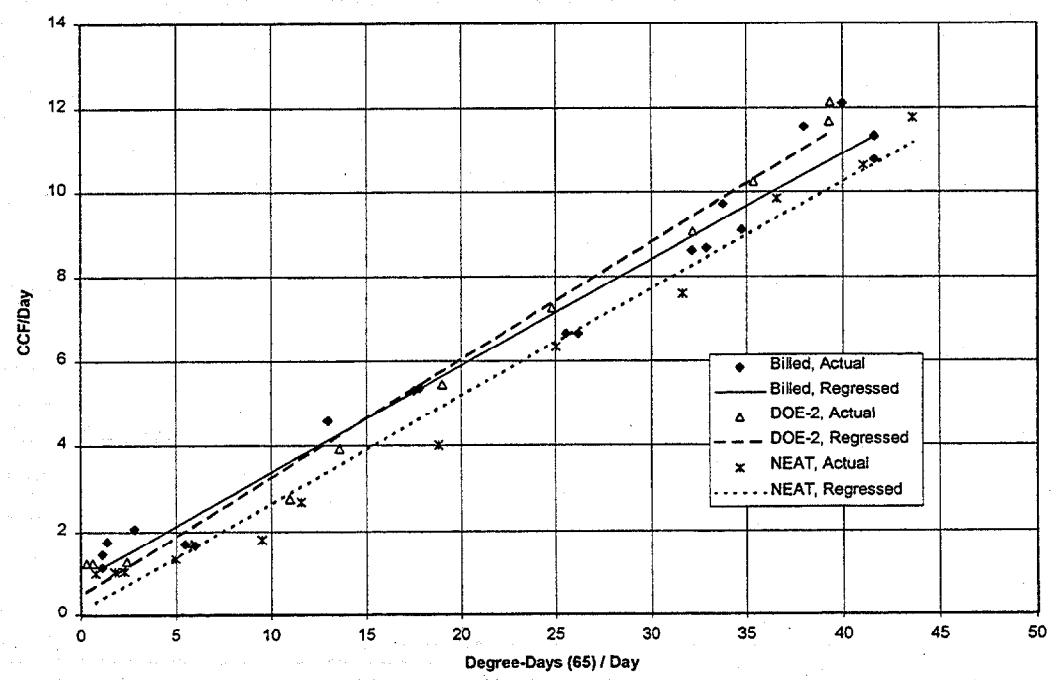

Figure 4.4. Comparison of Slopes of CCF/HDD Lines Before Retrofit from Billing Data, DOE-2, and NEAT for House Two.

result is significant because it demonstrates that the basic algorithms and default data used by NEAT are reasonable and that the collected data are sufficient to define a building's energy use characteristics, at least for these two case study buildings. This also means that savings estimates calculated later should not be biased by patterned high- or low-use baseline models. The differing y-intercepts are not cause for concern. They simply reflect differing assumed amounts of natural gas cooking and other non-HVAC equipment.

The component conductances reported by NEAT and DOE-2 were compared for the two houses. In some cases, component conductances varied significantly between NEAT and DOE-2. The conduction heat loss coefficients for NEAT and DOE-2 for the two buildings are shown in Table 4.1 below.

Table 4.1. Conduction Design Heat Loss by Component from NEAT and DOE-2 for House One and House Two

	House One	House One	House Two	House Two
Building Component	**NEAT**	**DOE 2.1**	**NEAT**	**DOE 2.1**
Wall UA (Btuh/F)	145	145	201	298
Subspace UA	15	--	100	--
Window UA	73	133	116	74
Door UA	17	15	148	125
Roof-Ceiling UA	61	66	78	42
Subtotal	311	359	643	539
Infiltration	151	355	324	770
Total	462	714	967	1309

Window losses differed the most, probably due to different framing definitions. The variation was high for one building and low for the other, suggesting that the difference is not due to pattern error. However, because this analysis compares results for only two houses this could use more investigation. Convective losses due to infiltration varied more than conduction, 151 versus 355 Btuh/F for the first house and 324 versus 770 Btuh/F for the second house for NEAT and DOE-2 respectively. These are dramatically different losses. Because the monthly results did not differ tremendously, these values probably do not represent exactly the same heat loss effects, but again this may warrant further investigation. Radiant heat gain differences were not compared because NEAT does not produce data to explore this phenomenon.

For one of the two houses, the pre-retrofit DOE-2 model calculated that over 85% of the annual heating energy was devoted to meeting infiltration loads or as heating

system inefficiency. Two of the three implemented measures were nine hours of blower door leakage reduction and heating system replacement. Energy use at this house was cut almost in half after the retrofit (See Sections 4.3.2 and 4.3.3.1 for additional discussion).

4.3 METHODS OF SIMULATION FOR MEASURE-SPECIFIC AND MONTHLY ENERGY USE

In several iterations with NEAT and DOE-2 simulations were performed that represented the installation of recommended measures one at a time. Sections 4.3.1 and 4.3.2 describe the methods and their limitations. Section 4.3.3 presents the analysis of this exercise.

4.3.1 NEAT Measure-Specific and Monthly Energy Use Simulation Procedure

An important objective of this analysis was to validate NEAT savings estimates at the measure level. Because heating system measures so often interact with one another, it was important to simulate just the measures actually installed, to simulate them one at a time, and in the correct and same order for both models. A second objective was to perform an evaluation of monthly energy use by computing a ccf/HDD regression line. These two objectives were accomplished by following the NEAT model measure and post-retrofit simulation procedure described in the steps below.

Step 1. In the NEAT Set-Up screens, select as candidate measures only those actually installed, based on the utility database.

Step 2. Simulate the base model with the *before* retrofit blower door test (BDT) results input as the *after* test flow rate. (NEAT uses the BDT results specified in the *after* retrofit fields in computing the effects of the measures so as not to recommend measures based on pre-retrofit conditions that will change after retrofit). Store the monthly energy use results as the pre-retrofit condition.

Step 3. Re-simulate the model with the post-BDT airflow entered as the after test flow rate. Store the results and note the order of recommendation of installed measures.

Step 4. Subtract the results of step 3 from the results of step 2 to get the annual energy savings due to the BDT.

Step 5. Add the first recommended measure to the post-BDT model as part of the new *base* model, re-run NEAT, and store the new base model's annual energy use. Subtract the results of step 5 from the results of step 4 to get the annual energy savings due to the first non-infiltration reduction measure.

Step 6. Repeat step 5 until the pre-retrofit house finally becomes the post-retrofit house with all measures installed at the same time. Again store monthly energy use performance for later comparison with DOE-2. **It is interesting to note that for both houses all installed measures were recommended in these NEAT runs based on the default input parameters.**

There was one exception to this protocol. NEAT's base model does not simulate uninsulated ducts. Since the base model can't duplicate this particular condition, for this measure we used NEAT's estimated savings for duct insulation and distributed the annual savings into monthly portions of the final post-retrofit model.

4.3.2 DOE-2 Model Measure and Post-Retrofit Simulation Procedure

After NEAT modeling was completed and the proper measure order was established, DOE-2 modeling was relatively straightforward. The base model simulation was performed, then the base model with reduced infiltration per the BDT results was simulated, then measure two was added, then three, and so forth.

As with the base model, whenever possible, DOE-2 measure variable changes reflected the same inputs as NEAT. For example, adding R-13 wall insulation has the net effect of increasing the wall R-value by 11.9 between studs after static air space removal is accounted for. In DOE-2, a similar net effect was modeled by removing air and adding exactly R-3.71 per inch insulation in the 3.5 inch wall cavity.

For infiltration modeling in DOE-2, the Air Change (ACH) method was chosen over the Sherman-Grimsrud method to reduce and control the number of infiltration-dependent variables. Using the ACH approach rather than the Sherman-Grimsrud also had the effect of reducing infiltration losses to what was judged to be a more reasonable level based on the bills. This methodological choice was the only "reconciliation" effort performed in DOE-2.

As with NEAT, DOE-2's duct insulation measure could only be simulated indirectly, although for a different reason. In DOE-2, duct conduction losses are expressed as a percentage of system load rather than as a unique thermal model of their own. Reduction in duct losses was indirectly approximated by reducing losses from 5.4% to 2.54% of the building system-supplied air.

4.3.3 Measure-by-Measure Comparisons

Billing analysis cannot separate the effects of individual measures. However, the approach described above performs this separation of effects by measure explicitly for NEAT and DOE-2. This section compares the findings for the measure-by-measure savings predicted by the models in each of the two case study houses.

4.3.3.1 Infiltration Reduction. The blower door test (BDT) data for the first house promised tremendous savings -- a reduction of 3,139 cfm_{50} for a house of only 1,036 square feet. NEAT predicted infiltration reduction savings of 58% of total savings, or 561 ccf, while DOE-2 predicted savings of 47% of the total, or 484 ccf. (Table 4.2). The second house started out much tighter and leakage reduction was less, 603 cfm_{50}. NEAT-predicted savings of 31%, or 81 ccf, matched well with DOE-2 predictions of 34%, or 85 ccf. (Table 4.3). Overall, model predictions of infiltration reduction savings are not significantly different. There are two primary reasons for the differences that are seen. First, NEAT simulates on a monthly basis, while DOE-2 simulates on an hourly basis. Second, more than any other construction characteristic, natural infiltration depends upon and is sensitive to other variables such as house exposure or crack type. Some variables to which natural infiltration is sensitive are not part of the NEAT input file. Therefore, the NEAT and DOE-2 models may make different assumptions about such variables.

Table 4.2. Savings by Measure for House One.

	Air Infiltration		Wall Insulation		Replace Heating System		Total	
Model	% Total	ccf	% Total	ccf	% Total	ccf	% Total	ccf
NEAT 5.2	58%	561	18%	172	24%	233	100%	966
DOE-2.1	47%	484	14%	144	39%	397	100%	1,024

Table 4.3. Savings by Measure for House Two.

	Air Infiltration		Rim Joist		Attic Insulation		Duct Insulation		Crawl space		Total	
Model	% Total	ccf	% Total	ccf	% Total	ccf	% Total	ccf	% Total	ccf	% Total	ccf
NEAT	31%	81	10%	26	19%	48	11%	29	29%	77	100%	260
DOE-2.1	34%	85	4%	10	22%	56	20%	50	21%	53	100%	253

4.3.3.2 Wall Insulation. Increasing the wall R-value for NEAT and for DOE-2 produced predicted savings of 18%, or 172 ccf, and 14%, or 144 ccf, respectively (Table 4.2). Neither the method nor the results of the conduction measure deviated substantially. To the extent that the results do differ, thermal mass effects probably contribute to variations in the building load.

4.3.3.3 Heating System Replacement. NEAT predicts 24%, or 233 ccf, savings, while DOE-2 predicts 39%, or 397 ccf, savings (Table 4.2). NEAT and DOE-2's savings estimates for heating system replacement were modeled by the following procedure: 1) increasing the heating system efficiency from 60% to 85%; 2) switching to an intermittent ignition device; and 3) adding a vent damper. NEAT's model actually is more complicated than DOE-2's for this measure. DOE-2 was modified by decreasing the

heating system equipment Heating Input Ratio and removing auxiliary equipment (the standing pilot light). Neither of the changes affects the space load. NEAT, however, recognized that an unintentionally conditioned subspace had been specified and assumed that this was due to the furnace. Combined with the fact that the furnace was oversized (at least at its original design capacity), NEAT computed the heating system efficiency variable to significantly affect the effective UA loss coefficient. In one scenario not used for the final analysis, NEAT's interpretation of improving heating system efficiency actually cost money because the subspace UA increased so much. The algorithm used to link furnace waste heat to subspace UA may overstate the correlation and explains why NEAT's predicted savings were lower than DOE-2's. The relation may be overstated because rated furnace efficiency losses tend to be due to incomplete combustion/excess oxygen, poor heat exchanger effectiveness, or part load losses. For all of these causes, most of the lost energy goes up the stack and out of the house rather than being released as subspace heat gain. (These algorithms in NEAT have since been modified, in part due to the findings presented here.)

4.3.3.4 Crawl Space Insulation. Crawl space insulation involves insulating walls that rest between exposed air and unconditioned or unintentionally conditioned space. NEAT estimated 29%, or 77 ccf, savings for this measure, while DOE-2 estimated 21%, or 53 ccf, savings (Table 4.3). The difference between the two likely is due to the different ways that the buildings are modeled. DOE-2's two zone configuration tempers the effect of the measure compared to NEAT, which uses the effective subspace UA.

4.3.3.5 Duct Insulation. NEAT estimated 11%, or 29 ccf, of savings due to duct insulation, while DOE-2 estimated 20%, or 50 ccf (Table 4.3). As described above, duct insulation had to be simulated indirectly in both NEAT and DOE-2. In lieu of comparative analysis that would have been of dubious value given the necessary DOE-2 approximation already described, NEAT's performance and sensitivity to change was studied more carefully. In the base model, duct losses interacted with unintentionally conditioned subspaces much in the same way as heating systems. In the measure savings section of NEAT, the duct insulation savings estimates were reasonable in their order of magnitude and positive direction. This is the most important assessment because it is these results upon which measure installation decisions are made. These data were used for the main model. However, the before-after base model comparison gave counterintuitive results: duct losses were more than compensated for by increased subspace losses. Further investigation did reveal a weakness in NEAT's model for the subspaces. This weakness has been corrected and will not appear in future versions of NEAT.

4.3.3.6 Attic Insulation. The models produced very similar results for this measure. NEAT predicted 19%, or 48 ccf, savings for attic insulation, while DOE-2 predicted 22%, or 56 ccf, savings (Table 4.3).

4.3.3.7 Measure Analysis Summary. Generally the measure-by-measure examination revealed comparable savings, with a 28% standard deviation of the absolute difference

between the models. Variation was lower for straightforward measures (infiltration reduction, envelope insulation) and higher for measures with complex interactions among building components (duct insulation, heating system replacement, rim joist insulation). This was most likely due to the fact that the two modeling systems dealt with the complex interactions in somewhat different ways.

4.4 COMPARISONS OF NEAT-PREDICTED SAVINGS TO DOE-2 SIMULATED SAVINGS AND OF MODELED SAVINGS TO ACTUAL USE AS MEASURED BY BILLING DATA

In this section, the pre-and post-retrofit consumption estimates and the savings estimates obtained with three different methods (NEAT, DOE2.1, and billing data analysis) are presented for the two test houses. In addition, a comparison of the results obtained from the billing analysis that assumes a constant balance point temperature (which is used in Section 4.4.1) and one that assumes a changing balance point temperature is presented in Section 4.4.2.

4.4.1 Comparisons of Pre-Retrofit Consumption, Post-Retrofit Consumption, and Savings Estimates Obtained with the Three Methods

As was discussed in Section 4.2.2, the two engineering models and the billing analysis technique produced very similar estimates of pre-retrofit consumption in both test houses. This good agreement among the pre-retrofit consumption estimates demonstrated that the basic algorithms, default values, and data used by NEAT are sufficient to define a building's energy use characteristics with reasonable accuracy.

Although the pre-retrofit consumption estimates agreed well across methods, there was less agreement among the savings estimates (Table 4.4). In particular, for the first house, NEAT predicted savings of 966 ccf, while DOE-2 predicted savings of 1,024 ccf, which is a good match. The billing analysis estimate of savings, however, was lower than either model's prediction at 795 ccf. In the second house, the NEAT estimate of savings was 260 ccf and the DOE-2 estimate was 253 ccf, again showing excellent agreement. The savings estimate from the billing analysis, however, at 163 ccf was much lower than either of the model predictions.

In the first house, the two engineering models produced similar estimates of post-retrofit consumption (938 ccf and 1102 ccf), while the billing analysis estimated higher post-retrofit consumption (1252 ccf) than either of the models. Here, the billing analysis estimate of savings was lower than the savings estimated by the models mainly because of the higher post-retrofit consumption shown by the billing analysis. For the pre-retrofit consumption estimates, the billing analysis results (2047 ccf) fell between the model estimates of 1904 ccf and 2126 ccf.

In the second house, the two engineering models produced estimates of post-retrofit consumption (1456 ccf and 1625 ccf) that fell on either side of the billing analysis estimate of 1536 ccf. The pre-retrofit consumption estimates from the billing analysis (1699 ccf) also fell between the model estimates of 1716 ccf and 1878 ccf. In this house, the billing analysis estimate of savings was lower than the ones from the models because both its pre- and post-consumption estimates had intermediate values, making the difference between these two estimates smaller.

In summary, the similarity of NEAT and DOE-2 results suggests that typical NEAT realization rates of less than one are probably not due to inaccuracies in its internal engineering algorithms. The lower savings estimated by the billing analysis, as compared to both of the engineering models, seem likely to be due to implementation and behavioral factors such as those discussed in Section 2.

Table 4.4. Comparisons of Pre-Retrofit Consumption, Post-Retrofit Consumption, and Savings Predictions Obtained from NEAT, DOE 2.1, and Billing Data.

	NEAT	DOE 2.1	Billing Data
House 1 (ccf/year)			
Pre-Retrofit Consumption	1904	2126	2047
Post-Retrofit Consumption	938	1102	1252
Savings	966	1024	795
House 2 (ccf/year)			
Pre-Retrofit Consumption	1716	1878	1699
Post-Retrofit Consumption	1456	1625	1536
Savings	260	253	163

The post-retrofit slopes of the billing, NEAT, and DOE-2 lines shown in Figures 4.5 and 4.6 reflect weather-normalized sensitivity to temperature for the first and second house, respectively. In the first house, there is considerable variation across the methods, with differences in slope that exceed 30%. In the second house, on the other hand, the post-retrofit slopes do not vary from one another by more than 10% across the methods. The differing y-intercepts, shown in Figures 4.5 and 4.6, simply reflect differing assumed amounts of natural gas cooking and other non-HVAC equipment.

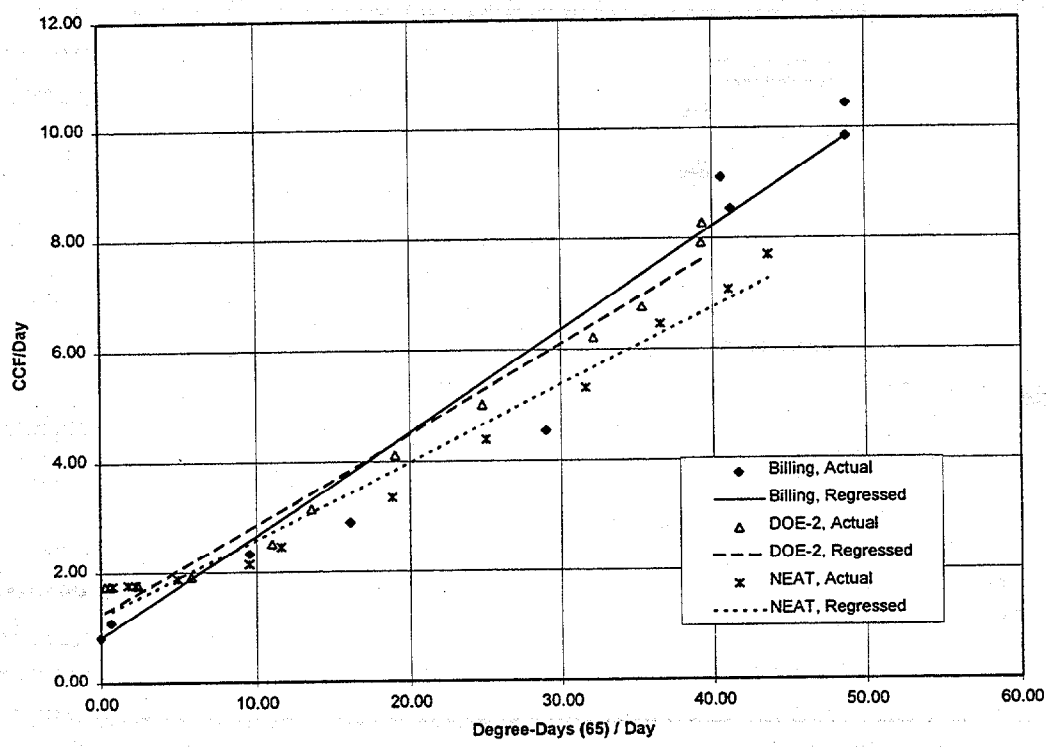

Figure 4.5. Post-Retrofit Slopes of CCF/HDD Lines for House One.

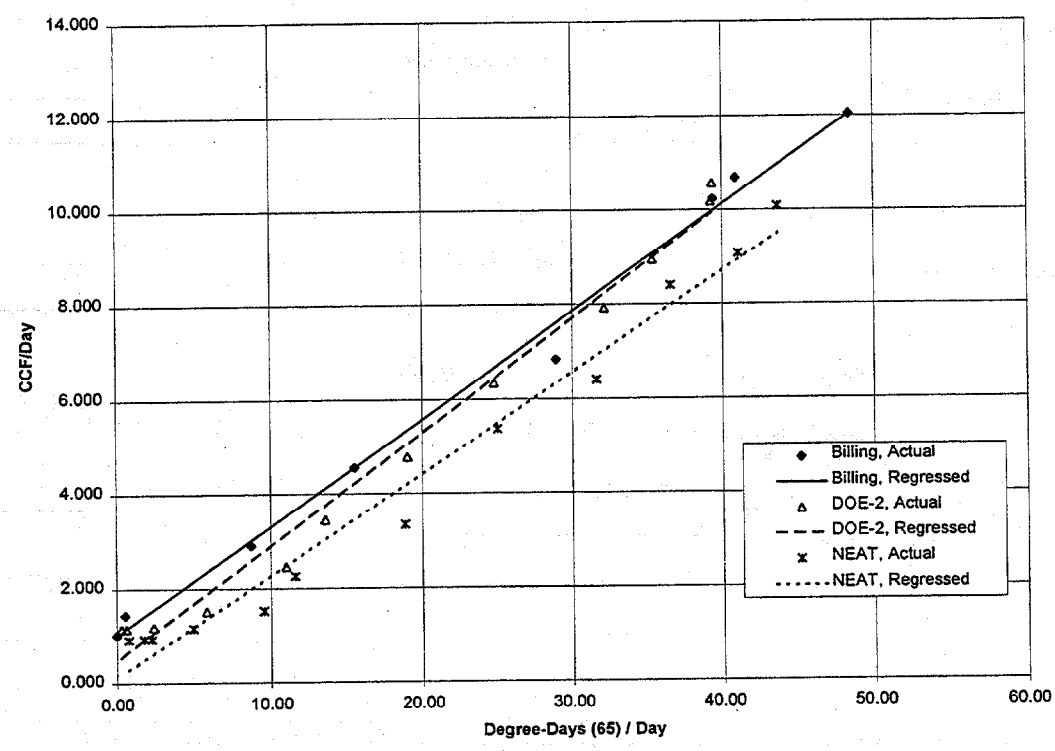

Figure 4.6. Post-Retrofit Slopes of CCF/HDD Lines for House Two.

4.4.2 Influence of Balance Point Temperature Assumptions

The analysis procedure used up to this point for comparing the savings predictions of NEAT and DOE-2 against actual billing data for the two houses is convenient in that it reconciles the differences in weather associated with the three sources of data. However, it assumes a constant "balance point temperature" of 65^0 F. The balance point temperature of a house is that outdoor air temperature for which the house requires no heating or cooling from its HVAC equipment to keep it at the desired thermostat setpoint. Normally, as a house becomes more energy efficient, this temperature drops, since internal sources of heat and solar heating can supply the required heat at lower and lower outdoor air temperatures without use of the heating equipment. The analysis used thus far for the two houses ignores this effect.[4]

Because of a concern that assuming a constant balance point would affect the results, an effort was made to check the significance of this potential shortcoming by applying PRISM to the two houses. PRISM provides estimates of savings that incorporate the effects of a changing balance point temperature. For the first house, PRISM estimated heating savings of 1,120 ccf/year. This value can be compared to the savings estimated by the billing analysis that used the assumption of a constant balance point and compared pre-and post- values of ccf/HDD_{65}, which produced a savings estimate of 795 ccf/year, significantly less than the PRISM results. The predictions from NEAT and DOE-2 (966 and 1,024 ccf/year) fall in between the estimate based on the assumption of a constant balance point and the PRISM estimate that incorporates a changing balance point (Table 4.5). Use of the PRISM results yields a realization rate of 120% instead of the 80% previously obtained with the fixed balance point billing analysis.

Table 4.5. Comparisons of Savings Predictions of NEAT and DOE 2.1 to Savings Estimated from Billing Data Using a Fixed versus a Variable Balance Point Temperature

	NEAT Predicted Savings	DOE 2.1 Predicted Savings	Fixed Balance Point Billing Analysis Estimate of Savings	PRISM Estimate of Savings with Variable Balance Point
House 1 (ccf/year)	966	1,024	795	1,120
House 2 (ccf/year)	260	253	163	81

[4] Both NEAT and DOE-2 provide for varying balance point temperatures, it is only the analysis procedure being used to compare the results from these programs that lacks this ability.

The PRISM analysis of this first house indicates a balance point that changes from 65^0 F to 51^0 F as pre-retrofit conditions move to post-retrofit conditions. This change is consistent with the concept of the balance point temperature explained above and the substantial heating savings seen.

For the second house, using PRISM estimates produces an opposite effect. PRISM estimates heating savings of 81 ccf/year for this house. This is about half of the 163 ccf/year estimate produced by the previous billing analyses that used the assumption of a constant balance point (Table 4.5). In this case, incorporating a changing balance point moves the estimate based on billing data even further from the results predicted by NEAT and DOE-2, 260 and 253 ccf/year, respectively. The realization rate using the PRISM result drops from the 64% reported earlier to 32%.

There are indications that unusual circumstances could present problems in applying PRISM analysis techniques to the data for the second house. Examination of the points and regression lines for ccf/day versus degree-days/day from the billing data show little difference between pre- and post-retrofit conditions with scatter that would introduce significant error. Compared to the first house, the pre- and post-retrofit regression lines for the second house are nearly indistinguishable. PRISM recognizes the situation in reporting a post-retrofit balance point slightly higher than the pre-retrofit value, 61.25^0 F compared to 60.00^0 F. This is despite retrofit measures that should have substantially reduced the balance point (infiltration reduction and attic, duct, and foundation space insulation).

The additional analysis that uses PRISM's variable balance point feature, reaffirms fair agreement between modeled and measured savings for house one. For house two, the analysis possibly lends some insight into reasons for the poor agreement between modeled and measured savings.

5. SUMMARY AND CONCLUSIONS

5.1 REALIZATION RATES

The purposes of this study were to determine realization rates for NEAT and to identify reasons, if possible, for realization rates that differed from 100%. Because the calculation of realization rates is very sensitive to decisions made by the analyst concerning which cases to include or exclude, and how to deal with outlying data, realization rates were calculated for several subsets of the total sample of 49 dwellings.

Because it was apparent that there was high variability in data quality (in both the NEAT inputs and the billing data), a scheme was developed for rating the overall quality of the data upon which a dwelling's realization rate would be calculated. Using this rating scheme, three groups of houses were defined: dwellings with high, medium, and low data quality. As Table 5.1 shows, the realization rates, with adjustment to billing data, are about the same for the high data quality group (n=12) and for the total sample of 49 houses (69% and 66%, respectively), while the realization rate for the combined set of high and medium data quality houses is somewhat lower at 57%. This occurs because realization rates above 100% are concentrated in the houses with the poorest data quality.

Table 5.1. Realization Rates Calculated with Data Sets with Varying Levels of Data Quality

Data Sets Based on NEAT Input and Billing Data Quality	Average Realization Rate with Billing Data Adjustment
High Data Quality (n=12)	69%
High and Medium Data Quality (n=31)	57%
All houses- High, Medium and Poor Data Quality (n=49)	66%

It is unclear which set of houses provides the best estimate of NEAT's realization rate. The fact that previous validation studies reported realization rates between 50% and 60%, suggests that our result of 57% for the high and medium data quality subsets of homes is most consistent with previous findings.

The discrepancy between measured and NEAT-predicted savings occurred mainly because of higher than predicted post-weatherization consumption. There was little discrepancy between measured and predicted pre-weatherization consumption. The average measured pre-weatherization consumption, based on a PRISM analysis, for the New York gas utility program houses was 266 Mbtu. NEAT-predictions of average pre-

weatherization consumption were very close to the PRISM values, at 254 Mbtu. Although Dalhoff (1996) reported that NEAT tended to overpredict pre-weatherization consumption in his Iowa validation study, in our New York data set there was little tendency to overpredict the pre-weatherization consumption. This probably occurred because the New York gas utility program targeted such high-usage customers. In our data set, as in Dalhoff's, post-weatherization consumption, was often underpredicted. The PRISM-based value for post-weatherization consumption was 213 Mbtu, while the average NEAT prediction for post-weatherization consumption was 149 Mbtu.

5.2 VARIATIONS IN REALIZATION RATES BY SUBGROUPS

In this study, several subgroups of dwellings were examined to see if certain subgroups had higher or lower realization rates than others. Few significant relationships were found. Realization rates did not appear to be associated with pre-weatherization consumption levels or with the pre-weatherization energy use index of Btu/square foot/ heating degree day, nor were they correlated with the amount or type of insulation installed. Some additional factors that showed little or no relationship to realization rates included the heated floor area, the number of occupants, and the thermostat setpoint. The one factor that appeared to have a strong relationship to the realization rates was the use of secondary heating fuels. Among the 41 houses with no use of secondary heating fuels, realization rates averaged 63%. Among the eight houses that used a secondary heating fuel, the average realization rate was 36%. A small subset of homes with the highest data quality also had a lower standard deviation than the other groups of homes as was discussed in Section 3.4.

5.3 IDENTIFIED SOURCES OF ERROR

NEAT, like other building energy simulation models, contains hundreds of variables and parameters. Because of the complexity of these models, there are many possible levels of validation, depending upon how much control over potential sources of error one implements. The simplest form of validation is to compare actual building energy use to model predictions with no attempt to control sources of error. Even if good agreement is obtained in such a comparison, however, it is difficult to interpret the results because all of the possible sources of error are operating simultaneously and may be offsetting in their effects.

This study, as well as previous research, showed that the National Energy Audit (NEAT), like most residential engineering energy audits, has typical realization rates well below 100%. This study attempted to explore the reasons for the observed discrepancy between the NEAT-predicted and measured savings in several ways, each of which is discussed below.

5.3.1 Implementation Errors

Because a complete understanding of and strict control of implementation practices is necessary if one is to measure the accuracy of audit predictions with any precision, an effort was made to identify and correct potential sources of error in NEAT's implementation. Obviously, the completeness and accuracy with which input data are measured and recorded will affect the correctness of NEAT predictions. Because of the time and expense that would have been involved, it was not feasible to directly check the accuracy of the input measurements for the dwellings by reexamining or remeasuring them. It would be desirable in future studies to have the resources to reexamine and remeasure input data so that better control of this source of error could be achieved.

In this study, the effort to control for implementation errors consisted of a careful check of NEAT input files for their reasonableness and consistency. Building descriptions that were not completely consistent, reasonable, or understandable were found in almost one-third of the cases. When a more correct value could be determined, it was substituted for the initial value to create a corrected input data file, and NEAT was run again on the corrected input file. Missing values for infiltration rates or furnace efficiency measurements were replaced with average values in the corrected input data files. Overall, problems related to inconsistent or missing building input data were fairly common. These building description input errors introduce an uncertain amount of error into the NEAT predictions of savings. Sometimes the correction of input errors led to an increase in a dwelling's predicted savings, and sometimes it led to a decrease. Overall, predictions based on the corrected input files did not differ greatly from those based on the uncorrected input files. Predictions based on the uncorrected input files averaged 97.6 Mbtu per dwelling per year, while predictions based on the corrected input files averaged 104.7 Mbtu per dwelling. Realization rates calculated with the uncorrected and corrected input files also did not differ greatly. These findings led to the conclusion that, at least for this study's sample of dwellings, input data errors in the building descriptions were probably not the dominant reason for low realization rates.

In the ideal study, all building descriptions, as well as the characteristics of measure installations, would be checked by on-site inspections. There were no such inspections in this study. As a result, it is possible that installation characteristics differ from the assumptions made by NEAT. Engineering estimates normally assume perfect installation of measures, whereas reality may not reflect such perfection. For example, voids in wall insulation may occur, not simply because of worker error, but because of building construction details that make it difficult or impossible to insulate the entire envelope. Such abnormalities will produce lower savings than predicted by the engineering estimates. Discrepancies between the model assumptions and the actual on-site conditions may well be a significant cause of realization rates below 100%.

5.3.2 Inaccuracies in Model Assumptions and Algorithms

A second potential source of error could be inaccuracies in NEAT's assumptions and internal algorithms. To test the importance of this source of error, comparisons of NEAT's predictions of energy consumption and savings with the predictions of the DOE-2.1E model, which is an industry standard, were made. For pre-retrofit consumption, the two engineering models and the billing analyses all produced very similar results. This high similarity of results for pre-retrofit consumption demonstrates that the basic algorithms, default values, and data used by NEAT are sufficient to define a building's energy use characteristics with considerable accuracy. Furthermore, the fact that the two engineering model simulations of pre-retrofit consumption match so well, both with each other and with the billing data, suggest that the basic engineering approaches used in NEAT are sound.

Although pre-retrofit consumption estimates agreed well for all three methods (NEAT, DOE-2, and billing analysis), there was less agreement in the savings estimates. The two engineering models produced similar estimates of savings. The billing analysis results, however, showed lower savings than either of the engineering models. In summary, the similarity of NEAT and DOE-2 results suggests that typical NEAT realization rates of less than 100% are probably not primarily due to inaccuracies in the model's internal engineering algorithms. Other factors seem more likely to be the reasons for the realization of lower than predicted savings.

The process of comparing NEAT and DOE-2 assumptions led to the identification of three assumptions in NEAT which were inaccurate. Two of these inaccurate assumptions were related to the interactions of heating system efficiency and duct insulation with the effective heat loss from the living space to the unintentionally heated subspace, where the heating system and ducts are assumed to be located. With increased system efficiency and insulation of the ducts, less heat is added to the subspace from these sources, making the subspace colder. NEAT appeared to overestimate the effect this had on the overall heating requirements of the house because duct losses were more than compensated for by increased subspace losses. Further investigation did reveal a weakness in NEAT's model for the subspaces. This weakness has been corrected and will not appear in future versions of NEAT.

The third inaccurate assumption is that NEAT, like most engineering audits, essentially models an empty house and assumes that walls and floors are bare. In reality, homeowners often have a portion of their walls or floors covered by furniture, shelving, cupboards, drapes, decorations, storage (particularly in attics or basements), etc. Some of these items may add resistance to heat flow through these surfaces which is not accounted for in the engineering calculations. Thus, actual savings from insulating building components can be less than those predicted based on an assumption of empty houses. Although household items may add resistance to heat flow, we do not consider it worthwhile to modify NEAT to reduce this source of error because it would be too

difficult to measure, quantify, and model such effects. In addition, such a demanding effort would be likely to produce only a small improvement in accuracy.

5.3.3 Behavioral Factors

Behavioral factors account for much of the variation in residential energy use. They also may account for the typical finding of realization rates that are less than 100%. It is usual in an evaluation of energy savings to see at least a small percentage of treatment houses that increase their consumption after weatherization. Such increases in consumption occur when behavioral factors overwhelm the effects of the energy-efficiency improvements. This study did not control for most of the behavioral factors that affect realization rates. The only method of controlling for this source of error was to restrict the sample to dwellings with no occupancy changes. One behavioral factor that our results suggest is an important influence is the use of secondary fuels. A number of previous validation studies also have identified changes in secondary fuel use as an important reason for lower than expected savings (Nadel and Keating, 1991). Clearly, an ideal validation study would have extensive measurement of, and controls for, a broad range of behavioral influences.

5.4 NEAT'S VALUE AS A TOOL FOR SELECTING MEASURES

In all the different NEAT runs that were conducted, with and without various modifications to the input data and the assumptions, the same measures were almost always recommended in the same order. In addition, when a test was conducted to see if a modified NEAT set-up configuration (with SIRs set at zero and all measures except those reported to have been installed switched off), would produce the same recommended measures as runs conducted without such modifications, the lists of recommended measures were almost always identical. Furthermore, during the process of comparing NEAT and DOE-2, it became clear that in both test houses all installed measures were recommended in the NEAT runs that were based on the default input parameters.

The fact that the set and sequence of recommended measures is insensitive to variations in input parameters and set-up configurations, suggests that NEAT's ability to select the most cost-effective set of measures, and to rank them in order of cost effectiveness, is not usually compromised by many of the sources of error that affect realization rates. Realization rates of less than 100% do not appear to affect NEAT's value for selecting an optimal set and sequence of measures. Recommendations concerning the intensity with which measures should be applied, however, can be affected. For example, recommendations concerning how many additional inches of insulation to install may be too high if the predicted savings from insulation are much higher than the actual savings will be. To reduce this problem, the SIR can be set above 1.0 in the input screen. This will eliminate the recommendation of levels of insulation

that are higher than the level that may be cost justified with realization rates of less than 100%. With an assumed realization rate of 50%, the SIR should be set to 2.0. With an assumed realization rate of 70%, the SIR should be set to 1.4.

5.5 NEAT PREDICTIONS ARE NOT A SUBSTITUTE FOR MEASUREMENTS OF PROGRAM PERFORMANCE

This NEAT validation study, as well as those conducted by Sharp (1994) and Dalhoff (1996) found that NEAT's typical realization rates fall between 50% and 70%. Another important finding was that realization rates vary a great deal depending upon the exact subset of houses that is examined. These findings mean that unvalidated NEAT predictions of savings, as well as predictions from other engineering models, should not be used to characterize program performance. Such predictions cannot be used as accurate substitutes for evaluations based on measured fuel consumption data.

REFERENCES

Applied Energy Group, Inc. 1996. *Evaluation of Niagara Mohawk Power Corporation's ULIEEP Power Partnership Program: Report of Process and Impact Evaluation Findings for Program Year Two (July, 1993-June, 1994),* prepared for Niagara Mohawk Power Corporation.

1993 ASHRAE Handbook: Fundamentals, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Inc., Atlanta, Georgia, pp. 25.10-25.12.

Barakat and Chamberlin, Inc., 1995. *Impact and Process Evaluation of the Consolidated Edison Utility Low-Income Energy Efficiency Program (ULIEEP),* prepared for Consolidated Edison, New York, New York.

Brown, Marilyn A. and Philip E. Mihlmester, 1994. *Summary of California DSM Impact Evaluation Studies,* ORNL/CON-403, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Cramer, James C., Nancy Miller, Paul Craig, Bruce Hackett, Thomas Dietz, Edward Vine, Mark Levine and Dan Kowalczyk, 1985. "Social and Engineering Determinants and Their Equity Implications in Residential Electricity Use," *Energy,* Vol. 10, No. 12, pp. 1283-1289.

Dalhoff, Gregory, 1996. *An Evaluation of thePerformance of the NEAT Audit as Implemented in the Iowa Weatherization Program,* prepared by Dalhoff and Associates for Wisconsin Energy Conservation Corporation.

Fels, Margaret F., Kelly Kissock, Michelle A. Marean, and Cathy Reynolds, 1995. *PRISM (Advanced Version 1.0) Users' Guide,* Center for Energy and Environmental Studies, Princeton University, Princeton, New Jersey.

Gettings, Michael B., 1990. *The North Carolina Field Test Audit: Description and Operation, Draft Version,* Oak Ridge National Laboratory.

Judkoff, R.D., 1988. "Validation of Building Energy Analysis Simulation Programs at the Solar Energy Research Institute," *Energy and Buildings,* Vol. 10, pp. 221-239.

Krigger, John, Nate Adams, and Michael Gettings, 1997. *National Energy AudiT (NEAT) Users Manual,* Oak Ridge National Laboratory, ORNL/Sub/91-SK078/2.

McQuiston, Faye and Jerald Parker, 1988. *Heating, Ventilation, and Air Conditioning: Analysis and Design,* 3rd Edition, pp. 161-167, John Wiley and Sons, New York.

Nadel, Steven and Kenneth Keating, 1991. "Engineering Estimates vs. Impact Evaluation Results: How to They Compare and Why," pp. 24-33 in *Energy Program Evaluation: Uses, Methods, and Results, Proceedings of the 1991 International Energy Program Evaluation Conference,* Chicago, Illinois.

New York State Energy Research and Development Authority and New York Department of State, 1993. *Instrumented Audits Technology Transfer Project,* prepared by Synertech Systems Corporation. Energy Authority Report 94-6.

Sharp, Terry, 1994. *The North Carolina Field Test: Field Performance of the Preliminary Version of an Advanced Weatherization Audit for the Department of Energy's Weatherization Assistance Program,* ORNL/CON-362, Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Socolow, R. H. (editor), 1978. *Saving Energy in the Home: Princeton's Experiments at Twin Rivers*, Ballinger, Cambridge, Massachusetts.

INTERNAL DISTRIBUTION

1-5. K. R. Ballew
6-25. L. G. Berry
26. J. E. Christian
27. G. E. Courville
28. T. R. Curlee
29. A. O. Desjarlais
30-39. M. B. Gettings
40. S. G. Hildebrand
41. C. I. Moser
42. T. R. Sharp
43. R. B. Shelton
44. R. L. Shelton
45. Central Research Library
46. Document Reference Section
47. Laboratory Records-RC

EXTERNAL DISTRIBUTION

48. Lilia A. Abron, President, PEER Consultants, P.C., 1000 N. Ashley Drive, Suite 312, Tampa, Florida 33602
49. Gregory Dalhoff, Dalhoff and Associates, 533 Marshall Circle, Verona, WI 53593-1727
50. Jean Diggs, Office of State and Community Programs, 5E-066, U.S. Department of Energy, 1000 Independence Ave., S.W., Washington, D.C. 20585
51. Thomas E. Drabek, Professor, Department of Sociology, University of Denver, Denver, Colorado 80208-0209
52-54. J. F. Eisenberg, ORNL, Capitol Gallery Bldg, Suite 306W, 600 Maryland Ave., S.W., Washington, D.C. 20024
55. Jim Golden, Energy Efficient Building Association, 2950 Metro Drive, Suite 108, Minneapolis, Minnesota 55424
56. Benjamin F. Hobbs, The Johns Hopkins University, Department of Geography and Environmental Engineering, Arees Hall, Baltimore Maryland 21218
57. Philip Hoyes, Office of State and Community Programs, 5E-066, U.S. Department of Energy, 1000 Independence Ave., S.W., Washington, D.C. 20585
58. Jonathan B. Maxwell, Aspen Systems, 575 Oak Ridge Turnpike, Oak Ridge, Tennessee 37830
59. Gail McKinley, Office of State and Community Programs, 5E-066, U.S. Department of Energy, 1000 Independence Ave., S.W., Washington, D.C. 20585
60-61. OSTI, U.S. Department of Energy, P.O. Box 62, Oak Ridge, TN 37831
62. ORNL Site Manager, U.S. Department of Energy, Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee 37831-6269
63. P. Richard Rittelmann, FAIA, Executive Vice President, Burt Hill Kosar Rittelmann Associates, 400 Morgan Center, Butler, Pennsylvania 16001-5977
64. Ron Santoro, Office of State and Community Programs, 5E-066, U.S. Department of Energy, 1000 Independence Ave., S.W., Washington, D.C. 20585
65. Susan F. Tierney, The Economic Resource Group, Inc., One Mifflin Place, Cambridge, Massachuseets 02138
66. C. Michael Walton, Ernest H. Cockrell Centennial Chair in Engineering and Chairman, Department of Civil Engineering, University of Texas at Austin, Austin, Texas 78712-1076