PRISM: A Standardized Method for Keeping Score on Energy Saved in Houses*

Margaret F. Fels
Center for Energy and Environmental Studies
Princeton University
Princeton, NJ 08544

* From PU/CEES Report No. 163, for distribution at Energy Conservation Program Evaluation: Practical Methods, Useful Results, Chicago, IL, 19-21 August 1985.

FOREWORD

The Princeton Scorekeeping Method, or "PRISM", is a statistical procedure which uses available billing and weather data to produce accurate estimates of saved energy. Derived from simple physical principles, the method provides physical descriptors as well as a weather-normalized index of consumption for each house analyzed. The long-range objective is a standardized approach which is equally applicable to all fuels, and which can be used for a wide range of climates and building types. This report summarizes the scorekeeping methodology -- its motivation, its current status, and needed areas of expansion.

Until recently, the emphasis was on houses heated by natural gas: the individual-house approach was developed for the Modular Retrofit Experiment under funding by the Buildings Systems Division of the US Department of Energy (see PU/CEES Reports No. 130 and 131), while the aggregate approach, funded by the Ford Foundation and the New Jersey Department of Energy, was developed to monitor statewide conservation trends in New Jersey (PU/CEES Report No. 156). We have explored the extension of the method to oil-heated houses (PU/CEES Report No. 139), and are currently working on a project funded by the Electric Power Research Institute (EPRI Project RP2034-4) to extend it to houses heated and cooled by electricity (PU/CEES Reports No. 155, 160, 165 and 166, and forthcoming reports).

None of this work would have been possible without the continued counsel and enthusiasm of Robert Socolow. Under his direction more than a decade ago, the scorekeeping approach was originated by Thomas Schrader, Yoav Benjamini, Lawrence Mayer and Thomas Woteki. Special thanks go to Miriam Goldberg and Michael Lavine for bringing new ideas and statistical rigor to the approach, and, most recently, to Daniel Stram for his creative contributions in statistics and software development.

I. Introduction

There Is No Substitute for Real Scorekeeping

To date, a frequent failing of commercial and government conservation enterprises has been a lack of accounting to "keep score" on the value and magnitude of the energy saved by the measures implemented. At Princeton, years of related research have convinced us that serious scorekeeping is essential to the success of all conservation ventures. Without it, the importance of conservation cannot be effectively communicated to homeowners, the best programs cannot be distinguished from ineffective ones, and the credibility of conservation is being threatened.

Many utilities in the U.S. have undertaken extensive retrofit assistance programs for their customers, not only because of the federal Residential Conservation Service (RCS), which mandates nearly free energy audits for customers, but also because of a growing commitment to energy conservation as a utility investment strategy. RCS audits have reached some two million homes. In addition, the Low-Income Weatherization Program, federally funded but managed at the community level, is reaching many additional homes, not only with an audit but with extensive, often costly, retrofits as well. With rising fuel prices, one may expect these retrofit programs to become even more popular.

Missing in almost all these programs has been an accurate evaluation of how much energy is actually being saved by specific actions taken. The program's yardstick of success is often the number of participants, with no regard for the number of kilowatt-hours of electricity, barrels of oil, or cubic feet of gas saved. Estimates of savings, if they exist, are based on engineering models typically without calibration to real-world experience. Such estimates, though useful for planning purposes, are notoriously higher than the actual savings realized, in part because they do not accurately take into account either human behavior or the irregularities in the complex heat flows of real buildings.

On the private side, companies which sell conservation services invariably omit feedback to the customer on how much energy -- and money -- the purchase is saving. Furthermore, without records of actual savings achieved, companies deny themselves a readily available source of information from which they can understand -- and project -- the value of the services they sell. The resulting picture can be one of dissatisfied, confused customers dealing with a company unable to convey accurate estimates of the value of its own services.

It's Easier Than It May Seem

Perhaps surprisingly, it is extremely straightforward to obtain accurate estimates of actual energy savings, and the required data, utility bills and daily average temperatures, are readily available. As depicted in Figure 1(a), the Princeton scorekeeping method, "PRISM", uses utility bills from before and after the retrofit installation, together with average daily temperatures from a nearby weather station for the same periods, to determine a weather-adjusted index of consumption, Normalized Annual Consumption or NAC, for each period. Analogous to (and, based on field measurements, clearly more accurate than) the EPA miles-per-gallon rating, the NAC index provides a measure of what energy consumption would be during a year under typical weather conditions. The total energy savings are derived as the difference between the NAC in the pre- and post-periods. A conservation effect is thus neither masked by a cold winter nor exaggerated by a warm one, nor is it obscured if the time covered by billing periods in one "year" is longer than in another.

In order to adjust for the influences of occupant behavior and externalities such as energy price changes, and in effect to isolate the savings due to the program from savings that would otherwise have occurred, scorekeeping often requires a set of untreated, "control" houses. The same procedure applied to both the treatment and control houses, as shown in Fig. 1(b), gives a measure of control-adjusted savings for the treatment group. The analysis can then be updated for succeeding years, to track the durability of the savings.

A more complete evaluation is often desired, to determine the cost-effectiveness of various tried approaches to conservation, for example, or the effect of program participation and other explanatory variables. The savings estimates, along with other PRISM outputs, provide reliable input to such analyses. Thus the PRISM analysis depicted in Fig. 1 may be thought of as a standardized first stage of an evaluation, with subsequent analyses, limited by available data and shaped by the specific needs of the project being evaluated, comprising the second stage.

PRISM differs from other weather-normalization procedures in that the house's break-even temperature is treated as a variable, rather than a constant such as 65°F. Three physical parameters result from the model applied to the billing data for the heating fuel of an individual house: base level consumption, corresponding to the amount of fuel used per day (for appliances including water heaters) independent of outside temperature; the reference temperature, approximating the average daily outside temperature above which no fuel is required for heating; and the heating slope, corresponding to the amount of fuel required per degree drop in outside temperature below the reference temperature. These parameters can provide indications of the sources of conservation: insulating, turning down thermostats, more efficient appliance usage, etc., and thus define an "energy signature" of the house. The Normalized Annual Consumption (NAC) index is derived from these parameters applied to a long-term (say, ten-year) annual average of heating degree-days.

It turns out that NAC is extremely well determined (its standard errors are typically 3-4% of the estimate), so that savings of 6% or more may generally be considered significant. Furthermore, NAC is quite insensitive to which periods are included, or their length. We know of no more reliable index for monitoring conservation.

Toward a Standardized Approach

Among the evaluations of retrofit programs that have been performed, the haphazard array of approaches often makes it impossible to compare savings from one program with another, or to aggregate the effects across programs. The first "scores" are in from selected RCS and weatherization programs, and many of them are disappointing. Nevertheless, the lack of a coordinated approach makes it impossible to learn from our mistakes and plan for more effective programs in the future.

The long-range goal of our scorekeeping research at Princeton is to produce a standardized, easy-to-use approach which utilities, communities and others throughout the country may adopt for measuring the savings achieved by their retrofit programs. Over the past several years, the PRISM tools have been enormously valuable to our own buildings research program (for example, in the Modular Retrofit Experiment, a collaborative conservation project between Princeton and the natural gas utilities in the New Jersey area (Dutt et al., 1982) and for monitoring statewide conservation trends in New Jersey (Fels and Goldberg, 1984)). There is now increasing interest in the scorekeeping method on the part of outsiders. Recently, it was used for the evaluation of Wisconsin's low-income weatherization program (Goldberg et al., 1984). For their evaluation of Residential Conservation Service and other utility conservation programs, staff at Oak Ridge National Laboratory are using PRISM as stage one of their two-stage evaluation approach (see, for example, their scorekeeping of Bonneville Power Administration's weatherization pilot program, in Hirst et al., 1985). The method is being used extensively in Minnesota to monitor the success of a variety of city and state programs (see, for example, Hewett et al., 1984).

There is much more to be learned before PRISM will work equally well for all major fuels, over a wide range of climates and building types. While the initial emphasis of the methodology development was on gas-heated single-family houses, we are focusing our current research in three areas: 1) the inclusion of cooling for electrically heated houses -- a nasty problem because the demand for cooling is far more erratic (people-dependent) than it is for heating; 2) the treatment of "bad" houses that don't respond predictably to weather; and 3) the applicability of the approach to large multi-family buildings, to understand its limitations as well as its strengths. With the benefit of the wealth of real-world experiences embodied in ongoing scorekeeping projects such as the above, we are optimistic that these advancements in the methodology are feasible.

II. Summary of the Scorekeeping Method

The Princeton scorekeeping method involves a straightforward procedure for calculating energy savings between two time periods. For each house being analyzed, the procedure requires meter readings (or for fuel oil, delivery records) for approximately one year in each period. The consumption data are then corrected for the effects of weather, which of course is never the same for two different years, and also for differences in the time spanned by the two periods. From these results the weather-normalized consumption index, called Normalized Annual Consumption or NAC, is calculated.

The Physical Basis for the Model

We start by describing the method developed for fuels used for heating but not cooling. Generally, whether for natural gas, oil or electricity, a house's heating system is first required when the outside temperature (T) drops below a certain level (the heating reference temperature τ), and for each additional degree drop in temperature a constant amount of heating fuel (the heating slope β) is required.* Thus, the required heating fuel is linearly proportional to $\tau - T$, and the proportional constant β represents the house's effective heat-loss rate. In addition, the house may use a fixed amount of heating fuel (the base level α) which is independent of outside temperature T. Formally, the expected fuel consumption f, as illustrated in Fig. 2 for an idealized house, is given by

$$f = \alpha + \beta\ (\tau - T)_{+} \qquad (1)$$

where the term in parentheses is the heating degree-days to base τ and the "+" indicates zero if the term is negative.

The physical justification for assuming that both the base level α and the heat-loss rate β are constant has been carefully analyzed in previous work (Schrader, 1978). The derivation leads to a simple physical interpretation for each of the three parameters. The reference temperature τ represents the outdoor temperature below which the heating system is required. The value of τ is influenced primarily by the indoor temperature (thermostat setting) and, in addition, an offsetting contribution from the free heat (i.e., heat generated by appliances and occupants). The heat-loss rate β depends on the conductive and infiltration heat losses, while the base level α represents the fuel requirements of appliances (including lights, for electricity, and the water heater if fueled by the heating fuel).

If τ is not accurately determined, or if it changes significantly over the time periods studied, the error or change in τ will directly affect α, with an opposite sign. In fact, the slope β will be affected as well. Fig. 3 illustrates this for the idealized house by plotting f vs. h for one correct and two incorrect values of τ. A straight-line fit through each

---

* Here "fuel" includes electricity as well as natural gas, fuel oil, etc.

set of points will have a different slope and intercept. Therefore, an assumed (incorrect) reference temperature, such as the value of 65°F so commonly used, is likely to lead to less physically meaningful values of the base level and the heat-loss rate. As discussed below, it turns out that estimates of total consumption, over a year, for example, are much less sensitive to choice of τ.

Individual-house Analysis

Based on this physical interpretation, the two data requirements for the analysis are actual meter readings, from which consumption is calculated, and daily average temperatures, from which heating degree-days to different reference temperatures are computed in exact correspondence to the consumption periods. The input to the procedure is then F_i and H_i where:

F_i = average daily consumption (e.g., in kwh/day) in time interval i

$H_i(\tau)$ = heating degree days per day to reference temperature τ in time interval i.

Here $H_i(\tau)$ is computed from average daily temperatures T_{ij} for the N_i days in interval i, i.e.,

$$H_i(\tau) = \sum_{j=1}^{N_i} (\tau - T_{ij})_+ / N_i \tag{2}$$

where "+" indicates that the term in parentheses is set to zero if T_{ij} is above τ. Fig. 4 shows a plot of F_i against H_i for the 1978-79 heating year, for a sample house from the Modular Retrofit Experiment (MRE). A straight-line relationship is clearly suggested.

The set of data points $\{F_i\}$ and $\{H_i\}$ for an approximately year-long period are then fit to a linear model:

$$F_i = \alpha + \beta\, H_i\, (\tau) + \varepsilon_i \tag{3}$$

where ε_i is the error term. For a guessed value of reference temperature τ, the base level and heating slope parameters α and β are found by standard statistical techniques (ordinary least-squares linear regression). The parameters α and β are calculated in this way for several different values of τ. "Best τ" is the one for which a plot of F_i vs. H_i (τ), such as the one shown in Fig. 4, is most nearly a straight line. Formally, τ is determined as the value for which the mean squared error is minimized, or equivalently for which the R^2 statistic is highest. The corresponding values of α and β are the best estimates of base level and heating slope.

In our model, the term α characterizes the temperature-independent component of consumption (in units/day, where units may be ccf or therms for gas, kwh for electricity or gallons for fuel oil), which is dominated by appliance and water heater usage. The parameter β represents the incremental amount of gas required for each degree drop in temperature below the reference temperature. Referring to β as the heating slope (in units/°F-day), the term $\beta\Sigma H_i(\tau)$ gives an estimate of temperature-dependent demand, which is dominated by space-heating. The reference temperature τ (in °F), which varies from house to house, represents the average outside temperature below which a house's heating system is required.

The parameters α, β and τ resulting from the model are used to calculate Normalized Annual Consumption, the overall index of consumption which we call NAC. The NAC index represents consumption which would occur in a year with typical weather conditions, and is calculated as follows:

$$NAC = 365\ \alpha + \beta\ H_o\ (\tau) \tag{4}$$

where $H_o(\tau)$ is the heating degree-days (base τ) in a "typical" year. The values of H_o used in our recent New Jersey analyses are based on the twelve-year normalization period from 1970 through 1981.*

To illustrate, the model in Eq. 3 applied to the house data in Fig. 4 gives the following results for the best-τ approach:

τ = 68.1 (±2.7) °F
α = 0.90 (±0.26) ccf/day
β = 0.18 (±0.01) ccf/°F-day
NAC = 1324 (±27) ccf/year
R^2 = 0.985.

The numbers in parentheses represent the standard errors. The relatively small standard error for NAC as compared with α and β (2% vs. 12% and 6%) is typical of results from this model. The heating component $\beta H_o(\tau)$ represents 63% of the total consumption. The R^2 statistic indicates a very good straight-line fit, corresponding to the line drawn in Fig. 4.

The methodology is directly extendable to electrically heated houses without cooling. For example, the heating-only model in Eq. 3 applied to the house data in Fig. 5 gives the following results:

τ = 59.5 (±1.6) °F
α = 29.0 (±1.6) kwh/day
β = 2.73 (±0.17) kwh/°F-day
NAC = 20,700 (±375) kwh/year
R^2 = 0.990.

Again, the NAC, with a standard error of 2%, is extremely well determined.

---

* Weather data to compute H_i for each period and H_o, to any reference temperature, were collected from National Oceanic and Atmospheric Administration (NOAA) data for the Newark, NJ, weather station. For example, values of H_o for τ = 60 , 65 and 70°F are 3807, 4917 and 6181 °F-days/year respectively.

In general the NAC estimate provides a reliable conservation index from which energy savings and conservation trends may be accurately estimated. On the other hand, the three parameters α, β and τ comprising the energy signature of the house, and the estimate of annual heating consumption $\beta H_o(\tau)$ derived from them, are less well determined, and their changes over time are often difficult to interpret due to the interference of physical and statistical effects. While it is tempting to attribute a change in the base level to water heater wrap or more efficient appliances, for example, or a drop in the heating-consumption estimate to added ceiling insulation or other measures to tighten the structure, such physical inferences are often not statistically valid. We feel that these parameters provide physically meaningful indicators, whose changes may not be statistically significant but whose behavior can often suggest the reason for a consumption change.

A frequently mentioned shortcut is the used of fixed τ, at $65^{o}F$. Although the individual parameters (α more than β) are highly sensitive to the τ value used, the NAC results are not, especially when the best-τ values are fairly close to $65^{o}F$. (The median τ value for several samples analyzed by this method has been close to $60^{o}F$.) Nevertheless, our studies indicates that α and β are considerably more meaningful when estimated for best τ than when estimated at a fixed value. We strongly recommend that the best-τ approach be used when the results for α, β and τ are of interest, as they usually are in a conservation analysis, or when there is reason to believe that the true τ value is quite different from the assumed value.

The Measurement of Savings

The NAC index provides the basic parameter for monitoring energy savings resulting from retrofit programs. Using billing and weather data for approximately year-long periods before and after (and not including) the period during which the retrofits were performed, $NAC_{pre}(T)$ and $NAC_{post}(T)$ are calculated as averages (medians or means*) over houses in the treatment group, for the pre- and post-periods respectively. The raw, weather-adjusted change in energy consumption is then given by

$$S_{raw} = NAC_{pre}(T) - NAC_{post}(T) . \qquad (5)$$

If a control group is included, analogous control indices $NAC_{pre}(C)$ and $NAC_{post}(C)$ are calculated as averages over the control houses, for the same pre- and post-periods. For an estimate of the savings attributable to the program of interest, the raw savings can then be adjusted as follows:

$$S_{adj} = NAC_{pre}(T)\ [NAC_{post}(C)\ /\ NAC_{pre}(C)] - NAC_{post}(T) , \qquad (6a)$$

* The median is generally preferred as a more "robust" (i.e., insensitive to outliers) measure of the center of the group's distribution.

or, in percentage terms,

$$S_{adj,\%} = [NAC_{post}(C) \;/\; NAC_{pre}(C)] - [NAC_{post}(T) \;/\; NAC_{pre}(T)] \;. \quad (6b)$$

The raw savings for the treatment group (Eq. 5), the control savings and the savings adjusted by the control (Eq. 6) are all quantities of interest in scorekeeping.

For the MRE house in Fig. 4, the raw savings were 325 ccf/year, or 25% of pre-period consumption. This house belonged to the "House Doctor" group, for which the median savings may be summarized as follows:

raw savings, treatment group: $S_{raw}(T)$ = 200 ccf/year, or 15% of pre-NAC
raw savings, control group: $S_{raw}(C)$ = 130 ccf/year, or 9%
control-adjusted savings: S_{adj} = 139 ccf/year, or 10%.

Thus the savings are highly sensitive to whether they are adjusted by a control, with the net effect of a 30% deflation in this experiment's raw savings due to the control adjustment.

Inclusion of Electric Cooling

The methodology presented thus far has been applied extensively to gas- and oil-heated houses, and electrically heated houses without cooling. For all fuel types, R^2-values are typically 0.97 or better, and the accuracy of the estimates corresponding to Figs. 4 and 5 is typical of the individual houses studied. Thus, direct extension of the methodology to electrically heated houses without cooling has been straightforward.

If electricity is used for cooling but not heating, a model analogous to Eq. 3 applies, with cooling degree-days $C_i(\tau_c)$ computed to a cooling reference temperature τ_c replacing $H_i(\tau)$. If the house is electrically heated and cooled, the model becomes:

$$F_i = \alpha + \beta_h H_i(\tau_h) + \beta_c C_i(\tau_c) + \epsilon_i \quad (7)$$

where β_c is the cooling rate. The corresponding weather-normalized index is given by

$$NAC = 365\alpha + \beta_h H_o(\tau_h) + \beta_c C_o(\tau_c) \quad (8)$$

where cooling degree-days C_o are computed for the same normalization period establishing H_o.

Even in a heating-dominated climate, summer consumption not uncommonly tracks cooling degree-days, as Fig. 6 illustrates. The results of applying the heating-plus-cooling model in Eq. 7 to the data in Fig. 6 are shown with the figure. Once again, the NAC estimate, with a relative standard error of 2%, is extremely well determined.

Not surprisingly, not all houses behave as predictably as this example does. In our current project for EPRI, we have been exploring several modifications of the above equation, such as holding τ_c at $70^{\circ}F$, or estimating summer consumption in excess of base level. Our experience to date suggests that the R^2 values for houses heated and cooled by electricity will be somewhat lower than they are for a heating-only fuel, but generally high enough for an accurate measurement of savings.

Extension to Utility Aggregates

The above methodology is designed to be applied to individual-house billing data for large numbers of houses, in utility conservation programs, for example, or retrofit projects such as MRE. An analogous approach has been demonstrated to work well for utility aggregate sales of natural gas to gas-heating customers. To account for the billing lag, a simple function of this month's and last month's heating degree-day, AH_i, replaces H_i in Eq. 3 (see Fels and Goldberg, 1984).

As for the single-house example in the previous section, the error bars for the aggregate NAC, at approximately $\pm 3\%$ of NAC for single years, are considerably narrower than the bars associated with the individual parameters. The resulting ability to measure small changes in consumption makes the NAC parameter a valuable conservation index for monitoring purposes.

References

G.S. Dutt, M.L. Lavine, B.G. Levi and R.H. Socolow (1982) "The Modular Retrofit Experiment: Exploring the House Doctor Concept", Princeton University Center for Energy and Environmental Studies Report No. 130, Princeton, NJ.

M.F. Fels and M.L. Goldberg (1984) "Using Billing and Weather Data to Separate Thermostat from Insulation Effects," Energy 9, #5, 439-445.

M.L. Goldberg, A. Jaworski and I. Tallis (1984) "Evaluation of Wisconsin's Low-Income Weatherization Program", Final Report and Appendices of Final Report to Wisconsin Department of Administration, Madison, WI.

M. Hewett, T. Dunsworth, T. Miller and M. Koehler (1984) "Measured vs. Predicted Savings from Single Retrofits: A Sample Study", Minneapolis Energy Coordination Office, draft report, Minneapolis, MN.

E. Hirst, D. White and R. Goeltz (1985) "Three Years after Participation: Electricity Savings Due to the BPA Residential Weatherization Pilot Program", Report ORNL/CON-166, Oak Ridge National Laboratory, Oak Ridge, TN.

T.F. Schrader (1978) "A Two-Parameter Model for Assessing the Determinants of Residential Space Heating", MS Thesis for Department of Aerospace and Mechanical Engineering, Princeton University, Princeton, NJ.

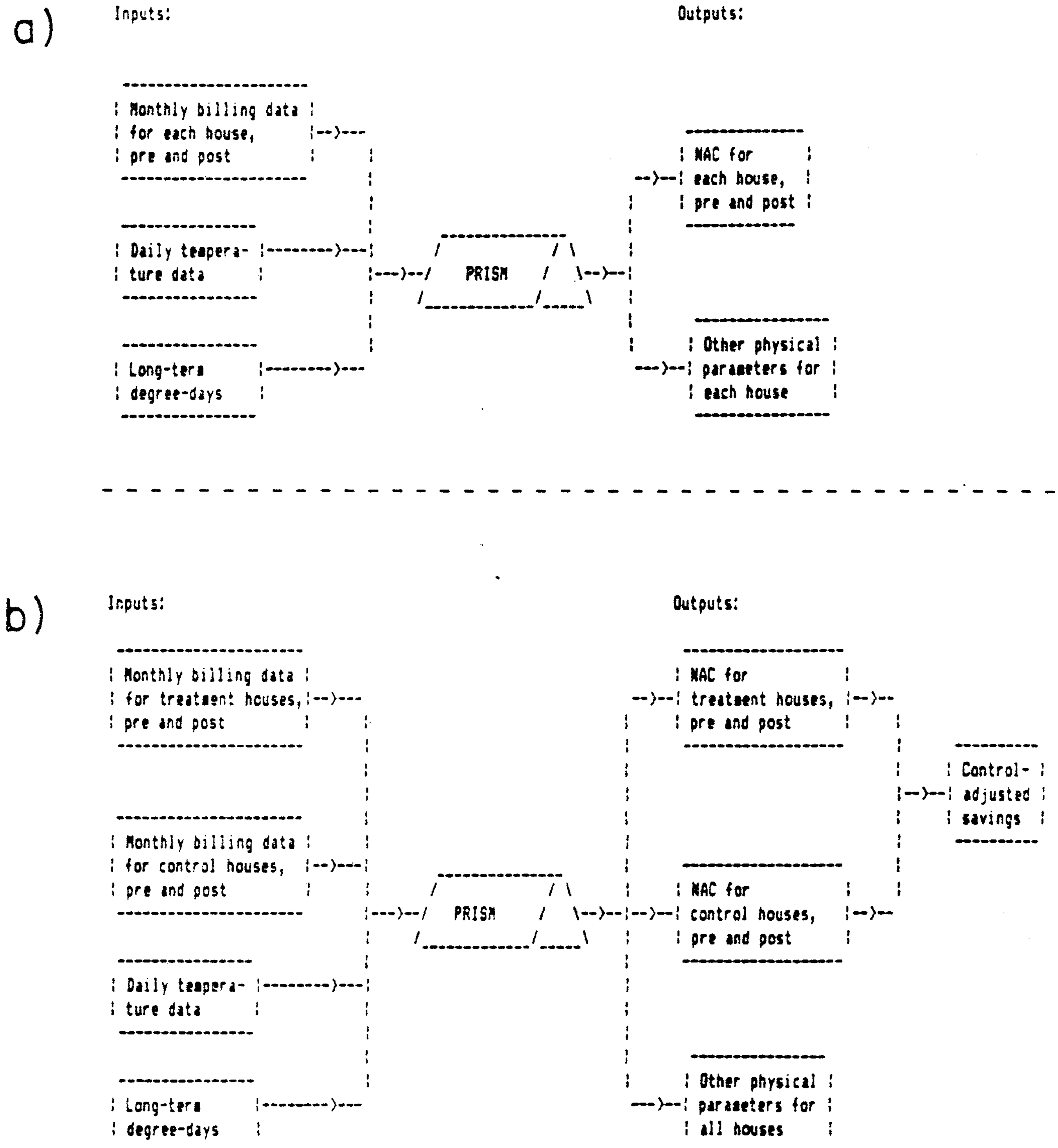

Figure 1. Schematic diagram showing the data requirements for the Princeton scorekeeping method and the estimates that result from it: a) the basic procedure for a set of houses; b) the procedure when a control group is included.

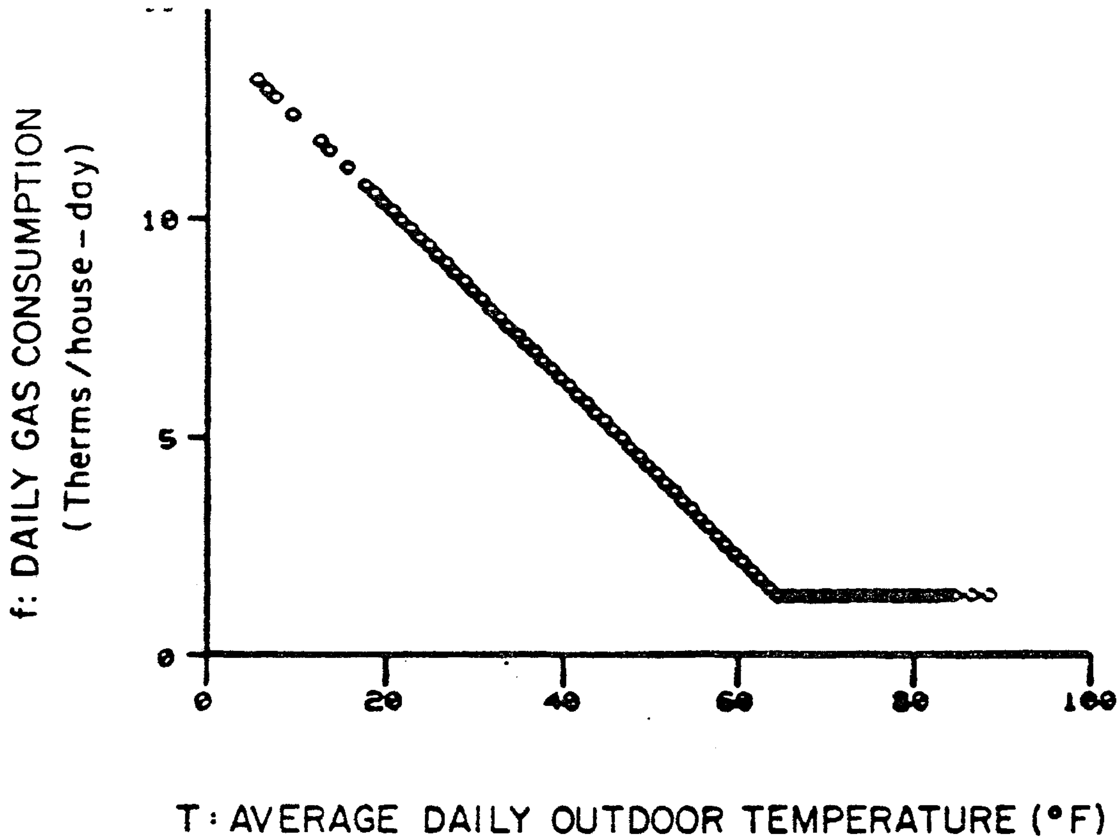

Figure 2. Daily gas consumption (f) as a function of outside temperature (T), for a single idealized house.

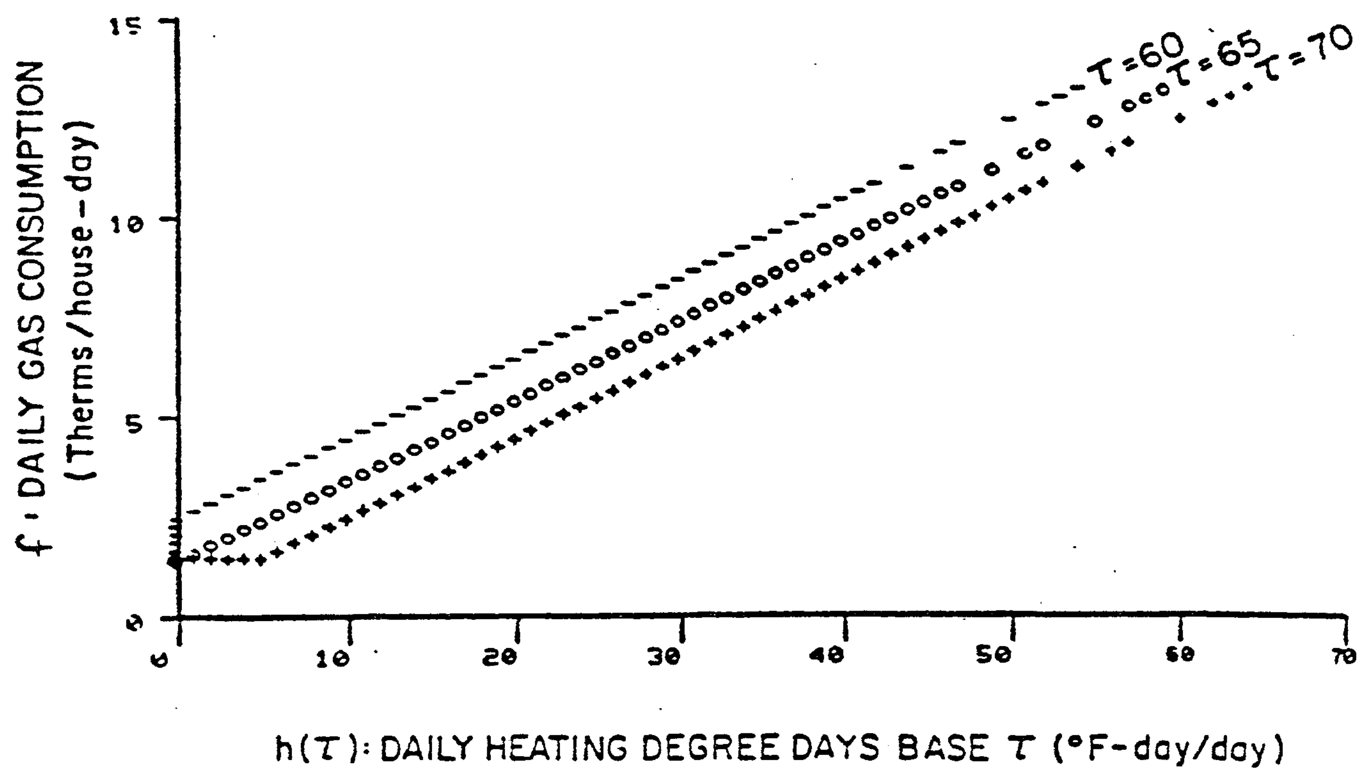

Figure 3. Daily gas consumption (f) as a function of degree-days base τ, for a single idealized house. In these plots, the three curves correspond to the same consumption and temperature data, with degree-days calculated at different bases τ.

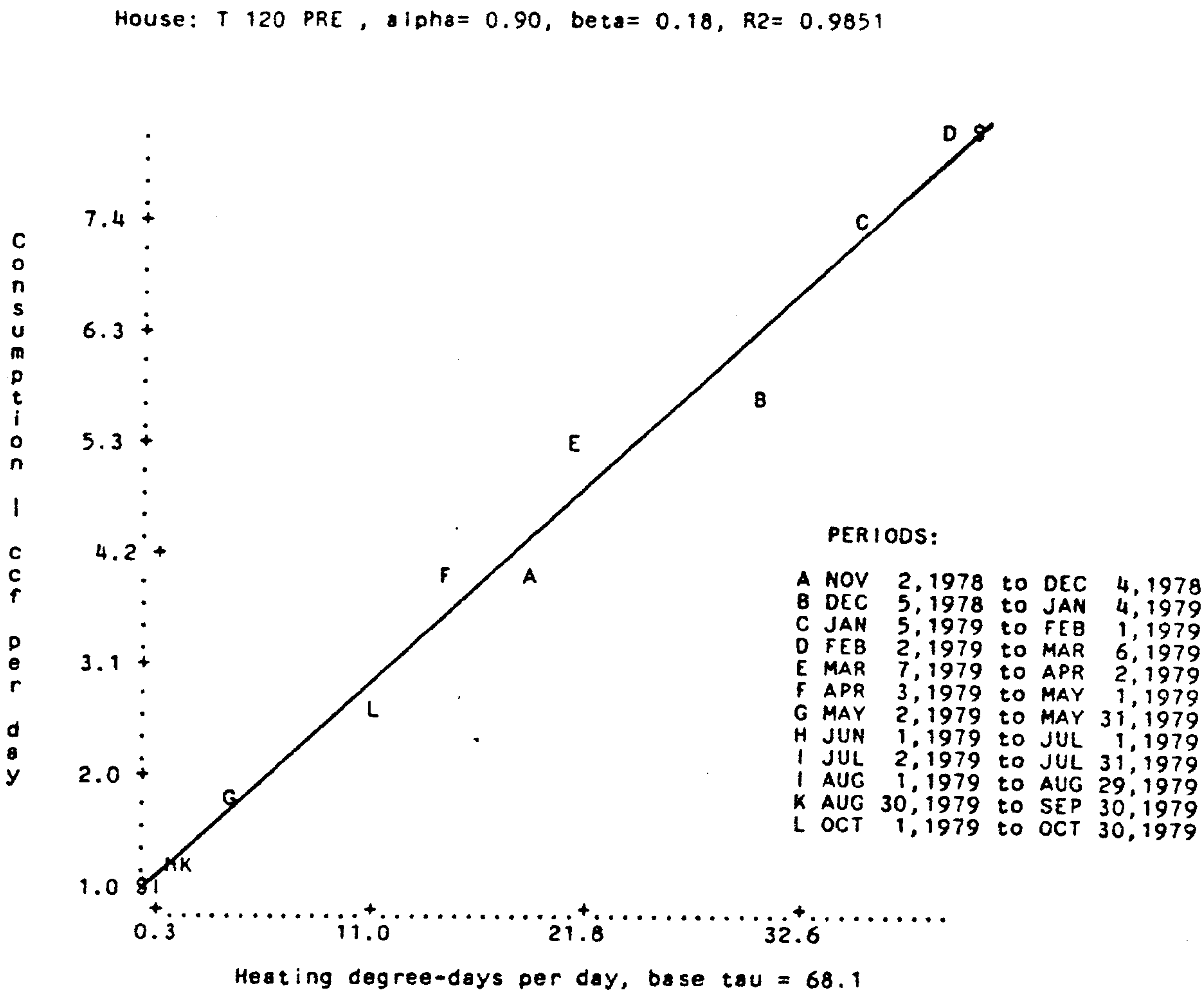

Figure 4. Consumption data (F_i) plotted against heating degree-days base τ, i.e., $H_i(\tau)$, for sample gas-heated house in New Jersey. Heating degree-days to best τ ($68^{\circ}F$) are shown. The straight line results from fitting the model to Eq. 3.

GPU

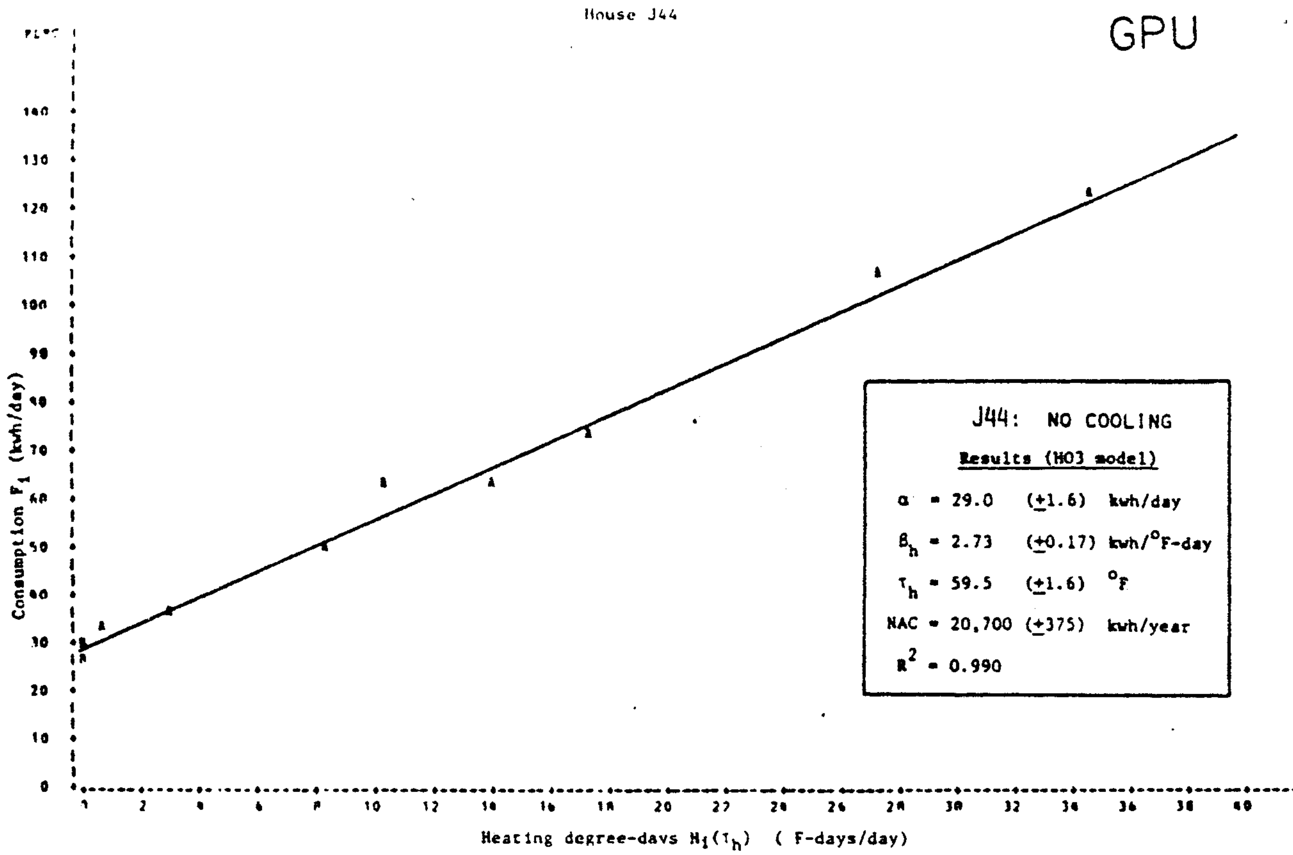

Figure 5. Consumption data plotted against heating degree-days, base τ, for sample electrically heated house in New Jersey. Estimates shown are obtained from the heating-only model in Eq. 3. The straight line represents the best-fit model, whose parameters are indicated.

GPU

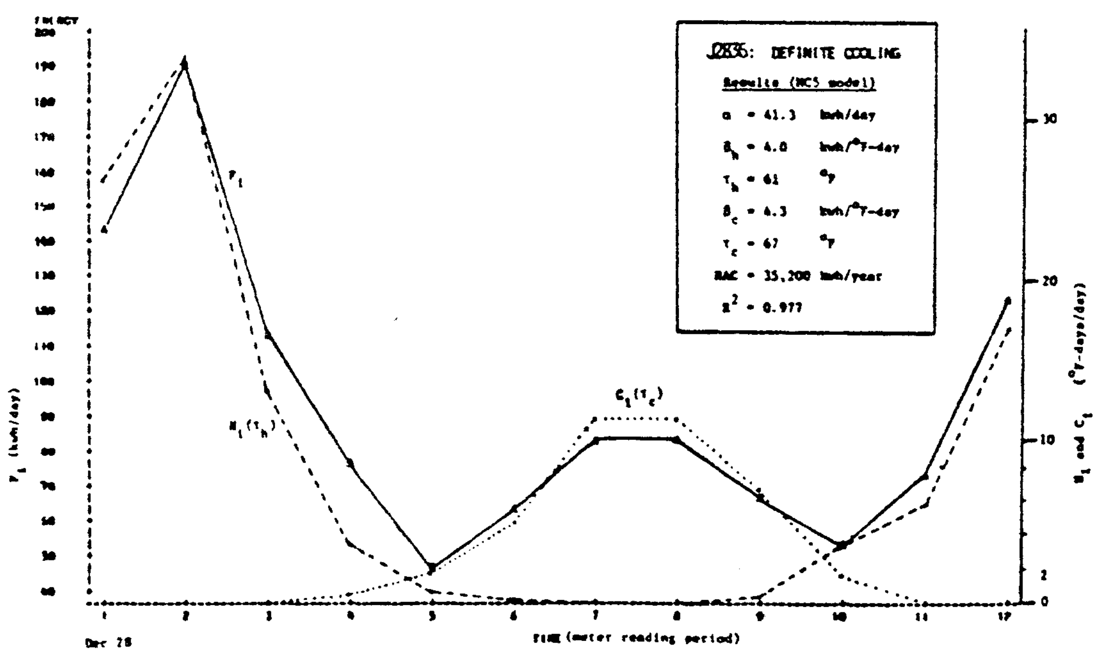

Figure 6. Superposition of degree-day data on consumption data, as a function of meter reading period, for sample electrically heated and cooled house in New Jersey. Estimates shown are obtained from the heating-plus-cooling model. H_i [---] and C_i [...] are computed respectively to base τ_h and τ_c. Correspondence of degree-day scale on the right with consumption scale on the left was set by eye.