Lawrence Berkeley National Laboratory
Recent Work

Title

Predictability and Persistance of Demand Response Load Shed in Buildings:
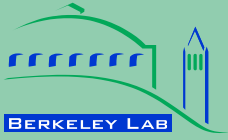
Permalink

https://escholarship.org/uc/item/5sc04804

Authors

Price, Phillip N.
Addy, Nathan
Kiliccote, Sila

Publication Date

2015-07-24

Predictability and Persistance of Demand Response Load Shed in Buildings

Phillip N. Price, Nathan Addy, Sila Kiliccote

Energy Technologies Area

July 24, 2015

DISCLAIMER

Predictability and Persistence of Demand Response Load Shed in Buildings

Phillip N. Price, Nathan Addy, Sila Kiliccote
Lawrence Berkeley National Laboratory
Berkeley CA 94720

July 24, 2015

Abstract

We analyze data from 36 commercial and government buildings that participated in a Demand Response (DR) program in California, to investigate the extent to which DR load shed in each building depends on outdoor air temperature, and whether the load shed varies systematically from year to year. Our baseline model has substantially lower error than other standard models but uncertainty in the load shed is still an impediment to addressing these questions. The model is accurate enough in 29 buildings to be used to investigate the relationship between outdoor temperature and the DR load shed, and data availability and accuracy are sufficient to investigate year-to-year persistence of load shed in 19 buildings. We find that for buildings in this dataset, most buildings shed several percent of their load during DR events. In about two thirds of buildings, higher outdoor air temperature lead to slightly reduced load shed. Year-to-year changes in load shed were generally small, except that in several buildings the load shed was small or nonexistent in the first year of participation in the program and increased subsequently.

1   Introduction

"Demand Response" (DR) means a temporary change in building electric power usage, called "load," in response to a signal from an electric utility or aggregator. There are several types of DR programs, most of which involve the utility or aggregator declaring a "DR event" for a period of a few hours, a few times per year, during which program participants are required or expected to reduce their electric load compared to normal operations. In this paper we report results from one specific program and consider only the energy used during the DR event, making no distinction between cases in which load was shifted to other times outside the DR period and cases in which load was eliminated altogether. The amount by which the load is reduced during the DR period, compared to what the load would have been under normal building operations, is called the "load shed".

We analyze data from 36 buildings that participated in a specific Demand Response (DR) program in California at some time from 2006–2009 inclusive. All of the buildings

responded automatically to the DR signal, using a building-specific pre-programmed response that did not require human intervention.

Most of these are office buildings, or combine office space with manufacturing or laboratory space, although there are also a few retail stores and a bakery. We have data from at least two years for 25 of the buildings, although, as we discuss below, only 19 of these are useful for investigating year-to-year changes in load shed. In this "critical peak pricing" program up to 12 DR events were called each calendar year, each event including two time periods: (1) a period from 12–3 PM during which the electricity price was three times the normal price, and (2) a period from 3–6 PM during which the electricity price was five times the normal price.

The DR strategy in all but one of the buildings included either a "global temperature adjustment" — setting the thermostats to allow the indoor temperature to rise to a higher level than normal throughout the building — or some other manipulation of the cooling and ventilating system. Some of the buildings also shut off some lighting or took other actions. The only building that did not manipulate indoor temperatures or cooling equipment operation as part of the DR response was the bakery, whose response was to shut off the "pan washer" (essentially a large dishwasher).

We investigate the following questions for buildings that participated in this DR program:

1. To what extent, if at all, can the DR load shed be predicted from outdoor temperature?

2. Does the DR load shed in a building tend to change systematically with time; more specifically, does the load shed tend to decrease, increase, or stay about the same from year to year?

3. Does the relationship between outdoor air temperature and building electric load tend to change systematically with time?

The latter two of these questions are motivated by the possibility that DR methods may eventually be adopted as part of normal building operations. This would likely be a benefit to both the utility and the customer: the building's peak load will be reduced, as desired by the utility, and the customer will likely save money through reduced energy costs during normal operations. However, such a migration of DR responses into normal building operation will generate smaller load sheds than if the same actions were taken only during DR events. For example, most of the buildings shed load largely by increasing the thermostat set point so that less air conditioning is provided. If building operators discover that occupants are tolerant of the resulting indoor temperature, which is higher than had previously been normal in the building, they may permanently increase the thermostat set point. In this case, the building will use less energy on hot days because the relationship between load and outdoor temperature will become weaker, but the DR load shed will also decrease because (1) the reduced indoor-outdoor temperature difference will reduce the impact of a given change in thermostat setpoint, and (2) the setpoint of normal operation is now closer to a temperature that occupants are unwilling to tolerate, so the magnitude of the setpoint change during a DR event will have to be reduced.

2 Data

2.1 Full Data Set

We acquired load data for each of the 36 building from Pacific Gas and Electric (PG&E). The data are a from a convenience sample of large commercial customers who participated in at least one year of an automated DR program from 2006–2009 inclusive. The dataset comprises 12 office buildings; 8 buildings that combine office and manufacturing space; 6 that combine office and medical laboratory space; 2 that combine offices and other uses; 5 large retail stores; and a museum, a jail, and an industrial bakery.

We obtained hourly outdoor air temperatures for the ZIP code for each building's location from the company Weather Underground.

2.2 Reduced Data Sets

As we discuss in the Results section, we find that the baseline uncertainty in some buildings is large, which leads to large uncertainties in the load sheds for the individual DR events. Data from those buildings are not useful for quantifying load shed, or for quantifying the relationship between load shed and temperature, so most of our analyses will exclude the seven buildings for which the standard error of the baseline predictions exceeds 8% of the baseline. The remaining 29 buildings are used to examine the relationship between load shed and temperature, but ten of them include data for only a single year and are therefore not useful for investigating year-to-year changes.

3 Methods

A statistical model was used to estimate the load shed for each DR event in each building. Results were used to quantify the extent to which the load shed can be predicted from outdoor air temperature. Finally, we investigated whether the load shed tends to decrease from year to year.

3.1 Baseline statistical model

The load shed is the difference between how much energy the building *would have used if not for the DR event* and the energy the building actually used. The energy the building would have used in the absence of the event must be predicted from a statistical model. See Coughlin et al. (2008) for an overview and discussion of statistical models used for baseline calculations.

We fit a statistical model to data from non-holiday weekdays that are not DR event days, and then applied the model to predict the load during the DR periods. A separate model was fit for each building and for each year, using only data from May-September (inclusive). We refer to non-holiday weekdays as "eligible days," and exclude ineligible days from the model.

The model is fit in four steps, which we summarize here briefly and then explain in detail below. Parameters were estimated separately for the 12 - 3 PM and 3 - 6 PM time periods, using the following algorithm.

1. Using only the eligible days that are *not* DR days, fit a model to predict the load as a function of the outdoor air temperature and the time of week. By "time of the week" we mean a cyclical index that counts the interval of the week, analogous to "hour of the day." Every Monday at 3 PM is the same time of the week.

2. The difference between the actual load and the predicted load is called the *residual*. Empirically, if yesterday's he prediction was too low (i.e. the residual was positive), today's prediction is also likely to be too low. Determine the relationship, on average, between the residual at a given day and time and the residual for the previous eligible non-DR day. Do this separately for days that follow an eligible non-DR day by one or two days, and those that follow such a day by three days (e.g. every Monday is at least three days removed from the previous eligible non-DR day).

3. Similarly, determine the relationship between the residual on a given day and the residual for the *subsequent* eligible day that was not a DR day: if tomorrow's predicted load is too high, then the prediction for today is also probably too high.

4. Use the relationships determined in steps (2) and (3) to predict the residual for each day from the residual for the previous eligible non-DR day, and from the residual on the subsequent eligible non-DR day. Add the average of these predicted residuals to the predicted load from step (1) to determine the final prediction. Once the statistical models have been determined by fitting to data for eligible non-DR days, the same models (and parameters) can be applied to DR days in order to predict the baseline for those days.

The model outlined above is described in detail below. We first introduce some notation. We first describe the model for predicting average load from 12 - 3 PM; the same model is used for 3 - 6 PM. The index d counts the number of days since the start of the dataset. Additionally, we need notation to refer to the day of the week: $DOW(d)$ is 1 if day d is a Monday, 2 if it's a Tuesday, and so on. For each building, $y(d)$ is the mean load from 12 - 3 PM on that day, and $T(d)$ is the mean temperature from 12 - 3 PM.

1. Using only data that are from eligible days that are *not* DR days, fit a model to calculate a load prediction $\hat{z}(d)$. The prediction is

$$\hat{z}(t) = \alpha_{DOW(d)} + \beta_L T(d) + \beta_M(T(d) - T_0) + \beta_H(T(d) - T_1) \qquad (1)$$

Here $\alpha_{DOW(d)}$ is the "day of week effect", an upward or downward shift that allows the prediction on one day of the week to differ from another day of the week even if the temperature is the same. T_0 and T_1 are temperature "change points": electric load is assumed to be linear above, below, and between the change points, but the form of the model allows the slope to change at each change point: β_L is the change in load per unit change in outdoor temperature for "low" temperatures (below T_0); β_M is an additional slope that is added to β_L for temperatures above T_0; and β_H is an additional slope that is added to $\beta_L + \beta_M$ for "high" temperatures above T_1. The errors, also called residuals, are

$$\epsilon(d) = y(d) - \hat{z}(d) \qquad (2)$$

and all parameters are chosen to minimize the sum of squared errors subject to the constraints that the change point temperatures T_i are at least 2.2 C (4 F) apart, and that at least 10% of the temperature measurements exceed T_1 and at least 10% are below T_0.

2. The predictions $\hat{z}(d)$ take weekly patterns of energy use into account, as well as the effect of outdoor air temperature. But there is autocorrelation of the residuals: if the prediction for yesterday afternoon was too high, then the prediction for this afternoon is also likely to be too high. We remove this effect, on average, by adding an adjustment to the model. For each day d, let d_- identify the previous *eligible non-DR day*, and k_- be the integer number of days since that day. (In all of the following notation, the subscript '$-$' indicates that we are looking back in time; a $+$ will indicate that we are looking ahead). So, for instance, if $d = 15$ (the 15th day of the dataset), which happens to be a Monday, then $d = 15$, $d_- = 12$ (that is the previous Friday), and $k_- = 3$. Perform separate linear regressions, with no intercept term, for days with $k_- \leq 2$ and days with $k = 3$, to determine the average relationship between the past residual and the current residual. (So, for instance, if yesterday's load prediction was 22 kW too high, what is the expected error in today's load prediction?). This amounts to determining value of $\gamma_- s$ in the equation

$$\widehat{\epsilon}_- = \gamma_{-s}\epsilon(d_-) \tag{3}$$

where $s = 1$ if $k_- \leq 2$, and $s = 2$ if $k_- = 3$. The $k_- = 3$ case occurs almost every Monday, and $k_- = 1$ occurs almost every other day of the week. The $k_- = 2$ case occurs following a mid-week holiday or DR day. We do not use data from $k_- \geq 4$ since (1) there are very few such days so regression coefficients would be very uncertain, and (2) we expect the predictive value to be small at such a large temporal separation; for these days we let $\widehat{\epsilon}_- = 0$ (and similarly for those with $k^+ \geq 4$).

3. Similar to the above, but looking into the future instead of the past: For each day, let k_+ be the number of days until the next eligible non-DR day. Define $s = 1$ if $k_+ \leq 2$ and $s = 2$ if $k_+ = 3$. Perform no-intercept regressions for days with $s = 1$ and $s = 2$ to predict the residual on day d from the residual on the next eligible non-DR day:

$$\widehat{\epsilon}_+(d) = \gamma_{+s}\epsilon(d_+) \tag{4}$$

4. Calculate the final estimate for the load: start with the prediction from the change-point model, then adjust upward or downward based on the mean of the predicted residuals based on preceding and subsequent days.

$$\hat{y}(d) = \hat{z}(d) + \frac{1}{2}(\widehat{\epsilon}_-(d) + \widehat{\epsilon}_+(d)) \tag{5}$$

The procedure above is fit to data from all of the eligible non-DR days, and then used to predict the baseline load on the DR days. The estimated load shed on day d is then

$$S(d) = \hat{d}(t) - y(d) \tag{6}$$

which is defined such that a positive value means the building successfully shed load. The same procedure is used to estimate the load shed from 3 - 6 PM, and the results are summed to estimate the load shed for the entire DR period.

There is no direct way to measure the errors in the predictions on DR days since the baseline load cannot be known for those days: the whole point of DR is that the load differs from the baseline. Instead, to quantify the statistical distribution of errors we use cross-validation. For each building we use the model to predict the baseline load for non-DR days, and then calculate the error in the prediction by comparing the prediction to the actual load, which is the baseline load by definition. This procedure is applied separately for each of the 20 eligible non-DR days with highest peak outdoor air temperature and the standard deviation of the errors is determined, yielding 20 error terms e_i. Hot days are used for cross-validation because DR days usually have high outdoor temperature, so the hot non-DR days are expected to be the days with loads most similar to those on DR days. The errors on hot days may differ systematically from the errors on cooler days: specifically, the errors on hot days are probably larger in magnitude, on average, because the small number of hot days may lead to poor estimate of the β parameters in Equation 1.

The first stage of the statistical model described above is a variant of a simple regression model that has been used previously and is described in Mathieu et al. (2011). The current model differs from that model in three ways, not counting changes in notation:

1. The current model estimates a different relationship between load and outdoor temperature for the two time periods of the day. The temperature relationship varies with time of day because the internal heat load in the building differs with occupancy and other time-dependent variables, and because insolation also varies with time and is not perfectly correlated with outdoor temperature.

2. The current model uses two temperature change points, rather than a larger number of fixed "knots," in the relationship between outdoor air temperature and building load. For the present paper we are fitting only summer afternoon data so there is no need to span a large range of temperatures, and thus no need or justification for a large number of segments in the relationship.

3. The final step of adjusting for the residual from nearby days helps reduce common problems of the previous models, such as the effect of a trend or a persistent shift in load during the summer. This could also be achieved with a time-series model in which the regression parameters and autoregression parameters are estimated together, but the present model is more stable — more robust to outliers, for example — and is immune to numerical problems with parameter estimation that could be a problem for a model that combines multiple temperature change points with an autoregressive model.

The present model adjusts the initial baseline prediction upward or downward based only on residuals from earlier and subsequent days. In contrast, some previous baseline DR predictions have used a "morning adjustment", in which an initial prediction for the load during the afternoon is increased or decreased on the basis of the residual earlier on the same day. We considered that approach but rejected it because (1) some buildings

deliberately change their energy use before the DR period, usually through "pre-cooling" to reduce indoor temperatures so that the building will remain more comfortable during the DR period when less cooling is provided, and (2) in many buildings the load during the morning is very different from one day to the next, so a morning adjustment can yield poor results in some buildings even though it helps on average.

4 Results

4.1 Error in Baseline Estimates and Load Sheds

We compared our model (described above) to two other models that have been used to estimate load shed: (1) a "10/10" model in which the baseline load during the DR period is estimated to be the mean load during that time of day over the past ten eligible non-DR days, and (2) a regression model, described in Mathieu et al. (2011) and applied to DR load shed estimation in Kiliccote et al. (2010), that includes "time of week" effects as well as a piecewise-linear temperature relationship. The latter model is similar to the first stage of our model, although it assumes that the relationship between load and temperature is the same throughout the workday whereas ours estimates a different relationship for each of the two DR periods.

To quantify the statistical distribution of baseline errors we used cross-validation. For each building, load predictions were made for the period from 12–6 PM for each of the 20 hottest eligible non-DR days of each summer (that is, the days with the highest peak temperature). The primary parameter of interest is the "baseline percent error", defined as the the difference between the predicted average load and the actual average load during the time period, divided by the actual average load during the period, times 100. An error of 2% means the prediction was 2% too high; -2% means the prediction was 2% too low. When we discuss "baseline error" in the rest of this report we are referring to the percent error unless we explicitly say otherwise.

The present model outperformed both of the other models. For instance, the median error (for all of the hot days in all of the buildings and all of the years) is less than 4%, compared to 6% for the 10/10 model and 5% for the regression model of Mathieu et al. Considering each prediction individually, the absolute baseline error is lower with the present model for 65%–70% of the cases. The situation is even more extreme when considering the predicted load summed over all 20 hot days of each year, which is the error in the total energy used in the afternoons of those days: the present model had a lower error than the 10/10 model in almost 95% of the building-years, and the error is lower by a factor of two in about a third of the building-years.

Figure 1 shows the baseline error for each of the 20 hottest days each summer, for each building and each year for which we have data in the building. Buildings are numbered according to an arbitrary scheme, and are ordered in the plot so that the buildings with largest error are at the top. As can be seen in buildings near the top of the figure, buildings that have large errors in one year also have large errors in other years.
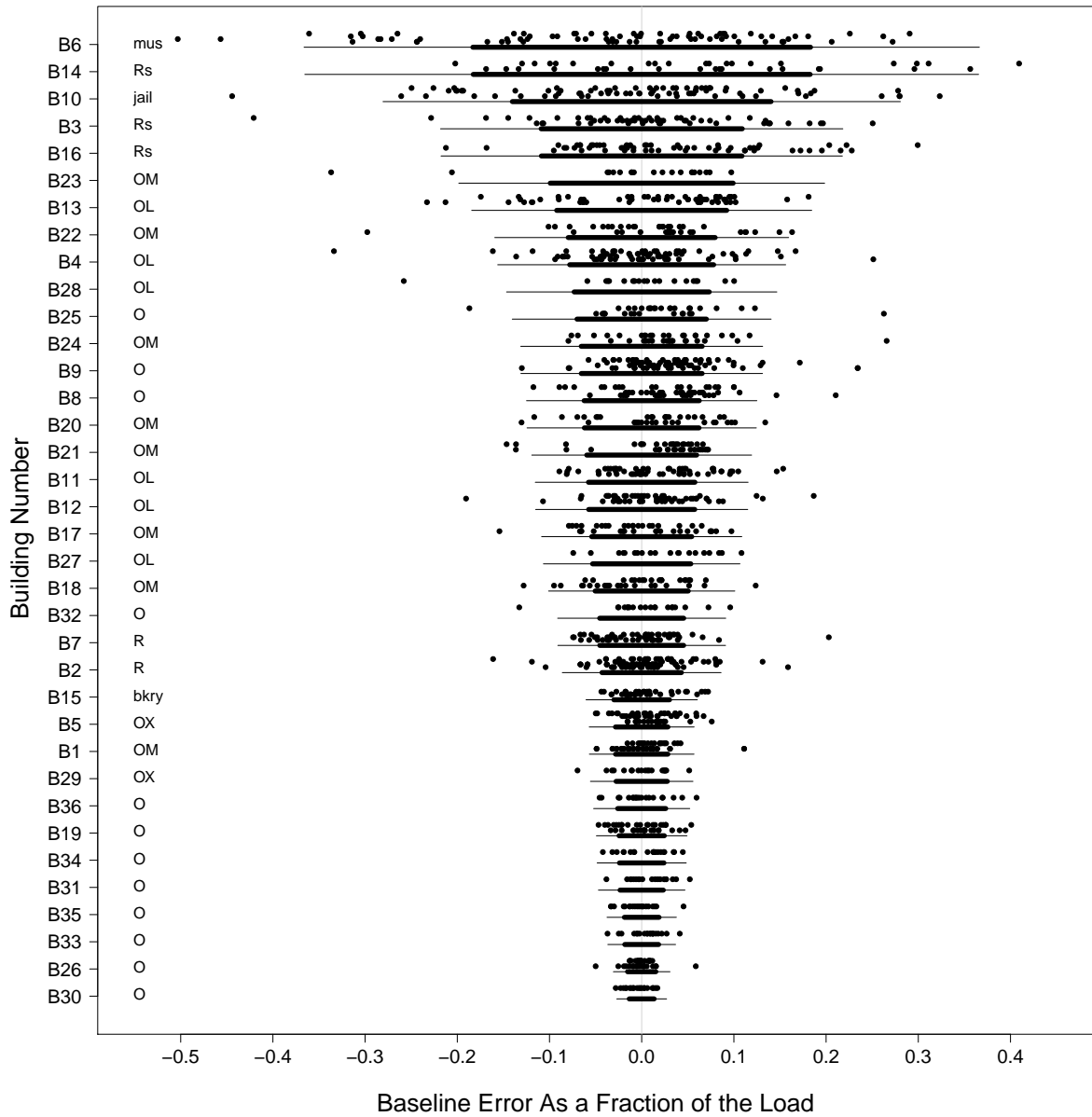
Figure 1: Error in the baseline estimate for each of the 20 hottest days of the year in each building. Building types, shown along the y axis, are: O = office, OM = office + manufacturing, OL = office + labs (medical), OX = office + other, R = retail, Rs = retail with solar panels, bkry = bakery, mus = museum, and jail = county jail. Each year's results are plotted with a different vertical offset, so a building with data from four years (2006–2009) will have four rows of points, with the lowest representing 2006. Also shown for each building is a thick line spanning ±1 root-mean-squared error (RMSE) for that building, and a thin line spanning ±2 RMSE.

The baseline percent error tends to be larger for buildings with low baseline load. Figure 2 plots the root-mean-squared percent error in the baseline versus the mean load during the afternoons of the hot days. Buildings with low load tend to be less predictable.
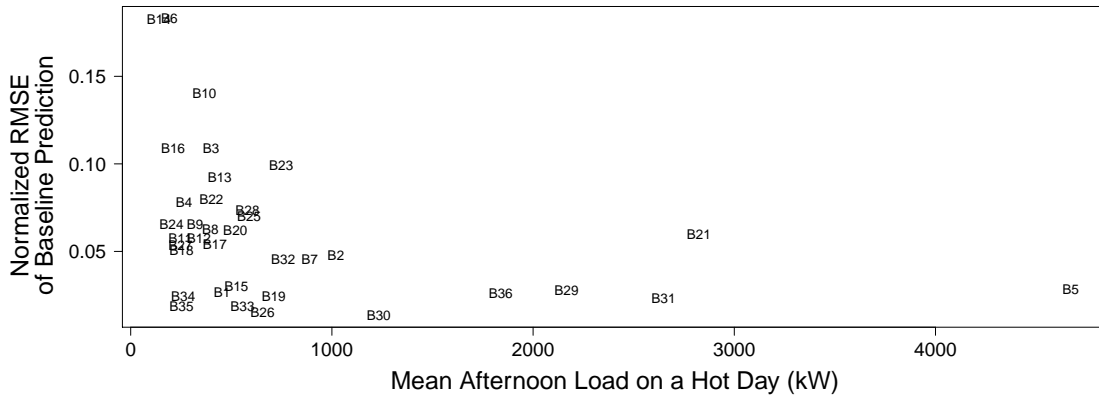

Figure 2: Normalized RMSE of baseline prediction

In addition to the RMSE in the baseline predictions, a summary statistic that is relevant is the 'bias': over all of the hot days experienced by a building in a given year, how well does the baseline model predict the average energy used? Figure 3 shows the fractional error in the baseline prediction as a function of outdoor air temperature, for each of the buildings for which the RMSE is less than 8%. (8% was chosen because larger errors make it impossible to draw firm conclusions about year-to-year variability and temperature dependence). Bias would show up in the figure as an upward or downward shift of the points away from a horizontal line.

Finally, we consider the relationship between baseline error and outdoor temperature, as shown in Figure 3 for each of the buildings with baseline RMSE < 8%. The axis scales are the same as in Figure 6, which shows estimated load shed as a function of temperature, to facilitate comparison. If the error in the baseline estimate depends systematically on temperature then the error in the estimated load shed will also depend on temperature, which will either mask or amplify the actual relationship between load shed and temperature and could even lead to an error in the sign of the relationship. The piecewise-continuous temperature dependence of the model, and the adjustment for residual error, should help eliminate any systematic effects, but it is still possible for systematic errors to creep in. For example, the temperature dependence could change during the summer in such a way that the functional dependence estimated by the model, which averages over the entire modeled period, is wrong during the hottest or coolest days.
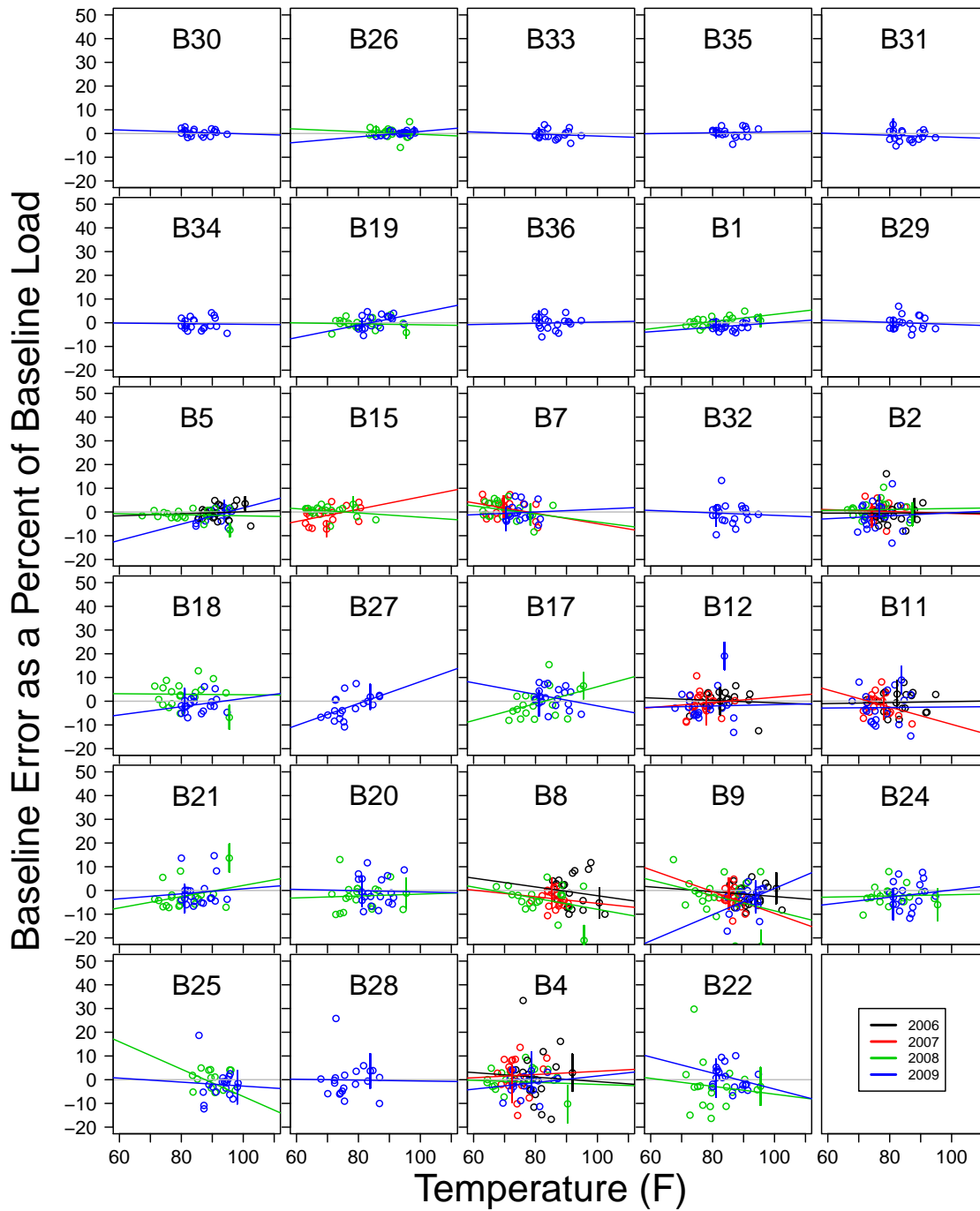
Figure 3: Percent Error in the Predicted Baseline vs Outdoor Temperature for each building for which the uncertainty in the individual load sheds is less than 8 with a regression line for each year. A representative error bar is shown for each year and each building (one standard error). Different colors are used for different years (see key at bottom right).

As Figure 3 shows, in almost all of the buildings there is no evidence for a substantial systematic relationship between baseline error and temperature. The slopes of some of the linear regression lines differ noticeably from 0, but only to an extent that is consistent with stochastic variability. For example, out of the 58 building-years of data shown in the plot the slope is "statistically significant" at the 5 percent level in only three cases, all from the year 2009: Building 5 (left side of third row), Building 27 (second from left in fourth row), and Building 9 (second from right in the fifth row). Given 58 building-years of data, a statistically significant slope is expected in three of them even if the true slopes are all zero, simply due to stochastic variability, so this result is consistent with chance.

4.2 Uncertainty in Load Shed

The load shed is the difference between how much energy the building would have used under normal operation and the actual energy used by the building. The energy that would have been used under normal operation is given by the baseline, which is not known precisely but which can be estimated, as discussed above; the actual energy used is known exactly. The error in the load shed is therefore equal to the error in the baseline estimate.

In a previous section we discussed the errors in the baseline estimates on hot summer non-holiday non-DR days. Those errors are known because the buildings were operating under baseline conditions by definition. When the same model is applied to the DR days, the statistical distribution of possible errors represents the uncertainty in the baseline prediction. We use the statistical distribution of known errors for the 20 hot days in a given building-year to estimate the statistical distribution of possible errors on DR days in the same building. We summarize the distribution by assuming it is "normal" ("Gaussian") with mean 0 and a standard deviation given by the RMSE of the load shed predictions for the 20 cross-validation days in that building year. The resulting standard deviation is graphically illustrated in Figure 1, which displays a horizontal black bar extending 1 standard deviation on each side of 0. A thinner line extends to ± 2 standard deviations, which would encompass 95% of the data — or all but one data point per building-year on average — if the assumption of normality were perfectly true. In fact the number of points outside ± 2 standard deviations is about what is expected from the assumption of normality, but there is a slight excess of outliers that are more than 3 standard deviations from zero. In short, the standard error is a useful statistic, but it is not a "sufficient statistic" to characterize the statistical distribution of errors, as it would be if the errors were normally distributed.

4.3 Load Shed in Early and Late Afternoon

Figure 4 shows the estimated load shed in the 3-6 PM DR period versus the estimated load shed in the 12-3 PM period. It appears that in almost all of the buildings the buildings responded nearly identically in both periods. The exception is Building 15 (third row, second column), which shed load only during the 3-6 PM period. In the rest of this paper we ignore the distinction between the two periods and focus on the total load shed during the entire 12 - 6 PM DR event.
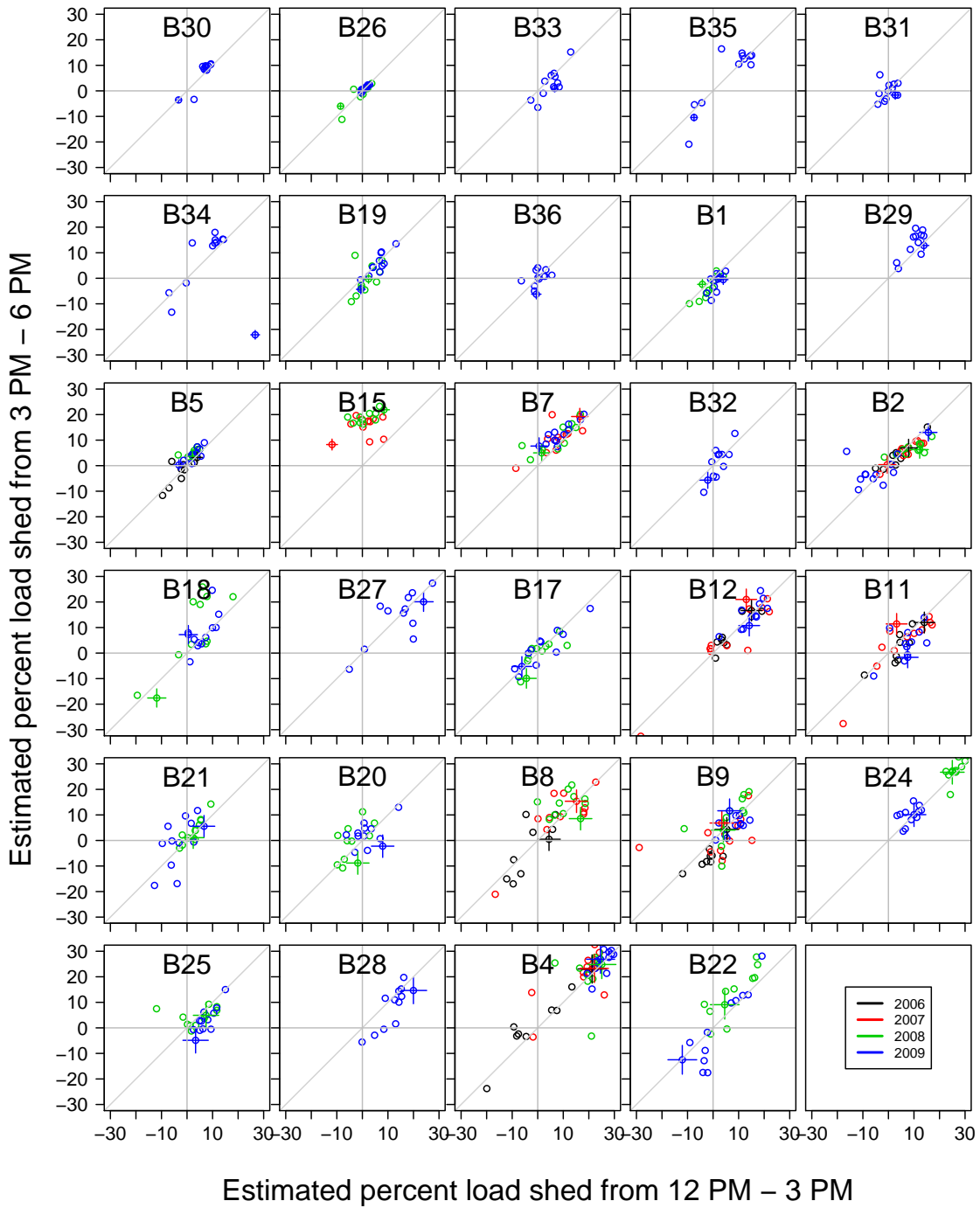
Figure 4: Estimated load shed from 3-6 PM vs estimated load shed from 12-3 PM, for each DR event, for each building for which the uncertainty in the individual load sheds is less than 8%. Representative error bars shown for one point in each year.

4.4   Variation of load shed by calendar year

As discussed above, the load shed in a building usually varies with outdoor air temperature and is also subject to variability that is not temperature-related; additionally, the load shed estimates are uncertain.

One possible approach to comparing load shed across buildings and years would be to calculate the average load shed for each building over each year. But DR events took place at different average temperatures in different years, and different buildings experience different temperatures even in the same year. A building that implements the same DR measures in different years will attain different average load sheds, because the years have different temperatures during the DR events. A simple comparison of load sheds will therefore not suffice to determine whether the effectiveness of a building's DR has changed from year to year. Instead, Figure 5 shows the load shed at an outdoor temperature of 80 F for each building in each year, predicted from the regression relationship between load shed and temperature. The uncertainty in the estimate is also shown.

Some buildings display upward or downward shifts in load shed from one year to the next. Given the fairly small number of buildings with multiple years of data, and the fairly large uncertainties in the load shed at 80F in most buildings, we think a formal analysis is not merited. Instead, we note a few facts related to the variation of load shed with time in this dataset.

1. In several of the buildings with multiple years of data, load shed was small or nonexistent in the first year but was larger in later years. These are (from left to right) Buildings 19, 5, 8, 9 and 4.

2. Only two buildings definitely experienced a substantial decrease in load shed at 80F from one year to the next: Buildings 2 and 24.

3. In two buildings with data only from 2009, load sheds were small or nonexistent. These buildings, like those mentioned in the previous items, may have attained load sheds in later years, for which we do not have data. These are Building 31 and 36.

4. Several buildings with multiple years of data provided little or no estimated load shed (or, indeed, negative estimated load shed) at 80F in any year. These are Buildings 26, 17, 21, and 20. We do not know why these buildings failed to provide load shed; perhaps they failed to configure their control systems to respond correctly to the DR signal, whether unintentionally or intentionally (for example, they may have decided not to participate because occupants objected to elevated indoor temperatures on DR days).

5. Two buildings are predicted to shed load at 80F in every year for which we have data, but did not always shed load at the higher temperatures that they actually experienced; these are Buildings 1 and 5. These buildings appear in Figures 6 and 8 in the second row, fourth column; and third row, first column. As shown in the figures, these buildings experienced temperatures substantially higher than 80F in most DR periods, and in these buildings the load shed decreases with temperature.
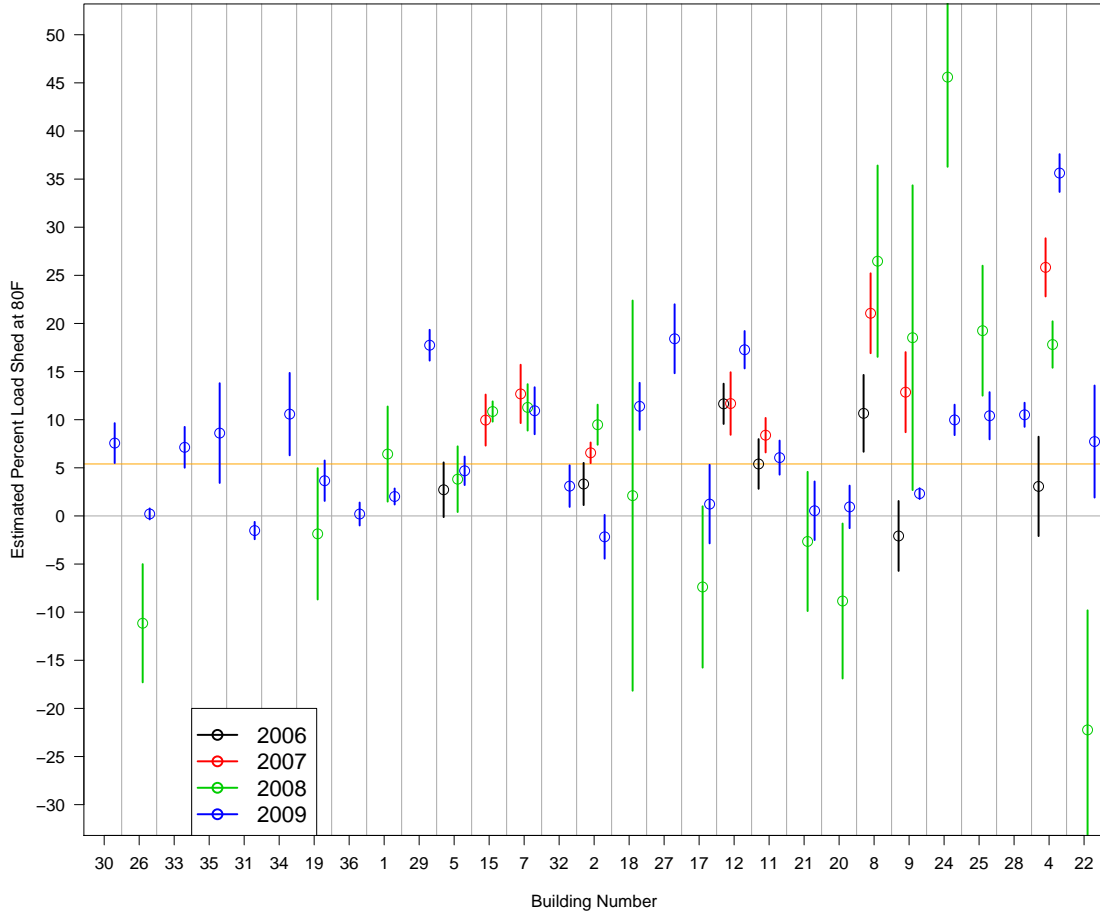
Figure 5: Estimated percent load shed at 80F, for each building for which the uncertainty in the individual load sheds is less than 8%, with bars indicating ±1 standard error. An orange horizontal line shows the median, which is about 5%.

4.5   Relationship between Load Shed and Temperature

All buildings use more electricity on hot days than on cool ones; that is the reason for the temperature-dependent terms in the baseline model. But is the magnitude of the load *shed* dependent on temperature?

Figure 6 shows the estimated load shed versus temperature for each DR event in each of the buildings with RMSE < 8%. The temperature is the average outdoor temperature during the DR period. A representative error bar is shown on one point for each building-year. Different colors indicate different years, and regression lines are superimposed. Buildings with the lowest uncertainty in the load sheds are shown at the top, with uncertainly increasing toward the bottom right, and all plots use the same scales on both axes.
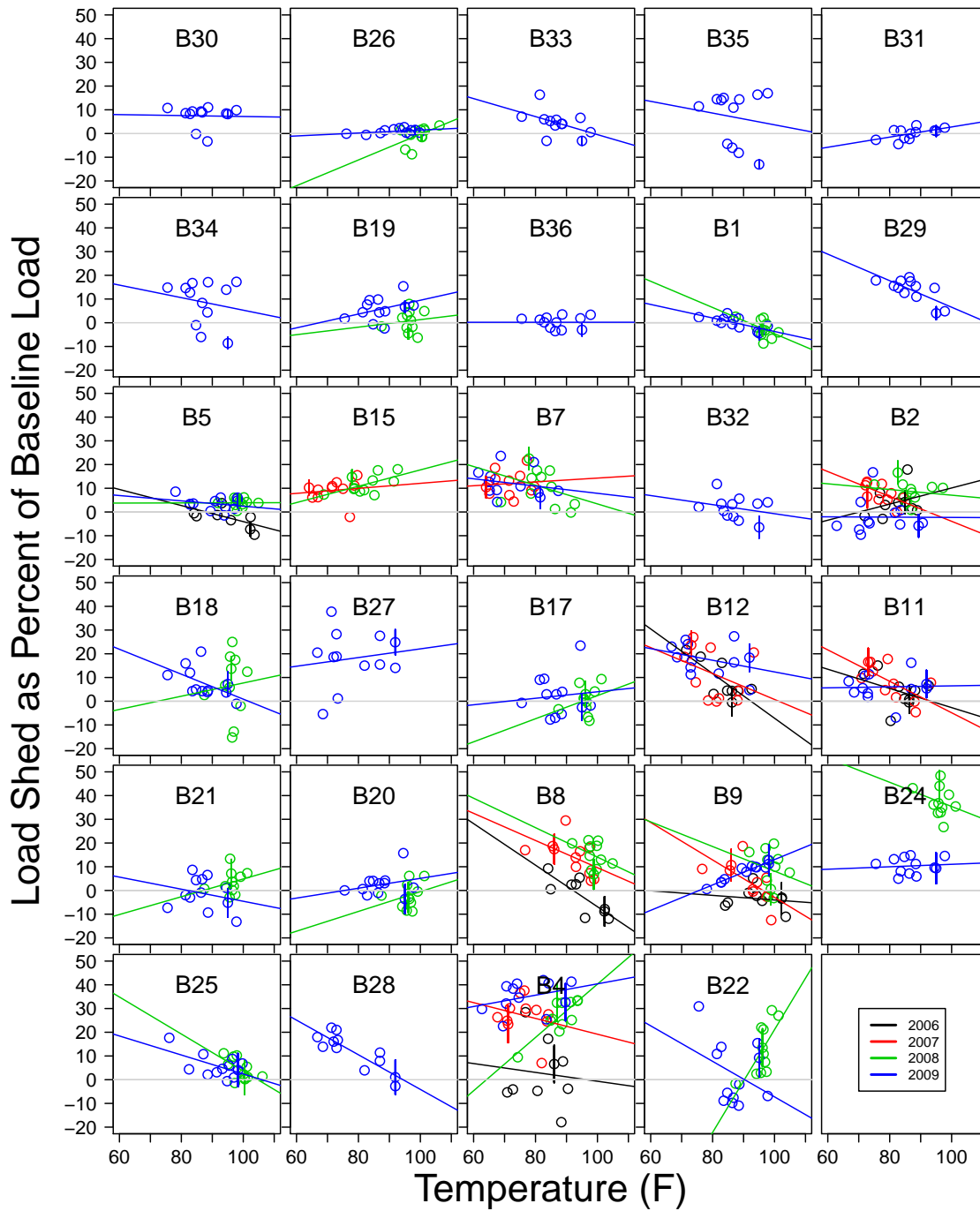
Figure 6: Percent Load Shed vs Outdoor Temperature for each building for which the uncertainty in the individual load sheds is less than 8%. A representative error bar is shown for each year (one standard error); for buildings near the top of the figure the plotting symbol is larger than the error bar. Different colors are used for different years.

The uncertainty in the regression slope varies substantially from building-year to building-year, so some of the steep lines have low uncertainty while others have high uncertainty. Figure 7 shows the slope and the uncertainty for each year, for each building whose baseline nRMSE is less than 0.08. The median slope over all of the building-years is shown with an orange horizontal line. Overall, in this dataset the estimated load shed as a percent of the baseline load decreases with temperature in more buildings than it increases.

In addition to the systematic relationship between load shed and outdoor air temperature, there are clearly other factors that affect load shed. Returning to Figure 6, it appears that some buildings may have participated in some DR events in some years, but failed to participate in others, or may even have taken actions that increased their load (thereby leading to negative load shed). As discussed earlier, the baseline model adjusts for outdoor air temperature, and cross-validation confirms there is no systematic error in the baseline predictions as a function temperature, so the negative load shed estimates are not an artifact of systematic baseline errors. However, each baseline prediction is subject to some uncertainty and it's likely that in some of the cases in which the estimated load shed is slightly negative the actual load shed was zero but the baseline prediction was slightly too high; this would lead to an estimated negative load shed. In other cases, though, the building might actually have increased rather than decreased its load in response to the DR signal. For example, Building 35 (top row of the figure, near the right side) provided substantial load shed for most DR events but shows a temperature-dependent negative load shed for four events. There are various ways such behavior can occur, including incorrect programming of the building's DR strategy, or a person overriding the DR in a way that is less efficient than simply returning the building to normal (non-DR) operation.

Some buildings appear not to have taken action during DR events in some entire years, in spite of participating in the program (and therefore paying higher electricity prices during DR events). Out of the total of 59 building-years shown in the figure, it appears that 11 may not have involved any load shed:

1. In 2006, Building 5 (third row, first column) may not have taken effective action in any DR event.

2. In 2008, Building 19 (second row, second column) may not have taken action in any DR event.

3. In 2008 and 2009, Building 26 (first row, second column) may not have taken action in any DR event. The same may be true of Building 17 (fourth row, third column).

4. In 2009, Building 31 (first row, fifth column) may not have taken any action during any of the DR events. The same may be true of Buildings 36 (second row, third column), 1 (second row, fourth column), 32 (third row, fourth column), and 2 (third row, fifth column).

Additionally, two buildings may not have taken action during some DR events in some years: In 2009, Building 30 (first row, first column) may not have taken action during two of the DR events; and In 2009, Building 35 (first row, fourth column) may

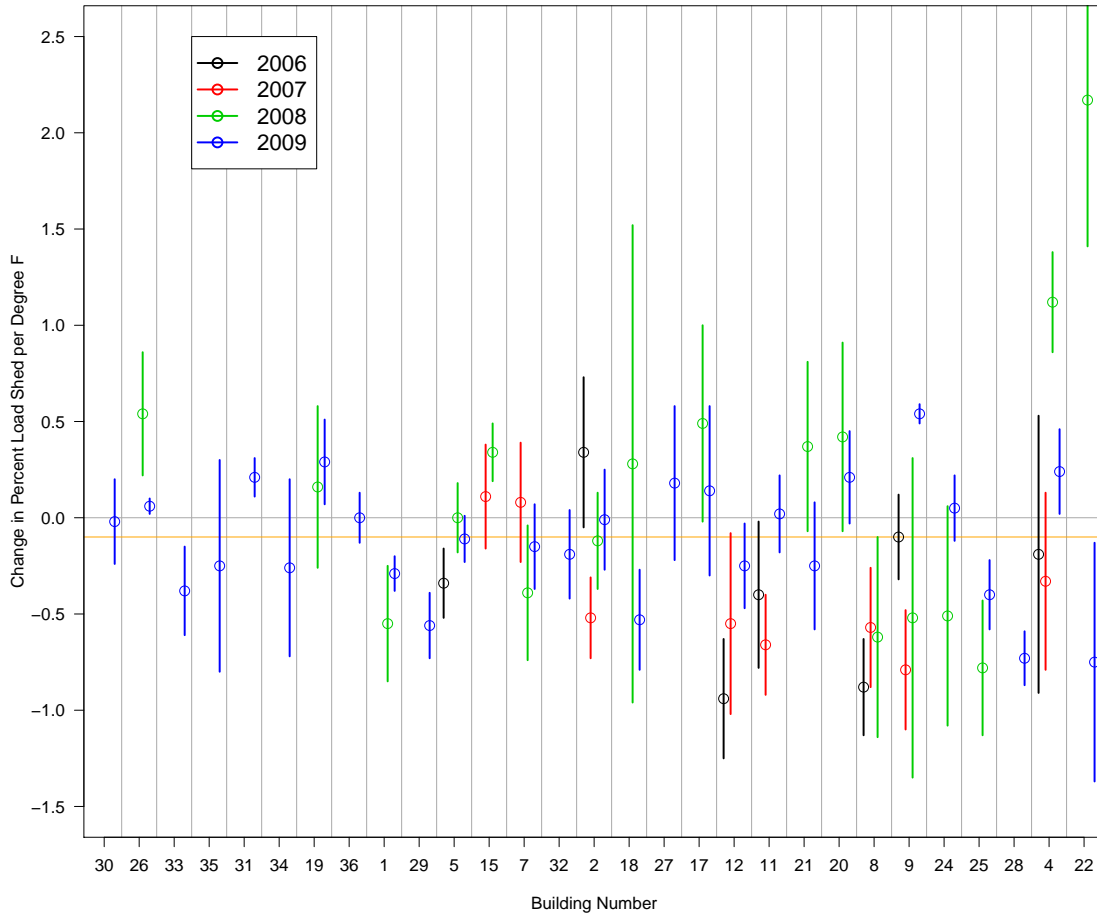not have taken action, or may have taken counterproductive action, during four DR events.

Figure 7: Estimated change in percent load shed per degree F, for each building for which the uncertainty in the individual load sheds is less than 8%, with bars indicating ± 1 standard error. An orange horizontal line shows the median, which is about -1 percent per degree F.

All of the previous discussion of load sheds expresses the load shed as a percentage of the building's estimated baseline load, which seems natural: "During this DR event, this building used 4% less energy than if there had been no DR event." One effect of expressing the shed in this way is that shedding a given amount of load — a certain number of kW — will yield a different percentage depending on the temperature and on the day of week of the event. The estimated baseline load in all of the buildings increases with temperature, so a given number of kW will represent a larger percentage of the baseline on a cool day than on a hot day. Also, in some buildings the baseline load depends on the day of the week, so shedding a given number of kW will result in a different percent decrease on Monday than on Tuesday. This behavior of the "percent load shed" metric may be desirable or undesirable depending on one's research interest,

although, as we see below, the effect is small and can probably be ignored in most cases.
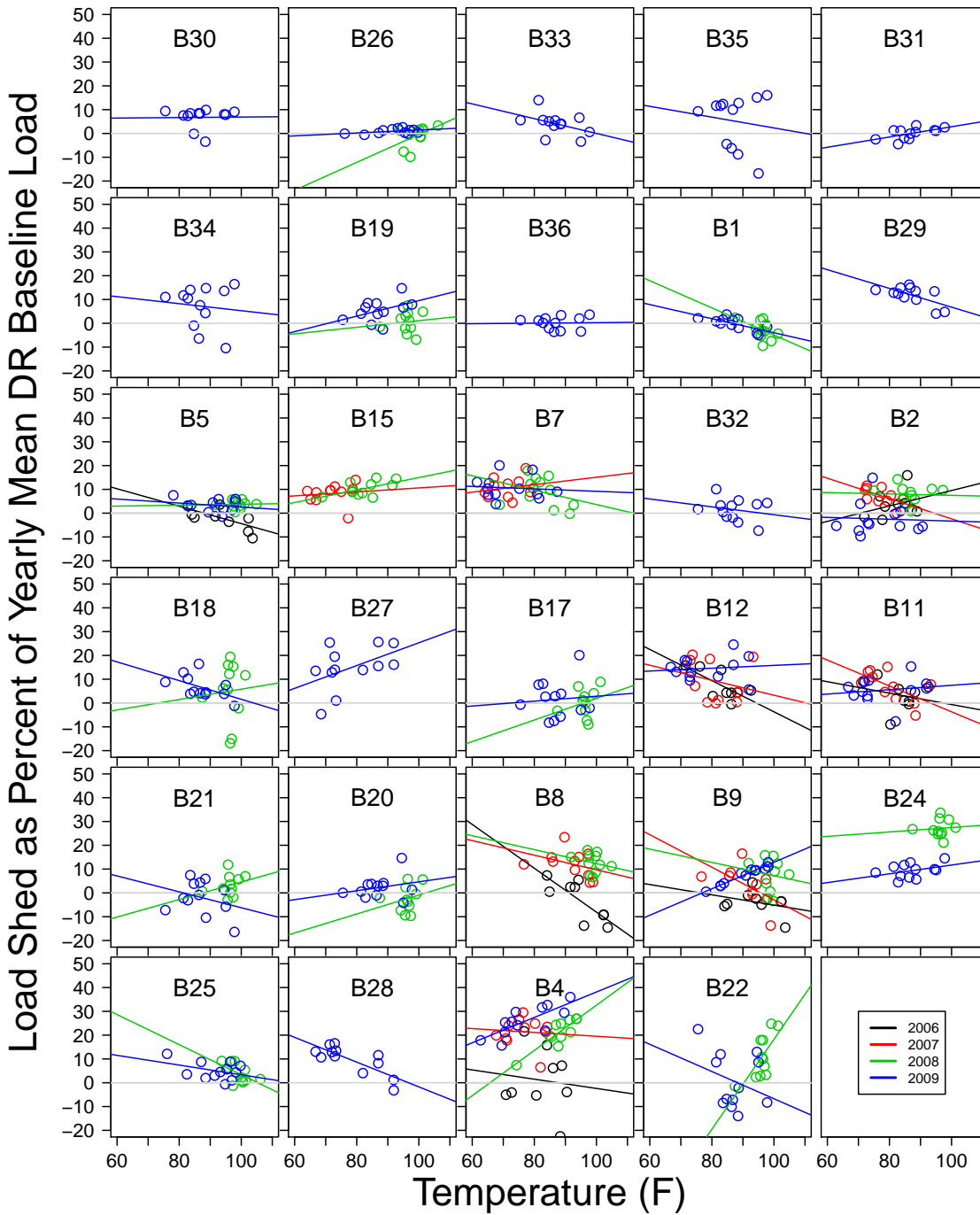

Figure 8: Load Shed as a Percent of the Mean Baseline Load during DR periods vs Outdoor Temperature, for each building for which the uncertainty in the individual load sheds is less than 8%. Different colors are used for different years.

An alternative approach would be to look at the absolute rather than relative load shed — that is, to work with the number of kW shed — but this would make it hard to compare the DR performance of buildings that vary greatly in baseline energy consumption. Instead, we define a normalized percent load shed by dividing the estimated load shed during an event (the mean kW over the course of the event) by the mean baseline load during all of the DR events experienced by that building in that year. Thus the same denominator is used for all of the events in a building during a given year, no matter what the outdoor temperature or the day of the week. The result is shown in Figure 8. Differences between Figure 8 and 6 are inconsequential in most buildings. The plots are not identical, however: compare Building 24 on both figures (fifth row, last column).

5   Summary and Conclusions

We have used a new baseline model to estimate the load shed in 36 commercial and government buildings, focusing on the 29 buildings for which the baseline predictions are most accurate. The model accounts for outdoor air temperature and for the day of the week, allows a different relationship between temperature and load for the early and late afternoon, and adjusts for temporal autocorrelation of errors.

In this dataset there is no evidence that DR effectiveness tends to decline from year to year as some have speculated: for most buildings with more than one year of participation, the load shed at 80 F is as high or higher in later years as in the earliest year of participation.

About two thirds of the buildings in this dataset have a slightly negative relationship between load shed and outdoor air temperature; that is, higher outdoor temperature leads to slightly lower load shed. This is true both absolutely and relative to the baseline load for the event.

6   Thanks

References

[1] Coughlin K, Piette MA, Goldman C, and Kiliccote S. Estimating Demand Response Load Impacts: Evaluation of Baseline Load Models for Non-Residential Buildings in California. Lawrence Berkeley National Laboratory report LBNL-63728, January 2008.

[2] Kiliccote S, Piette MA, Mathieu J, and Parrish K. Findings from Seven Years of Field Performance Data for Automated Demand Response in Commercial Buildings. Lawrence Berkeley National Laboratory report LBNL-3643E, May 2010.

[3] Mathieu JL, Price PN, Kiliccote S, and Piette MA. Quantifying changes in building energy use, with application to Demand Response. IEEE Transactions on Smart Grid 2:507-518, 2011.