NorthWrite Energy Modeling

Introduction and Summary

This paper describes NorthWrite's approaches to energy modeling, including but not limited to linear regression.

NorthWrite's models are very accurate. The following table compares a typical high-quality regression model with a NorthWrite model for a reference data set. The NorthWrite model meets or exceeds the performance of a typical regression model for accuracy.

Time Period	Statistic	Typical Regression	NorthWrite
Occupied Hours	R^2 (higher is better)	0.776	0.833
	CV-RMSE (lower is better)	6.0%	5.2%
Unoccupied Hours	R^2 (higher is better)	0.211	0.645
	CV-RMSE (lower is better)	22.5%	15.1%

In addition to the accuracy, NorthWrite's models require almost no configuration, reducing setup time and cost while also minimizing the chances for errors.

NorthWrite Energy Modeling Experience

NorthWrite personnel have great understanding of the strengths and weakness of various models through experience developing the NorthWrite platform, and also through participation in professional committees and significant prior experience. Two NorthWrite employees independently programmed the ASHRAE regression approaches on different software platforms. NorthWrite personnel participate in ASHRAE committees for energy modeling and measurement and verification. The Bonneville Power Administration (BPA, a division of the United States Department of Energy—USDOE) updated their measurement and verification protocols in 2010 and 2011. A current NorthWrite manager was the primary author for five of the ten documents developed for the BPA, including the *Verification by Energy Modeling Protocol*, which described the use of linear regression models for measurement and verification, and the *Regression for M&V: Reference Guide*, which went deeper into regression statistics and the requirements for linear regression to be valid. That manager also wrote the technical portions of a successful proposal to the California Energy Commission to develop a desktop-based measurement and verification tool for a larger project managed by Lawrence Berkeley National Laboratory.

Baseline Model Purpose

There are at least two important uses for a baseline model. These uses are related, but distinct:

1. Ongoing tracking of performance, looking at short-term intervals
2. Measurement and Verification (M&V) of the impact of energy savings improvements, whether retrofit, operational, or behavior-based.

In general, these two purposes shouldn't necessarily use the same type of model. A different model type is optimal for each use. To handle these two types of applications—performance tracking and M&V—NorthWrite has available multiple types of models. NorthWrite's platform is designed to easily incorporate any type of data-driven model.

Energy Modeling Approaches

Independent Variables

Baselines can be developed using a variety of approaches. In most cases, there will be a relationship between energy use and one or more of the following parameters:

- Weather. In typical order of importance, the weather variables are
  1. ambient dry bulb temperature
  2. ambient wet bulb temperature
  3. other weather conditions (sun, wind)
- Time Factors. In typical order of importance, the time factors are
  1. Day of the week (or day of the week groupings, i.e. "daytypes")
  2. Hour of the day (or hour of day groupings, i.e. occupied and unoccupied hours)
  3. Season of the year

For the weather variables, dry bulb temperature is by far the most important variable. Wet bulb temperature tends be related to dry bulb temperature, and hence the *additional* information it contributes is relatively minor.

For the time variables, depending upon how the baseline is to be used, Hour of the day or Occupancy may be very important, or not important at all. If the baseline is to be used for measurement and verification, sub-daily data may not be necessary at all. Models can be simplified, and in many cases be more accurate, by aggregating hourly or sub-hourly data to daily totals. However, this limits the utility of the baseline for performance tracking and also can lengthen the time (days of data) needed to create an accurate baseline.

These two types of independent variables—weather and time—can be more broadly described as continuous or categorical (or bin). Good energy modeling approaches usually incorporate bin-type variables.

Time Categorization

Linear regression models use bins as well as continuous variables. An example is creating different regressions for different day types, since weekday and weekend energy use are often different. This table shows how bins can be used to create models requiring and providing varying levels of detail:

Table 1. Examples of Time Categories and Models

Reference. Categorization Level	Number of Regressions	An Individual Regression Includes	Each Data Point in a Regression Represents
1	1	all data	Daily Energy Use
2	1	all data	Hourly Energy Use

3	1 to 7	data for a single Daytype	Daily Energy Use
4	1 to 4	data for a single Occupancy period (Occupied, Unoccupied, startup, shutdown)	Hourly Energy Use
5	7	data for a single Day of the Week	Daily Energy Use
6	24 to 7*24	data for a single Hour of the Day, for a single Daytype	Hourly Energy Use
7	7*24	data for a single Hour of the Day, for a single Day of the Week	Hourly Energy Use

As mentioned above, seasonal changes are also possible because of equipment being enabled or disabled seasonally. This is usually not an issue because equipment is typically enabled during times it could be needed.

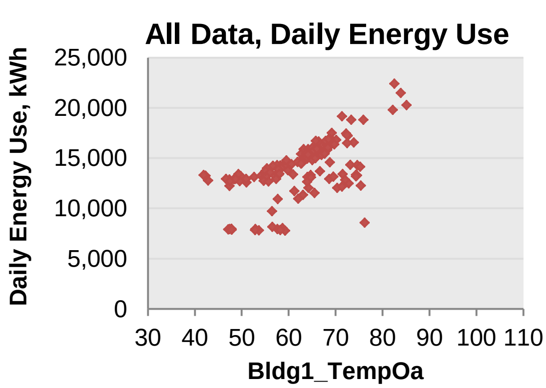
Note that by using multiple regression, it may not be necessary to have the number of regressions shown in the first column. However, some approaches to multiple regression—including the approach documented by ASHRAE—result in all models having the same slope. This is usually not correct. For example, models for weekdays and weekends should usually have different slopes.

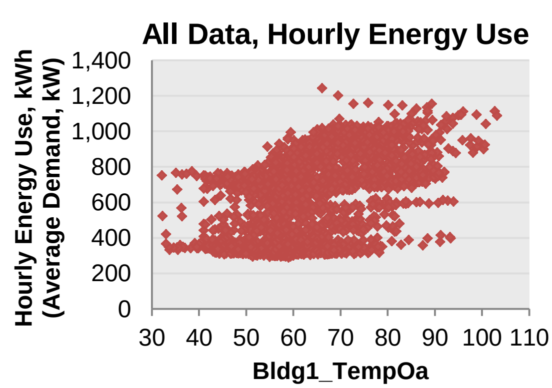
NorthWrite's default modeling approach is a model based on data bins, developed for the USDOE by scientists at the Pacific Northwest National Laboratory (PNNL). It uses regression between bins and for extrapolation outside the range of known data. The common independent variables are ambient temperature, day of the week or daytype, and hour of the day.

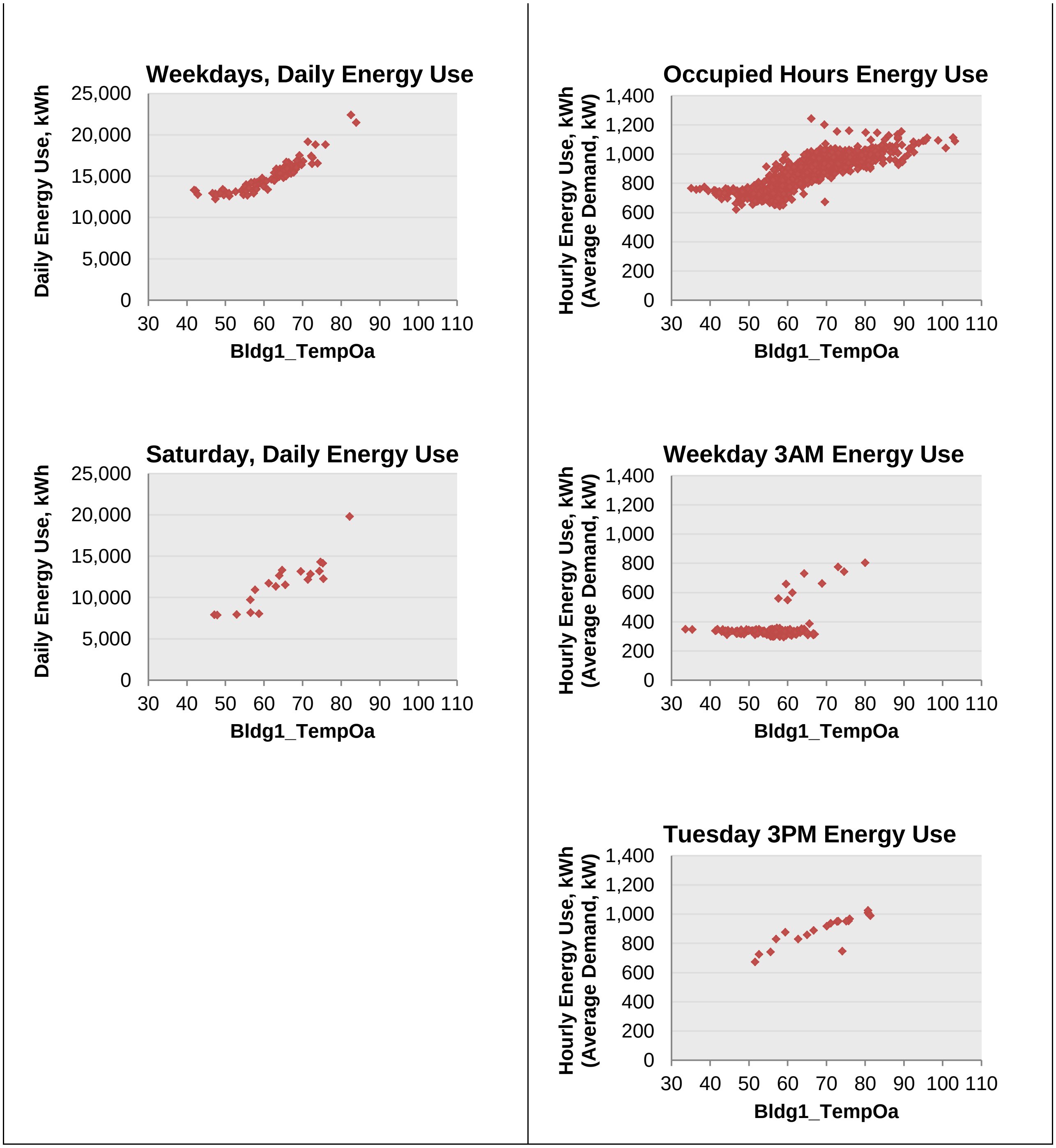
As implied by Table 1, there is usually some minimal level of time categorization that is needed for a good model. Since the NorthWrite models have a high level of categorization, the models are highly accurate. Here are some graphs from a typical office building that shows how the increasing levels of categorization are used to segment the data and improve a model:

Figure 1. Increasing Segmentation of Data by Time Categories

Daily Energy Use Models	Hourly Energy Use Models

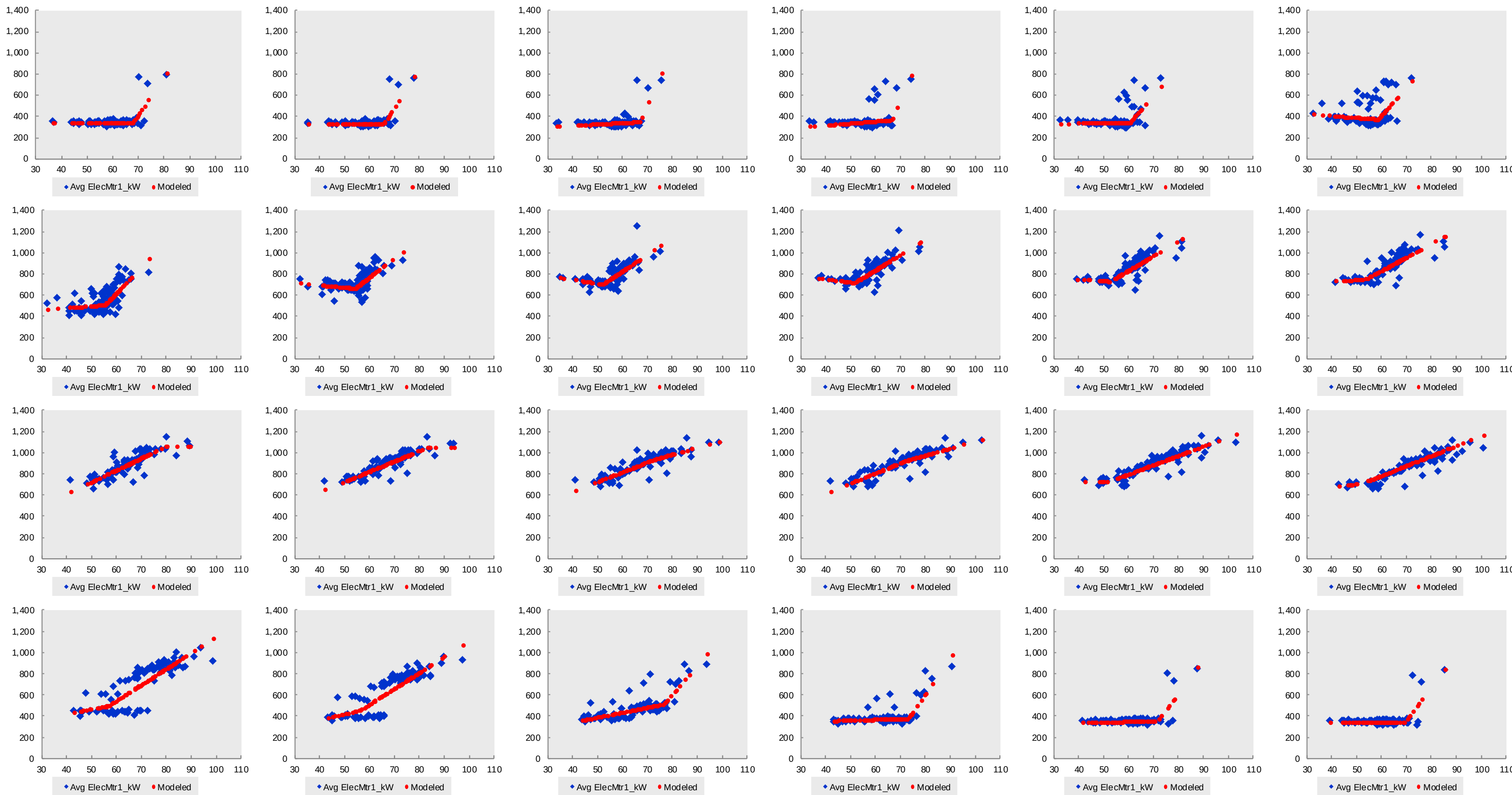






Of course, there are limitations to the categorization. With increasing categories there is a corresponding increase in the number of ‘sub-models’ required, and each sub-model will have fewer points. Figure 2 shows data and models, where each individual regression shows the data for a single hour of the day, for a single daytype. Data shown is for Weekdays; another 24 sub-models would be needed for the Weekend daytype. Data is shown in blue; the change-point regression model (see following paragraphs) predictions are shown in red.

Figure 2. 4-Parameter Change Point Regression Models for Each Hour, for Weekdays



Measurement and Verification Applications

For M&V purposes, the proper use of linear regression can be an appropriate way of estimating a baseline. However, linear regression should not be confused with a model that is a straight line relationship. Even such well-known approaches as plotting energy use vs. degree-days are appropriate only in limited circumstances. The best applications of linear regression at the building or meter level are with the approaches documented in ASHRAE Research Project 1050, *Development of a Toolkit for Calculating Linear, Change-point Linear and Multiple-Linear Inverse Building Energy Analysis Models.* These models typically use average daily temperature and daytype as the independent variables, and total daily energy use as the dependent variable.

ASHRAE 1050-RP describes five linear regression and change-point linear regression models that use a single independent variable, plus variations of these for multiple independent variables.

These regression models are the best accepted of regression models for M&V purposes. However, they do have a number of weaknesses:

1. They require user interaction to select the model type
2. They don't accurately represent the behavior of HVAC economizers, especially with sub-daily data
3. They don't accurately represent the behavior of variable speed cooling (or likely variable refrigerant flow) equipment.
4. Energy data from buildings doesn't meet the statistical requirements for regression. This does not impact the ability of the model to predict total energy use over time, but it can decrease its ability to identify to detect slight increases in energy use.

NorthWrite's default models address all of these issues. (1) A model is directly built on past data from similar conditions, and doesn't require user interaction. (2) and (3) The model follows the data, not any a priori assumption about what shape the model should take. (4) NorthWrite creates two models, one for measurement and verification and one for performance tracking.

M&V looks at energy use after a change and compares it to what the average energy use would have been, under the same conditions (of weather, occupancy, production, etc. as appropriate) in the absence of the change. The projection of "what the energy use would have been" is essentially an average of the baseline data corresponding to those conditions. Regression models include data far from the corresponding conditions to try and fit a pre-specified shape; NorthWrite's models only include data from near the corresponding conditions, and don't use a pre-specified shape.

Performance Tracking Applications

NorthWrite uses a different statistical parameter for performance tracking than for M&V. For performance tracking, NorthWrite uses the median instead of the average. This means that the model is based on the typical energy use—the value that is in the middle of all the values, rather than the average, which is biased by unusual conditions. Using regression to do this would require either a 2-pass approach where the first pass is used to eliminate the outliers or unusual energy use, or a different method of "fitting" the model to the data than ordinary regression.

Consider a baseline condition where lights or equipment are occasionally left on during an unoccupied period. A model of the average energy use over time must include these times of higher energy use. But these times should not be included when looking at *expected* energy use. After all, we shouldn't *expect* that unneeded equipment is left on.

Figures 3 and 4 show load profiles of electricity demand for the example office building, for July and August, respectively. These charts show the average and median values each hour, as well as the distribution of all values. In Figure 3, note that the distribution of demand for the overnight hours shows that some nights had much higher demand than other nights, as a result of equipment or lights being left on all night. Consequently the average value is higher than the typical, or median, value. The extra equipment was consistently turned off overnight in August, as shown in Figure 4.

Figure 3. July Load Profile of Electricity Demand

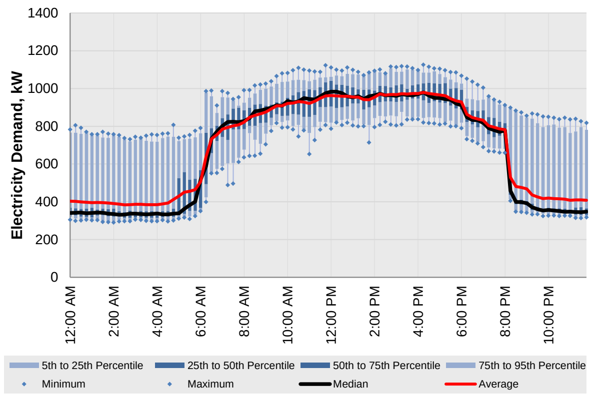
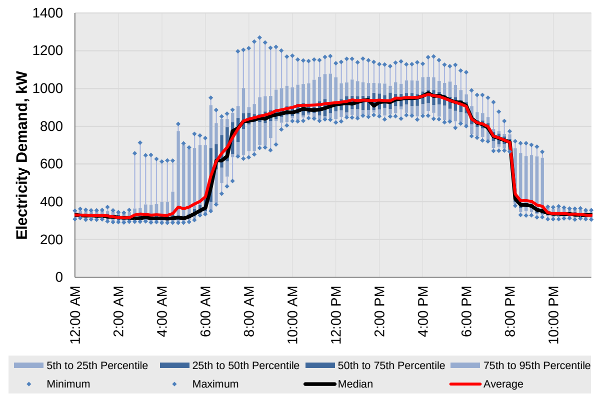


Figure 4. August Load Profile of Electricity Demand



The improved operation saved about 9600 kWh, or 2.9% of monthly weekday energy use. An appropriate regression or bin model would reach this conclusion.

However, reflect again on Figure 3. A regression model would predict the power levels shown by the red line. If we wished to predict what the power is most likely to be, or should be, at a given time, the red line would give an estimate that is too high during the overnight hours by 47 kW, or about 17%. This would compromise the ability of performance tracking to identify increasing energy use, and correct it.

NorthWrite's approach is to use the best model type for each intended purpose. This overcomes issues with simple regression models, provides accurate long-term average models for M&V, and the best short-term expected value models for performance tracking.

So how good are the NorthWrite models? Comparisons with two pure regression models were made, looking at three partial-day time categories. The regression models used categorizations 4 and 6 from Table 1, and the NorthWrite models fit categorization 6.The results show that the NorthWrite Energy Expert models, developed by PNNL using funding from USDOE, more closely follow the actual data and provide a better overall fit than the regression models.

Figure 5 shows the model Correlation Coefficient, R^2, for the 3 time periods and 4 model types. For the occupied time period, all models showed good correlation. The reference type 4 regression had the poorest fit, which is not surprising since it uses the least time period categorization. However, the NorthWrite models both showed better correlation than the regression model with level 6 categorization.

The results are dramatic for the unoccupied time period, and if looking at a single hour in the unoccupied period, with the NorthWrite models providing a much better fit. (There was no hourly regression model developed for the overall unoccupied period, so that data bar is missing from the figure.)

For energy use that has a low correlation to the independent variable (ambient temperature), it is better to compare models on the basis of the coefficient of variation of the root-mean-squared-error, CV(RMSE), which is the RMSE divided by the average power. Here again, the NorthWrite models outperform the pure regression model, as shown in Figure 6.

Figure 5. Comparison of Correlation Coefficient for Different Models

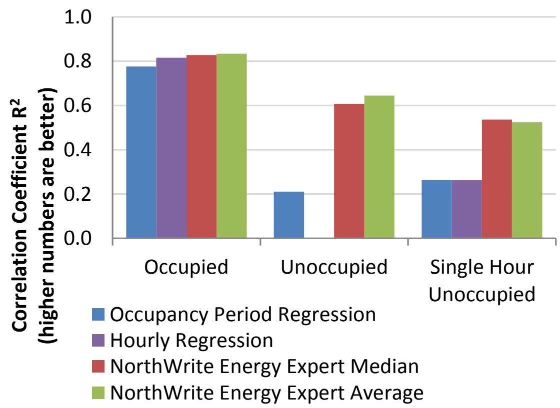
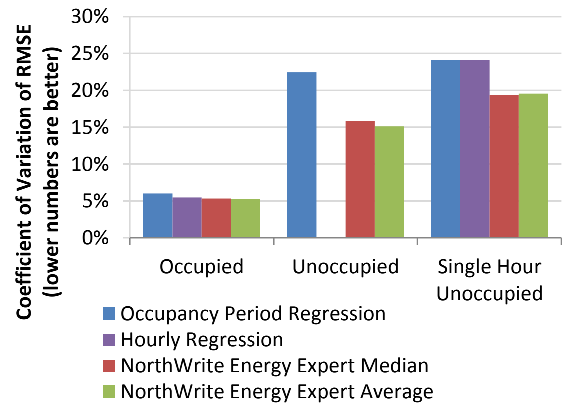


Figure 6. Comparison of CV(RMSE) for Different Models



Note that the NorthWrite Average model provides a better fit than the Median model. This is as expected and as designed: since the median model essentially ignores times of unusual operation, it provides a poor fit during those times, and a poorer overall fit, even though the fit during typical operation is better.

Another difference in the models is shown in Figure 7: The regression models and NorthWrite average model have no bias—the sum of the modeled values equals the sum of the actual values. However, the median model can have bias. Again, this is as expected, since this model represents typical operation rather than average operation. The bias can be in either direction, depending on whether unusual operation causes higher or lower energy use.

Figure 7. Comparison of Model Bias

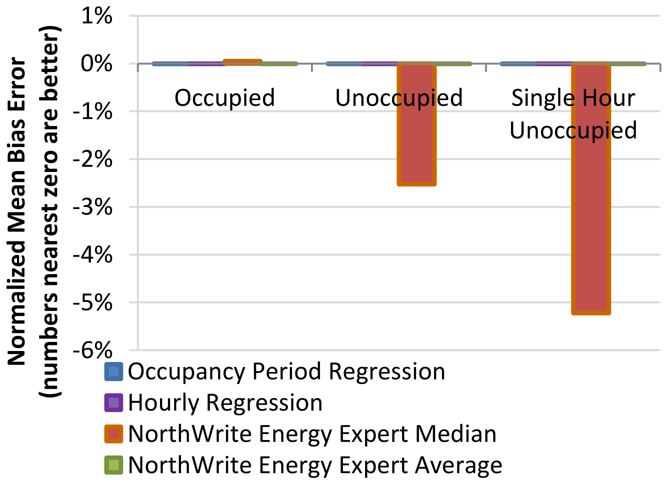
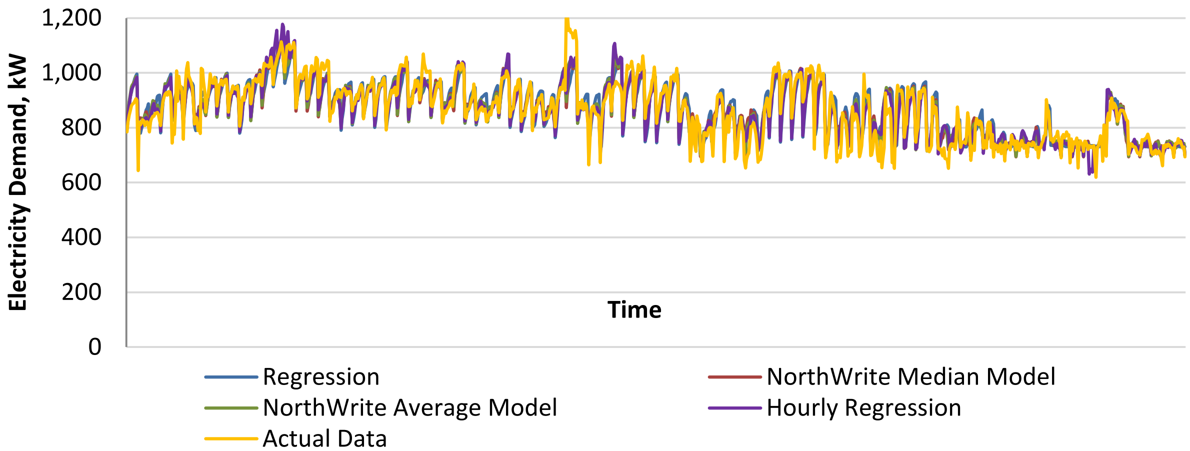


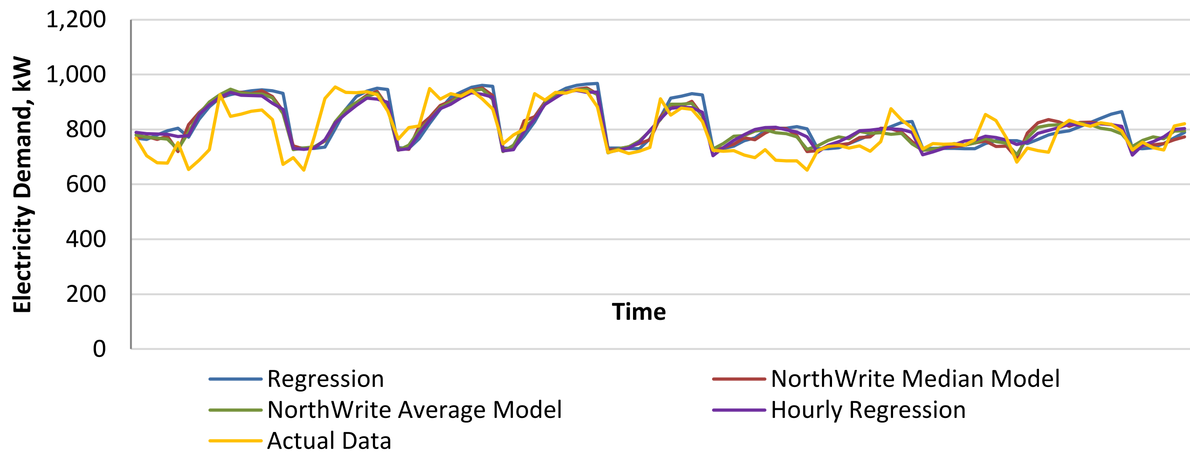
Figure 8 provides a look at the models' performance using a common time-series view. This is only for the Occupied time period, where all the models performed well.

Figure 8 Time Series Chart of Model Predictions and Actual Data



Showing this much data, with the models showing similar performance, it is difficult to distinguish the models. So let's zoom in on the data a bit.

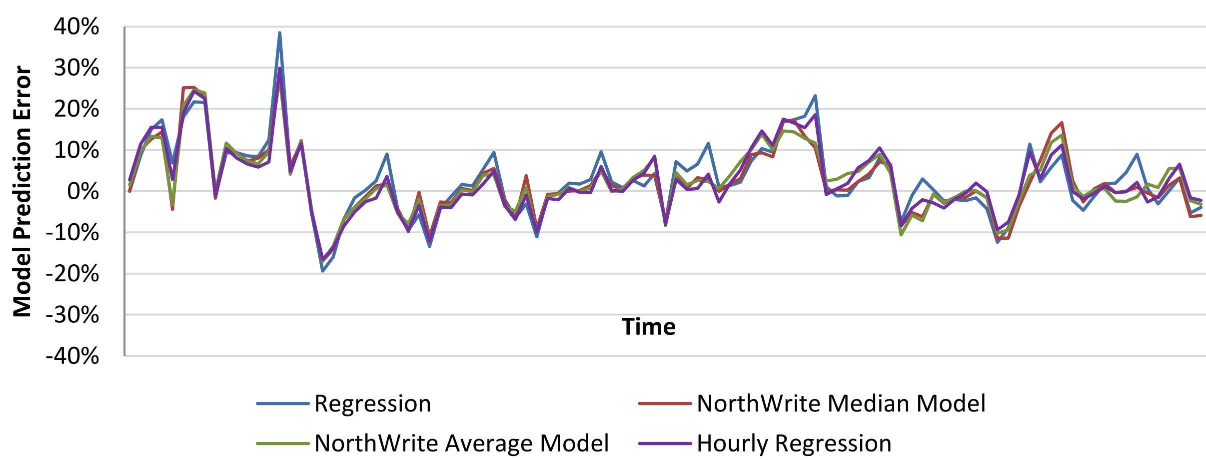
Figure 9 Time Series Chart of Model Predictions and Actual Data, a Closer Look



In Figure 9, we can better distinguish the models from the actual data. The models show similar results, but we can start to see that the NorthWrite models provide predictions that are often a bit closer to the actual data.

This can be made a bit clearer by charting the residual or prediction errors—the difference between the modeled and actual values. Figure 10 shows the modeled prediction errors for each hour, including only the occupied hours.

Figure 10 Time Series Chart of Model Prediction Errors, a Closer Look



As previously shown in Figures 5 and 6, the models exhibit little difference when modeling the occupied hours. But Figure 10 shows that the regression model using hourly data for the whole occupied period has the greatest errors, and that most of the time the NorthWrite models are closest to the true data, qualitatively verifying the quantitative information provided by Figures 5 and 6. The NorthWrite models show greatly increased superiority when modeling the unoccupied period.