Quantifying Uncertainty In Grey-box Building Models Arising From Smart Meter Energy Data Sampling Frequency

Jonathan Chambers[1], Tadj Oreszczyn[1], and David Shipworth[1]

[1]Energy Institute, University College London
Contact: jonathan.chambers.13@ucl.ac.uk

Abstract

The increasing availability of high resolution smart meter data is an opportunity to develop data-driven models of residential buildings. However, model performance may be undermined by the very density of this data - using models which take advantage of steady state approximations could reduce uncertainty in results. This requires a better understanding on the relation between model uncertainty and the time resolution of the input data used.

We apply a grey-box model to smart meter data from 25 dwellings. Model uncertainty metrics demonstrate the transition from dynamic to steady state approximations as a function of time resolution. An optimal sampling frequency of 24 hours is found, providing quantitative evidence for assumptions made in previous research. It is shown that multiple uncertainty metrics are required to properly characterise the uncertainty profile, and that using the maximum resolution of data available may counter-intuitively increase model uncertainty.

Introduction

The EU Energy End-use Efficiency and Energy Services directive mandates smart meter installation across member states (EU 2006), with the UK aiming to install gas and electricity meters to all homes and small businesses by 2020 (Ofgem 2013). The resulting data could resolve a number of long standing questions in energy analysis, as well as support evidence based policy. There has been particular interest in using empirical models to address the shortcomings of standard dwelling assessment methods (A. Summerfield et al. 2015). However techniques for analysis are relatively immature and best practices for domestic data analysis are yet to be established (Kavousian, Rajagopal, and Fischer 2013; Flach et al. 2014).

Most work to date focused on commercial buildings. In 1994, ASHRAE encouraged researchers to compare prediction algorithms through a competition to predict energy consumption based on synthetic commercial property data (Haberl and Kreider 1994). A variety of machine learning approaches were tested, including multi-regression analysis (MRA), principal component analysis (PCA), and artificial neural networks (ANN) (Reddy and Claridge 1994; González and Zamareño 2005). Some similar analysis was performed on the residential sector (Kolter and Ferreira 2011; Amina et al. 2012). ANN generated some of the most accurate consumption models, however ANN model coefficients have no physical significance making interpretation difficult. (Edwards, New, and Parker 2012; Swan and Ugursal 2009; Tso and Yau 2007; Coakley, Raftery, and Keane 2014).

An alternative modelling approach, known as "grey box" modelling, is to create physically-based models and fit them to data (Coakley, Raftery, and Keane 2014). Rabl (1988) reviewed several physically-based models for estimating model parameters from consumption. General deterministic models were considered where energy consumption obeyed causality, in that the present value of interior temperature T_{int} must be a unique function of the past values of T_{int}, the exterior temperature T_{ext}, auxiliary heat input Q_{aux} and solar heat input Q_{sol} (eq. 1). A discrete time step formulation was shown to be equivalent to an ARMA model. An alternative differential equation form of the heat balance equation was formulated, and from it an expression of the thermal network as a matrix of differential equations was derived. The authors sought to establish the fundamental equivalence of these approaches analytically.

$$\int_0^\infty dt'[a_{int}(t')T_{int}(t-t') - a_{ext}(t')T_{out}(t-t') - a_{aux}(t')Q_{aux}(t-t') - a_{aux}(t')Q_{aux}(t-t')] = 0 \tag{1}$$

Importantly, Rabl considered the effects of making steady state approximations. He demonstrated that by choosing appropriate limits for the integral periodic and transient terms cancel out, significantly decreasing data requirements in the steady state case. When applying this to hourly data from a large commercial building, a 24hr base period was chosen on the assumption that most drivers of energy consumption in buildings follow a daily pattern - however no attempt was made to demonstrate formally that this time period was optimal. Given that coheating tests of buildings found that considerable residual energy could remain in a building due to solar irradiation on the previous day (Stamp, Lowe, and Altamirano-Medina 2013), 24 hours might not meet the steady-state approximation requirement. However, since these coheating tests were performed on unoccupied dwellings and were concerned only with thermal losses, they do not constitute firm evidence that multiple days should be used to analyse consumption in occupied dwellings.

The PRInceton Scorekeeping Method (PRISM) was introduced by M. Fels (1986) as a method for deriving a Normalised Annual Consumption (NAC) from monthly billing data. Total consumption was modelled as baseload plus heating demand as a function of external temperature, expressed in its piecewise form in eq. 2:

$$P(T_{ex}) = \begin{cases} PTG(T_{ex} - T_h) + P_b, & \text{if } T_{ex} < T_h \\ P_b, & \text{if } T_{ex} >= T_h \end{cases} \tag{2}$$

where the building-specific constants are:

- PTG - Power Temperature Gradient (kW/ $^\circ$C)
- T_h - external temperature below which heating is used ($^\circ$C)
- P_b - baseload power demand (kW)

Since daily energy consumption was not available, the term $(T_h - T_{ex})$ was substituted with the heating degree-days (HDD) to the base T_h. The NAC was calculated as

$$NAC = 365 * P_b + PTG * H(T_h) \tag{3}$$

where $H(T_h)$ was the average HDD for 1970 to 1981. The NAC was presented as a reliable index of energy consumption, while the model fit parameters PTG, P_b, and T_h were not considered appropriate indexes in their own right due to their large uncertainties.

M. Goldberg and Fels (1986) note that while the errors in PTG, P_b, and T_h are large compared to the 3-4% typical for NAC, these parameters could be used to track changes in demand profiles if an average of a sufficiently large number of houses is used. They determined parameters changes and uncertainty in 243 gas heated dwellings in Wisconsin, USA. A key finding was that while aggregate results were stable their large interquartile range suggested that some households changed significantly, indicating that while aggregate analysis is beneficial to reducing parameter uncertainties, important information about individual dwellings is lost.

PRISM is recommended in ASHRAE's Guideline 14-2002 for measurement of energy and demand savings as a basic index of energy consumption (ASHRAE 2002). ASHRAE outlines extensions to PRISM to model different consumption patterns such as dwellings with air conditioning, introducing additional parameters to do so. However, the guide advises against use of these models because the fit will becomes less well determined when increasing the number of fit parameters for the same quantity of data.

Ruch and Claridge (1993) explored variants of PRISM with more parameters to achieve better fitting to daily data from a range of commercial buildings. Error and goodness-of-fit statistics were applied to allow robust comparison between model variants. However it was noted that that traditional fit measures such as RMSE do not take into account how well the fit performs over a range of temperatures.

PRISM is particularly interesting as a starting point for analysing smart meter data since it provides a simple model with physically interpretable parameters. It has not been used in the UK even in its monthly form in the past due to the lack of monthly billing data to match with heating degree days. A. Summerfield et al. (2015) applied a variant of PRISM to smart meter data from the Energy Demand Research Project (EDRP), which ran a range of behavioural feedback energy saving trials. Half-hourly gas and electricity meter data from the control group was used. Summerfield focused on using the gradient term PTG to characterise heating demand and thereby effective building thermal losses. Unlike PRISM, the T_h parameter (external temperature below which heating is used) was set to a constant 15 $^\circ$C. External temperature data was drawn from a grid of daily mean temperatures - therefore all analysis was performed at daily frequency.

Birt et al. (2012) fit a model consisting of 3 separate linear sections to hourly electricity data to identify a range of consumption properties. This resembled the PRISM in fitting temperature-consumption profiles with piecewise linear models. When applying this method to hourly residential electricity data, the authors suggest the accuracy of the method might be improved with higher resolution data. Their analysis did not account for the difference between steady state and dynamic modelling despite using a linear model with no dynamic elements.

Problem statement

The selection of grey-box models shown demonstrate the importance of understanding their physical implications and how these relate to the data to which they are applied, in particular the selection of an appropriate sampling frequency to match implicit assumptions, such as a steady state approximation. However, these considerations have not been prominent in the literature. In the UK, the lack of billing data at less than yearly frequency has made such considerations moot. However, the advent of smart meters enables a free choice of sampling period. In order to use steady-state physically-based models, a method for choosing the appropriate sampling period must be considered.

PRISM and its variants are an interesting starting point because of their relative simplicity and physically interpreted fit parameters. Previous work has highlighted the use of daily sampling periods - indirectly through the use of heating degree-days, and directly through averaging consumption data to match the daily temperature data (M. F. Fels, Goldberg, and Lavine 1986; A. Summerfield et al. 2015). Given the multi-day thermal lags observed by Stamp et al. on the one hand, and the insistence by Birt et al. (2012) that higher resolution should give better results on the other, there is a need for a systematic exploration of the issue.

This work aims to select and apply appropriate error estimation methods to a PRISM-derived model. It will evaluate the change in model error and uncertainty in model parameters as a function of the data resampling frequency, and thereby determine an optimal resampling frequency for the model. This resampling frequency will demonstrate the steady-state limit, and determine whether thermal lags are a sufficient factor in occupied dwellings to require that more than one day be considered for a steady state approximation.

This study will make use of a sample of gas-heated residential buildings which are equipped with smart meters on both electricity and gas supplies, allowing total consumption to be monitored.

Method

A variation of PRISM is used. The consumption is modelled as in eq. 2. Instead of HDD, total power demand from smart meters and real temperature data is used. The model parameters are fit for each house using the least-squares `curve_fit` function of the `numpy` Python programming package (Jones et al., n.d.).

The model is fit at half-hourly, hourly, 6-hourly, 12-hourly, daily, 2-days, 3-days, weekly, and monthly frequencies. The original data is downsampled by taking the mean of values within each time range. When resampling to daily, weekly, and monthly frequencies the data was averaged from the start of the day (00:00), week (Monday at 00:00), and month (start day of month at 00:00) (McKinney, n.d.). As the resampling reduces the number of data points, certain houses do not have enough data to perform the regression at the longer frequencies. Results were only discarded for a site for the frequencies with too few data points, results at other frequencies were retained.

Residual error metrics are well established in literature. Recommended methods for validating model calibrations are Root-Mean-Square-Error (RMSE), Coefficient of Variance of RMSE (CVRMSE) (eq. 4) and the Normalised Mean Bias Error (NMBE) (eq. 5) (ASHRAE 2002; Im and Bhandari 2014; Coakley, Raftery, and Keane 2014). The ASHRAE guidelines consider a building model calibrated if it achieves a CVRMSE under 15% and an NMBE under 5% on a monthly basis.

$$CVRMSE(\%) = \frac{\sqrt{\sum_{i=1}^{N}(m_i - p_i)^2/N}}{\bar{m}} \quad (4)$$

$$NMBE(\%) = \frac{\sum_{i=1}^{N}(m_i - p_i)/(N-p)}{\bar{m}} \quad (5)$$

Where :

- m_i - measured data point
- p_i - predicted data point
- N - number of data points
- $\bar{m}$ - mean of dependant variable
- p - total number of regression parameters in the model.

The CVRMSE and NMBE metrics describe the regression residuals but give no indication of uncertainty in regression parameters. It is sometimes assumed that acceptable values for CVRMSE and NMBE are sufficient conditions for considering a model to be calibrated. However in order to use fit parameters as model outputs, uncertainty estimates for these parameters are required.

The model parameters have uncertainties arising from measurement uncertainties in the input data and deviations between the idealised model and reality. In the case of least-squares fitting, the standard error of parameters can be derived analytically under the assumption that input data errors are normally and independent and identically distributed (iid) (c.f Richter (1995)). The standard error in a model parameter is given by the square root of the corresponding diagonal element of the model fit covariance matrix (eq. 6).

$$SE_{a_j} = \sqrt{C_{jj}} \tag{6}$$

Variance is given by:

$$\sigma^2_{a_j} = C_{jj} \tag{7}$$

C_{jj} is the jth element on the diagonal of covariance matrix $\mathbf{C}$. The covariance matrix is obtained from the `curve_fit` function (Jones et al., n.d.).

Dataset

The data is drawn from the micro combined heat and power (micro CHP) field trials. The micro CHP accelerator project was a large demonstration program running from June 2004 to April 2008 (CarbonTrust 2011). The residential sites included 69 buildings equipped with mCHP units and a control group of 27 houses with condensing gas boilers and central heating. Only the 27 control sites were used, none of which were equipped with air conditioning units.

Electric and gas smart meter energy consumption readings (in kW) are included together with external temperatures measured at the same intervals. The dataset had been previously cleaned and data rows with errors flagged according to error type, no additional cleaning was performed (Carbon Trust 2008). These rows were omitted from the analysis.

Results and Discussion

Removal of error-flagged rows resulted in 25 sites for analysis. The mean yearly total consumption of 16991kWh across sites was 12.9% lower than the national average 19508kWh (DECC 2013). The dwellings had an average Energy Performance Certificate (EPC) rating in the "C" band in the A (best) to G (worst) rating scale - 40% of UK homes are in the "C" band, while the UK average is "D" band (DCLG 2013). Although these homes were somewhat more efficient than average, they display energy consumption patterns similar to that observed in other studies (A. Summerfield et al. 2015). This suggests that the method should be applicable to a broad range of UK dwellings.

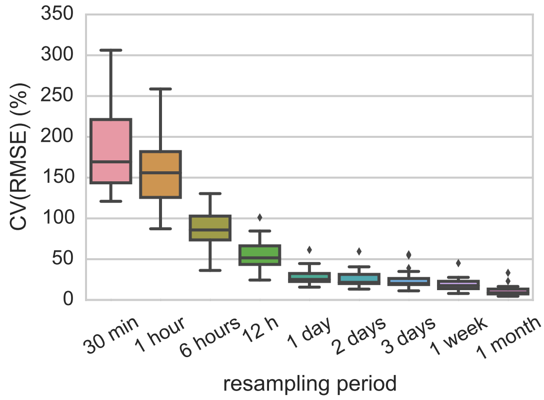

Figure 1: Plot of CVRMSE as a function of resampling frequency demonstrating rapid decrease in residual error with decreasing sampling frequency (longer averaging periods).

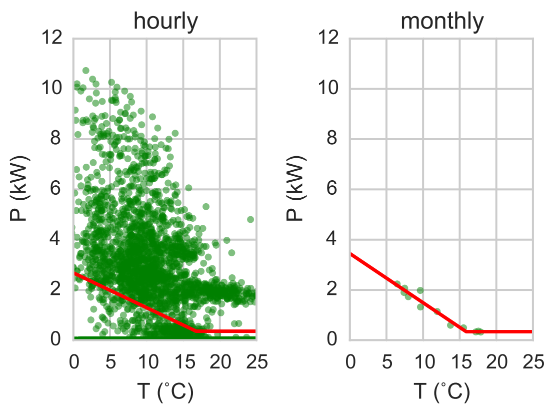

Figure 2: Power curves for site CTBLR01749 at hourly and monthly resampling frequencies.

The change of CVRMSE as a function of resampling frequency is shown in Figure 1. As expected, the CVRMSE decreased rapidly as the sampling frequency decreased from half-hourly to monthly. The residual error for the data at hourly sampling was much greater than the recommended threshold of 30% for hourly data recommended by ASHRAE. The 'well behaved' site shown in Figure 2 illustrates the large spread of consumption at the hourly time scale for a single sample site, with the maximum consumption being much greater than the mean consumption. In this site, using monthly average data results in a well-behaved dataset and a good fit. However, inspecting another site in Figure 3, the monthly resampling results in an under-defined model, with an unphysical negative value for the baseload consumption. Using intermediate daily averaging achieves a much better defined fit.

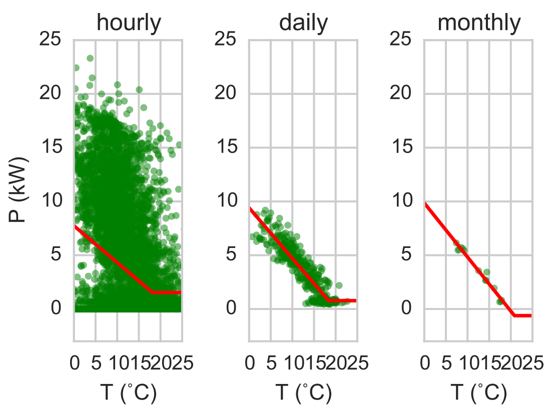

Figure 3: Power curves for site CTBLR01723 at hourly, daily, and monthly resampling frequencies.

Very little bias was observed in the model at any resampling frequency, with no values greater than 0.01%, suggesting that consumption data points for a given temperature are symmetrically distributed about the model predicted value. This is consistent across sites and resampling frequencies, and demonstrates that the model error distribution is symmetric. However, as can be seen in Figure 2 at hourly resampling the large number of zero values tends to downwards-bias the heating regimen slope compared to monthly sampling. This is not revealed in the NMBE metric since the values still lie symmetrically about the curve.

Outliers were removed before plotting since a small number of very extreme results (in excess of 1x10^9 %) would otherwise heavily skew the distribution. A simple threshold of 100% was used as an outlier filter as the results fell into two classes - well-behaved cases with uncertainties below 100% and aberrations with uncertainties much higher than this limit.

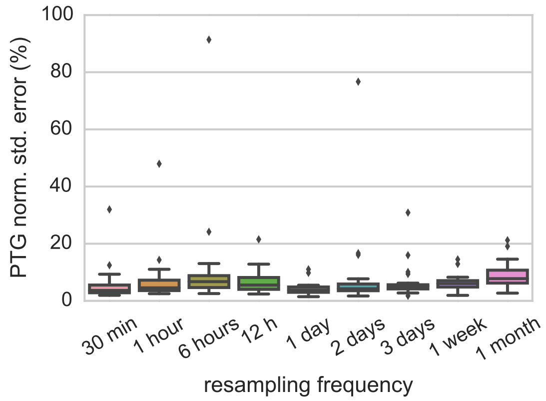

Figure 4: Power Temperature Gradient (PTG) standard error as a percentage of the parameter value for a range of resampling frequencies

Standard errors in the PTG are minimised at daily resampling, as can be seen in Figure 4. Errors in P_b (Figure 5) and T_h (Figure 6) follow a similar pattern, although errors were larger overall for P_b. PTG and T_h a less clear advantage of using daily consumption compared to half-hourly. If only these parameters were used, such as in A. J. Summerfield et al. (2007) where only the PTG, it may seem reasonable to use the maximum data density available. However, cross-referencing with the CVRMSE demonstrates the necessity of using longer time scales. This demonstrates the importance of using multiple criteria for accepting a model fit.

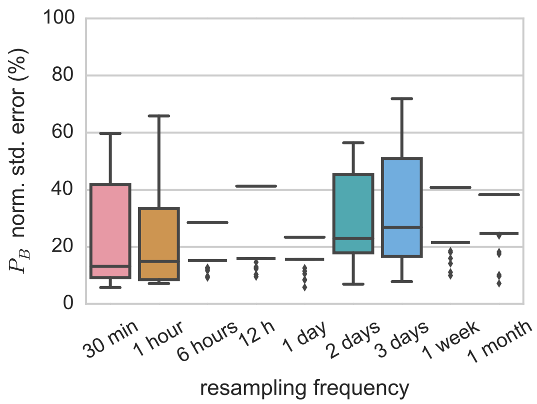

Figure 5: Baseload power (P_B) standard errors as a percentage of parameter value for a range of resampling frequencies

Using 2 or 3 day resampling consistently increases the uncertainty in the parameters - dramatically so in the case of P_b. This suggests that the thermal time-lag effects observed by Stamp, Lowe, and Altamirano-Medina (2013) do not have a significant effect in these occupied dwellings. The

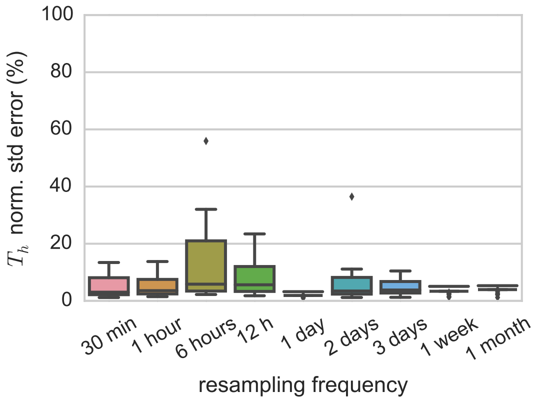

Figure 6: Heating reference temperature (T_h) standard errors as a percentage of parameter value for a range of resampling frequencies

baseload power demand in particular sees no benefit, as one would expect use of appliances that drive baseload to follow a daily cycle and be unaffected by thermal lags. Daily sampling is therefore optimal.

Conclusion

Comparing residual error with model parameter uncertainties reveals the optimal sampling frequency in a way that would not be possible using only one of these metrics. Parameter uncertainties are small at both half-hour and daily frequency but do not indicate a clear choice between the two. Comparing these with the CVRMSE uncertainties however clearly shows the advantage of daily averaging periods. This suggests that multiple uncertainty metrics must be considered when evaluating model fits to empirical data.

Sampling frequencies longer than one day only increased model uncertainty, suggesting that the multi-day thermal lags observed by Stamp et al. in unoccupied dwelling thermal testing do not translate into multi-day steady state limits in occupied dwellings. Conversely, higher frequency data increased residual error due to additional noise.

Daily sampling was shown to be optimal, providing empirical support for assumptions made in previous studies. Daily sampling displayed the lowest uncertainties because the steady state assumptions of the model were met at this frequency. This suggests that any occupied dwelling model which operates with a steady state assumption should use daily sampling.

Future work could cross-check the error metrics by performing a bootstrap uncertainty estimation which does not depend on the assumptions of the covariance standard error method. This would would also enable uncertainty estimation of fit algorithms which do not produce covariance matrices, such as parameter search or brute force algorithms. Jack-knife and bootstrapping are well documented techniques to estimate uncertainties which do not make assumptions on the system error distribution (Diciccio and Efron 1996; Faber 2002). However these are much more computationally intensive - initial tests for this study found that the default implementations were not fast enough. Ongoing work aims to develop fast codes for running bootstrap uncertainty estimation on the consumption model.

The dataset used was relatively small, as it was the best-quality data available at the time of analysis. Although previous research indicates that the consumption patterns should be similar to UK normal, the dwellings were not selected to be nationally representative nor is the dataset large enough to expect to capture a meaningful cross-section of typical UK dwellings. Fortunately, an anonymized version of the Energy Demand Research Project (EDRP) dataset including electricity and gas data for over 14000 homes has been recently released. Research is ongoing to apply this method to this much larger dataset, with the aim of further validating the approach.

This work uses a simple dwelling energy consumption model to demonstrate the impact of data sampling rate on result uncertainty, and illuminates the physical assumptions implicit in the model. It suggests that other, more complex models must take care to understand the temporal ranges over which they operate. Using the full resolution of available data may counter-intuitively increase model uncertainty - the work undertaken in this paper demonstrates a reproducible method for mitigating this risk.

Acknowledgements

This research was made possible by support from: the UK Engineering and Physical Sciences Research Council (EPSRC); EDF Energy R&D UK Centre; the RCUK Centre for Energy Epidemiology (EPSRC ref EP/K011839/1); the London-Loughborough Centre for Doctoral Training in Energy Demand (LoLo) at UCL Energy Institute (EPSRC ref EP/L01517X/1); and the Carbon-Trust under the ‘Micro-CHP accelerator’ project.

References

Amina, M., V.S. Kodogiannis, I. Petrounias, and D. Tomtsis. 2012. “A hybrid intelligent approach for the prediction of electricity consumption.” *International Journal of Electrical Power & Energy Systems* 43 (1): 99–108.

doi:10.1016/j.ijepes.2012.05.027.

ASHRAE. 2002. "ASHRAE savings calculation." In *ASHRAE Guideline 14*. Vol. 1986. Ashrae 1990.

Birt, Benjamin J., Guy R. Newsham, Ian Beausoleil-Morrison, Marianne M. Armstrong, Neil Saldanha, and Ian H. Rowlands. 2012. "Disaggregating categories of electrical energy end-use from whole-house hourly data." *Energy and Buildings* 50 (July): 93–102. doi:10.1016/j.enbuild.2012.03.025.

Carbon Trust. 2008. "Micro-CHP Accelerator Guide to Field Trial Data." Carbon Trust.

CarbonTrust. 2011. "Micro-CHP Accelerator: Final Report." March. London.

Coakley, Daniel, Paul Raftery, and Marcus Keane. 2014. "A review of methods to match building energy simulation models to measured data." *Renewable and Sustainable Energy Reviews* 37 (September): 123–41. doi:10.1016/j.rser.2014.05.007.

DCLG. 2013. *English Housing Survey Headline report*. Department for Communities & Local Government.

DECC. 2013. "Domestic energy consumption in the UK between 1970 and 2012." In *Energy Consumption in the UK*, 1–10. London: DECC.

Diciccio, Thomas J, and Bradley Efron. 1996. "Bootstrap Confidence Intervals." *Statistical Science* 11 (3): 189–228.

Edwards, Richard E., Joshua New, and Lynne E. Parker. 2012. "Predicting future hourly residential electrical consumption: A machine learning case study." *Energy and Buildings* 49: 591–603. http://www.sciencedirect.com/science/article/pii/S0378778812001582.

EU. 2006. "Energy end-use efficiency and energy services." Brussels: European Comission. http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32006L0032.

Faber, Nicolaas. 2002. "Uncertainty estimation for multivariate regression coefficients." *Chemometrics and Intelligent Laboratory Systems* 64 (2): 169–79. doi:10.1016/S0169-7439(02)00102-8.

Fels, Margaret. 1986. "PRISM: An introduction." *Energy and Buildings* 9 (1-2): 5–18. doi:10.1016/0378-7788(86)90003-4.

Fels, Margaret F., Miriam L. Goldberg, and Michael L. Lavine. 1986. "Exploratory scorekeeping for oil-heated houses." *Energy and Buildings* 9 (1-2): 127–36. doi:10.1016/0378-7788(86)90014-9.

Flach, Peter, Tom Wilcox, Royston Sellman, Joshua Thumim, and Arno Knobbe. 2014. "Subgroup Discovery in Smart Electricity Meter Data." *IEEE Transactions on Industrial Informatics* 10 (2): 1327–36. doi:10.1109/TII.2014.2311968.

Goldberg, ML, and MF Fels. 1986. "Refraction of PRISM Results into Components of Saved Energy." *Energy and Buildings* 9: 169–80. http://www.sciencedirect.com/science/article/pii/0378778886900186.

González, Pedro a., and Jesús M. Zamarreño. 2005. "Prediction of hourly energy consumption in buildings based on a feedback artificial neural network." *Energy and Buildings* 37 (6): 595–601. doi:10.1016/j.enbuild.2004.09.006.

Haberl, J. S., and J. F. Kreider. 1994. "Instructions for 'The Great Energy Predictor Shootout II: Measuring Retrofit Energy Savings'." Energy Systems Laboratory, Texas A&M University. http://hdl.handle.net/1969.1/2137.

Im, Piljae (Oak Ridge National Laboratory), and Mahabir (Oak Ridge National Laboratory) Bhandari. 2014. "Is Monthly and Whole Buildings Level Calibration Enough?" In *Building Simulation Conference*, 177–86. Atlanta.

Jones, Eric, Travis Oliphant, Pearu Peterson, and Others. n.d. "SciPy: Open source scientific tools for Python." http://www.scipy.org/.

Kavousian, Amir, Ram Rajagopal, and Martin Fischer. 2013. "Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior." *Energy* 55 (June): 184–94. doi:10.1016/j.energy.2013.03.086.

Kolter, JZ, and J Ferreira. 2011. "A Large-Scale Study on Predicting and Contextualizing Building Energy Usage." *AAAI*, 1349–56. http://www.aaai.org/ocs/index.php/AAAI/AAAI11/paper/download/3759/4088.

McKinney, Wes. n.d. "pandas: a Foundational Python Library for Data Analysis and Statistics."

Ofgem. 2013. "Transition to smart meters | Ofgem." https://www.ofgem.gov.uk/electricity/retail-market/metering/transition-smart-meters.

Rabl, A. 1988. "Parameter Estimation in Buildings: Methods for Dynamic Analysis of Measured Energy Use." *Journal of Solar Energy Engineering* 110: 52–66.

Reddy, T.A., and D.E. Claridge. 1994. "Using synthetic data to evaluate multiple regression and principal component analyses for statistical modeling of daily building energy consumption." *Energy and Buildings* 21 (1): 35–44. doi:10.1016/0378-7788(94)90014-0.

Richter, P. H. 1995. "Estimating Errors in Least-Squares Fitting." *The Telecommunications and Data Acquisition Progress Report*, 107–37. http://cdsads.u-strasbg.fr/abs/1995TDAPR.122..107R.

Ruch, David K., and David E. Claridge. 1993. "A development and comparison of NAC estimates for linear and change-point energy models for commercial buildings." *Energy and Buildings* 20 (1): 87–95. doi:10.1016/0378-7788(93)90041-R.

Stamp, Samuel, Robert Lowe, and Hector Altamirano-Medina. 2013. "An Investigation Into The Role Of Thermal Mass On The Accuracy Of Co-heating Tests Through Simulations & Field Results." In *13th Conferenece of International Building Performance Simulation Association*, 39–46. Chambéry, France: IBPSA. http://www.ibpsa.org/proceedings/BS2013/p/_1384.pdf.

Summerfield, A.J., T. Oreszczyn, I.G. Hamilton, D. Shipworth, G.M. Huebner, R.J. Lowe, and P. Ruyssevelt. 2015. "Empirical variation in 24-h profiles of delivered power for a sample of UK dwellings: Implications for evaluating energy savings." *Energy and Buildings* 88 (February): 193–202. doi:10.1016/j.enbuild.2014.11.075.

Summerfield, Alex J., R.J. Lowe, H.R. Bruhns, J.A. Caeiro, J.P. Steadman, and T. Oreszczyn. 2007. "Milton Keynes Energy Park revisited: Changes in internal temperatures and energy usage." *Energy and Buildings* 39 (7): 783–91. http://www.sciencedirect.com/science/article/pii/S0378778807000576.

Swan, Lukas G., and V. Ismet Ugursal. 2009. "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques." *Renewable and Sustainable Energy Reviews* 13 (8): 1819–35. http://www.sciencedirect.com/science/article/pii/S1364032108001949.

Tso, Geoffrey K.F., and Kelvin K.W. Yau. 2007. "Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks." *Energy* 32 (9): 1761–68. http://www.sciencedirect.com/science/article/pii/S0360544206003288.