Tools and Methods for Addressing Autocorrelation in Energy Modeling

The NW Industrial Strategic Energy Management (SEM) Collaborative was formed in 2012 by the Bonneville Power Administration (BPA), Energy Trust of Oregon (ETO), and the Northwest Energy Efficiency Alliance (NEEA) for the purpose of identifying and addressing market barriers to SEM adoption in the industrial sector. Three teams were subsequently formed to focus on: 1) Small-to-Medium Industrial Solutions, 2) Market Analysis and Planning, and 3) Energy Tracking and Savings Protocols (ETSP).

The members of the ETSP team include representatives from BPA, ETO, NEEA, Idaho Power, U.S. Department of Energy, Consortium for Energy Efficiency, Puget Sound Energy, BC Hydro, and the Northwest Food Processors Association. This paper is the initial work product of the ETSP team, which was tasked with the identification of consistent and defensible methodologies for measuring and verifying SEM energy savings. The examples in this paper were drawn from recent modeling efforts in BPA's Energy Smart Industrial and the Energy Trust of Oregon's Production Efficiency programs.

Introduction

After reviewing SEM measurement and verification (M&V) protocols from different programs, the ETSP team identified autocorrelation as a common statistical issue in industrial data sets. Understanding that autocorrelation has the potential to negatively affect the predictive capability of regression-based energy models used to create adjusted energy baselines, more consistent treatment of this issue may help improve confidence in SEM-based savings, and thereby address a potential market barrier. The ETSP team compiled this paper for the purposes of outlining the implications of autocorrelation in the context of SEM measurement and verification, and providing examples of how program implementers have successfully identified and treated the presence of autocorrelation in regression-based energy models.

Statistical Definition

Autocorrelation is present when the error term in period *t* is related to the error term in period *t-1*. More simply stated, autocorrelation is characterized by a systematic pattern in the error term (residuals) of the model. This systematic pattern may be due to omitted variables from the model, serially correlated predictors coupled with model specification error, or correlated disturbances that are beyond the scope of predictor variables. The pattern in residuals also violates the ordinary least squares (OLS) assumption of independent error terms and may cause misinterpretation of regression results due to understated standard errors (apparent accuracy/precision better than it really is). The degree of autocorrelation is measured by the autocorrelation coefficient, also called the autocorrelation parameter, and can take on values between zero and one. Autocorrelation is said to be present when the autocorrelation coefficient (ρ) is significantly greater than zero. The Durbin-Watson test is another common approach for determining if the regression model errors are autocorrelated. For uncorrelated errors, the value of the Durbin-Watson statistic should be approximately 2. Montgomery (2008) provides guidance on calculating lower and upper bounds for the Durbin-Watson statistic, which are a function of sample size, the number of predictor variables and the desired confidence level. Most statistical regression packages will calculate the autocorrelation coefficient and the Durbin-Watson test statistic.

Implication for Model Results

In cases where the residuals of a regression-based model are autocorrelated, the coefficients in the model remain unbiased but their standard errors may be significantly underestimated. This could lead to models that include insignificant independent variables and underestimated error bounds around model predictions. In the context of SEM measurement and verification, a significant implication of autocorrelation is an overrepresentation of the confidence level associated with reported savings.

When autocorrelation is present and accounted for with an energy model, it can negatively impact the effective number of data points needed to meet a required confidence level. Simply stated, if a time series of length N exhibits positive autocorrelation, the effective number of independent observations N' is given by:

Equation (1): $N' = N \cdot \frac{(1-\rho)}{(1+\rho)}$

At a fundamental level, if the autocorrelation coefficient (ρ) is known, the true standard error may be estimated by multiplying the calculated standard error by factor *f*:

Equation (2): $f = \sqrt{\frac{1+\rho}{1-\rho}}$

The estimate of the true standard error can then be used to assess the impact on the statistical significance of a predictor variable. Statistical significance is normally based on the t-statistic, which is the estimated value of the beta coefficient divided by its standard error. Thus, the inflation of the standard error by factor *f* will have the effect of reducing the value of the t-statistic.

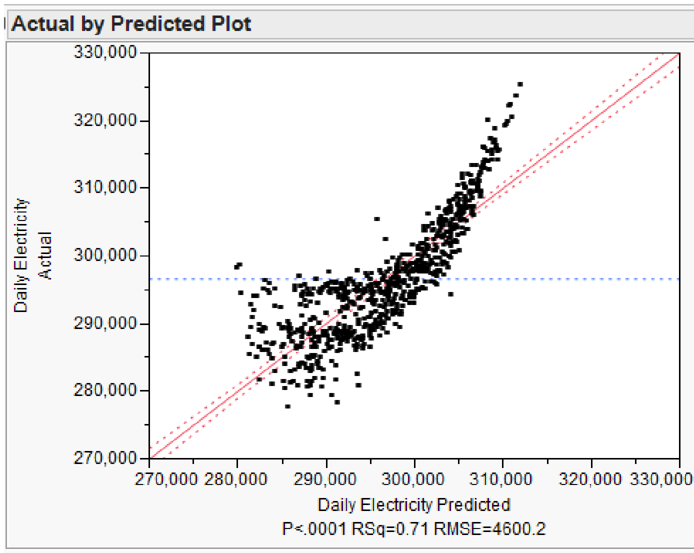
Alternatively, a comparison of Fractional Savings Uncertainty of two or more competing models may be used to assess the impact of autocorrelation on inflation of the standard errors.

Example #1: Integrated Circuit Manufacturing

Autocorrelation creates uncertainty when using an energy model to estimate energy savings. If the residuals from the baseline model are significantly autocorrelated, there are a few things to watch out for that may be causing the autocorrelation. First, the model may be missing a variable – for instance, if weather were a significant driver of energy consumption but an ambient temperature variable was missing from the model, this will cause significant autocorrelation. Next, there could be major changes during the baseline period that would cause a major trend in the residuals before and after the change; this would result in significant autocorrelation. Finally, autocorrelation, by nature, increases as the data granularity increases. If the data interval is too small, increasing the interval and therefore decreasing the sample size will most likely reduce the autocorrelation. However, reducing the sample size also reduces the statistical power of the model, so this method should be used with caution.

Model Misspecification: Missing Variable

First, it is possible that a missing variable is causing the baseline trends. In the example of a large integrated circuit manufacturer, a baseline period of two years was looked at with a satisfactory model, but unsatisfactory autocorrelation. The output of the model is shown below:



Summary of Fit

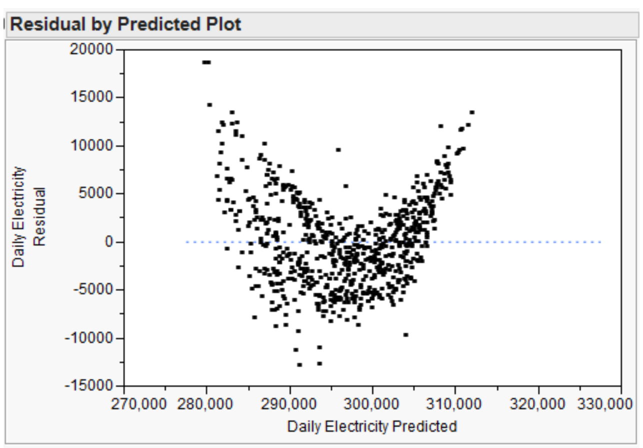
RSquare	0.71
RSquare Adj	0.71
Root Mean Square Error	4,600.21
Mean of Response	296,611.03
Observations (or Sum Wgts)	729.00

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	249,565.98	1,269.65	196.56	0.00
Wet Bulb Temp	704.98	20.17	34.95	0.00
Production Integrated Circuits	0.12	0.01	11.44	0.00

Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
0.47	729	0.7600	<.0001*



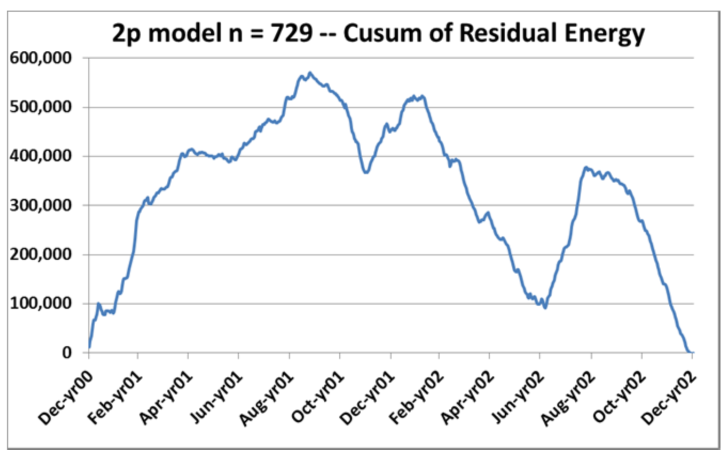
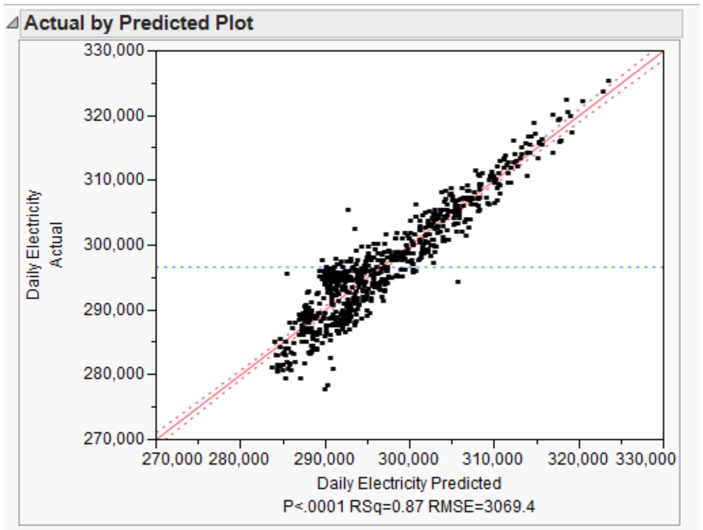


Figure 1. Graphical and Tabular Statistical Analysis Outputs of Two-Parameter Daily Regression Model – Two-Year Baseline – Significant Autocorrelation

This model has an R-squared of 0.71 and all of the t-statistics are significant, but the autocorrelation coefficient and Durbin-Watson show significant autocorrelation. Investigation of the relationship between the electricity usage and wet-bulb temp shows a non-linear relationship. This can also be seen in the scatter plot between the residuals and the predicted electricity. In this case, the autocorrelation is probably a symptom of missing a variable or a non-linear relationship between energy and one or more predictor variable. A temperature change point model and polynomial model were considered. After evaluating the fit of both, a polynomial term for the wet-bulb temperature variable was selected over the two-year baseline. The results of the new model are shown below:



Summary of Fit

RSquare	0.87
RSquare Adj	0.87
Root Mean Square Error	3,069.35
Mean of Response	296,611.03
Observations (or Sum Wgts)	729.00

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	334,554.31	2,948.21	113.48	<.0001*
Wet Bulb Temp	-3,066.47	126.03	-24.33	<.0001*
Production Integrated Circuits	0.13	0.01	18.24	<.0001*
Wet Bulb Temp Squared	39.94	1.33	30.10	<.0001*

Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation
0.65	729	0.67

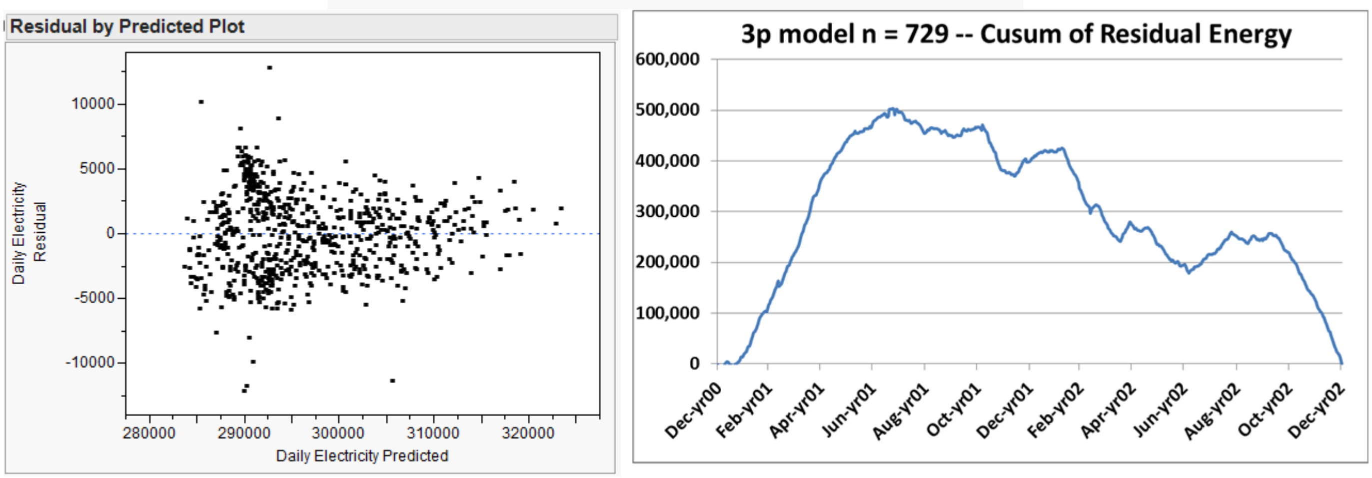


Figure 2. Graphical and Tabular Statistical Analysis Outputs of Three-Parameter Daily Regression Model – Two-Year Baseline – Significant Autocorrelation

This regression output shows significant improvement over the two-parameter model, with an R-squared of 0.87 and an improvement in autocorrelation. However, with an autocorrelation coefficient of 0.67 and a significant Durbin-Watson of 0.65, this model is still autocorrelated.

Model Misspecification: Discontinuity or Change in Mode

More concerning, the cumulative sum of the residuals (CUSUM) time series plot appears to be in a downward trend going into the intervention period, which, if it continues, could be misinterpreted as energy savings. In addition, the maximum CUSUM peak is more than 0.5% of the annual usage, which isn't overly high but still a clear trend. After further discussion with the site, it was discovered that a large capital improvement was installed right at the time the downward trend in the CUSUM started, and therefore the site was using less energy to produce the same amount of product. This capital project is likely the reason there is such a strong trend, which gives the modeler two choices:

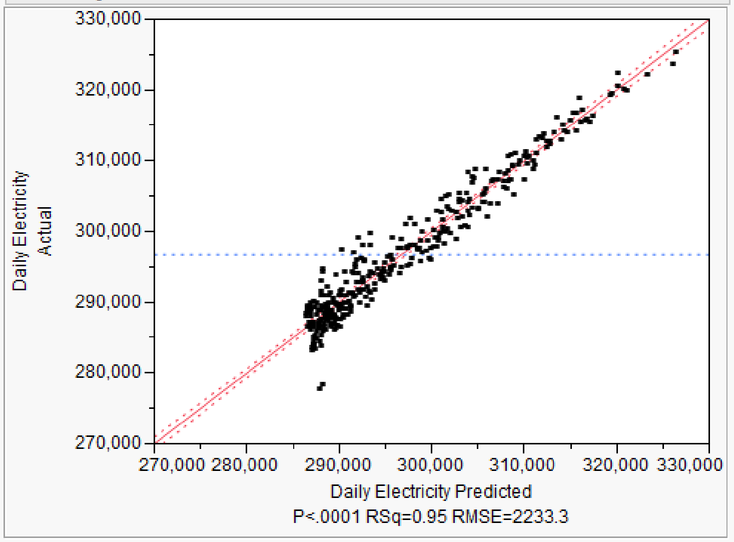
1. Reduce the baseline period to a year, so that the baseline only includes operation with the capital project, or
2. Create a categorical variable which is 0 before the capital project and 1 after, to allow the model to solve for the downward trend.

In this case, the modeler selected option 1 based on the fact that there were enough data points to cut out a year's worth of data, the model would be simpler with fewer variables, plus program-specific considerations.

After excluding the first year of baseline data to eliminate the significant change in the facility's energy usage, the resulting model is shown below:

Whole Model

Actual by Predicted Plot



Summary of Fit

RSquare	0.95
RSquare Adj	0.95
Root Mean Square Error	2,233.27
Mean of Response	296,699.48
Observations (or Sum Wgts)	365.00

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	342,185.26	3,690.77	92.71	<.0001*
Wet Bulb Temp	-3,263.51	166.20	-19.64	<.0001*
Production Integrated Circuits	0.05	0.01	4.84	<.0001*
Wet Bulb Temp Squared	44.00	1.68	26.16	<.0001*

Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
0.81	365	0.59	<.0001*

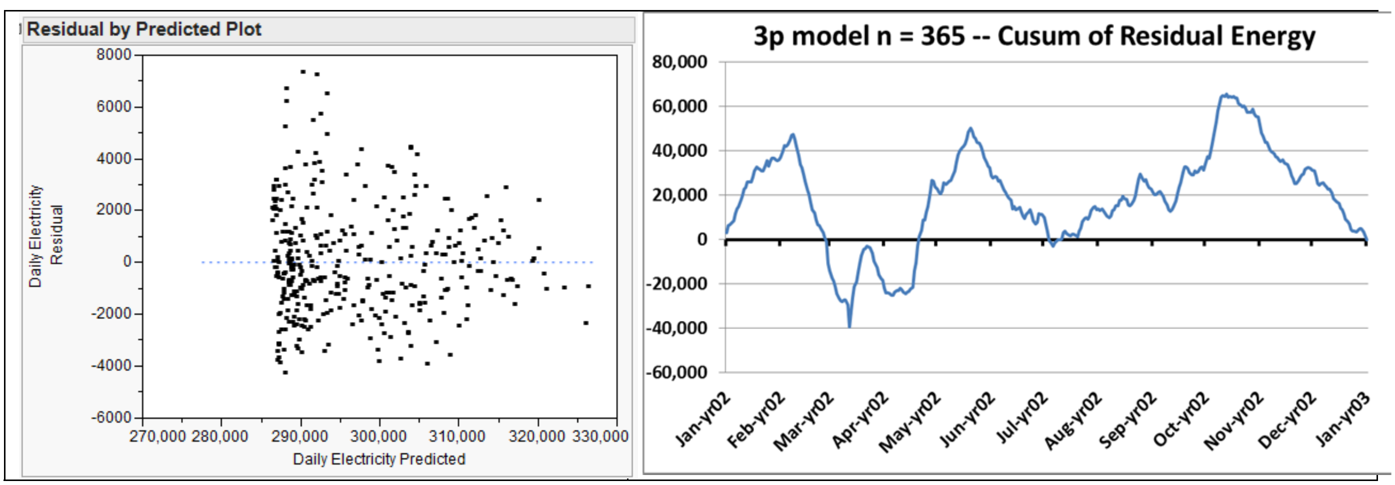


Figure 3. Graphical and Tabular Statistical Analysis Outputs of Three-Parameter Daily Regression Model – One-Year Baseline – Improved Autocorrelation

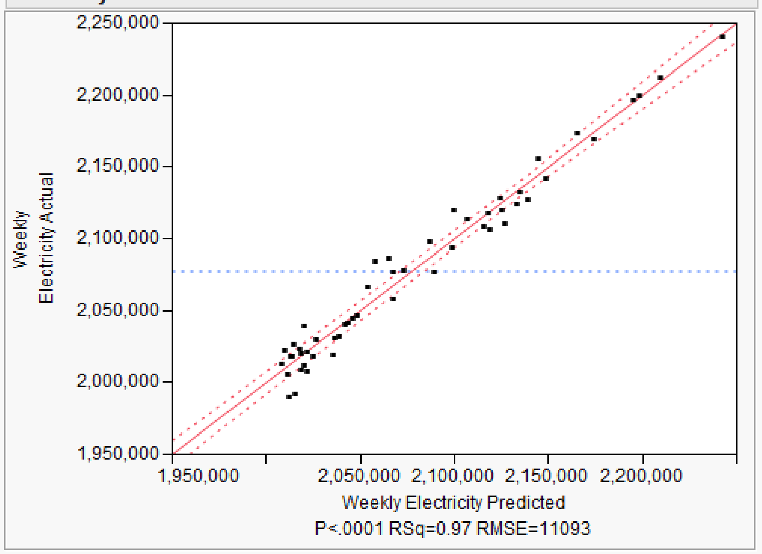
By eliminating a year of baseline data and therefore reducing the major trend that could have appeared to be savings during the intervention period, the new model's autocorrelation has improved, moving from an autocorrelation coefficient of 0.69 to 0.59 and a significant Durbin-Watson from 0.63 to 0.81. While this is an improvement in the autocorrelation, it is does not address the concern that the value of the standard errors is underrepresented. Looking at the CUSUM graph, the trends are smaller in magnitude, with a maximum CUSUM peak of only 0.05% of their annual usage as well as showing both positive and negative trend at a much higher frequency. Overall, the new one-year baseline model is far superior to the two-year model with the significant downward trend.

Data Granularity

Because autocorrelation is still an issue, it is worth reviewing the data interval to determine whether consolidating the data into weekly data points would help reduce the autocorrelation. Generally, the shorter the interval, the more autocorrelated the data. However, one should be careful about attempting to improve model fitness or address autocorrelation by aggregating to a longer time unit (e.g., from daily to weekly), since what is accomplished is masking the noise in the data set. The resulting weekly model is shown below:

Whole Model

Actual by Predicted Plot



Summary of Fit

RSquare	0.97
RSquare Adj	0.97
Root Mean Square Error	11,093.18
Mean of Response	2,077,654.44
Observations (or Sum Wgts)	52.00

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	2,369,817.27	67,195.05	35.27	<.0001*
Wet Bulb Temp	-3,071.87	451.26	-6.81	<.0001*
Wet Bulb Temp Squared	42.32	4.49	9.42	<.0001*
Production Integrated Circuits	0.04	0.03	1.60	0.1154

Durbin-Watson

Durbin-Watson	Number of Obs.	AutoCorrelation	Prob<DW
1.29	52	0.34	0.0012*

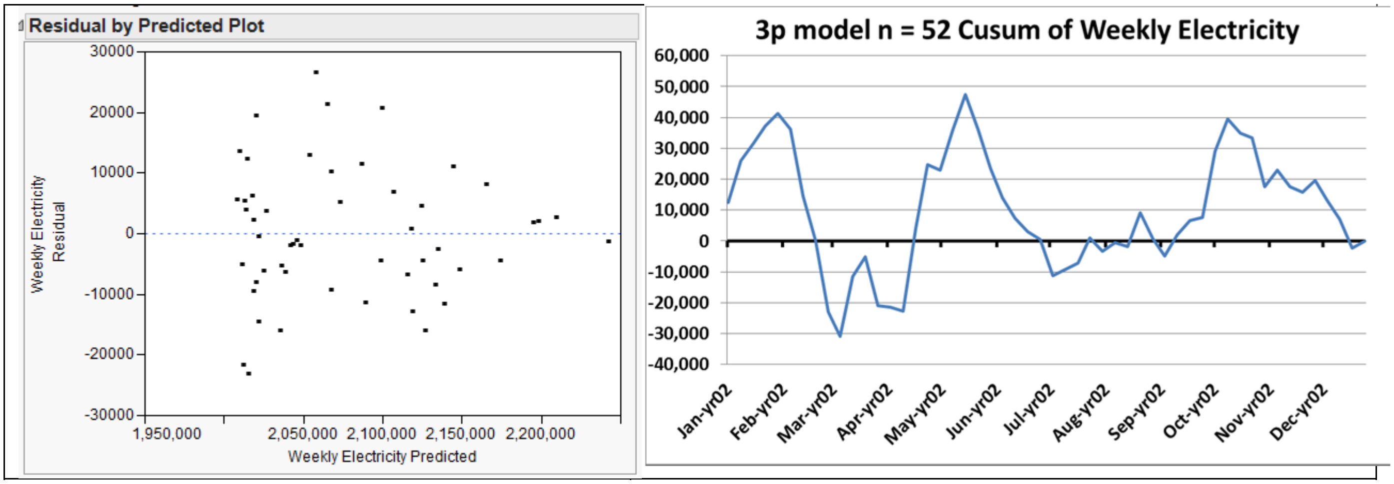


Figure 4. Graphical and Tabular Statistical Analysis Outputs of Three-Parameter Weekly Regression Model – One-Year Baseline – Improved Autocorrelation

As predicted, going to a weekly model does improve the autocorrelation, moving from an autocorrelation coefficient of 0.59 to 0.34 and a significant Durbin-Watson from 0.81 to 1.29. In order to decide which model will be the most accurate to detect savings, the Fractional Savings Uncertainty was calculated for each of the candidate models, The Fractional Savings Uncertainty output is the uncertainty of the model's predicted value for a given confidence level, divided by the expected savings from the intervention based on the methodology described in ASHRAE Guideline 14-2002[1]. The results are shown in the following table:

Table 1. Fractional Savings Uncertainty Analysis of Competing Models

Model Description	n	Autocorrelation Coefficient (Durbin Watson)	Coefficient of Variation	Fractional Savings Uncertainty at 80% Confidence[1]
two-parameter daily model using production and wet-bulb temperature – two-year baseline	729	0.76 (0.47)	0.0155	14.35%
three-parameter daily model with second-order polynomial wet-bulb and production – two-year baseline	729	0.67 (0.65)	0.010	7.93%
three-parameter daily model using second-order polynomial wet bulb and production – one-year baseline	365	0.59 (0.81)	0.0075	5.81%
three-parameter weekly model using second-order polynomial wet-bulb and production – one-year baseline	52	0.34 (1.29)	0.0053	7.17%

(1) Fractional Savings Uncertainty analysis performed based on a 365-day baseline and treatment period, assuming a project savings of 2.5%.

From the Fractional Savings Uncertainty, the daily model using the second-order polynomial wet-bulb temperature and a one-year baseline is the best model. This model still does not have ideal

1

autocorrelation, but the larger data set associated with a daily interval results in the best predictive capability among the competing models.

Conclusions

Autocorrelation is a common issue encountered in the process of developing regression-based models for industrial SEM projects. Large systems tend to stay in the same state from one period to the next, thus the persistent nature of industrial processes makes them particularly vulnerable to autocorrelation. Moreover, ambient-dependent facilities modeled at a daily interval are inherently prone to autocorrelation due to the tendency for ambient patterns to persist over multiple days, combined with the model specification errors inherent in industrial applications.

Because the presence of autocorrelation impacts the predictive capability of the model, practitioners should be familiar with tools for characterizing autocorrelation, understand common methods of improving the issue, and understand methods for evaluating competing models with different levels of statistical fitness and autocorrelation. This paper begins to explain some of tools and methods, but a practitioner with deep interest in the background and theory of autocorrelation is encouraged to pursue a more in-depth review of statistical literature.

References

ASHRAE. 2002. *ASHRAE Guideline 14-2002 For the Measurement of Energy and Demand Savings*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. Atlanta: Ga.

Efficiency Evaluation Organization (EVO). *International Performance Measurement and Verification Protocol*. 2012. 10000-1.2012. www.evo-world.org.

ESI Energy Performance Tracking (EPT) Team. 2012. *Monitoring, Targeting and Reporting (MT&R) Reference Guide – Revision 3.1*. http://www.bpa.gov/energy/n/pdf/MTR_Reference_Guide-Rev3_1.pdf. Portland, Ore: Bonneville Power Administration.

Montgomery, Douglas C., Jennings, Kulahchi. *Introduction to Time Series Analysis and Forecasting*. 2008. Pp135-139, 406. John Wiley & Sons, Inc.

Neter, John and William Wasserman, *Applied Linear Statistical Models*. 1974. Pp 249-254, 339-347.

[SEP] Superior Energy Performance, 2012, *Superior Energy Performance Measurement and Verification Protocol for Industry*. http://www.superiorenergyperformance.net/pdfs/SEP_MV_Protocol.pdf. Oakland, Calif.: The Regents of the University of California.

Contributing Authors

Todd Amundson, Bonneville Power Administration
Mimi Goldberg, KEMA
Stephen King, Triple Point Energy
Aimie McKane, Lawrence Berkeley National Laboratory
Steve Martin, Energy Smart Industrial and Cascade Energy Inc.
Keri Macklin, Triple Point Energy
Mark Thompson, Northwest Energy Efficiency Alliance and Forefront Economics Inc.
John Thornton, Northwest Food Processors Association
Julia Vetromile, KEMA