



Understanding Modeling Errors for Module Models Using Residual Analysis and Application to the Results of the Model Intercomparison

**Joshua Stein
Sandia National Laboratories
Albuquerque, New Mexico, USA**

**PV Performance Modeling Workshop
Sandia National Laboratories, Albuquerque, NM
September 22 & 23, 2010**

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.





Goal

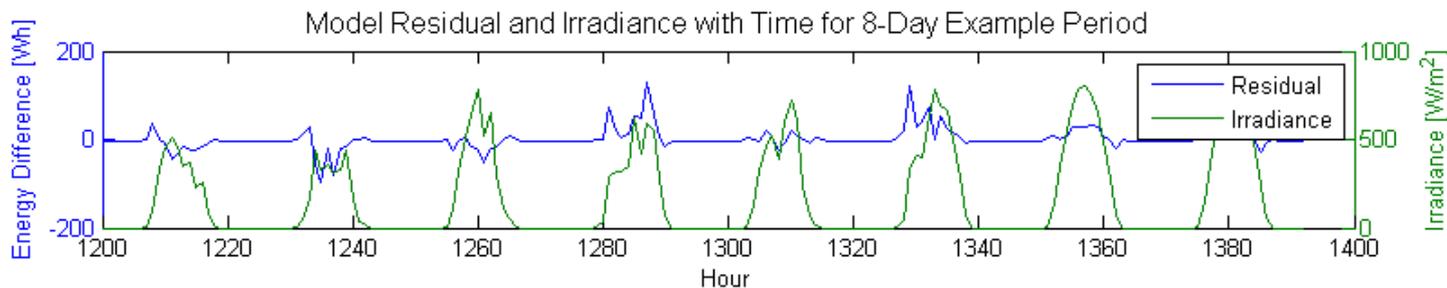
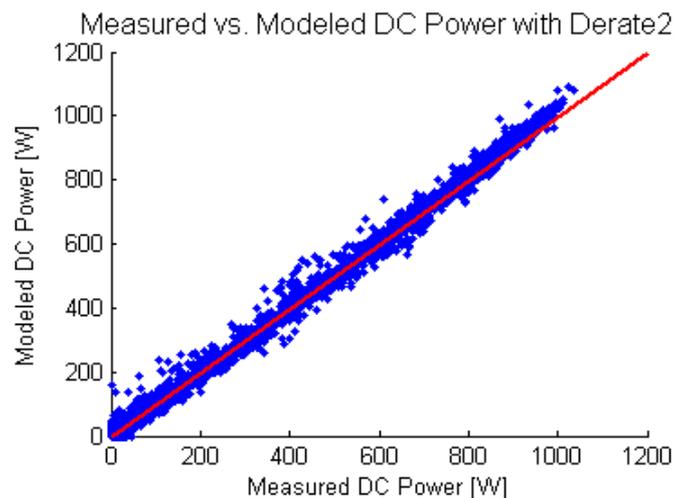
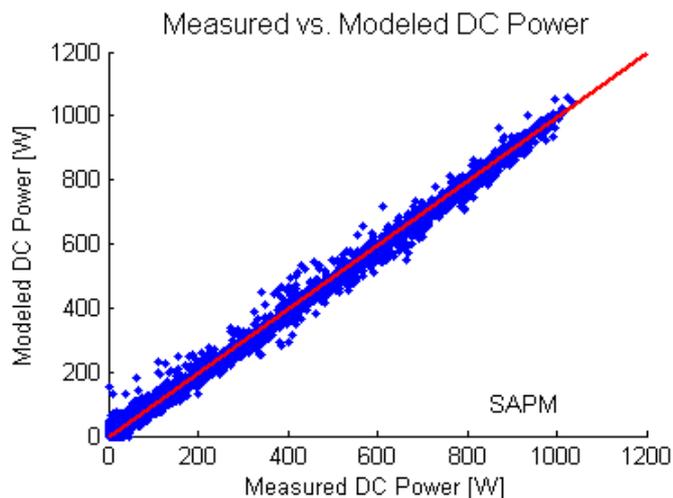
- **Examine selected participant model results using residual analysis**
 - Each participant will be given a personalized model validation report (hope to have this available within a month of the meeting's conclusion)
- **Models used as examples:**
 - Sandia PV Array Performance Model
 - PVSyst
 - 5-Par Model
 - PVWatts



What Can Be Learned? (Caveats)

- **Residual analysis examines both model and input parameters as a set.**
 - Non random residuals can reflect problems with the model and/or problems with the input parameters.
- **Used iteratively, residual analysis is a powerful tool for model development and validation.**
- **The examples shown here are meant as an example of the approach only.**
 - Do not generalize these results
 - They apply to the specific system and installation
 - When patterns are repeatable with multiple systems, generalizations may be appropriate.

Sandia PV Array Performance Model



Total Measured Energy = 1932830 Whr

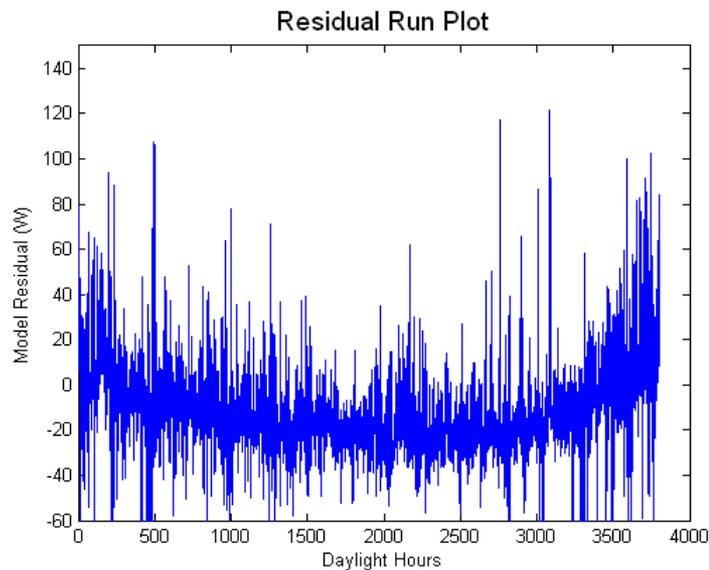
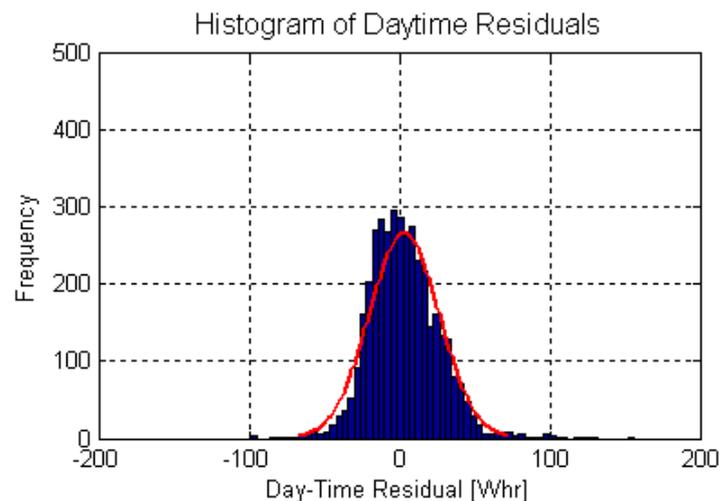
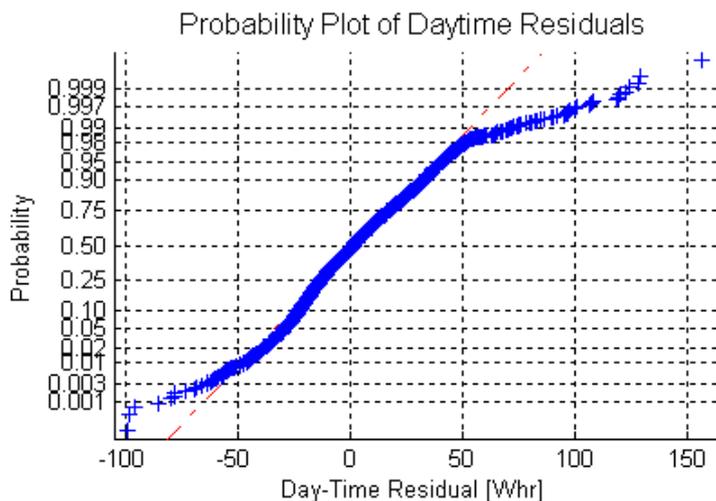
Total Modeled Energy = 1875467 Whr

Percent Energy Difference Before Derate = -2.9678%

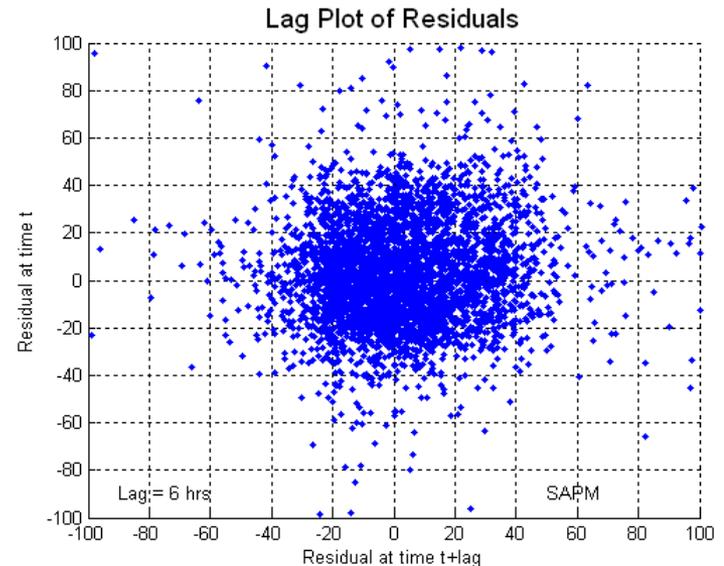
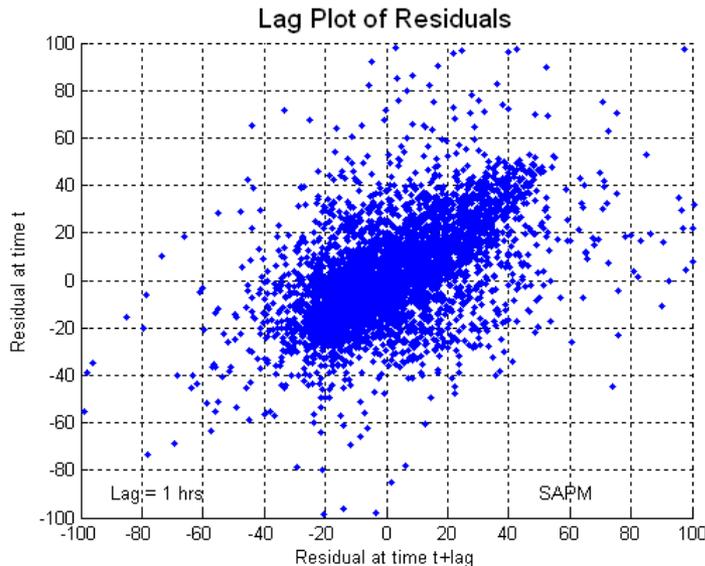
Derate1 (minimize Sum LSE) = -1.7682%

Derate2 (Sum of Residuals = 0) = -3.0586%

Distribution of Residuals (SAPM)

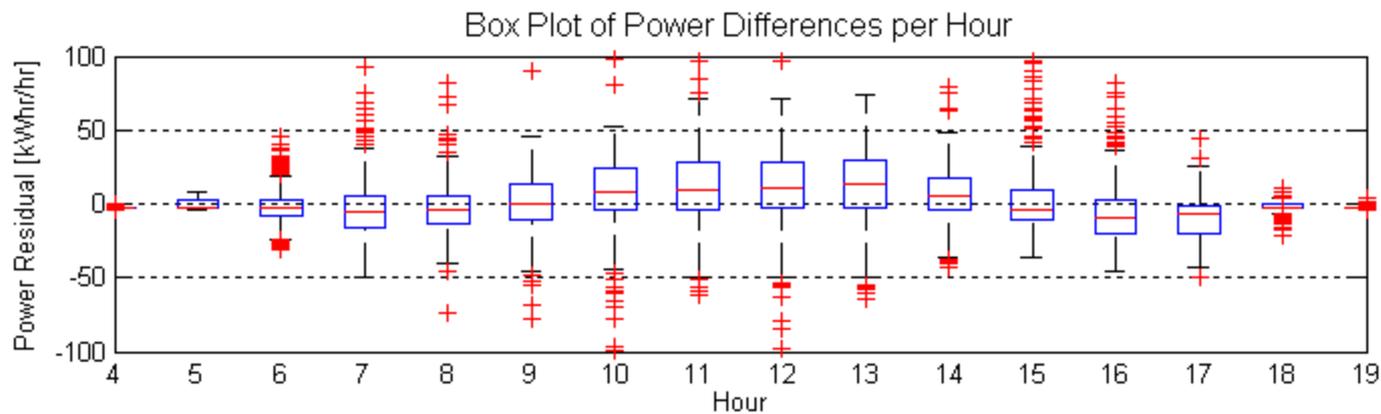
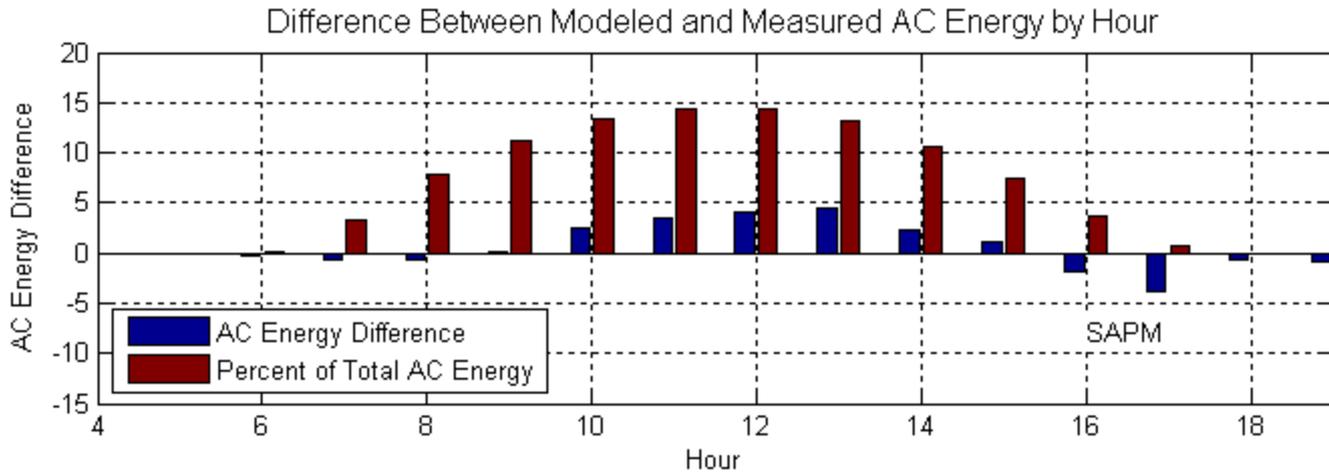


SAPM Lag Plots

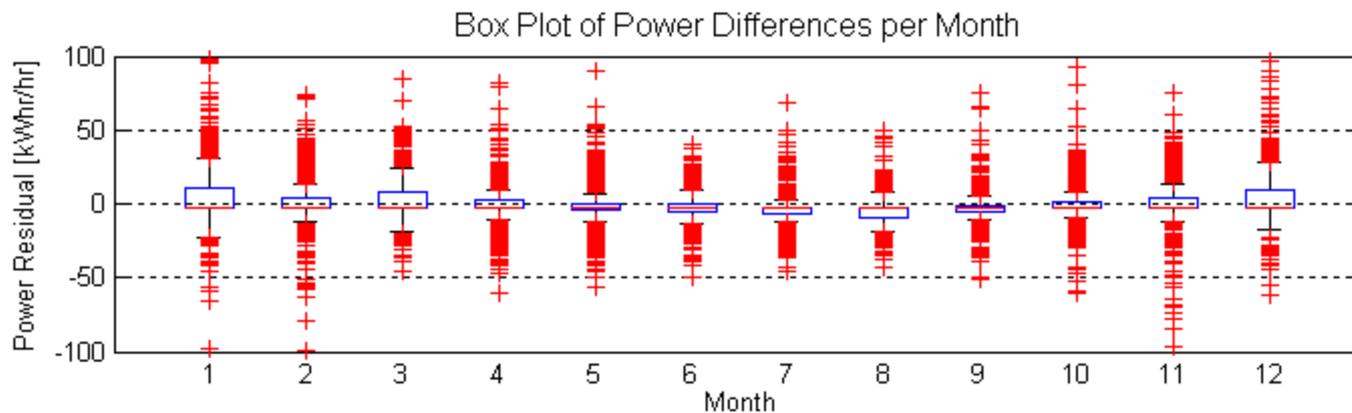
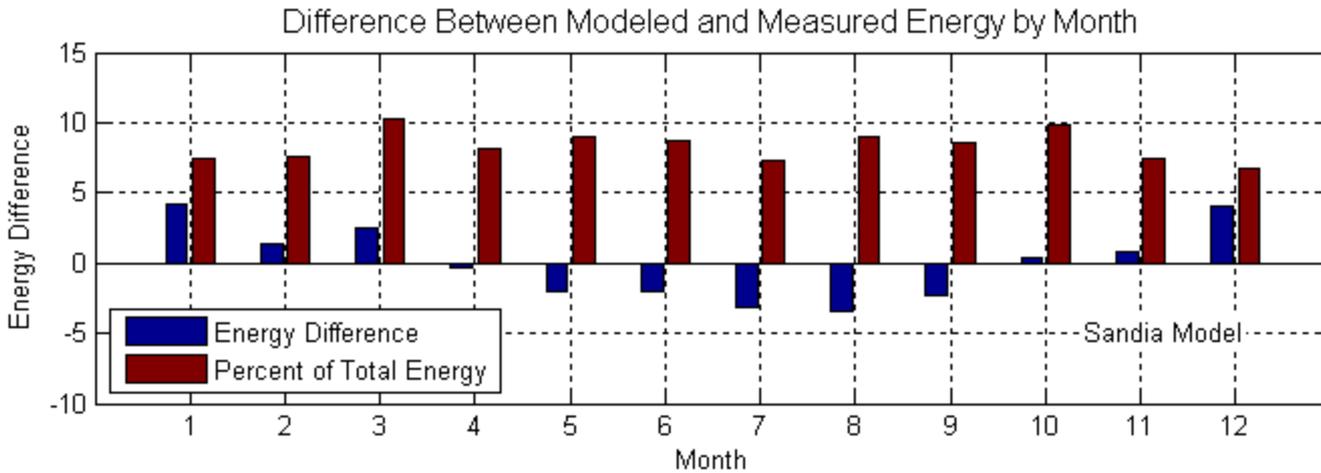


Slight correlation in 1-hr lag plot suggests residuals are autocorrelated at 1-hour. This pattern indicates a systematic bias associated with time in the model. Since it disappears at 6 hours, it is probably related to a diurnal process. Interesting but not all that informative because nearly every variable is diurnally-controlled.

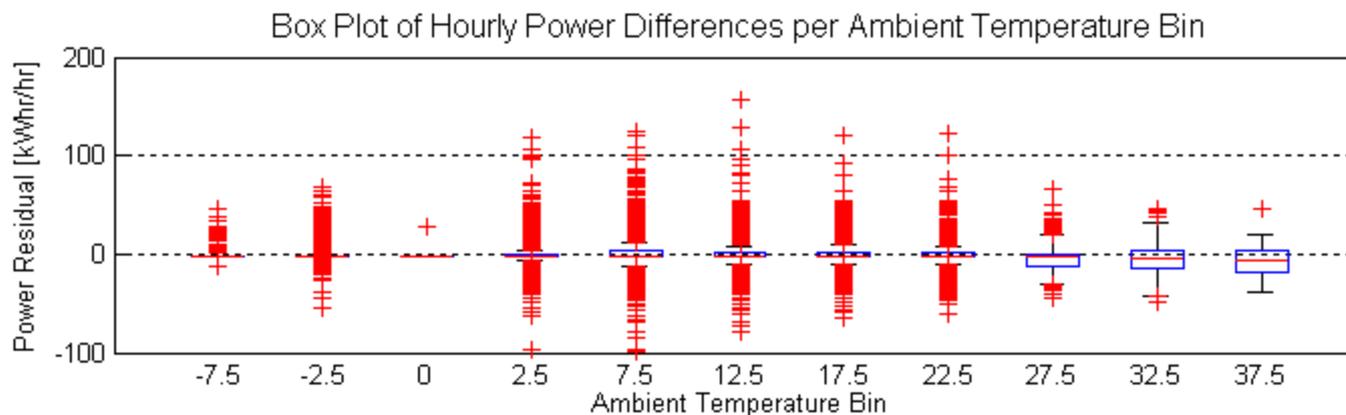
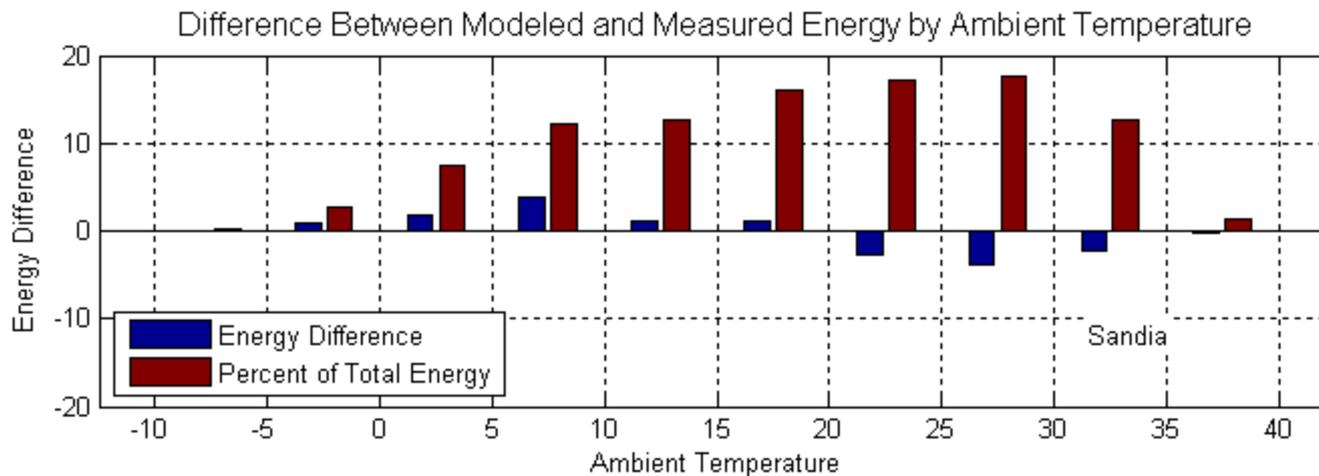
SAPM by Hour



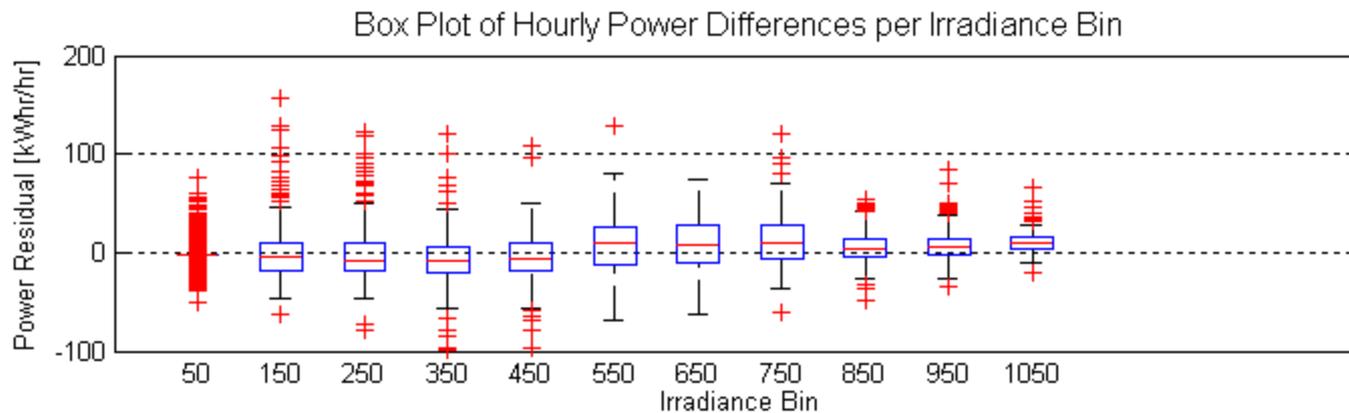
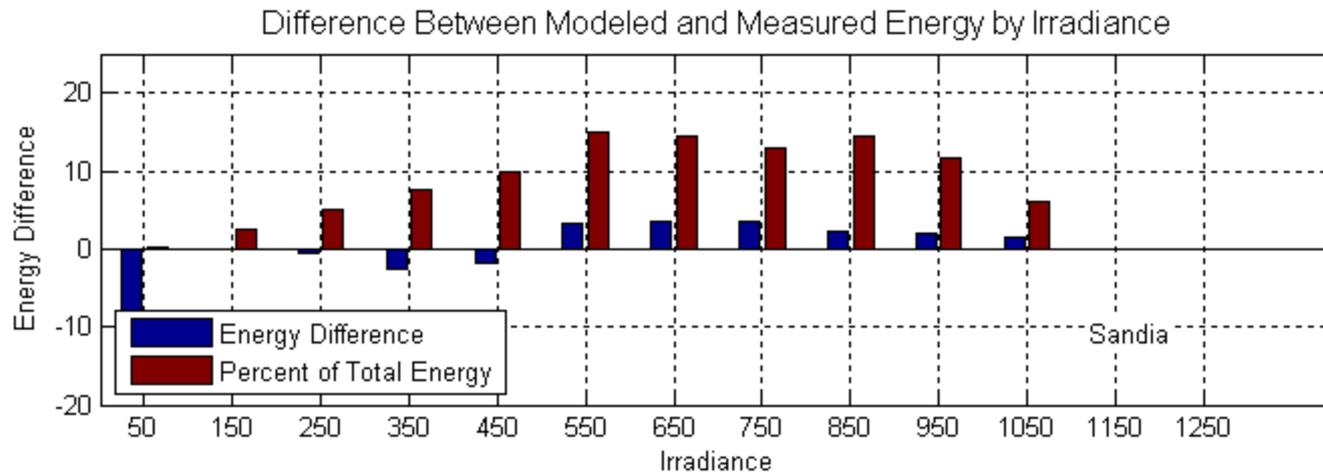
SAPM by Month



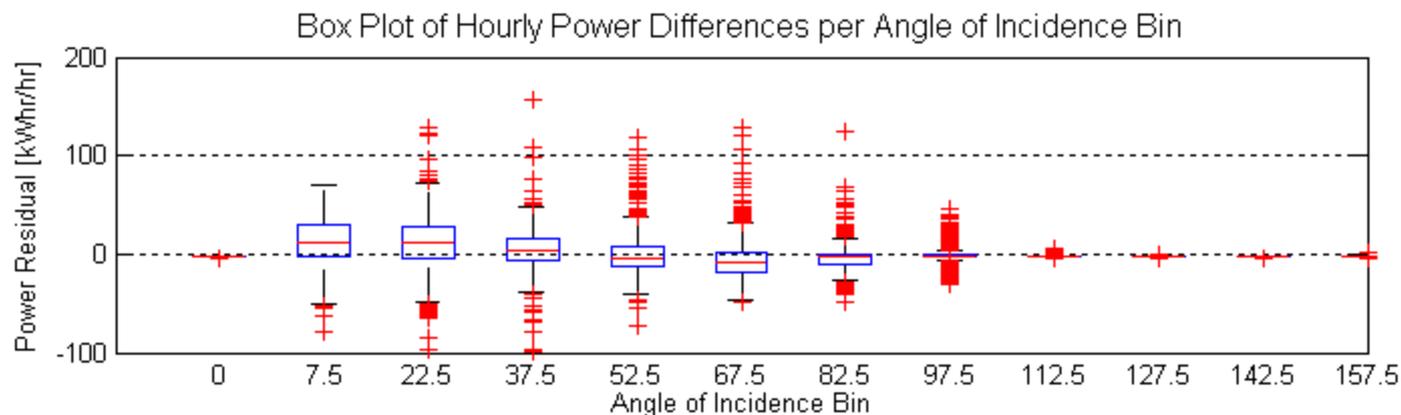
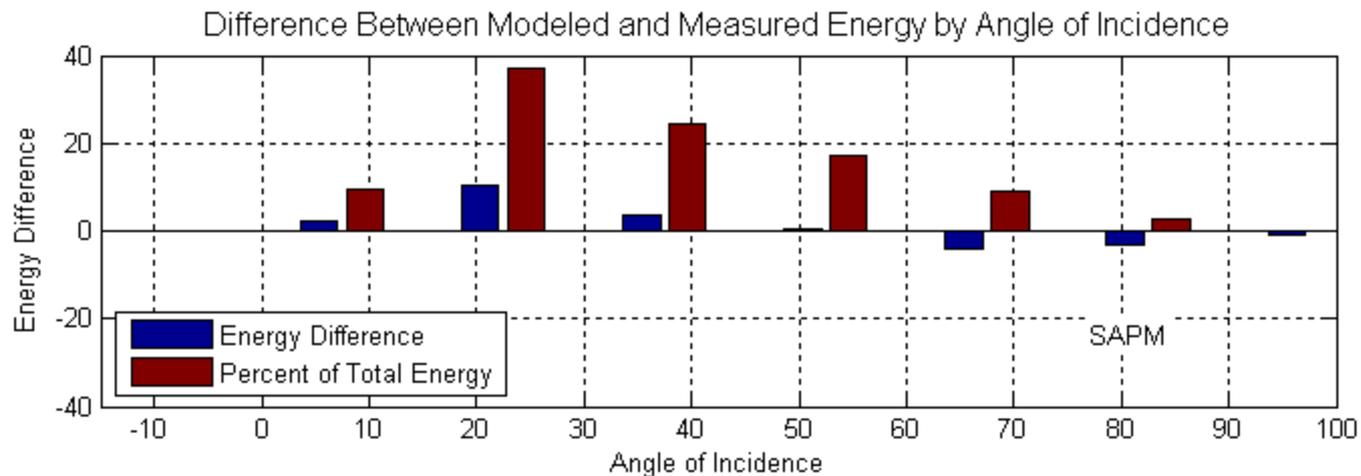
SAPM by Temperature



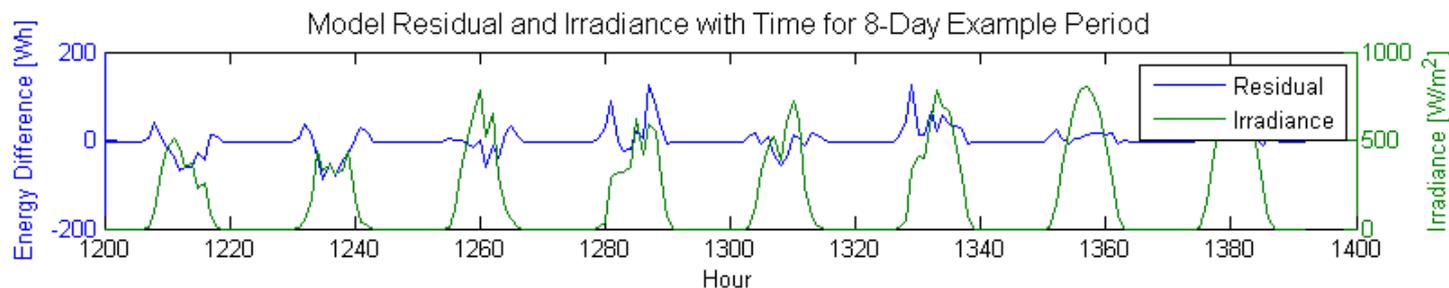
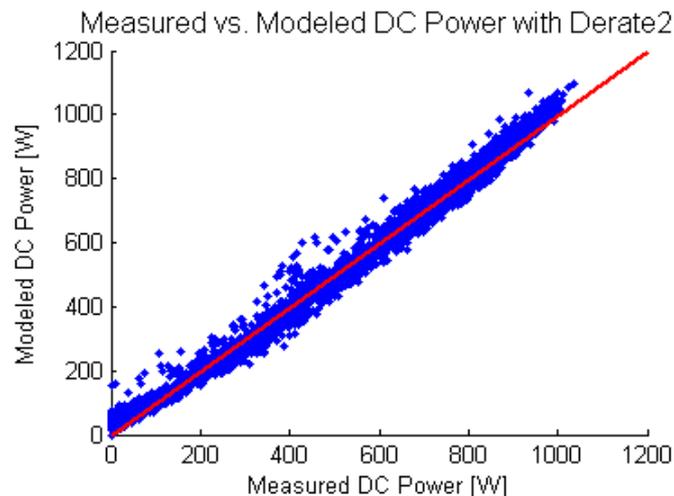
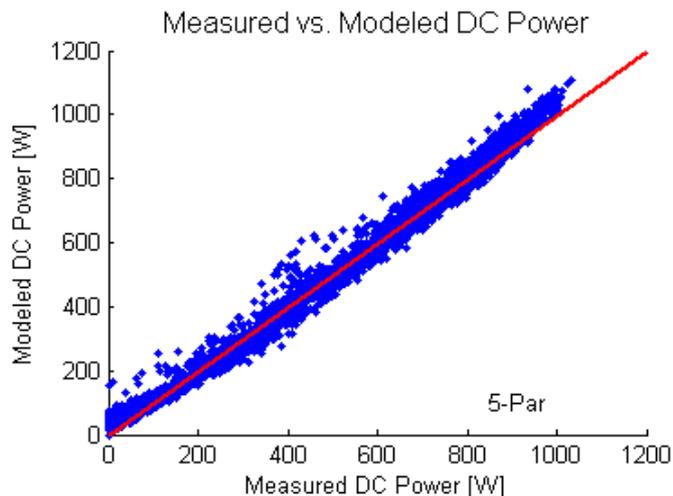
SAPM by GHI



SAPM by AOI



5-Par Model



Total Measured Energy = 1932830 Whr

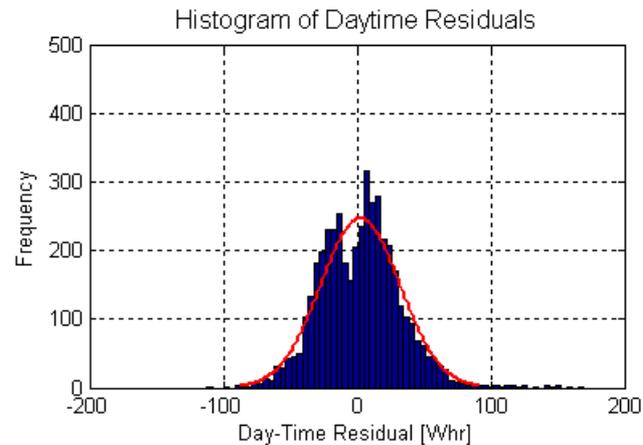
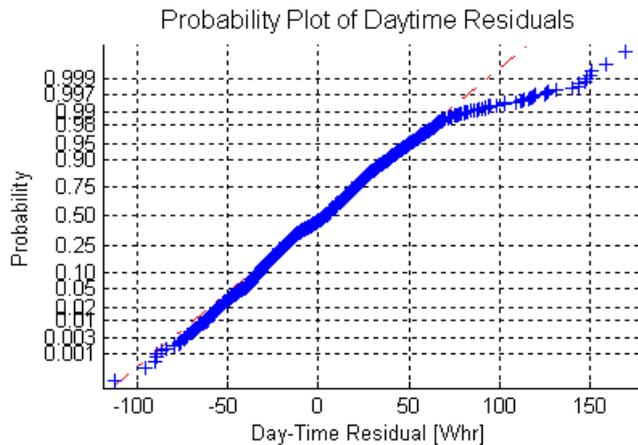
Total Modeled Energy = 1952765 Whr

Percent Energy Difference Before Derate = 1.0314%

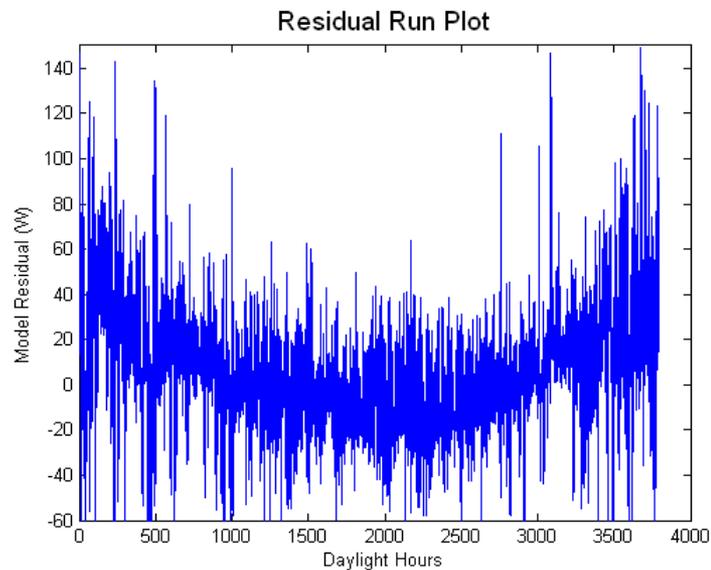
Derate1 (minimize Sum LSE) = 1.115%

Derate2 (Sum of Residuals = 0) = 1.0209%

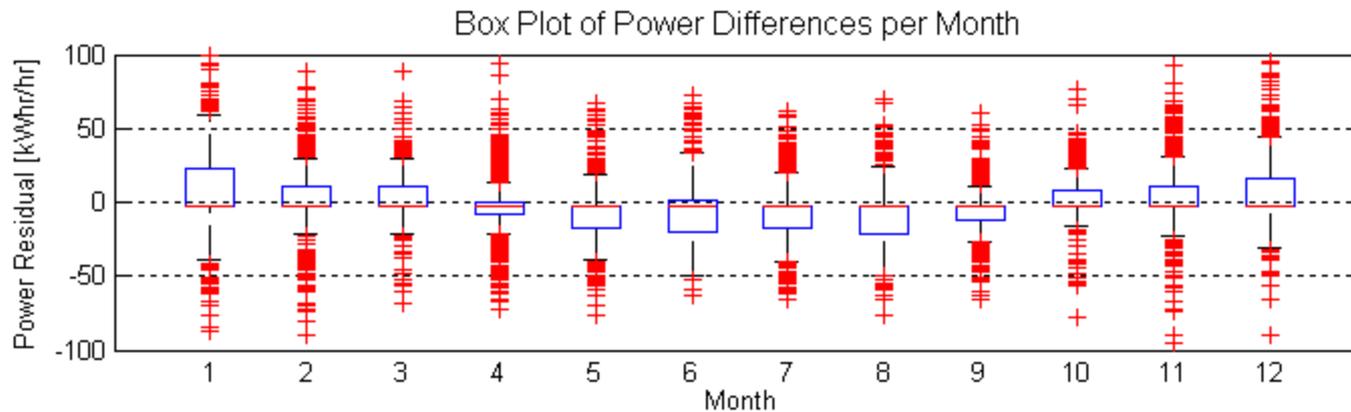
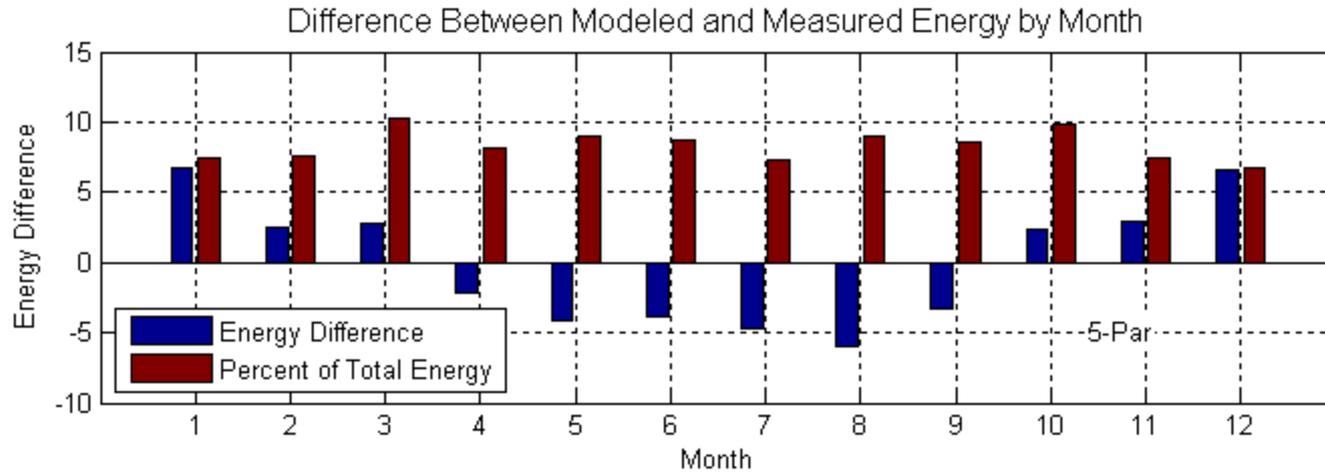
Distribution of Residuals (5-Par)



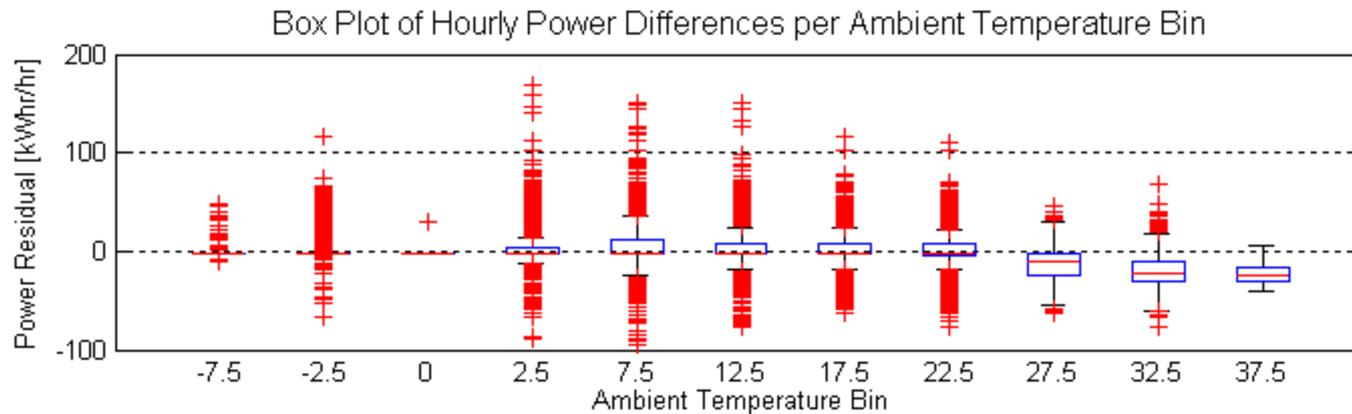
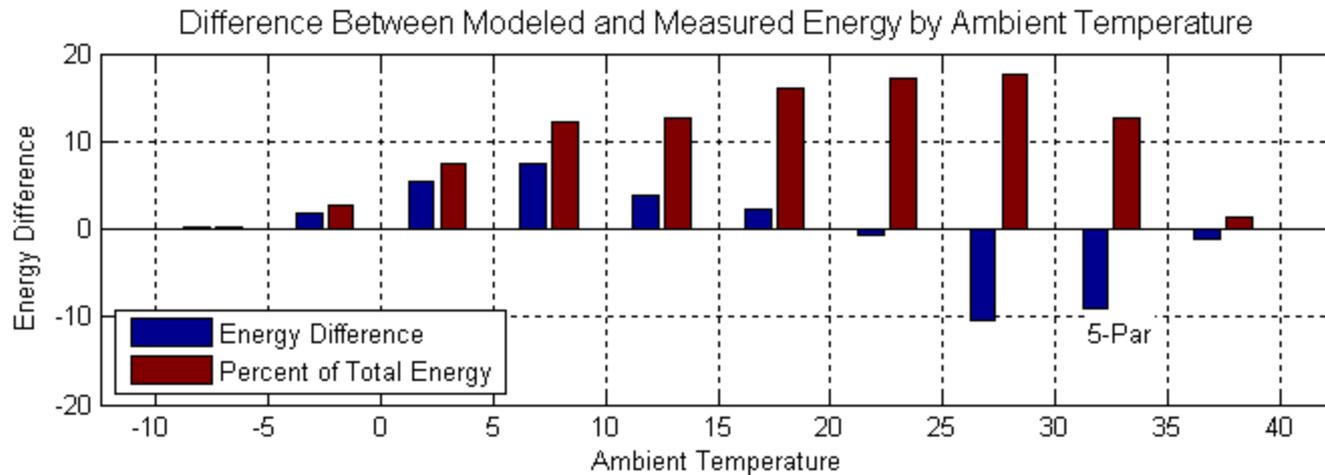
Root Mean Square Error = 29.9768 Whr
Mean Bias Error (derated) = 2.4074 Whr
Correlation Coefficient = 0.99775



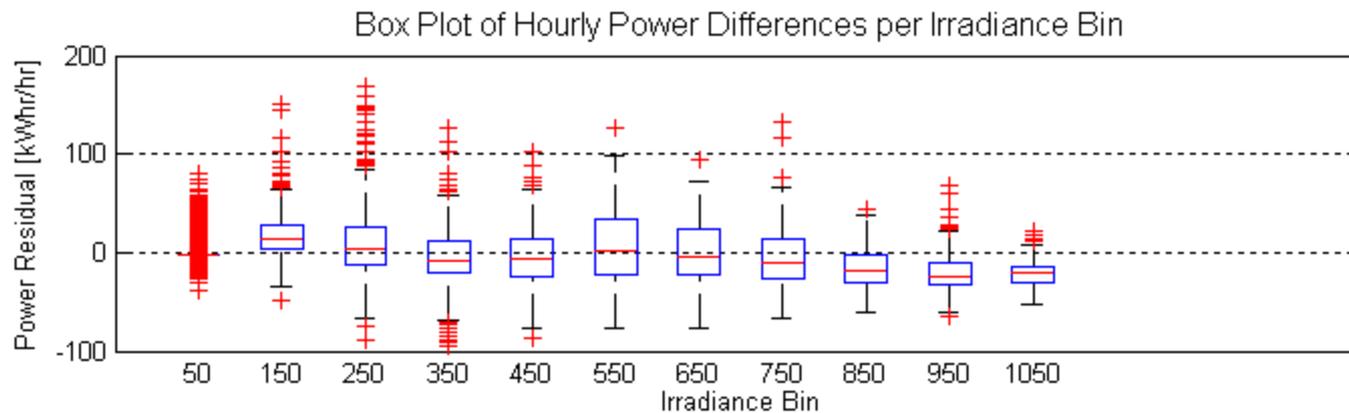
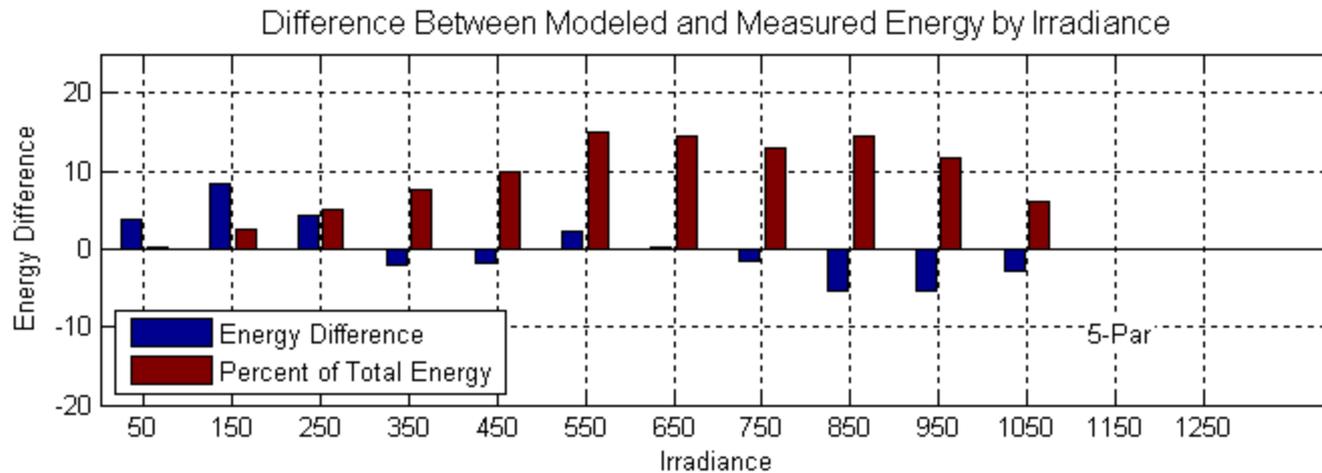
5-Par by Month



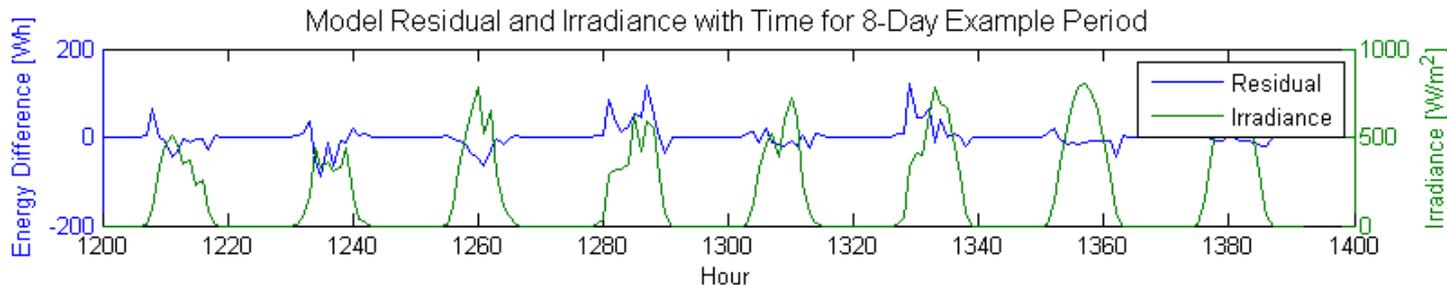
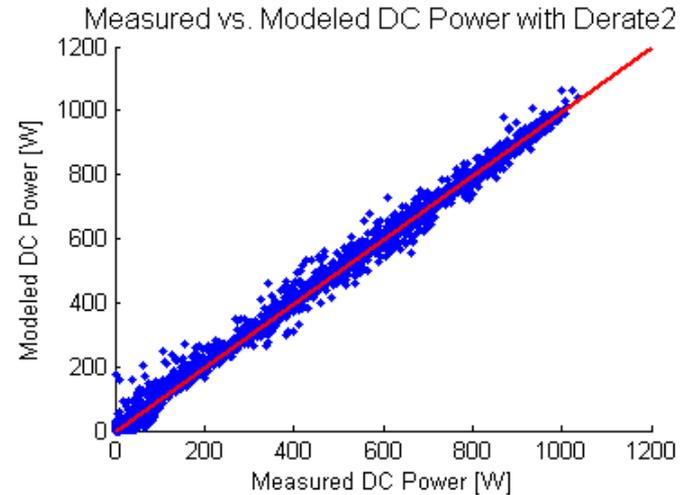
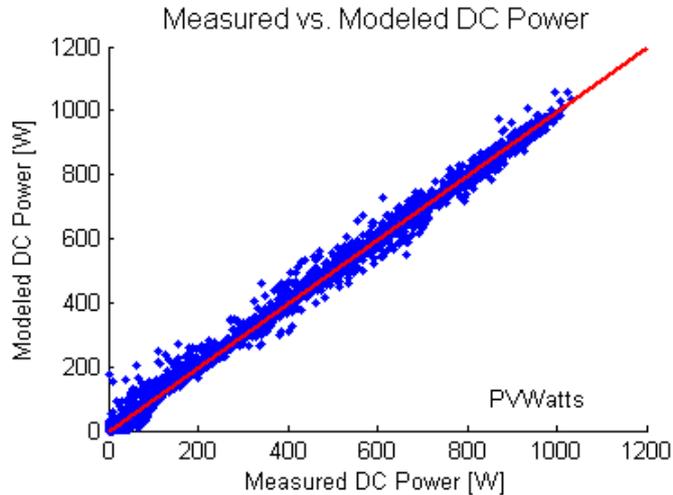
5-Par by Temperature



5-Par by GHI

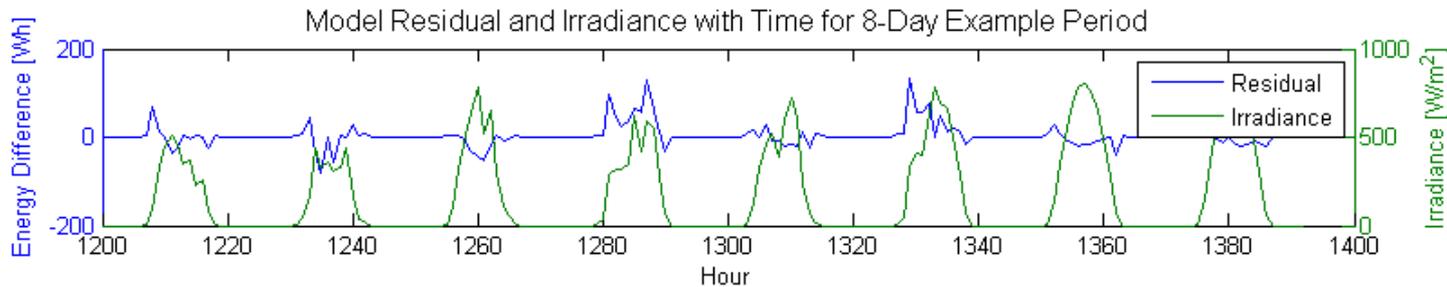
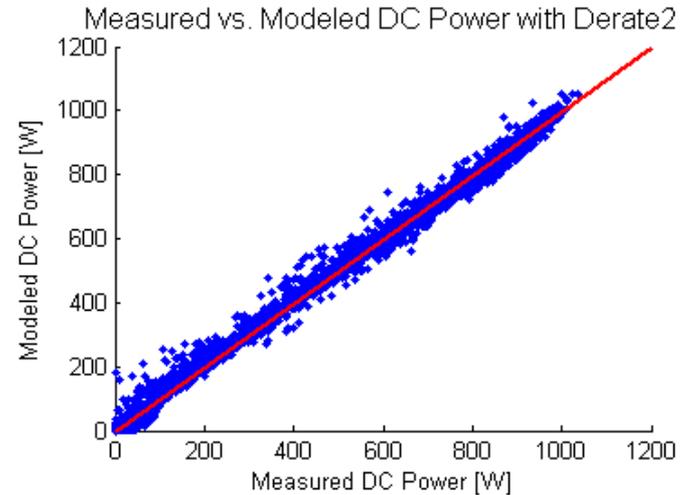
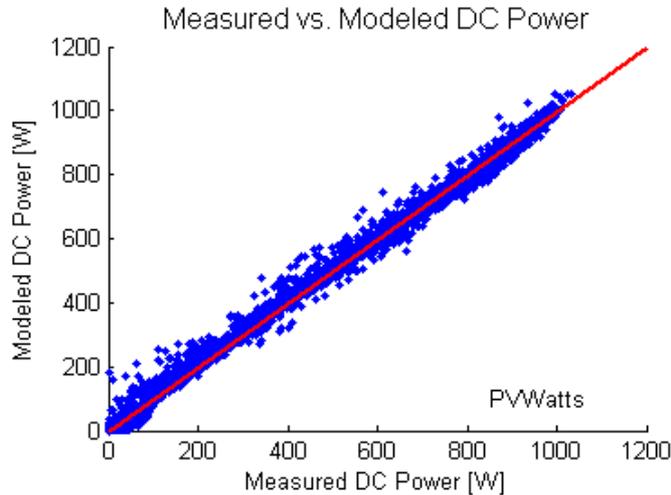


PVWatts(a) Model



Total Measured Energy = 1932830 Whr
Total Modeled Energy = 1924835 Whr
Percent Energy Difference Before Derate = -0.41365%
Derate1 (minimize Sum LSE) = -0.84158%
Derate2 (Sum of Residuals = 0) = -0.41537%

PVWatts(b) Model



Total Measured Energy = 1932830 Whr

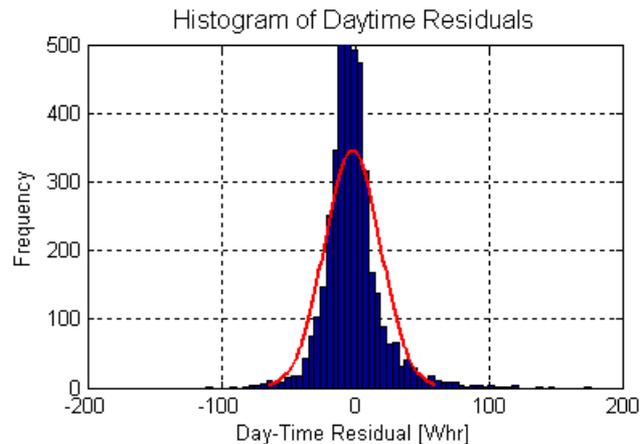
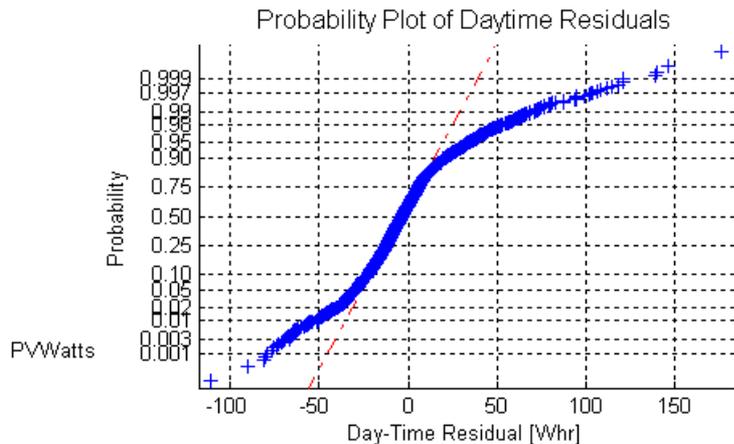
Total Modeled Energy = 1932876 Whr

Percent Energy Difference Before Derate = 0.0023865%

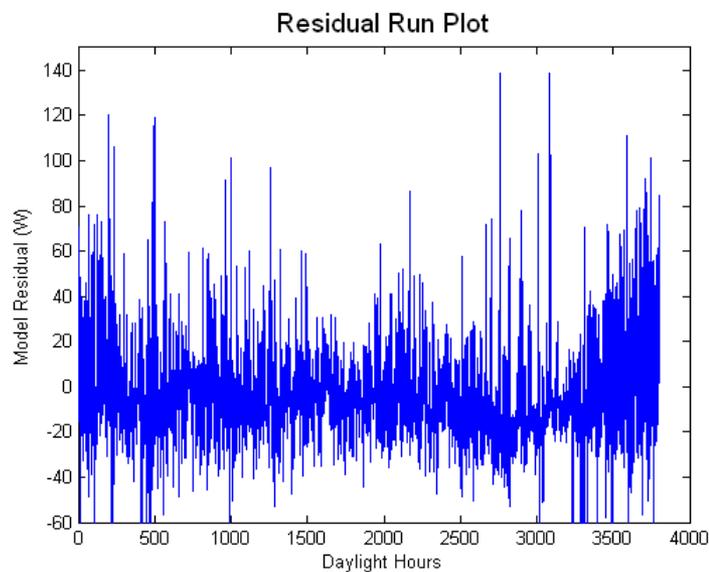
Derate1 (minimize Sum LSE) = -0.70584%

Derate2 (Sum of Residuals = 0) = 0.0023865%

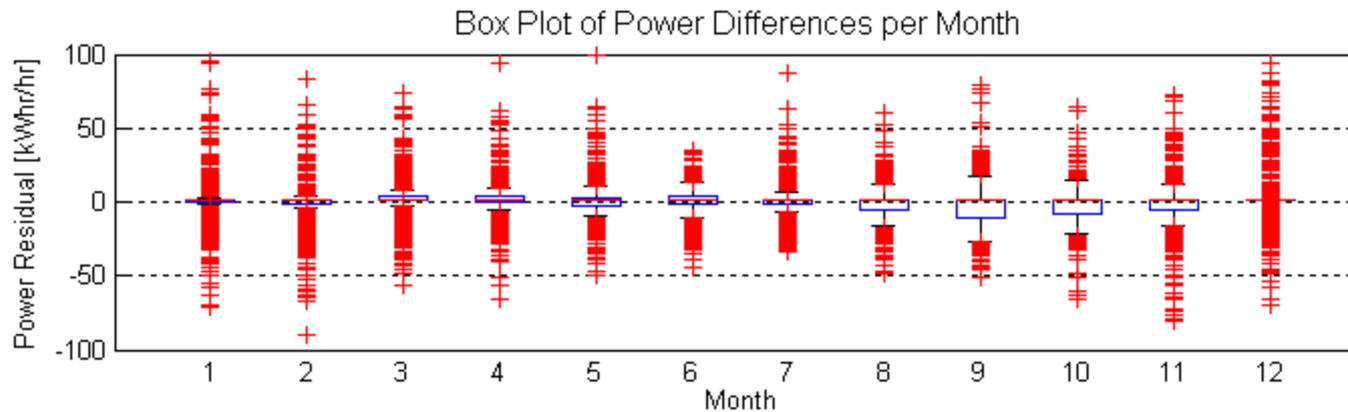
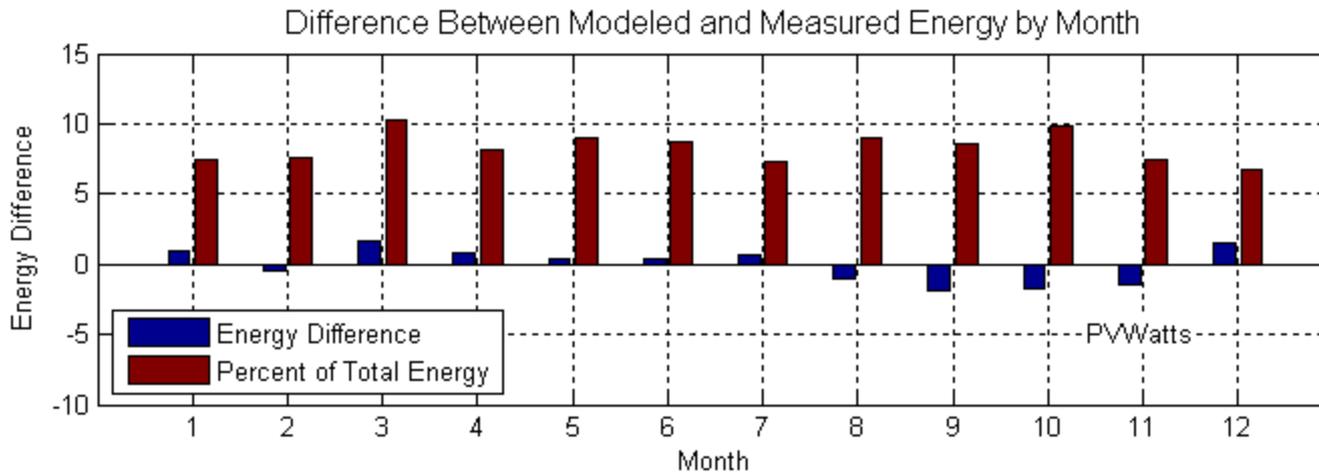
Distribution of Residuals (PVWatts)



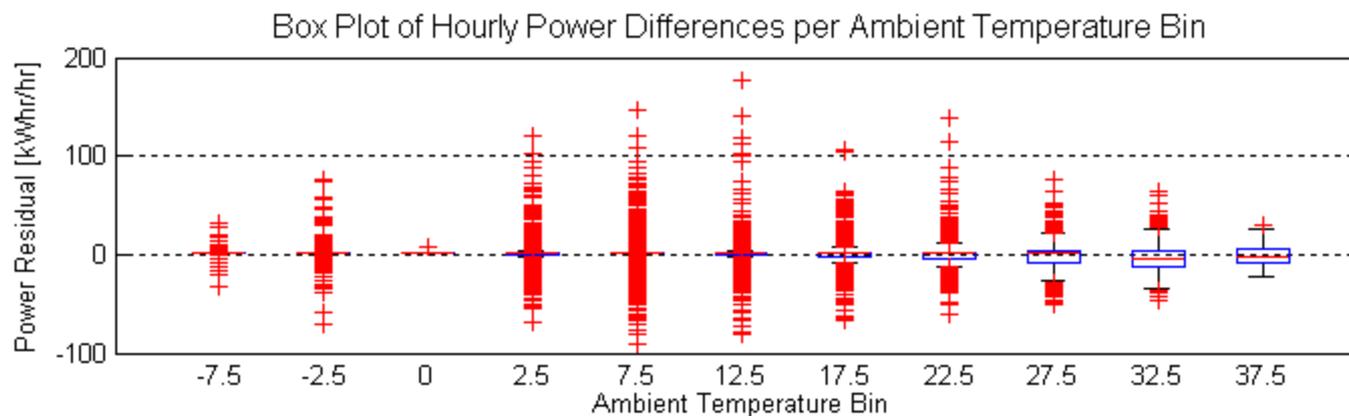
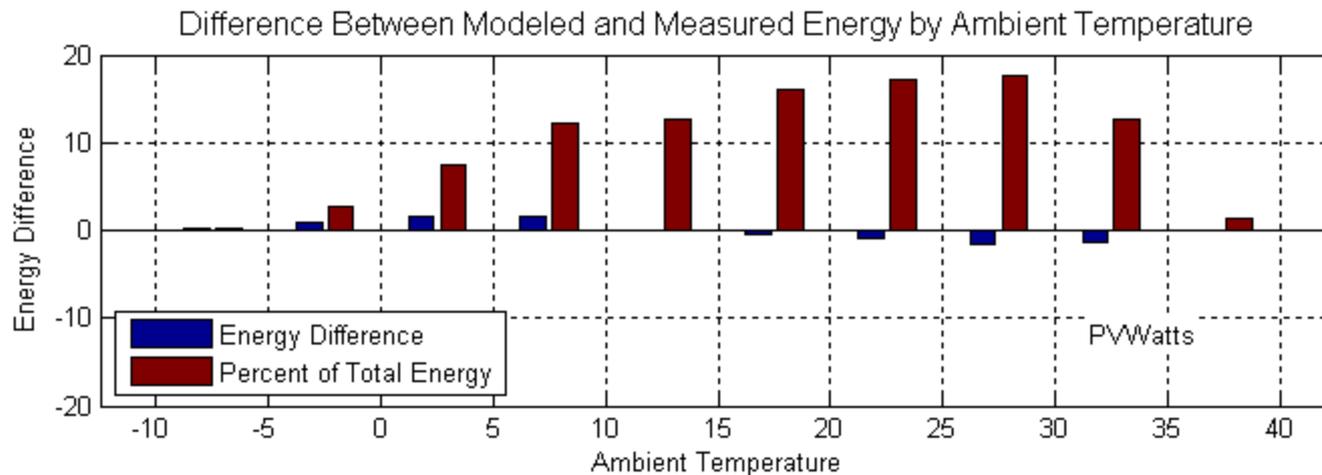
Root Mean Square Error = 20.9807 Whr
Mean Bias Error (derated) = -1.8025 Whr
Correlation Coefficient = 0.999



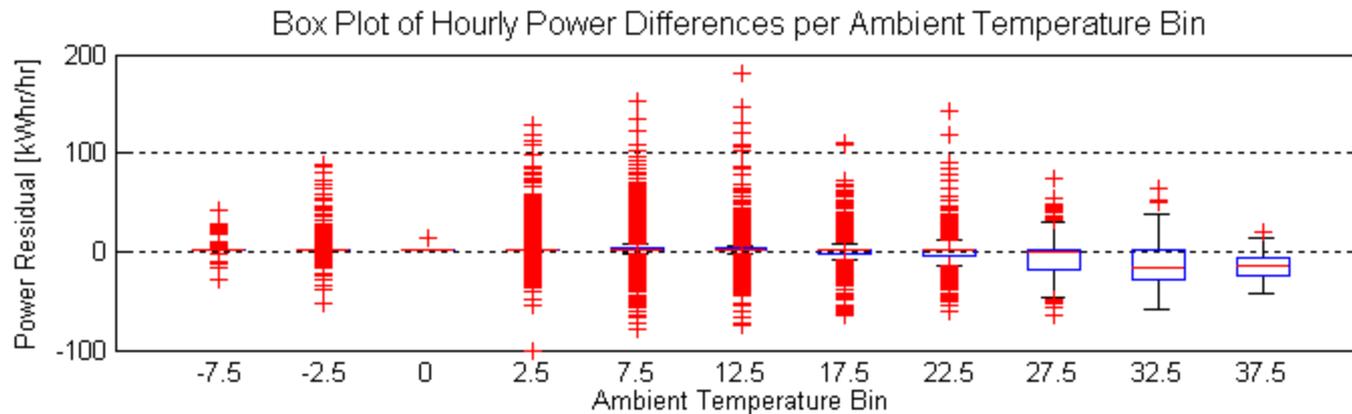
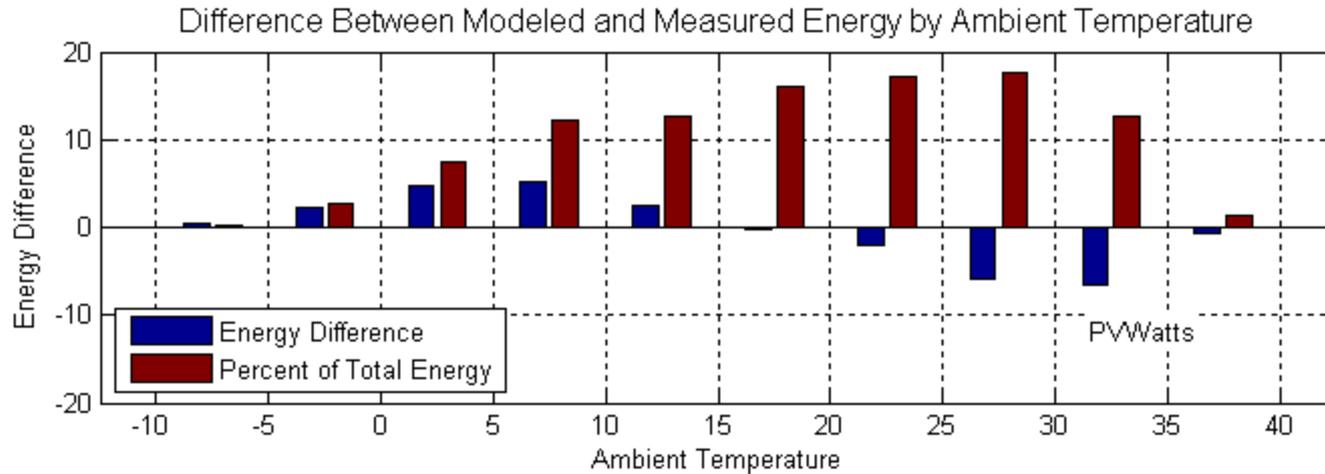
PVWatts by Month



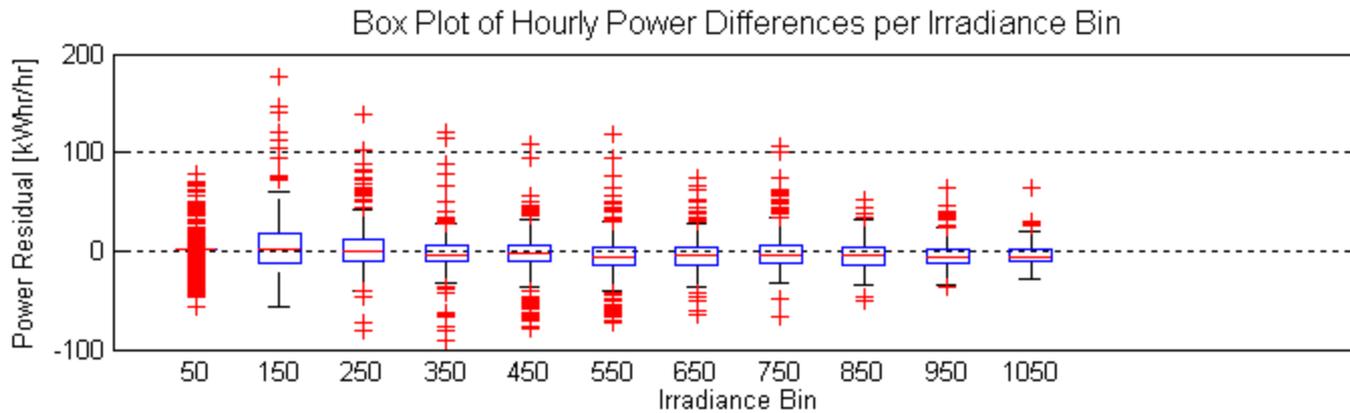
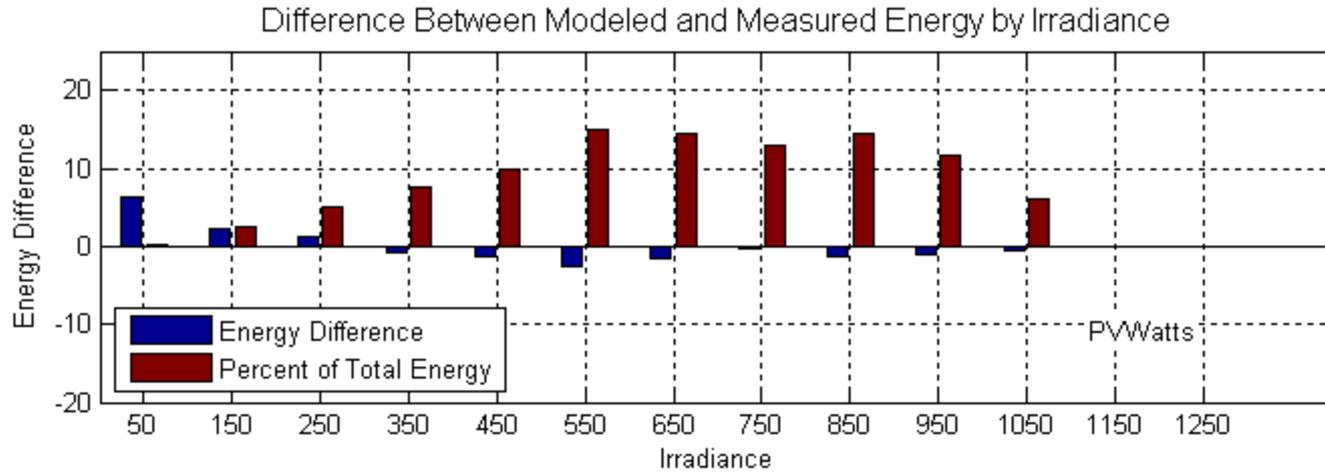
PVWatts(a) by Temperature



PVWatts(b) by Temperature



PVWatts by GHI





Conclusions

- **For the system examined, several models exhibited correlation between power residuals and temperature.**
 - SAPM, PVSyst, and 5-Par models overpredict power at low temperatures (below 20 C) and underpredict at high temperatures (above 20 C)
 - This might suggest the temperature coefficient is too large?
 - PVWatts did not exhibit this pattern (but not all PVWatts model runs are identical)
- **Residual Analysis provides important information that can be used for model refinement and model validation.**