FORECASTING SOLAR RADIATION FOR THE LOS ANGELES BASIN – PHASE II REPORT

James Hall
JHtech
PO Box 877
Divide, CO 80814

Jeffrey Hall
JHtech
PO Box 877
Divide, CO 80814

ABSTRACT

A prototype solar radiation forecast system for the Los Angeles basin has been developed utilizing data from existing solar radiation sensors, the National Digital Forecast Database and meteorological observations from local airports. The ensemble-style model contains two components: one predicts solar radiation using non-linear time series analysis of recent meteorological observations; the second predicts solar radiation based on recognition of daily and seasonal patterns. Both components are adaptive and re-calibrate themselves daily based on the training data presented to them. Outputs from the two components are weighted and combined to forecast solar radiation one and three hours into the future. Forecasts were generated for four locations: Beverly Hills, Santa Monica, Hacienda Heights and Long Beach. The average forecast errors in out-of-sample testing for all four sites were 25% and 33% for one hour and three hour forecasts respectively.

1. INTRODUCTION

Expanding the use of renewable energy sources, such as wind and solar, has become a US energy priority. California has legislated that 20% of their energy be generated by renewable sources within five years. However, the rapidly changing output of photovoltaic (PV) systems under cloudy conditions makes it difficult to utilize PV as a significant source in commercial power grids. Many feel that without reliable forecasts of solar energy, PV technology will never reach its potential as a significant energy resource.

Previous research has attempted to forecast solar radiation using numerical prediction, cloud movement or statistical time series models. Nearly all of this work is in the early stages of development, with only a few researchers reporting results from out-of-sample testing at specific locations. The best forecasting results we have found in the literature for true out-of-sample tests are around 35% relative mean average total error (rMAE) for one hour forecasts and 50% total error (rMAE) for three hour forecasts. [1]

One source of solar radiation data that has not been utilized by previous researchers comes from over 3600 medium-quality sites across the U.S. that measure solar radiation and make their hourly and daily observations publicly accessible. These sites are professionally operated and maintained by universities and government agencies for specific local purposes such as agriculture, water management and environmental monitoring. Average errors from these sites have been shown to be about half that from satellite-based observations.[2] Figure 1 shows that within the greater Los Angeles area there are over 150 such sites in current operation that provide hourly measurements of GHI (global horizontal irradiance).

2. PHASE I SUMMARY

Phase I of this study was the development of a prototype forecasting model for solar radiation at a single site (Fontana, CA). This site was challenging due to the complex interaction between ocean, winds, man-made haze and natural cloud formations.

Fig. 1: Hourly observations of solar radiation and meteorological parameters are available from over 150 locations in the greater Los Angeles region.

An ensemble-style forecasting model was developed for this prototype with two separate components. One component predicted solar radiation using non-linear regression on recent meteorological observations. The second component predicted solar radiation based on daily pattern recognition. Both components were adaptive and re-calibrated themselves daily based on the training data presented them, which in the initial testing was a 30 day sliding window of historical data. The component models were tuned for the greatest accuracy between the hours of 10 AM and 2 PM, this being a critical forecasting period for solar-electric utilities. Outputs from the two components were weighted and combined to forecast solar radiation one hour and three hours into the future. An advantage of the ensemble approach is that additional components can easily be incorporated into the model as they are developed.[3]

In Fontana, a large portion of the summer and fall days are clear, making solar forecasting relatively straightforward. However the rainy and rapidly changing weather of January thru March will challenge any forecasting system. The forecast accuracy in a blind test for May 2009 was 16% (rMAE) for one hour forecasts and 25% (rMAE) for three hour forecasts. The forecast accuracy for February 2010 was 27% (rMAE) for one hour forecasts and 44% (rMAE) for three hour forecasts. A sample of the model output for February 2010 is shown in Fig. 2.

3. <u>PHASE II DEVELOPMENT</u>

In Phase II of the development, the model was expanded to make forecasts for multiple sites over an area about 50 by 50 km (30 by 30 miles) surrounding Los Angeles proper. For simplicity, inputs to the Phase II model only used observations of solar radiation and the corresponding meteorological data available from the medium-quality observation sites in the region. Most of these observations were available hourly, and represented the average of the previous hour's readings. These inputs were asynchronous, i.e. the hourly averaging periods began at different times. In addition, the time delay before reporting the data was not constant from site to site. The asynchronous nature of the input data required the development of new techniques for processing, quality controlling and combining the observations before presenting them as inputs to the model.

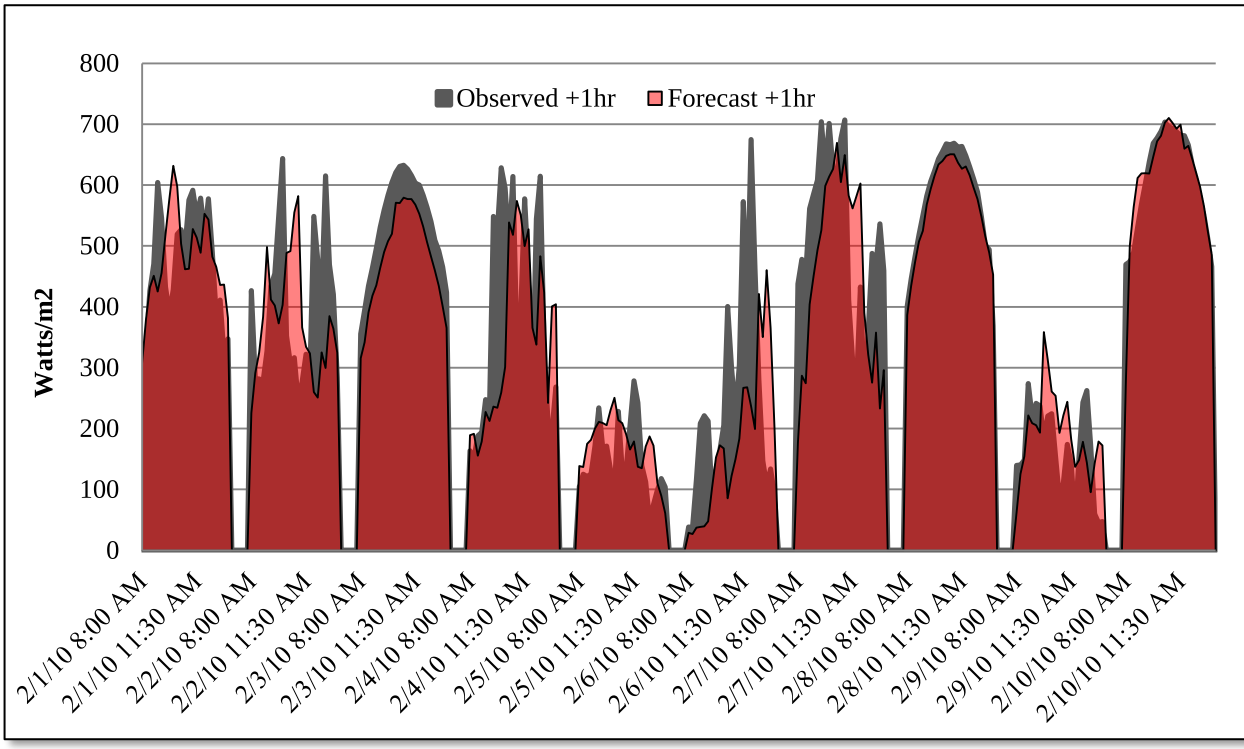

Fig. 2: Sample of February 2010 forecasts for solar radiation one hour ahead at Fontana, CA.

Now that asynchronous data management techniques have been developed, we will re-incorporate inputs from airport meteorological observations and the National Digital Forecast Database into future models, which should improve forecast accuracy even further.

In Phase I, during cloudy days we observed that the GHI measurements tended to oscillate, causing some over-correction in the model. This resulted in significantly increased errors whenever the forecasts and actual observations would oscillate out of phase. In the Phase II model, additional dampening was applied to the three hour forecasts to reduce this oscillation and the corresponding errors.

Fig. 3: Test locations for Phase II solar forecasts.

4. RESULTS

After all development work on the Phase II model was completed, an out-of-sample test was conducted to evaluate the performance. The accuracy was evaluated at four specific locations in the Los Angeles region having medium-quality sensors: Beverly Hills, Santa Monica, Hacienda Heights and Long Beach (Fig.3). Since the model is adaptive, it was first presented with hourly data from July thru August 2010, to allow for auto-calibration. Next, test data from September 1, 2010 through October 7, 2010 were presented one hour at a time and solar forecasts were generated. This test data had not been used in any previous phase of the model development, and it included four separate periods of cloudy and rainy weather which would be difficult for any solar forecasting system

The accuracy for Phase II tests were comparable to those reported for the Phase I model at Fontana, CA. The lowest average forecast errors were for Santa Monica with 22% average error (rMAE) for one hour forecasts (Fig. 4) and 33% average error (rMAE) for three hour forecasts (Fig. 5). The highest errors were for Long Beach with 27% average error (rMAE) for one hour forecasts (Fig. 6) and 35% average error (rMAE) for three hour forecasts (Fig. 7). The average forecast errors for all four sites were 25% and 33% for one hour and three hour forecasts respectively. Remembering that the medium-quality sensors used for the observed values in the error calculations have a total error of 5-10%[2], one can see that a significant portion of the total forecast error may come from uncertainty in the reference observations themselves.

Several researchers have published preliminary results for their solar forecasts; a typical range for the better models being 32-40% (total root mean square error) for one hour forecasts and 35-45% (total root mean square error) for three hour forecasts. However these results are not directly comparable to this study, as the researchers were fitting the data and not conducting out-of-sample tests.

The only published results found that were directly comparable were by Reikard[1], who conducted an out-of-sample comparison of several time-series models. The best models reported by Reikard had 35% total error (rMAE) on one hour forecasts and 51% total error (rMAE) on three hour forecasts. These errors were significantly higher than the results in this pilot study.

5. CONCLUSION

This pilot study has validated an adaptive modeling technique that can produce useful forecasts of solar radiation for multiple locations in the Los Angeles basin. To our knowledge, the results demonstrated significantly lower forecasting errors than reported in any previous work. In addition, these results were obtained in true blind tests with a model that was adapting and continuously re-calibrating itself without human intervention. These factors significantly increase the likelihood that the model will succeed in an actual deployment where conditions from year to year are seldom the same. With the large quantity of historical and near-real time observations available from medium-quality ground sites, it should be feasible to construct a reliable solar forecasting model for the greater Los Angeles area, or for any other region of the US.

6. FUTURE WORK

The next phase of development will be a pre-production system for solar forecasting suitable for use by electric utilities. A real-time data acquisition system for gathering, processing and quality controlling the solar radiation and meteorological data from the entire region will be implemented and on-going live forecasts of solar radiation for the Los Angeles basin will be generated.

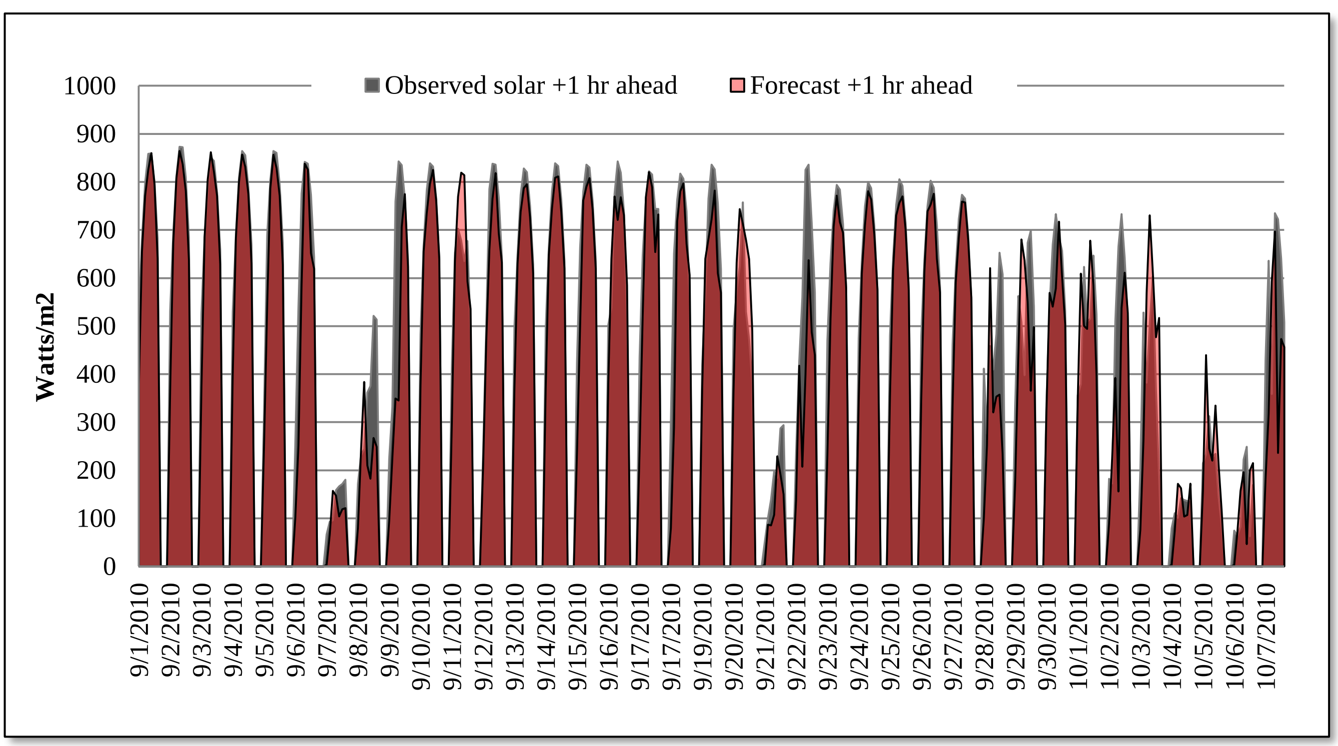

Fig. 4: Forecasts for solar radiation one hour ahead at Santa Monica, CA.

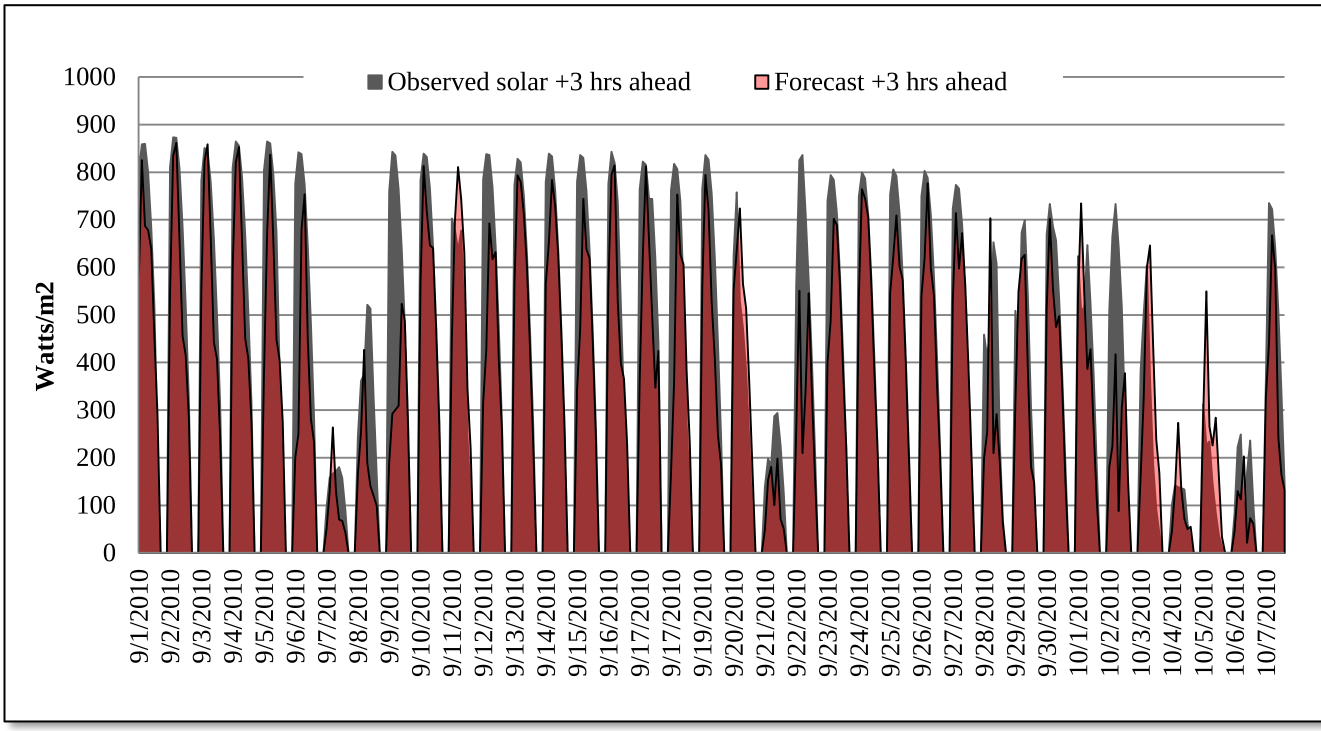

Fig. 5: Forecasts for solar radiation three hours ahead at Santa Monica, CA.

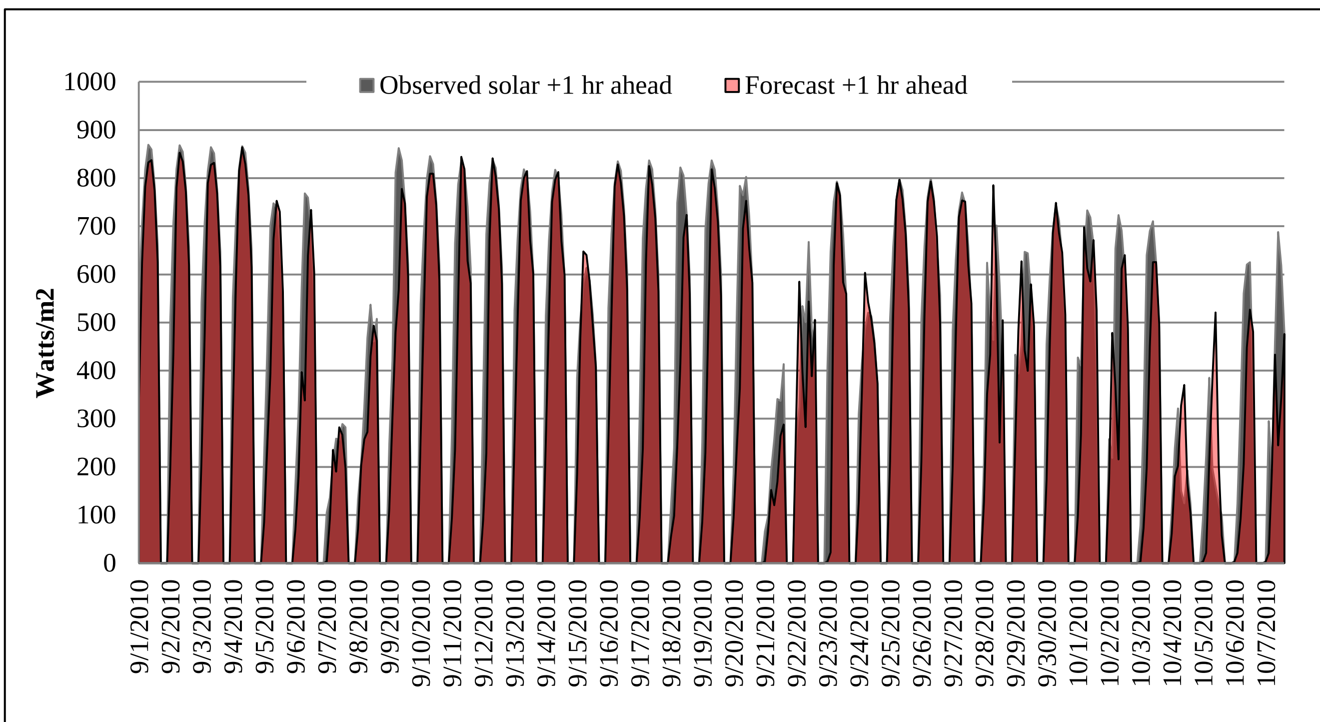

Fig. 6: Forecasts for solar radiation one hour ahead at Long Beach, CA.

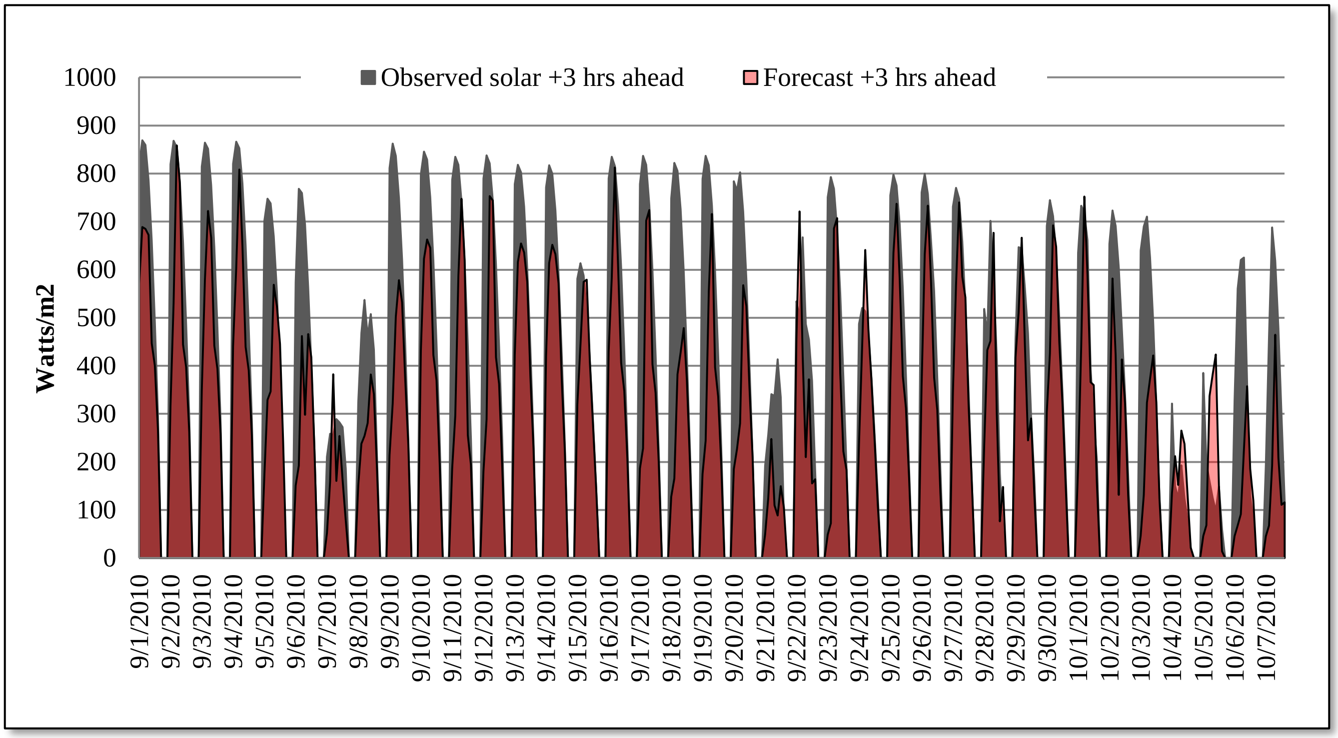

Fig. 7: Forecasts for solar radiation three hours ahead at Long Beach, CA.

7. REFERENCES

(1) G. Reikard, "Predicting Solar Radiation at High Resolutions: A Comparison of Time Series Forecasts". *Solar Energy*, vol. 83, issue 3, Mar. 2009, pp. 342-349

(2) James Hall & Jeffrey Hall, Quality Analysis of Global Horizontal Irradiance Data from 3500 US Ground-Based Weather Stations, Proceedings of the 2011 ASES Annual Conference, American Solar Energy Society, 2011.

(3) James Hall, Forecasting Solar Power with Adaptive Models – A Pilot Study, available online at: http://www.solardatawarehouse.com/forecasting%20solar%20power.pdf