Large Scale Ground Truth Database for Irradiance Forecast Validation

James M. Bing
NEO Virtus Engineering, Inc.
410 Great Road, B-6
Littleton, MA 01460
jbing@neovirtus.com

Obadiah Bartholomy
Sacramento Municipal Utility District
6201 S Street
Sacramento, CA 95817
OBartho@smud.org

1. INTRODUCTION

The Sacramento Municipal Utility District (SMUD), the country's 6th largest utility has teamed with NEO Virtus Engineering (NEO), a solar engineering, consulting and monitoring provider, to deploy a service-territory wide solar monitoring network for validating solar forecasting models. Thanks to a grant from the California Public Utilities Commission, under their California Solar Initiative RD&D program, SMUD has initiated a number of projects related to integrating high penetrations of solar PV.

NEO Virtus Engineering has developed a solar forecasting approach that makes use of the National Weather Service's National Digital Forecast Database (NDFD) and a cloud cover radiation model (Kasten-Czeplak and Gul-Muneer). The cloud cover radiation model or CRM, in conjunction with array geometry calculations, will provide the necessary conversion from percentage-of-sky-covered-by-clouds to incident irradiance on the module surface. This irradiance forecast, by extension, becomes a forecast of photovoltaic energy production within the spatial-temporal context of the utility service territory. To validate this model, NEO is developing and SMUD is deploying a network of 71 solar monitoring devices covering most of SMUD's 900 square mile service territory. The project was started in June, 2010, and will continue for 2 years, with at least 14 months of solar monitoring.

2. FOCUS ON SOLAR FORECASTING NEEDS

As PV penetrations increase, utilities like SMUD see an increasing need for tools to help schedule other resources around solar intermittency. Improved solar forecasting tools and validation of existing tool performance are both needed to enhance utilities' ability to manage such intermittency. At current solar installation rates, SMUD expects improved solar forecasting approaches will be necessary within the next 2-3 years. Recognizing this, SMUD and NEO decided to collaborate to provide a large scale validation of a specific forecasting approach which would include assessing appropriate performance metrics for judging forecasting success that were more relevant to time periods of interest.

Current approaches to solar forecasting have been benchmarked[1] and show that solar forecasting approaches available today have a typical RMSE on the order of 100 – 150 W/m2 for day ahead forecasts. However, given the reliable nature of the solar resource, aside from local cloud conditions, this level of error averaged over a long enough time may imply significantly greater error for short time periods of interest during inclement weather or light utility load vs. peak load performance. Further, understanding error on a % rather than W/m^2 basis would allow improved translation to other time periods of interest. To date, ground-based validation has not been done for large areas to confirm feeder level or service-territory level performance of solar forecasting approaches. This project aims to begin to improve the state of the art by creating a robust dataset against which to evaluate performance in conjunction with utility load information to evaluate performance during key periods of interest.

Beyond day-aead forecasts, there is also a need for forecasting short-term 1-3 hour ahead PV output, as well as likely variability around that output to be able to plan for appropriate regulating and load following resources. While knowing the absolute forecast is helpful for scheduling resources, understanding the variability around the absolute, in terms of ramp-rates and maximum and minimum values will be necessary for ensuring the right resources are scheduled. Part of this work will be developing a short-term 1-3 hour forecast that includes an assessment of variability based on real-time data from the sensor network. While this portion may not be broadly applicable without a significant proliferation of ground-based measurement or readily available or accurate satellite data, the work should inform theoretical methods[2] being developed to assess variability on top of forecasting approaches (Ref. Hoff, 2010)

3. DISCUSSION OF PROJECT

SMUD was awarded a grant from the CPUC in 2010 for a 2 year project to deploy hardware and software tools to model and mitigate impacts of high penetrations of PV on the distribution network. SMUD's grant partners and subcontractors on the project include Hawaiian Electric Company (HECO), BEW Engineering, Sunpower Corporation, and NEO Virtus Engineering. The full scope includes modeling and measuring high PV penetration circuits, developing utility interfaces to enance the understanding of the performance of intermittent resources, developing methods to utilize the smart meters to communicate with PV inverters, and finally this project, deploying a network of irradiance

sensors to monitor and validate solar forecasting approaches. Overall, these efforts will benefit the utilities involved as well as all California ratepayers by identifying solutions to integrating increasing amounts of PV onto the distribution grid.

As part of this research and to validate forecast accuracy, irradiance measurements will be made using a combination of approximately five Rotating Shadowband Radiometers (primary stations) and sixty six global horizontal measurements systems (secondary stations). This combination of primary and secondary monitoring stations will be deployed on the same five kilometer square grid as used by the NDFD for their skycover (cloud cover) forecasts. The monitored area will span almost 1775 square kilometers within SMUD's service territory. The sixty six secondary stations will measure global horizontal irradiance and ambient temperature. The primary stations will measure global horizontal, diffuse horizontal and direct normal irradiance, and ambient temperature (This pairing of primary and secondary stations mimics the format of the National Solar Resource Database [NSRDB]). All monitoring stations will record one minute averages. The monitoring stations will be located in the "nominal centroid" of each 25 square kilometer cell. The funded research will first establish the irradiance forecasts over the monitored region and then will quantify the error between measured and forecast irradiance over the term of the experiment.

A primary deliverable of this research will be the database of irradiance and temperature measurements. The monitoring regime calls for one-minute records of global horizontal and ambient temperature measured in locations roughly evenly distributed across the utility's service region. As noted previously, approximately five additional sites within the same grid will also contain diffuse horizontal and direct normal irradiance. The monitoring network will be deployed for approximately 14 months so that the database will cover a full 12 months with all 71 stations deployed. Once completed this database will be made available to researchers in the field of solar energy forecasting.

The installation of the network is being completed in the March-April 2011 timeframe, with the development of the forecasting models following closely behind.

4. FORECASTING APPROACH

4.1. Feed Forward Model: Overview

The approach to forecasting photovoltaic system electrical production used here is a version of a method developed and patented by NEO Virtus Engineering (under its previous company name, New Energy Options, Inc.) in 2003. This feed forward approach is a concatenation of models whose inputs are forecasts of cloud cover (sky cover in the terminology of the NDFD) and temperature and system specifications and site characteristics of the systems being modeled. A simplified rendering of the structure of the models used in the 20 to 36 hour in advance forecasting model can be summarized as follows:

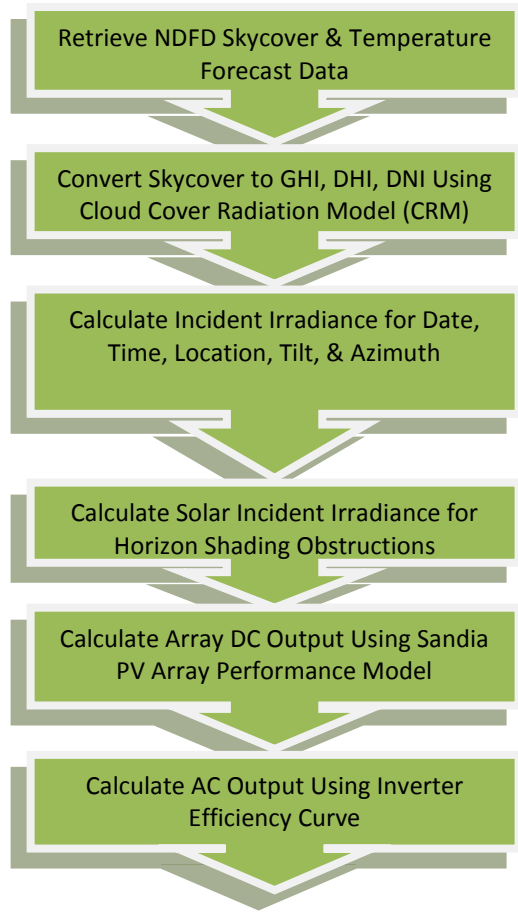

Figure1: Flow of Solar Forecasting Model

The four constituent models plus the inverter efficiency curve together form a transfer function the input of which is a prediction of skycover and the output of which is a prediction of electrical generation from a photovoltaic system.

4.2. National Weather Service National Digital Forecast Database (NDFD)

The National Digital Forecast Database or NDFD[3] is a National Weather Service product developed in the past decade which provides digital forecasts of a range of meteorological parameters in both numeric and graphical form. Grids for the continental United States are currently available from NDFD at 5 kilometer spatial resolution. The temporal resolution for skycover (cloud cover) and temperature is every three hours out to 72 hours and every six hours out to 168 hours.

The forecast data may be viewed with a web browser or retrieved via file transfer protocol in binary form. Figure 2 is an example of display of forecast skycover for an area of northern California which includes Sacramento County.

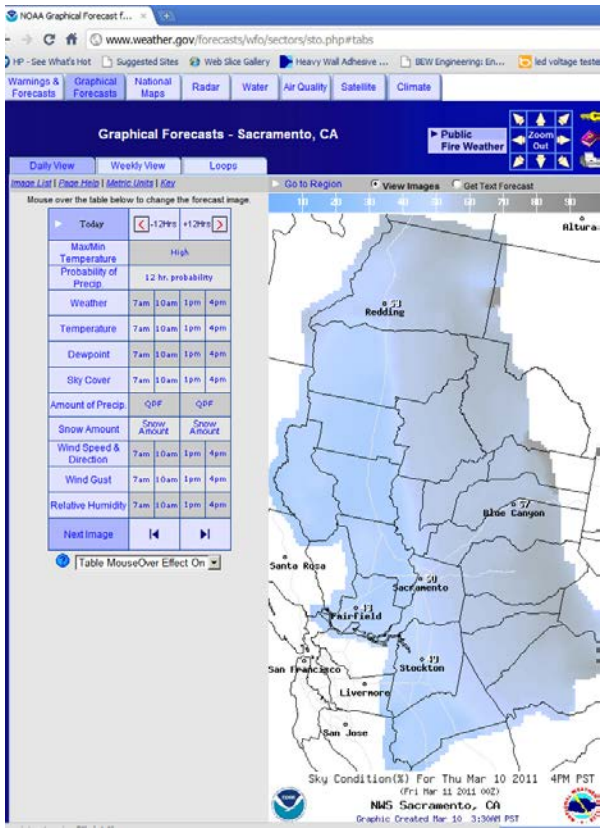

Figure 2: Skycover Variable Depiction from National Weather Service NDFD

As part of this project one of the functions which the MySQL database engine will perform is to poll the NDFD and retrieve and store forecast data for skycover and temperature for the Sacramento County region of northern California. During the validation process these forecasts will be converted to irradiance and PV electrical production and will be compared with measured data.

4.3. Cloud Cover Radiation Model (CRM)

The Kasten-Czeplak[4] and Gul-Muneer[5] Cloud cover Radiation Model (CRM)[6] provides a means of converting cloud cover amount or percentage into global horizontal, diffuse horizontal and direct normal irradiance values. The model can be summarized as:

$$I_{Gc} = \left(A \sin \mathrm{SOLALT} - B\right)$$

$$I_G / I_{Gc} = 1 - C\left(N/8\right)^D$$

$$I_{Diffuse} / I_{Gc} = 0.3 + 0.7\left(N/8\right)^2$$

Where:

SOLART is the solar elevation angle;

I_{Gc} is solar radiation under cloudless sky;

I_G is global horizontal solar radiation;

$I_{Diffuse}$ is diffuse horizontal solar radiation;

N is cloud cover amount, expressed in octa;

and ***A, B, C*** and ***D*** are geographic region specific empirically derived coefficients.

The direct normal component is obtained by taking the difference between global and diffuse irradiation.

4.4. Solar Geometry and Shading Model

There are two computations which will be performed for each array surface for which incident irradiance will be calculated. The first is a calculation based upon the tilt, azimuth orientation, geographic location (latitude and longitude), and date and time. The second is a calculation of the horizon and near field shading upon the array surface. These calculations will be done in MySQL using stored PV system parameters.

4.5. Photovoltaic Array Performance Model[7]

The Sandia Photovoltaic Array Performance Model takes the parameters of diffuse horizontal and direct normal irradiance which are incident upon a PV module's surface, and calculates a value known as “effective irradiance.” This value, in combination with temperature and a range of empirically derived module parameters is used in the model to simulate dc electrical production. This dc electrical energy production becomes the input for the last stage in the overall model, the system inverter efficiency curve.

4.6. Inverter Model

In the final stage in the model the dc electrical output of each array is applied to the efficiency curve of a utility interactive inverter. The result is a time-stamped prediction of ac power production for a specific point within the electrical distribution system.

5. NETWORK DEPLOYMENT

5.1. Device Design/components

Rotating Shadowband Radiometers (RSR) are solar monitoring devices capable of deriving diffuse

horizontal and direct normal irradiance from measured global horizontal irradiance. Five RSRs are being used for the primary monitoring stations. They will be located on roofs and in substations and at other locations within SMUD service territory. In most cases they will be mounted on tripods similar to weather monitoring stations.

Figure 3: Primary Station Rotating Shadowband Radiometer (RSR) and temperature probe

In order to locate the secondary stations in the centroids of the NDFD grids most of these units will be located on utility poles. These units have been designed specifically to meet this requirement. The durable, non-metallic, weather tight enclosure houses a data logger, PV power system, cellular modem, and an internal antenna. The system is designed for seven days of autonomy based upon a worst case winter irradiance scenario for the Sacramento area.

Figure4: Deployment of 1st Prototype Secondary Station

Each secondary station has a Licor 200SZ pyranometer and a thermister temperature probe. Data will be measured every second and averaged to one minute records which will be collected daily.

5.2. Database Configuration

NEO will create a database using the open source software, MySQL. The database will

- ftp retrieve NDFD forecast skycover & temp,
- store skycover & temp data,
- retrieve (daily) irradiance monitoring network of 5 primary and 66 secondary monitoring stations,
- store configuration of PV systems monitored by SMUD (array size, location tilt & azimuth angles, horizon shading, inverter type)
- compute global horizontal, diffuse horizontal and direct normal irradiance for each skycover forecast for each grid cell,
- simulate electrical production for each of the PV systems for which SMUD has production monitoring.
- Calculate short-term irradiance forecast (1-3 hours ahead) based upon measured GHI from primary and secondary monitoring stations.

After the full twelve month of irradiance measurements have been made the database will

- quantify the performance of the NEO model for calculating global horizontal irradiance from forecast skycover (day ahead, 20-36 hours in advance).
- calculate the error, for the one minute intervals over the full twelve months, in the forecast vs. measured production for PV systems for which SMUD has monitoring data.
- calculate the error in the 1-3 hour ahead forecast.

5.3. Communications

A cellular data network was selected for both primary and secondary systems for data retrieval. Spread spectrum solutions were evaluated but the line-of-site requirement could not be guaranteed and so they were eliminated from consideration. The service map for data plans for the selected cellular carrier, AT&T, indicates strong coverage across all of the territory being considered for deployment. The secondary monitoring stations will be installed at a height of approximately 40 feet above the ground. We believe that the installation height in combination with a 5dB gain antenna selected for the modems will provide good signal strength. The fall-back positions, should any of the grid cells have poor coverage, are to install a 10dB gain external antenna and, should that fail, use an external Yagi directional antenna. No formal county-wide survey of signal strength was conducted, however initial tests on SMUDs headquarters roof and a prototype monitoring station installed on a pole in one of the more remote rural grid cell indicate signal strength to be very good.

5.4. Site Selection

In order to validate an NDFD-based forecasting approach, ground based sites coinciding with the NDFD grid-cell centroids were desired. Considering security, accessibility, obstructions, and coordination time, SMUD's distribution pole network presented a unique opportunity to obtain the necessary coverage within reasonable distance from the grid centroids. Based on an overlay of SMUD's 120,000+ distribution poles with the NDFD grid centroids, suitable poles were identified within 1km (and in many cases within 100 meters) of the centroid for nearly every grid-cell in SMUD's 900 square mile service territory. For the area of interest, only 3 grid cells had no poles within 1km of the centroid, and only 3 were between 500m and 1km, the remaining 65 locations contained poles within 500m of a grid cell centroid.

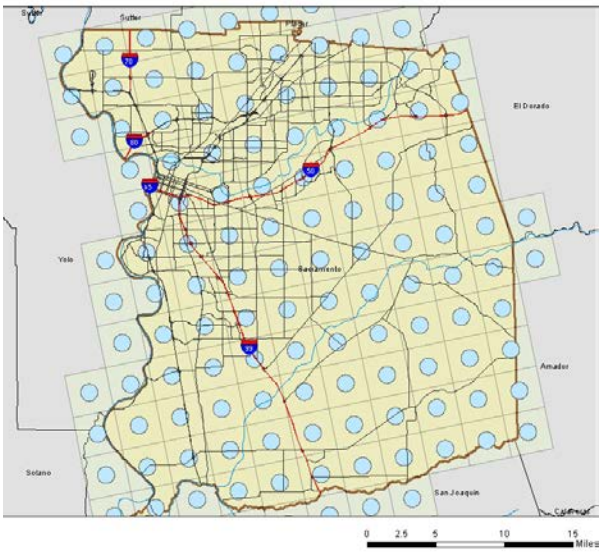

Figure5: Map of SMUD Service Territory, NDFD Grid cells and Centroids with 1km buffer

Specific pole selection was done by visually inspecting candidate poles near a centroid to determine whether there were any shading obstructions, and whether there was adequate climbing space on another quadrant of the pole for SMUD linemen.

In addition to siting secondary stations on distribution poles, siting of primary RSR based stations was done by identifying SMUD substation assets that coincided with grid centroids and that had no, or at least very limited, visual obstructions.

5.5. Installation

Installation of the devices is being done over a one month period by SMUD linemen and troubleshooters depending on availability and coincidence with other work in a given region. The devices arrive ready to be bolted to the poles, and require only a simple fuse activation to connect the battery. Communications are checked once the fuse is activated, and the device is mapped once it is installed on a given pole. Devices will remain untouched for the 14 month period, except for replacement if necessary due to any component failures. Evaluation of the effects of soiling will be done on primary stations due to better accessibility.

6. NEXT STEPS AND FOLLOW-ON WORK

Installation of the devices should be complete by the end of April, however data collection will begin with the initial deployment at the end of March. Database construction and forecasting software development have commenced, and validation using initial data should begin in late summer to early fall 2011. Model refinement and longer term validation will then follow.

Prior to completion of the validation work, some additional work relating to evaluation metrics will need to be done in order to ensure that the validation results are best use for utilities. This will involve development of new metrics, as well as clarification of terminologies and time windows of interest. To the extent that certain timeframes and weather patterns drive specific metrics that are unique to Sacramento, recommendations for multi-region validations will be developed. The following questions will guide that work:

- How does definition of "day ahead" best bridge the solar profile and the utility need?
- What is the "standard" time window during which the forward estimate is being considered? Or are there multiple time windows of interest?
- For the purpose of comparison of methods, at what time during the day should such a prediction occur?
- Should this window be defined based upon an ISO's Day Ahead market timing or upon day of scheduling demands?
- Are different validations for different regions required? How does worst case performance vary by region?
- What are specific timeframes of interest? Peak day availability, mid-spring and mid-fall variability?

7. CONCLUSIONS

The work described herein is intended to provide a robust ground-based validation of solar forecast approaches over a utility service territory. Such a validation has not yet been done to the authors' knowledge. The collaboration with and use of utility assets provides a unique opportunity for extended period ground-based validation covering a large area with a geometric distribution that matches the NDFD database. Results of the validation and public availability of the solar datasets will occur in 2012.

8. ACKNOWLEDGEMENTS

SMUD and NEO Virtus Engineering would like to acknowledge the grant funding support from the California Public Utilities Commission California Solar Initiative RD&D program and the ratepayers of California who funded it. We would also like to acknowledge the administration of that program by Itron.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the CPUC, Itron, Inc. or the CSI RD&D Program.

---

[1] R. Perez, S. Kivalov, D. Renne, J. Sclemmer, K. Hemker Jr., and T. Hoff, (2009) Validation of Short and Medium Term Operational Solar Radiation Forecasts in the US, American Solar Energy Society Annual Conference, Buffalo, NY

[2] T. Hoff, R. Perez, (2010) Quantifying PV Power Output Variability, Conditionally Accepted into Solar Energy

[3] The NDFD can be accessed at its home page at http://www.weather.gov/ndfd/

[4] Solar and terrestrial radiation dependent on the amount and type of cloud. F. Kasten and G. Czeplak, 1980, *Solar Energy* 24, 117-89.

[5] Models for obtaining solar radiation from other meteorological data. M. Gul and T. Muneer, 1998, *Solar Energy* 64, 99-108.

[6] Solar Radiation and Daylight Models; T. Muneer; 2004

[7] Photovoltaic Array Performance Model, D. King, W. Boyson, J. Kratochvil
Photovoltaic System R&D Department
Sandia National Laboratories, August 2004