PREPARED FOR
SUNPOWER

HIGH-RESOLUTION SOLAR RESOURCE DATA METHODOLOGY REPORT

APRIL 27, 2011

PREPARED FOR:
SunPower Corporation
1414 Harbour Way, South
Richmond, CA 94804

Attention:
Joe Surprenant
Joe.Surprenant@sunpowercorp.com
408-240-4215

SUBMITTED BY:
Marie Schnitzer, Director of Solar Services
MSchnitzer@awstruepower.com
Ph: 518-213-0044 ext. 1093
Fax: 518-213-0045

TABLE OF CONTENTS

List of Figures 3
List of Tables 3
1. Introduction 4
2. Obsevational Network Review 4
3. Mesoscale Modeling 6
4. Summary and Future Work 7
Figures 9
Tables 13

LIST OF FIGURES

Figure 1. MASS grid configuration. Project delivery regions are shown by red squares. SunPower 15-minute (green/black pins) and 1-second (blue triangles) monitoring stations..... 10

Figure 2. (a) Annual GHI for SunPower station 1 (blue) and SUNY/NSRDB (red); (b) Same as in (a), except for monthly averaged GHI; (c) Same as in (a), except for diurnal (hourly-averaged) GHI; (d) Scatter plot of SUNY/NSRDB versus SunPower observation station 1. Green line represents best fit linear regression. 11

Figure 3. Southern nested domain global horizontal irradiance difference between the 10.5-minute and 5-minute Surface Energy Budget-Radiation-Planetary Boundary Layer-Hydrology (SRPH) update frequencies (W/m^2). X-axis shows longitude, Y-axis corresponds to latitude. 12

LIST OF TABLES

Table 1. Table of observation stations delivered by SunPower. 14

Table 2. Model configuration for MASS runs. 15

1. INTRODUCTION

AWS Truepower (AWST) was engaged by SunPower to develop a high-resolution solar database as a part of the Planning & Modeling for High-Penetration Photovoltaic project funded through the California Public Utility Commission's California Solar Initiative RD&D Program. AWST was assigned to generate numerically predicted resource data at 15-minute (low frequency), 1-minute and 1-second (high frequency) temporal resolutions. The numerical data set aims to improve upon the freely available National Solar Resource Database (NSRDB) by increasing the temporal and spatial resolution while reducing the overall bias and root mean square error between actual observations. As a part of the project plan, the modeled data set are to be compared with the NSRDB and in-situ observations provided by SunPower using appropriate criteria to quantify its accuracy. The following tasks are to be completed in accordance with the proposed scope of work:

- Data acquisition and QC of SunPower/NSRDB data
- Model setup and run
- Validate 15-minute, 10-year time series data
- Derive and validate 1-second data using higher-resolution observed measurements
- Derive and validate 1-Minute data
- Deliver data to team
- Delver final technical report to SunPower

To date, the acquired SunPower and NSRDB data sets have been reviewed, and the 10-year modeled simulations have been completed over the target areas. AWST has also developed metrics to gauge the performance of the modeled data set. This report will describe the approach used to review the observation networks, the mesoscale model configuration and the model performance criterion.

2. OBSEVATIONAL NETWORK REVIEW

In February 2011 SunPower delivered data from a network of 50 observational stations to AWST. The network of stations included 15-minute irradiance measurements from 25 stations in the Los Angeles area and 25 stations in the San Francisco area. SunPower will also provide 1-second data from 3 stations in each region. Both sets of observations will be used to validate the numerical data generated by AWST. The 1-second irradiance measurements will also be used to derive the high frequency numerical resource data. The locations of the stations, along with outlined project areas, are shown on the map in Figure 1.

Each of the low-resolution data files has a header section containing: the met station identifier, latitudinal and longitudinal coordinates, the observational period, time interval, and the units of measure. Global horizontal irradiance (GHI) (W/m^2) is contained in all the files, with plane of array (POA) (W/m^2), ambient temperature (°C), wind speed (m/s), wind direction (degrees), and power (KW) recorded in some of the files. The data records are saved as the average of each preceding 15-minute period, i.e. the timestamp 1/8/2001 0:15 would be the average of the measured values from 1/8/2001 0:00 – 1/8/2001 0:15. The periods of record (POR) vary greatly within the observation network. The longest continuous POR delivered is from January 8, 2001 – February 9 2011 and the shortest is from March 13, 2009 – February 9, 2011. A complete table summarizing the delivered observation network is shown in Table 1.

After reviewing the low resolution data sets, it was found that the observations will be sufficient for validating the model. The observational network contains a large number of stations that have relative continuity and adequate PORs. The average percentage of missing records was found to be ≈1.5%. The following checks and corrections were applied to ensure the data were of good quality before comparing them with the modeled data:

- Verify that the observation data set POR has sufficient overlap with the modeled POR
- Verify correct timestamp
- Set all irradiance values to zero from 21:00 – 4:00 Local Standard Time
- Exclude unrealistic or missing values from comparison

After applying these correction procedures, an initial comparison with the hourly NSRDB SUNY 10 km gridded data set[1] (referred to as SUNY/NSRDB) was completed. It was necessary to remove intra-hour records from the modeled data for this purpose. The SUNY/NSRDB data set contains modeled irradiance values derived from satellite observations using methods developed by Perez et al.[2] Its period of record across all grid points is 1997-2005. Given that the observations provided by SunPower are mostly from recent years, a relatively low number of stations contain concurrent records. Given these constraints and the need for PORs of at least one year, only eight of the observation stations were used for the preliminary comparison.

Given the vast number of modeled data points (10 km spacing across the state), all observational stations were found to be within 5 km of the nearest SUNY/NSRDB grid point. For this reason, the modeled data was used directly, with no interpolation or distance weighting taken into account. Several statistical parameters were computed for each SunPower observation to determine the accuracy of the modeled data set. The mean bias was computed on an annual, monthly, and diurnal basis. These were each computed by subtracting the observations from the nearest appropriate NSRDB station. The root-mean-square error (RMSE) and the coefficient of determination (r^2) were computed using the hourly data from each corresponding station.

It was found that the hourly SUNY/NSRDB predicts the measured global irradiance moderately well over the long term, but tends to misrepresent the timing and occurrence of maximum daily irradiance values. The SUNY/NSRDB data tend to be systematically higher compared to all SunPower observation stations. The average annual bias for all stations is 4.8% of the annual mean GHI with an RMSE of 5.4%. The hourly bias for each individual station ranges considerably by site, from 3.8-15.3% with a RMSE range of 5.3-24.4% of the mean GHI.

The annual, monthly, and diurnal means for the observations and the SUNY/NSRDB were plotted and compared. A representative record containing the longest POR is shown in Figure 2. The coefficient of determination (r^2) for the overlapping five-year period is 0.84, a relatively good correlation between the two data sets. Although this station's r^2 value is relatively high, there are a number of outliers that reduce confidence in the SUNY/NSRDB data set. Through this review, it was determined that 10 km and one-hour spatial and temporal resolution do not adequately account for the region's irradiance variability.

---

[1] George R., *et al.,* National Solar Radiation Database (NSRDB) (1997) - 10 km gridded hourly solar database. *Solar 2007 Conf.*, Cleveland, OH, ASES.

[2] Perez, R.; Ineichen, P.; Moore, K.; Kmiecik, M.; Chain, C.; George, R.; Vignola, F. (2002). "A New Operational Satellite-to-Irradiance Model." Solar Energy 73(5):307–317.

To date, AWST is still waiting to receive high-resolution measurements needed to complete the high-frequency portion of the project. Once these are received, a similar review process will be completed and the one-minute/one-second data sets will be generated. Both the high- and low-frequency observations will be further analyzed against the mesoscale model output once available. The tendency for the solar resource to change with time, i.e. ramp rates, will be compared. This metric will determine how well the model captures large and/or sudden changes in the solar resource. Lastly, the power spectral density (PSD), which describes how the power of a signal or time series is distributed with frequency, will be computed for the modeled and observed data.

3. MESOSCALE MODELING

The numerical weather modeling was completed in February/March of 2011 using the Mesoscale Atmospheric Simulation System (MASS), a proprietary model developed by AWST partner MESO, Inc.[3] The simulations covered the period from January 1, 1999, to January 1, 2009, over the state of California. The project areas (San Francisco and Los Angeles) are shown in Figure 1.

MASS is a numerical weather prediction (NWP) model developed over the past 20 years as both a research tool and a commercial weather forecasting model. MASS simulates the fundamental physics of the atmosphere including conservation of mass, momentum, and energy, as well as the moisture phases. It contains a turbulent kinetic energy module that accounts for the effects of viscosity and thermal stability on wind shear as well as a surface energy budget parameterization that takes into account incoming, absorbed, reflected, and radiated solar energy. A dynamic model, MASS simulates the evolution of atmospheric conditions in time steps as short as a few seconds.

MASS uses a variety of geophysical and meteorological databases. The main meteorological inputs are reanalysis data, rawinsonde data, and land surface measurements. The most influential initial condition, the reanalysis database, is a gridded historical data set produced by the U.S. National Centers for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR)[4]. The data provide a snapshot of atmospheric conditions around the world at all levels of the atmosphere in intervals of six hours. Along with rawinsonde and surface data, the reanalysis data establish the initial conditions as well as lateral boundary conditions for the MASS runs. The MASS model itself determines the evolution of atmospheric conditions within the region based on the interactions among different elements in the atmosphere and between the atmosphere and the surface. The reanalysis data are on a relatively coarse grid (about 210 km spacing). To avoid generating noise at the boundaries that can result from large jumps in grid cell size, MASS is run in several nested grids of successively finer mesh size, each taking as input the output of the previous nest, until the desired grid scale is reached. In this simulation the outer grid extends 3000km in each direction.

The main geophysical inputs are elevation, land cover, vegetation greenness (normalized differential vegetation index, or NDVI), soil moisture, and sea-surface temperatures. The elevation data used by MASS are from the Shuttle Radar Topographical Mission 30 Arc-Second Data Set (SRTM30), which was

[3] Manobianco, J., J. W. Zack and G.E. Taylor (1996). "Workstation-based real-time mesoscale modeling designed for weather support to operations at the Kennedy Space Center and Cape Canaveral Air Station". Bull. Amer. Meteor. Soc., vol. 77, pp. 653-672.

[4] Kalnay, E., et al., (1996). "The NCEP/NCAR 40-Year Reanalysis Project". Bulletin of the American Meteorological Society. Vol. 77, pp. 437–471.

produced in an international project spearheaded by the National Geospatial-Intelligence Agency (NGA) and the National Aeronautics and Space Administration (NASA).[5] The NDVI data were derived from a predecessor of MODIS, the satellite-based Advanced Very High Resolution Radiometer (AVHRR).[6] The nominal spatial resolution of all of these data sets is 1 km.

The MASS model has a number of different options that can affect the simulations. Table 2 summarizes the high-level options chosen for this project. The runs employed a nested grid scheme with horizontal resolutions of 30 km, 8 km and 2 km. The grid configuration is shown in Figure 1. The MASS output was stored every 15 minutes. The output data files contained global, direct, and diffuse irradiance, wind speed, wind direction, and near-surface (2-m) temperature and humidity.

A preliminary study was completed to determine the optimal run configuration for the state of California. The MASS model was run for two years with a series of different options that affect the model output. One such option is the model dependency on the Surface Energy Budget-Radiation-Planetary Boundary Layer-Hydrology (SRPH) update frequency. All prior MASS simulations performed at AWST had applied an update frequency of 10.5-minutes. Considering the model outputs instantaneous data every 15-minutes, it was judged useful to know if calculating the SRPH indices more frequently would have an effect on the model results. To test this option, two simulations were completed; the first with an SRPH update frequency of 10.5 minutes, the second with an update frequency of 5 minutes. The differences were found to be small. The maximum bias between the two datasets was -3.5 W/m^2, representing 1.75% to 2.7% of the mean global irradiance for the 2-km grid, which varied from 130 W/m^2 to 200 W/m^2 from the north to south (Figure 3). Although the bias is small, the frequency update will also increase the model's ability to simulate solar ramp rates. The computational time for a simulation increases as the SRPH update frequency increases. The computational demands versus increased accuracy were taken into account during this comparison. It was found that the five-minute update frequency strikes a suitable balance, and it was therefore used for the ten-year simulation.

4. SUMMARY AND FUTURE WORK

A review of the observational data provided by SunPower has been completed. The preliminary comparison to the SUNY/SNRDB data shows that the network will provide an adequate baseline for validation. Quality-assurance checks and corrections have been applied to the observational data. The variables most pertinent to solar resource studies (i.e. direct, diffuse, and global irradiance, temperature, wind speed and direction, surface pressure, and humidity) from the MASS model have been archived and will be compared against the SunPower observational network using the outlined procedures.

In the coming months AWST will validate the quality of the MASS modeled data set using the metrics developed during the observational review. The mean annual, monthly, and diurnal cycle, ramp rates, and the power spectral density will be compared between the two data sets. This criterion will give a complete understanding of how accurate the MASS solar irradiance estimates are and if any modifications need to be made to the data set. Once the validation is complete, AWST will begin deriving the higher-frequency data. As defined in the project timeline, SunPower will deliver high-frequency observations that will be used to derive the final modeled data set for the region of interest.

[5]For more information, see http://www2.jpl.nasa.gov/srtm/.

[6]See http://edcwww.cr.usgs.gov/products/landcover/glcc.html.

A software application developed by AWST will be used to produce the one-minute and one-second output from the lower-frequency data. AWST will deliver the 15-minute, 10-year time series for all grid points in the study regions, as well as the application to extract the high-frequency data.

FIGURES

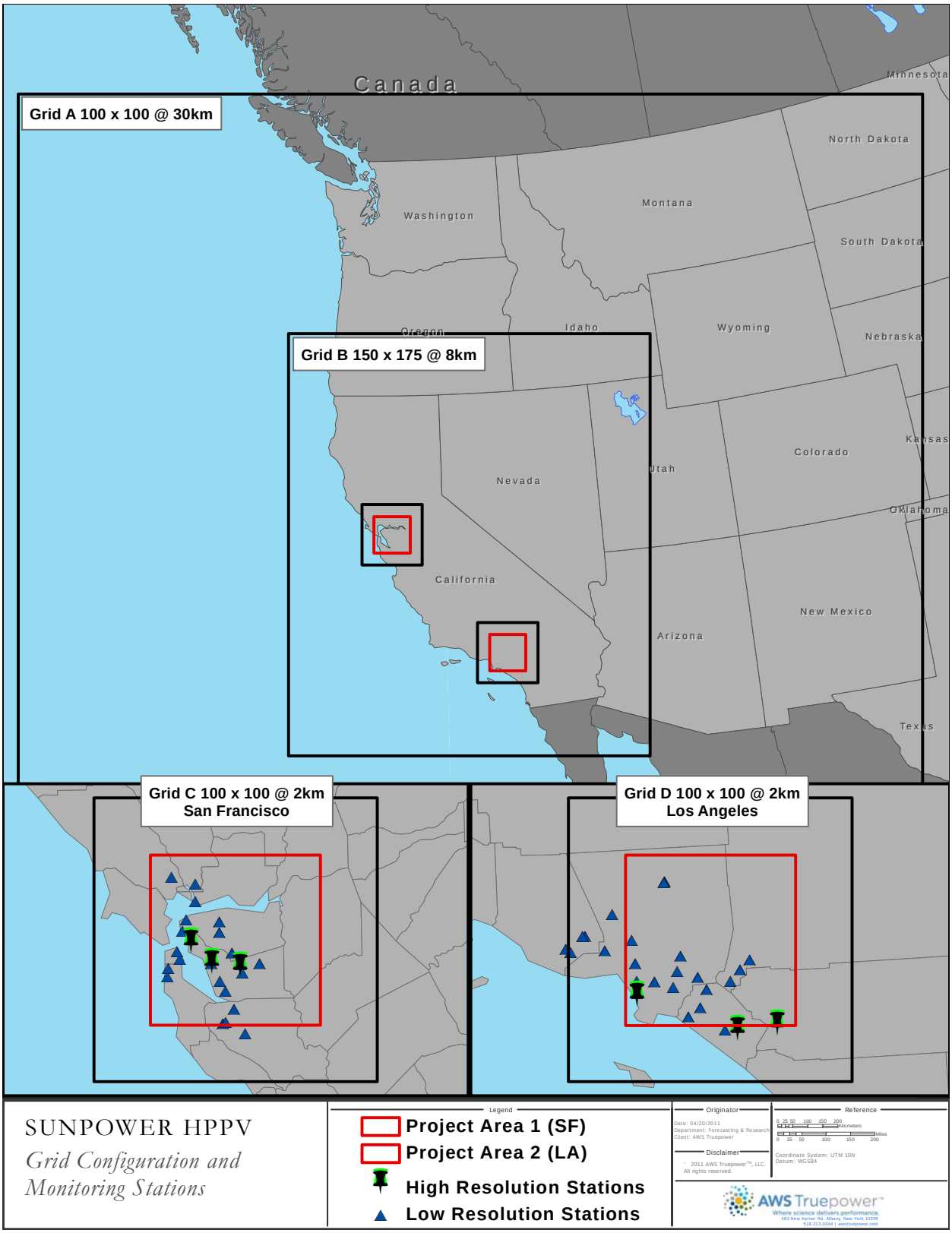

Figure 1. MASS grid configuration. Project delivery regions are shown by red squares. SunPower 15-minute (green/black pins) and 1-second (blue triangles) monitoring stations.

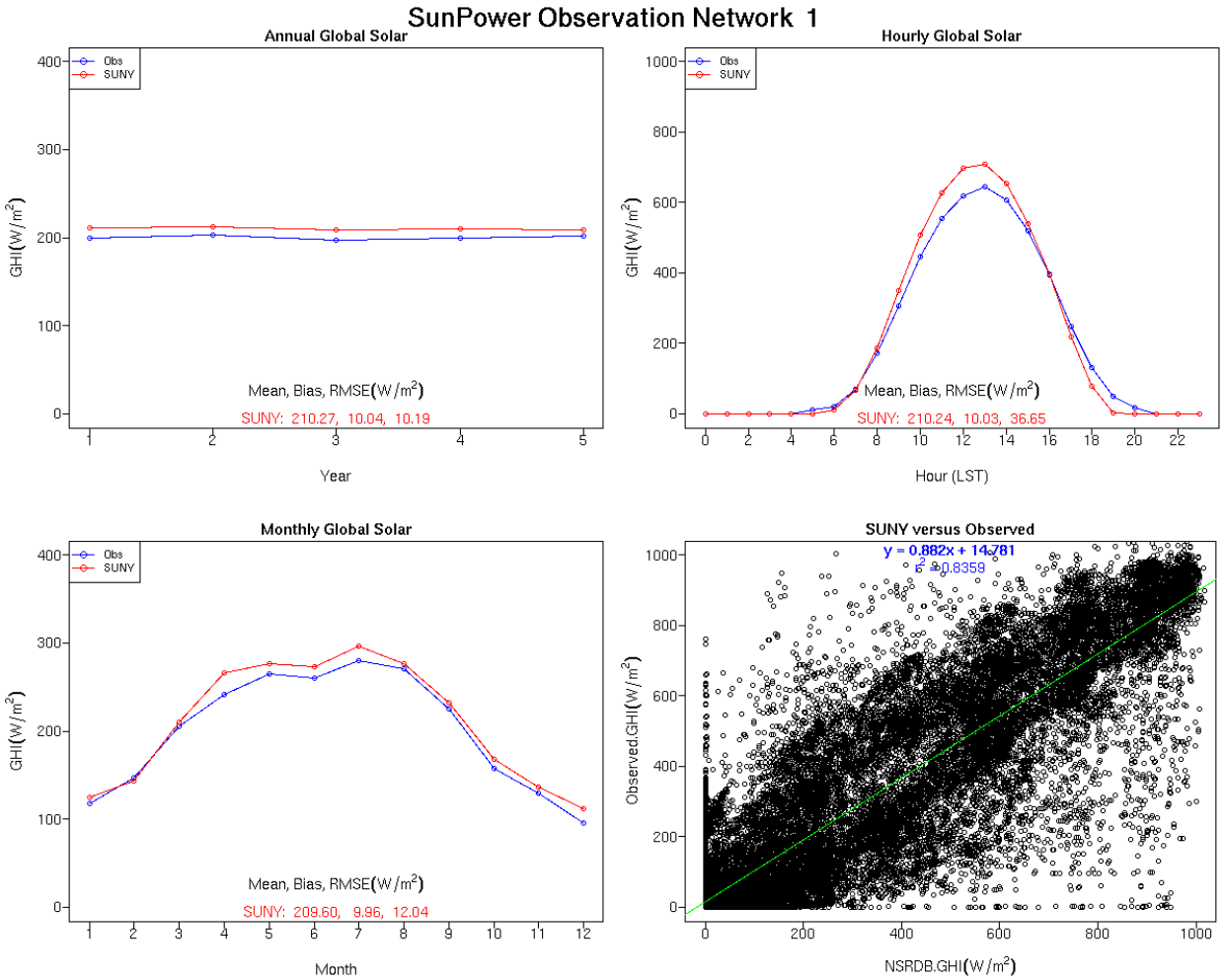

Figure 2. (a) Annual GHI for SunPower station 1 (blue) and SUNY/NSRDB (red); (b) Same as in (a), except for monthly averaged GHI; (c) Same as in (a), except for diurnal (hourly-averaged) GHI; (d) Scatter plot of SUNY/NSRDB versus SunPower observation station 1. Green line represents best fit linear regression.

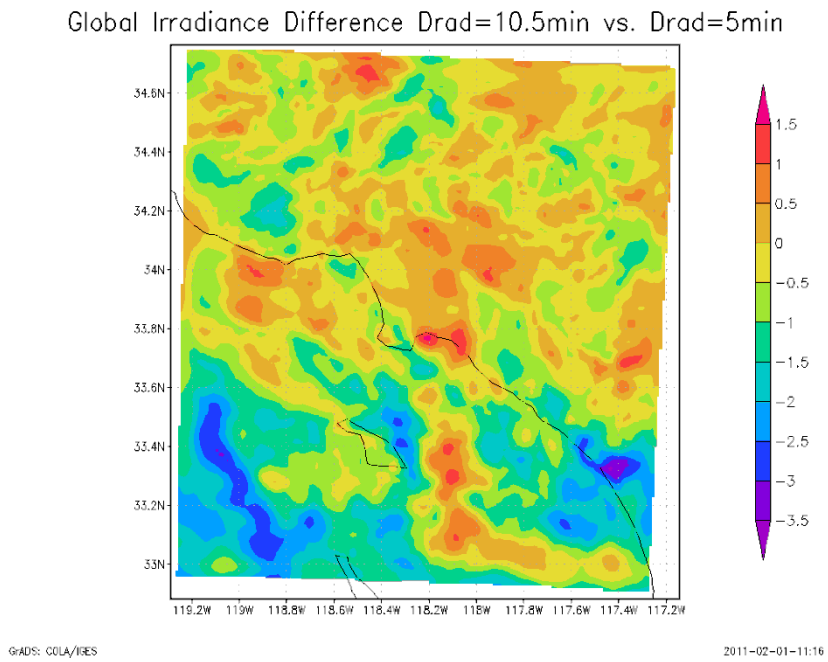

Figure 3. Southern nested domain global horizontal irradiance difference between the 10.5-minute and 5-minute Surface Energy Budget-Radiation-Planetary Boundary Layer-Hydrology (SRPH) update frequencies (W/m^2). X-axis shows longitude, Y-axis corresponds to latitude.

TABLES

Table 1. Table of observation stations delivered by SunPower.

ID	Latitude	Longitude	Start Date	End Date	POR (Months)	Project Area
1	33.800	-117.922	08-Jan-01	09-Feb-11	121	LA
7	34.100	-118.400	06-Jun-03	09-Feb-11	92	LA
17	37.411	-121.951	22-Aug-02	07-Feb-11	102	SF
65	37.929	-122.057	30-Jan-03	08-Feb-11	97	SF
71	37.552	-122.333	27-Jan-03	08-Feb-11	97	SF
86	33.978	-118.258	08-Dec-01	10-Feb-11	110	LA
96	37.780	-122.400	18-Oct-03	08-Feb-11	88	SF
99	34.190	-118.626	12-Nov-03	10-Feb-11	87	LA
106	34.245	-118.419	10-Feb-04	10-Feb-11	84	LA
111	37.640	-121.880	08-Mar-04	08-Feb-11	83	SF
141	37.590	-122.060	09-Apr-05	07-Feb-11	70	SF
146	37.700	-121.740	18-Mar-05	08-Feb-11	71	SF
147	37.460	-121.930	08-Jun-05	08-Feb-11	68	SF
174	37.730	-122.380	26-Jan-06	08-Feb-11	61	SF
176	37.320	-122.040	07-Mar-06	08-Feb-11	59	SF
186	37.620	-122.480	01-Jun-06	08-Feb-11	56	SF
201	38.210	-122.250	20-Sep-06	08-Feb-11	53	SF
214	33.908	-118.395	04-Dec-06	09-Feb-11	50	LA
223	33.664	-117.644	23-Jan-07	09-Feb-11	49	LA
230	34.210	-118.925	10-Apr-07	10-Feb-11	46	LA
232	38.253	-122.442	25-Apr-07	07-Feb-11	46	SF
245	37.900	-122.060	10-Jul-07	08-Feb-11	43	SF
248	37.694	-121.897	27-Aug-07	07-Feb-11	42	SF
258	33.747	-118.014	20-Nov-07	09-Feb-11	39	LA
261	37.703	-122.128	31-Oct-07	08-Feb-11	40	SF
263	37.525	-122.017	25-Jan-08	08-Feb-11	37	SF
270	33.993	-117.928	23-Feb-08	10-Feb-11	36	LA
272	34.284	-118.773	16-Jan-08	09-Feb-11	37	LA
282	34.186	-118.890	12-Feb-08	09-Feb-11	36	LA
291	34.133	-118.052	24-Apr-08	09-Feb-11	34	LA
292	34.036	-118.081	23-May-08	09-Feb-11	33	LA
293	37.981	-122.327	24-Apr-08	07-Feb-11	34	SF
297	37.254	-121.863	20-May-08	08-Feb-11	33	SF
298	34.282	-118.792	27-May-08	10-Feb-11	33	LA
304	33.914	-117.865	31-Jul-08	09-Feb-11	31	LA
305	37.327	-122.016	24-Oct-08	07-Feb-11	28	SF
307	37.967	-122.060	18-Sep-08	08-Feb-11	29	SF
317	33.935	-118.119	09-Oct-08	09-Feb-11	28	LA
318	33.986	-118.394	04-Nov-08	09-Feb-11	27	LA
328	37.674	-122.471	11-Oct-08	07-Feb-11	28	SF
338	33.649	-117.742	08-Nov-08	09-Feb-11	27	LA
340	38.097	-122.251	18-Nov-08	08-Feb-11	27	SF
341	33.956	-117.683	24-Nov-08	10-Feb-11	27	LA
343	34.027	-117.603	11-Jan-09	10-Feb-11	25	LA
354	34.612	-118.145	11-Jan-09	09-Feb-11	25	LA
361	37.768	-121.961	22-Jan-09	08-Feb-11	25	SF
363	34.416	-118.557	01-Jan-09	09-Feb-11	25	LA
367	34.604	-118.151	01-Jan-09	10-Feb-11	25	LA
375	34.086	-117.526	13-Mar-09	09-Feb-11	23	LA
377	37.910	-122.358	16-Feb-09	07-Feb-11	24	SF

Table 2. Model configuration for MASS runs.

Model	MASS v. 6.8
Initialization data source	NNGR
Data to be assimilated in the course of simulations (30 km, 8 km, and 2km grids)	Rawinsonde, surface observations (temperature, dew point, wind direction and speed, pressure)
Sea-surface temperatures	MODIS (Moderate Resolution Imaging Spectroradiometer)
High-resolution terrain and land cover (2 km grid only)	US Geological Survey National Elevation Dataset and Land-Use Land-Cover Dataset, both 30 m grid spacing
Cumulus scheme (30 km and 8 km grids only)	Kain-Fritsch
Spin-up	12 hours before start of valid run
Length of run	15-16 day series (e.g., Jan 1-15, Jan 16-31)
Frequency of data sampling	Every 15 min.
Data to be stored	Global horizontal irradiance, Direct irradiance, Diffuse irradiance, 10m earth relative wind speed, 10m wind direction, surface pressure, 2m temperature, 2m specific humidity