Preprocessing WRF initial conditions for coastal stratocumulus forecasting

https://escholarship.org/uc/item/7xt0358g

SOLAR ENERGY, 133

0038-092X

Yang, Handa
Kleissl, Jan

2016-08-01

10.1016/j.solener.2016.04.003

Peer reviewed
Preprocessing WRF initial conditions for coastal stratocumulus forecasting

Handa Yang\textsuperscript{a,}*\textsuperscript{,} Jan Kleissl\textsuperscript{a}

\textsuperscript{a}Center for Renewable Resources and Integration, Department of Mechanical and Aerospace Engineering, University of California, San Diego, 9500 Gilman Dr., EBU II - 304, La Jolla CA, 92093-0411, United States

Abstract

The impact of atmospheric liquid water content at model initialization in the Weather Research and Forecasting (WRF) model is explored through the application of two preprocessing schemes. The first scheme, the Well-mixed Preprocessor (WEMPP), was designed and developed based on a well-mixed boundary layer to provide an initial guess at liquid water content when initializing with data from the North American Mesoscale (NAM) model, as liquid water content is not present in NAM output. The second scheme was adapted from a satellite Cloud Data Assimilation (CLDDA) package intended to make the initial model cloud field consistent with observations, using input data from the CIMSS GOES sounder cloud product.

Preprocessed simulations were compared against baseline WRF simulations initialized with NAM and the Rapid Refresh (RAP) model (which contains liquid water output), as well as the raw parent model outputs. These intra-day forecasts were validated against both 5-min and 10-min averaged ground station and 30-min (hourly averaged) satellite irradiance observations over the course of a month.

Due to their extensive spatial coverage, optical thickness, and reflectivity, stratocumulus (Sc) clouds are responsible for much of the variability in available solar resource at the surface in coastal California, where most rooftop photovoltaic systems are located. Currently, the trend of numerical weather prediction models is to underpredict both the presence and thickness of Sc. Therefore, the validation is conducted for a summer month in southern California, when Sc are most prevalent.

Ground station validation showed average improvements by WEMPP in predicting surface irradiance over the baseline NAM (RAP) WRF initializations of 33\% (-3\%) MBE and 16\% (9\%) MAE, and by CLDDA of 47\% (18\%) MBE and 26\% (20\%) MAE. Additionally, simulations preprocessed by CLDDA were consistently able to outperform 24-hr persistence forecast at 3 out of 4 ground stations.

Validation against SolarAnywhere\textsuperscript{R} satellite irradiance observations showed that the combination of both preprocessors provided the most improvement in the prediction of Sc spatial coverage, thickness, and lifetime in coastal regions where marine layer stratocumulus is most frequently observed, but cloud cover over the ocean was overestimated by all preprocessors.

Keywords: numerical weather prediction; solar forecasting; cloud forecasting; solar irradiance

1. Introduction

The accurate prediction of coastal stratocumulus (Sc) in the southern California region is crucial for solar energy forecasting, as Sc is the dominant cloud type in the region, and their presence greatly attenuates the solar resource in the area with the majority of rooftop solar installations. Numerical weather prediction (NWP) models become increasingly important from hours-ahead to days-ahead forecasts, since non-linearities and global coupling in the atmospheric processes make other extrapolation methods less accurate for longer forecast horizons. Furthermore, the mesoscale and global processes which develop in these forecast horizons can only be modelled by physical models and not statistical models.

*Corresponding author: H. Yang
Email address: hay022@ucsd.edu (Handa Yang)

Preprint submitted to Elsevier March 29, 2016
Hence, NWP forecasting accuracy has been the subject of many studies. Comparisons between the North American Mesoscale (NAM), Global Forecast System (GFS), and European Centre for Medium-range Weather Forecasts (ECMWF) over the continental United States by Mathiesen and Kleissl (2011) found mean bias error (MBE) and root mean square error (RMSE) of hourly-averaged forecast surface irradiance to exceed 30 W m$^{-2}$ and 110 W m$^{-2}$, respectively, over a period of about a year. Similarly, a more comprehensive intercomparison of Global Environmental Multiscale model (GEM), Mesoscale Atmospheric Simulation System (MASS), and Advanced Multiscale Regional Prediction System (ARPS) models by Perez et al. (2013) found predominantly positive MBE over the United States. NWP models thus tend to overpredict expected solar power due to errors in cloud cover, aerosol optical depth, clear sky models, and the lack of consideration of shading effects from the horizon. Since regional or mesoscale NWP models are generally initialized from synoptic or global models, errors in the larger scale model are directly inherited. Additionally, a break in the simulation timeline occurs at the transition point from one model to another (e.g. NAM to WRF) due to the subsequent “spin-up” period, during which model dynamics stabilize. For some models, this discontinuity is further exacerbated by an imperfect transfer of atmospheric state caused by absent model output, such as missing atmospheric liquid water content in NAM.

While previous studies have highlighted the sensitivity of WRF forecasts to physics parameterizations—notably the planetary boundary layer (Hu et al., 2010), microphysics (Jankov et al., 2011), and cumulus (Jankov et al., 2005) schemes, model initial conditions also play an important role in forecast accuracy, and improvements are typically obtained through data assimilation techniques. While the WRF system contains its own data assimilation system in the form of WRFDA, external data assimilation systems are also available. For example, the Local Analysis and Prediction System (LAPS) (Albers et al., 1996) is a data assimilation system designed to incorporate a variety of observations from surface sites, satellites, Doppler radars, atmospheric profilers, and aircraft into analyzed grids intended for NWP initialization. Improvement in short-range WRF forecast accuracy was demonstrated by Etherton and Santos (2008).

Other researchers have developed methods to improve cloud field initialization, such as the cloud package used within the Gridpoint Statistical Interpolation (GSI) system (Hu et al., 2007). GSI blends surface, satellite, and radar observations into a 3-dimensional field describing cloud cover and precipitation, and it is currently implemented in the NAM and Rapid Refresh (RAP) models.

In this study, the impact of model initial conditions is evaluated by quantifying forecast errors of WRF 3.6 (with the Advanced Research WRF, or ARW, dynamical core) simulations with different initialization data sets. To take advantage of NAM's data assimilation system, WRF simulations were initialized at times coinciding with NAM initialization. This study focuses on targeted preprocessing of NAM output in order to facilitate transfer of cloud cover information and to supplement the cloud field with a high-resolution satellite cloud assimilation method. Standard initializations using data from the NAM and RAP are compared against two preprocessing schemes applied to the NAM data.

The first preprocessing scheme, the Well-mixed Preprocessor (WEMPP), was developed to provide an initial guess at liquid water content, which is lacking from NAM data. Since Sc is governed by a complex system of physics with feedbacks, the intent of WEMPP is to retain cloud water content in order to prevent radiative feedbacks caused by a sudden elimination of liquid water path (LWP) at initialization. Additionally, a more accurate initial guess at the cloud water field allows the microphysics scheme to produce a stable nocturnal Sc field in only 1-2 hr. The second preprocessing scheme is the satellite Cloud Data Assimilation (CLDDA) package developed by Mathiesen et al. (2013) (applied here at initialization rather than mid-simulation). A combination of the two preprocessing schemes was also explored. In order to isolate the effects of cloud liquid water initialization from the two preprocessors, additional data assimilation (e.g. through WRFDA, LAPS, or GSI) was not applied. WEMPP and CLDDA provide relatively simple and computationally inexpensive ways to maintain consistency with the parent NAM model and improve liquid water initialization. The implementation of the preprocessing is specific to boundary layer clouds and cannot be easily adapted to other cloud types such as altocumulus or cirrus. However, boundary layer clouds are the predominant cloud type, and due to their generally large optical depth they are of most interest to solar forecast applications. We therefore expect that similar algorithms could be applied elsewhere in North America (the NAM domain) for WEMPP and globally for CLDDA.

In the following sections, the methodology will be explained, beginning with the WRF domain and physics
configurations (2.1), followed by a description of the two preprocessing schemes used (2.2 and 2.3). Next, the study setup (time period and intercomparison setup in 2.4), validation methods, and data sources are presented (2.5). Validation results are presented and discussed in Section 3. Section 4 provides conclusions.

2. Methodology

2.1. WRF configuration

Two one-way nested domains at 8.1 km and 2.7 km resolution were initialized at 12Z (0400 PST) in WRF 3.6 with the physics configuration summarized in Table 1, which was based on a sensitivity study by López-Coto et al. (2013). Each domain contains 100 grid points in both horizontal directions and 75 vertical levels, with 50 levels below 3 km; their locations are shown in Figure 1. Output was recorded at 15-min intervals, and the time stepping interval was 10 s.

Figure 1: Outer and inner domains at 8.1 km (100x100x75 grid points) and 2.7 km (100x100x75) resolution. Satellite image ©2015 Google.

Table 1: Summary of WRF physics parameterizations

<table>
<thead>
<tr>
<th>Parameterization type</th>
<th>Namelist setting (outer/inner)</th>
<th>Scheme name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planetary boundary layer</td>
<td>5/5</td>
<td>MYNN2</td>
</tr>
<tr>
<td>Surface layer</td>
<td>5/5</td>
<td>MYNN</td>
</tr>
<tr>
<td>Land surface</td>
<td>2/2</td>
<td>Noah LSM</td>
</tr>
<tr>
<td>Radiation (LW &amp; SW)</td>
<td>5/5</td>
<td>New Goddard</td>
</tr>
<tr>
<td>Microphysics</td>
<td>10/10</td>
<td>Morrison 2-moment</td>
</tr>
<tr>
<td>Cumulus</td>
<td>14/0</td>
<td>New SAS</td>
</tr>
</tbody>
</table>

2.2. Well-mixed Preprocessor (WEMPP)

In a WRF simulation initialized with atmospheric state variables obtained from the North American Mesoscale (NAM) model, temperature and relative humidity (RH) fields are ingested, but atmospheric
liquid water is not present in NAM output. This leads to a cloud cover spin-up period of a few hours during which the microphysics scheme produces a cloud field based on temperature and relative humidity fields.

Since stratocumulus cloud decks play an important role in radiative and entrainment fluxes in the planetary boundary layer (PBL) (Wood [2012]), a simple preprocessing scheme was developed in order to preserve cloud cover information between NAM and WRF atmospheric states. In addition to preserving consistency between the NAM parent model and WRF, adding liquid water ab initio contributes to more accurate modelling of land surface fluxes and the Earth's energy budget by reducing both surface longwave cooling during nighttime (due to longwave emission from low clouds exceeding atmospheric longwave radiation) and shortwave heating during daytime.

The Well-mixed Preprocessor (WEMPP) was developed based on the well-mixed stratocumulus-topped boundary layer (STBL) framework developed by Lilly [1968]. A key characteristic of the STBL is the presence of a capping temperature inversion: a vertical layer in which temperature increases with height. The abrupt drop in ambient density caused by the strong positive temperature gradient at the inversion base height (IBH) acts as a lid for ascending air parcels, therefore approximately coinciding with the cloud top height. At nighttime, longwave radiative cooling at the cloud top drives turbulence within the boundary layer which strongly mixes the air within, leading to near-constant conserved variables (e.g. total water mixing ratio $q_t$ and equivalent potential temperature $\Theta_e$) (Wood [2012]). The conservation of $q_t$ can be used to predict the amount of liquid water within the boundary layer. Since the goal of WEMPP is only to facilitate the carryover of "missing" liquid water input, the temperature profile is not adjusted. Based on the above, the algorithm developed is as follows:

1. In every model column, the temperature inversion, if present, is detected by searching for adjacent layers in which temperature continuously increases with height. The layer with the largest difference in temperature between the top and base is defined to be the capping inversion, similarly to the method described by Iacobellis et al. [2009]. Only inversions with base height $< 3$ km above ground level (AGL) are considered.

2. Cloud top $z_{ctop}$ is then defined as the highest point below the IBH with RH $\geq 95\%$ (although physically cloudy air parcels should be 100\% RH, this criterion was chosen to correct observed dry biases in NAM).

3. Assuming a single cloud layer, compute mass-weighted cloud layer means of RH for subsequently lower grid points until layer RH $< 95\%$. Define this point to be cloud base $z_{cbase}$. This computation for the WRF-ARW grid is:

$$\text{RH} = \frac{\sum_{k=k_{z_{ctop}}}^{k_{z_{base}}} \text{RH}(z_k) \Delta \eta(z_k)}{\sum_{k=k_{z_{ctop}}}^{k_{z_{base}}} \Delta \eta(z_k)}, \quad \text{for } n = k_{z_{ctop}} - 1, k_{z_{ctop}} - 2, \ldots \text{ until } \text{RH} < 95\%, \quad (1)$$

where $\Delta \eta$ is the change in $\eta$ (the vertical coordinate in WRF-ARW) between full model levels.

4. At the cloud base, take the saturation water vapor mixing ratio $q_{sat}$ to be the constant conserved total water mixing ratio $q_t$ within the PBL:

$$q_t(0 \leq z \leq z_{ctop}) \equiv q_{sat}(z_{cbase}). \quad (2)$$

5. Set $q_c$ to $q_{sat}$ in all grid points within the cloud layer, and partition $q_t - q_{sat}$ to liquid water mixing ratio $q_l$:

$$\begin{cases} q_c(z) = q_{sat}(z), & \text{for } z_{cbase} \leq z \leq z_{ctop} \\ q_l(z) = q_t - q_{sat}(z) \end{cases} \quad (3)$$

An illustration of the procedure is shown in [Figure 2]. In the left panel, the detected cloud top and cloud base are overlaid on the original input relative humidity profile. In the right panel are adjusted vertical moisture profiles (see caption for detailed description). The lifting condensation level (LCL) was computed after Stull [2000] from the equation

$$z_{LCL} = a(T - T_d), \quad (4)$$
where \( a = 125 \text{ m K}^{-1} \), \( T \) is temperature, and \( T_d \) is the dew point temperature. Atmospheric state variables were obtained from the first vertical grid point, and the LCL is found to typically agree with the algorithmically-determined cloud base height to within a few tens of meters.

Figure 2: **Left:** Original initial temperature (red) and relative humidity (green) profiles above a land grid point at 12Z. **Right:** Preprocessed moisture profiles at initialization. Original water vapor mixing ratio \( q_v \) in green, preprocessed \( q_v \) as crosses, preprocessed liquid water mixing ratio \( q_l \) in solid blue, and saturated water vapor mixing ratio \( q_{sat} \) in dashed red. Lifting condensation level (LCL) is denoted by black triangles. The LCL does not necessarily equal \( z_{cbase} \), but supports the cloud base height determined by the algorithm.

2.3. **Cloud Data Assimilation (CLDDA)**

The Cloud Data Assimilation (CLDDA) preprocessor was designed by Mathiesen et al. (2013) to populate liquid cloud fields based on satellite observations. In short, the latitude and longitude positions of low clouds are first obtained from cloud top temperature data in the GOES Sounder cloud product (Li et al., 2001) maintained by CIMSS at the University of Wisconsin-Madison. In observed cloudy regions, cloud tops are defined at the temperature inversion base height or, in the absence of an inversion, at the intersection of the initial vertical temperature profile and satellite-derived cloud top temperature. Cloud base is then estimated from an empirical function of cloud top height (see Mathiesen et al. (2013)). Relative humidity is then set to 110% within all cloud decks. In order to avoid intense latent heating from the resulting condensation, microphysics heating is turned off for 1 hr following application. Lastly, in observed clear regions, RH is adjusted to a maximum of 75% in order to remove falsely predicted clouds. Figure 3 depicts the adjusted RH profile at initialization for the same time and grid point as in Figure 2 as well as the water vapor \( q_v \) and liquid water \( q_l \) (condensed by the microphysics scheme) mixing ratios 15 min after initialization. The LCL typically agrees within several tens of meters with the initially detected cloud base, though notable departure is observed after 15 min due to a decrease in the near-surface temperature (common in all NAM-initialized simulations).

2.4. **Setup**

Because both preprocessing schemes were designed to target stratocumulus clouds, the period of validation was chosen based on frequency of occurrence of stratocumulus in the San Diego region. Klein and Hartmann (1993) showed that the three-month seasonally- and spatially-averaged (10° x 10°) marine stratatus, stratocumulus, and fog amount off of the Californian coast over 30 years peaked in June, July, and August. Recently, Clemesha et al. (2016) showed that the seasonal cycle of coastal low cloudiness peaks in the San Diego region during the month of June (see Figure 4 in Clemesha et al. (2016)). Therefore, the month of June, 2013 was chosen for study, as maximal cloud coverage will coincide with maximal forecast error.

Five model outputs and/or preprocessing schemes were used to initialize WRF:

1. 12Z-cycle NAM (0th-hr forecast on 218 AWIPS CONUS grid at 12 km resolution)
2. 00Z-cycle RAP (hybrid levels at 13 km resolution)
3. 12Z-cycle NAM with WEMPP
4. 12Z-cycle NAM with CLDDA
5. 12Z-cycle NAM with CLDDA and WEMPP (COMBO)

All simulations were initialized at 12Z (0400 PST) and were run past sunset for 17 hr until 05Z (2100 PST). Average sunrise and sunset were at 0442 PST and 1858 PST, respectively. The first two configurations will function as baselines for comparison. The 00Z-cycle RAP was chosen rather than the 12Z-cycle because it is the only cycle in current archives containing liquid water information. It is therefore a reference case for an initialization that mitigates the lack of cloud water in NAM output. Additionally, the two original raw data forecasts (RAW-NAM and RAW-RAP) were included as performance baselines and were interpolated onto the same 8.1 km and 2.7 km WRF grids to facilitate intercomparison.

Configurations 3 and 4 were implemented as described in Sections 2.2 and 2.3. The COMBO configuration was designed to maximize initial cloud cover, and both preprocessing methods were applied. First, two independent simulations of configurations 3 and 4 were run for 15 min each. After completion, a composite moisture field was created by using the moisture profile from the simulation with the highest liquid water path (LWP) at each column, and the simulation was resumed.

Over the month, the average modification to the outer (inner) domain was 37.7% (45.6%) by WEMPP and 26.8% (52.1%) by CLDDA. Regions of modification in the inner domain for a typical day are shown in Figure 4.

2.5. Spinup Period

Although cloud water is expected to be accurate at initialization in preprocessed simulations, a spin-up period is still necessary to correctly simulate atmospheric dynamics. Skamarock (2004) quantified the spin-up period of model dynamics by analyzing model kinetic energy spectra. Using the same methods, kinetic energy spectral densities of all five WRF configurations is shown for the inner domain in Figure 5 for a day when the discrepancy between initialization and subsequent spectra is most pronounced. The initial spectra, which is lacking in energy at finer scales, becomes fully developed about 1-2 hr into the simulation, which is significantly shorter than that found by Skamarock (2004). We hypothesize that the finer spatial resolution of initialization data (12-13 km vs. 40-80 km) combined with the 10 s timestepping interval significantly reduced the spin-up period from the previously determined 6-12 hr.

2.6. Validation sites and data

2.6.1. Ground sites

Measurements from ground stations maintained by the San Diego Gas & Electric Company (SDGE) were used for validation. Four weather stations with LI-COR LI200 pyranometers were selected along a line
from west to east. Station locations are shown in Figure 6 and their geographical coordinates, elevation, distance from the coast, averaging interval, and mean observed GHI and clear sky index $k_c$ (see caption) are tabulated in Table 2.

Table 2: Summary of ground stations used for validation. $k_c$ is clear sky index, which is GHI normalized by modelled clear sky GHI. Mean GHI and $k_c$ are monthly values computed from daytime data collected in June, 2013.

<table>
<thead>
<tr>
<th>Station ID</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude [°]</td>
<td>33.12°N</td>
<td>33.14°N</td>
<td>33.13°N</td>
<td>33.16°N</td>
</tr>
<tr>
<td>Longitude [°]</td>
<td>117.20°W</td>
<td>117.24°W</td>
<td>117.20°W</td>
<td>117.03°W</td>
</tr>
<tr>
<td>Elevation [m MSL]</td>
<td>85</td>
<td>143</td>
<td>165</td>
<td>315</td>
</tr>
<tr>
<td>Distance from coast [m]</td>
<td>3,800</td>
<td>9,300</td>
<td>12,200</td>
<td>28,700</td>
</tr>
<tr>
<td>Averaging interval [min]</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>Mean GHI [W m⁻²]</td>
<td>449</td>
<td>481</td>
<td>509</td>
<td>518</td>
</tr>
<tr>
<td>Mean $k_c$ [-]</td>
<td>0.71</td>
<td>0.76</td>
<td>0.84</td>
<td>0.84</td>
</tr>
</tbody>
</table>
Figure 6: Locations of the four ground stations used for validation. The white dashed line marks the border of the coastal marine layer region, defined as the region where land elevation < 375 m MSL and monthly mean observed GHI is < 92% of clear sky GHI. The region contains 972 WRF grid points in the inner domain. Satellite image ©2015 Google.

All stations recorded averaged global horizontal irradiance (GHI) data at 5-min resolution (sampled every 10 s), with the exception of station 4 which recorded at 10-min resolution (sampled every 3 s). Measurements from all stations were aligned with instantaneous model output (hourly for raw data, 15-min intervals for WRF) for error computation through linear interpolation.

2.6.2. Satellite solar resource data

Additional validation was performed against Clean Power Research’s SolarAnywhere® (2013) satellite data at 0.02° spatial and 30-min (hourly averaged) temporal resolution, with the exception of the raw NAM and RAP forecasts which were at 1-hr instantaneous temporal resolution. Jamaly et al. (2012)’s validation against 52 California Irrigation Management Information System (CIMIS) ground stations at hourly-averaged temporal resolution over the year 2010 showed a previous version of SolarAnywhere to have MBE of 18.8 W m$^{-2}$, mean absolute error (MAE) of 45.6 W m$^{-2}$, and RMSE of 65.3 W m$^{-2}$—an accuracy which is comparable to that of typical ground stations.

Because the validation region contains the Pacific Ocean, regions affected by sunglint were computed after Gardashov and Eminov (2015), wherein the principal point of sunglint (PPS) could be determined as a function of geostationary satellite longitude and time of year. The PPS migrates farthest to the north and closest to the region of interest during the summer solstice. Computations of the PPS on June 21 show that the PPS resides a minimum of 17° latitude south of the southern edge of the outer WRF domain, thereby eliminating sunglint as a source of error.

2.6.3. Model output statistics (MOS) correction

A model output statistics (MOS) correction was applied in order to correct clear-sky bias, similar to that applied in Section 3.4 of Mathiesen and Kleissl (2011), except the MOS correction used in this paper is solely a function of solar zenith angle and was trained only during periods of clear sky. MOS correction functions were created for RAW-NAM, RAW-RAP, WRF-NAM, and WRF-RAP configurations from observations at solar zenith angle < 75° from May 2, 13, 19, 24, and 29, 2013. In validation against ground stations, MOS correction functions were derived at each station. Against satellite observations, the MOS function was derived at the grid point containing ground station 4 and applied to the entire domain. The WRF-NAM MOS function was applied to WRF-WEMPP, WRF-CLDDA, and WRF-COMBO since they share the same initialization data set and clear sky biases. The application of MOS is intended to match clear-sky irradiance predictions between all models to emphasize cloudy conditions in the evaluation of the preprocessing schemes.
Otherwise, since all tested configurations overpredicted clear-sky GHI, a systematic underprediction of GHI under cloudy conditions could be offset by the clear-sky bias and artificially reduce MBE.

The clear-sky irradiance biases are dependent on the physics parameterizations used, and therefore affect the entire domain. The highest clear-sky bias is observed in RAW-NAM (using the Geophysical Fluid Dynamics Laboratory (GFDL) radiation scheme at the time), while RAW-RAP shows almost no clear-sky bias (using the Goddard shortwave scheme at the time). The clear-sky bias in the WRF simulations has since been corrected by Zhong and Kleissl (2016) through a modification to the New Goddard shortwave radiation scheme.

2.6.4. Error metrics

The error metrics chosen here for validation are mean absolute error (MAE):

\[
\text{MAE} \equiv \frac{1}{N} \sum_{n=1}^{N} |x_n - x_n^{\text{obs}}|, \tag{5}
\]

and mean bias error (MBE):

\[
\text{MBE} \equiv \frac{1}{N} \sum_{n=1}^{N} x_n - x_n^{\text{obs}}, \tag{6}
\]

where \(x_n\) and \(x_n^{\text{obs}}\) are respectively the \(n^{\text{th}}\) forecast and observed variables over a total of \(N\) data points.

Errors were computed for GHI for SDGE ground stations, and for clear sky index \(k_c\) for SolarAnywhere satellite measurements, where

\[
k_c \equiv \frac{\text{GHI}}{\text{GHI}_{\text{clearsky}}}, \tag{7}
\]

with \(\text{GHI}_{\text{clearsky}}\) obtained from the Kasten clear sky model as modified by Ineichen and Perez (Ineichen and Perez, 2002; Perez et al., 2002).

The GHI errors will facilitate comparison of these preprocessing schemes against 24-hr persistence forecasts and other forecasting methods in the literature, while \(k_c\) errors will allow visualization of the spatial and temporal patterns of coastal stratocumulus in the San Diego region by normalizing out the time of day and locational dependence of solar irradiance. Since marine layer stratocumulus mostly occur during morning periods when the magnitude of GHI is low to medium, \(k_c\) errors more clearly show errors in cloud cover patterns than GHI errors, which are weighted towards solar noon. Overall errors over the coastal regions most frequently affected by marine layer cloud cover were computed by spatially averaging over land-only regions with satellite-observed monthly mean \(k_c < 0.92\) and at an elevation of < 375 m MSL. This region contains 972 WRF grid points in the inner domain.

Monthly averages of errors were computed at the finest available temporal resolution, constrained by either the forecast output interval or measured data availability. The time periods sampled, temporal resolution, number of daily samples, and total samples \(N\) are tabulated in Table 3 for WRF and raw data (NAM and RAP) forecasts.

<table>
<thead>
<tr>
<th>Type</th>
<th>Time period</th>
<th>Resolution</th>
<th>Daily samples</th>
<th>Total samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground stations [GHI]</td>
<td>0500 - 1900</td>
<td>15 min / 1 hr</td>
<td>57 / 15</td>
<td>1710 / 450</td>
</tr>
<tr>
<td>Satellite ([k_c])</td>
<td>0600 - 1700</td>
<td>30 min / 1 hr</td>
<td>23 / 12</td>
<td>690 / 360</td>
</tr>
</tbody>
</table>
To directly compare each preprocessing method, a forecast skill was calculated for each ground station. Kleissl (2013) defined a general forecast skill suitable for forecast result intercomparison as the ratio of forecast model RMSE to persistence model RMSE. Here, MAE was used instead of RMSE due to its linear nature. Forecast skill FS is therefore defined as

$$FS \equiv 1 - \frac{\text{MAE}}{\text{MAE}_{\text{persistence}}}.$$  

(8)

Positive values of forecast skill indicate superior performance to persistence forecast, with a maximum value of 1. The persistence forecast here was generated by taking the prior day’s measured irradiance at local times coinciding with WRF output times. Thus, the ground station (satellite) persistence forecast was at 15-min (30-min) resolution with 57 (23) samples per day and 1710 (690) over the month.

3. Results and Discussion

3.1. Validation against ground station measurements

An example time series at station 2 of pyranometer-measured and forecast GHI from all configurations as well as persistence forecast is shown in Figure 7 for June 2, 2013. Also shown is a time series of the WEMPP forecast prior to application of MOS for reference. Note that raw forecasts are at 1-hr resolution while WRF and persistence forecasts are at 15-min resolution.

![Figure 7: Time series at ground station 2 of pyranometer-measured (thick solid black) and forecast GHI of all forecast configurations as well as persistence forecast (dot-dashed gray with hexagonal markers) for June 2, 2013. WRF-WEMPP is shown both with and without MOS correction for reference. Note RAW-NAM and RAW-RAP are at 1-hr resolution while observations, WRF, and persistence forecasts are at 15-min resolution.](image_url)

Figure 8 shows monthly-averaged GHI MAE and MBE for all forecast configurations. Between the raw data forecasts, RAW-RAP consistently performs better than RAW-NAM in terms of both MAE and MBE. At all stations except for station 1, cloudier conditions are present in RAW-RAP as indicated by lower MBE. While the raw data forecast MAE’s are not directly comparable to those of the WRF forecasts due to a temporal resolution mismatch (the variability of irradiance increases with temporal resolution, thus increasing MAE), comparison against their children WRF forecasts reveals the importance of liquid water content at initialization through MBE, which is unaffected by the mismatch. WRF-NAM performs similarly to RAW-NAM, with a slight reduction in MAE and MBE across all stations. WRF-RAP, however, shows slightly increased MAE at all stations and similar MBE at stations 2-4 compared against RAW-RAP, but a significant decrease in MBE at station 1 of nearly 75% due to newly resolved cloud cover.

Also shown in Figure 8 are errors from cross-validation of the SolarAnywhere data at the ground station locations. Compared to Jamaly et al. (2012), MAE is higher by 13.4 W m$^{-2}$, while MBE is lower by 7.1 W m$^{-2}$, on average.

WRF-NAM is more positively biased than WRF-RAP, partly due to the proximity of the WRF-NAM initialization to sunrise as the microphysics scheme acts too slowly to generate a cloud field from zero initial liquid water content. In other words, without a preprocessing scheme it may have been beneficial...
Figure 8: **Top:** GHI mean absolute error and **bottom:** GHI mean bias error of ground stations grouped by station ID (1: coastal to 4: inland). Errors of 24 hr persistence forecasts are also shown. Note that RAW forecasts are at 1-hr resolution while all other forecasts are at 15-min resolution. Lastly, errors of the 30-min resolution SolarAnywhere observational data set are shown.

to initialize WRF-NAM earlier to allow cloud fields to spin up. Indeed, the introduction of liquid water at initialization through the application of WEMPP lowers errors to similar values as that of the WRF-RAP reference case (about 9% less MAE and 3% greater MBE). Considering all state variables at initialization in WRF-WEMPP are identical to those in WRF-NAM with exception of liquid water, this improvement in error metrics is significant, especially because RAW-NAM was shown to perform worse than RAW-RAP.

Assimilating satellite cloud information through CLDDA further lowers both MAE and MBE as compared with the WRF-RAP and WRF-WEMPP configurations. The additional application of WEMPP in the WRF-COMBO configuration slightly lowers MBE by further increasing liquid water content throughout the simulation domain, but the increase in MAE reveals the occurrence of overestimation of liquid water content.

For all stations, MAE decreases as the distance from the coast increases, as cloud cover is less frequent in inland regions. Defining clear periods to be times when \( k_c \) falls between 0.9 and 1.1 (i.e. approximately cloud-free and without cloud edge enhancement effects), the percentage of observed clear data points was, in order from most coastal (1) to most inland (4): 39.7%, 46.3%, 51.9%, and 61.2%. Because station 4 is just outside of the typical marine layer region and rarely observes much cloud cover except for early mornings, persistence forecast tends to perform very well. Overall, all configurations show only positive bias, indicating a systematic underprediction of cloud cover. At station 1, however, WRF-RAP and the preprocessed configurations occasionally overpredicted the optical thickness of present clouds or predicted cloud cover during clear conditions. Because of this, MBE at station 1 is lower than at stations 2 and 3, but the higher MAE reveals the occurrence of errors is greater.

Forecast skills were computed for each configuration at each ground station, indicating performance against the baseline persistence forecast. Positive values of forecast skill indicate superior performance compared to persistence forecast, and the maximum value is 1. According to Figure 9, the best performer is CLDDA with an average skill of -0.15. The remaining configurations rank as COMBO (-0.16), RAW-RAP (-0.18), WEMPP (-0.30), WRF-RAP (-0.35), WRF-NAM (-0.49), and RAW-NAM (-0.71). The main contributor to negative forecast skills is station 4, where the early morning cloud cover is typically not captured by any configuration, indicating a lack of simulated inland Sc penetration. At stations 1-3, CLDDA shows positive skill, as does the COMBO configuration for stations 2-3, and RAW-RAP at station 3.

### 3.2. Validation against satellite observations

A time series at the grid point containing station 2 of satellite-derived and forecast clear sky index \( k_c \) from all configurations as well as satellite persistence forecast is shown in Figure 10 for June 2, 2013. Also shown is a time series of the WEMPP forecast prior to application of MOS for reference. Note that raw forecasts are at 1-hr resolution while WRF and persistence forecasts are at 30-min resolution.
Figure 9: Forecast skill against ground station GHI persistence forecast for each configuration as a function of station ID. Positive values indicate performance superior to 24-hr persistence forecast.

Figure 10: Time series at ground station 2 of satellite-derived (thick solid black) and forecast $k_c$ of all forecast configurations as well as persistence forecast (dot-dashed gray with hexagonal markers) for June 2, 2013. WRF-WEMPP is shown both with and without MOS correction for reference. Note RAW-NAM and RAW-RAP are at 1-hr resolution while observations, WRF, and persistence forecasts are at 30-min resolution.

Figure 11 spatially depicts monthly-averaged MAE of each configuration alongside the satellite persistence forecast and monthly-average observed $k_c$ (bottom right). Of the raw data forecasts, RAW-NAM shows significant errors everywhere in the domain, and RAW-RAP performs similarly to WRF-RAP, with both showing highest errors over the ocean and near the coastline. The baseline WRF-NAM configuration shows highest MAE of all WRF configurations over the coast, and while WRF-RAP performs better than WRF-NAM, the coastal MAE still exceeds that of the preprocessed configurations. Among the preprocessors, WEMPP is the worst performer, and CLDDA shows the most significant improvement over WRF-NAM, especially farther inland. The COMBO configuration performs similarly to CLDDA in the San Diego region, but performs slightly better in the Los Angeles area, though errors over the ocean near the San Diego coast are highest of all configurations. The errors in the satellite persistence forecast are concentrated in the regions of lowest observed $k_c$: predominantly over the ocean and the coastline. This is mostly because persistence forecasts are unable to predict day-to-day variations from cloudy to clear conditions (and vice versa).

From Figure 12 all WRF configurations show a systematic positive bias over the coast and a systematic negative bias over the ocean. Both raw forecasts show positive bias everywhere. In the coastal marine layer region (bounded by the Californian border and black dashed line in Figure 12), the positive bias is caused by an underprediction of cloud cover. However, in typically clear regions (i.e. regions of $\approx 1.0$ $k_c$ in Figure 12 bottom right), a positive bias of $\approx 0.05$ to $\approx 0.1$ $k_c$ remains in the RAW-NAM despite application of MOS correction. This positive bias is due to the high clear-sky $k_c$ errors at high solar zenith angles in the
RAW-NAM which are not corrected by MOS (evident at 0600-0700 PST and at 1700 PST in Figure 10). While all WRF configurations show negative bias over the ocean, all preprocessing schemes further increase this bias, indicating an overprediction of cloud cover. In WEMPP, the overprediction is caused by the conservative 95% relative humidity criterion. Because slightly subsaturated air parcels are considered cloudy by WEMPP, preprocessed cloud bases are lower than would be in the original NAM data, thus increasing cloud optical thickness. The CLDDA algorithm appears to overestimate cloud thickness over the ocean because it was empirically developed from observations over land. The COMBO configuration shows the highest negative bias since it was designed to maximize liquid water content. The satellite persistence forecast shows near-zero MBE due to the day-to-day offsetting of bias errors.

Errors averaged over the coastal marine layer region are summarized in Figure 13. While the WRF-NAM configuration is the worst performer of all WRF configurations, application of WEMPP to NAM initialization data lowers MAE below that of the WRF-RAP configuration (though MBE remains slightly higher). Application of CLDDA lowers errors further, and the COMBO configuration is the best performer in terms of both MAE and MBE computed from both GHI and $k_c$. Of the raw forecasts, RAW-NAM is the worst performer overall, but RAW-RAP performs similarly to WRF-RAP and WRF-WEMPP in MAE.
(however, RAW-RAP is at hourly resolution which tends to decrease MAE), but with much higher MBE and coarser temporal resolution. Compared with ground station validation, spatially-averaged MBE is lower over the marine layer region due to the increased spatial coverage of Sc.

Shown in Figure 14 are forecast skills computed from GHI MAE at ground station locations as well as for the marine layer region. The configurations rank as follows in terms of average skill over the ground stations [marine layer region]: CLDDA (0.07) [0.05], COMBO (0.06) [0.06], WEMPP (-0.01) [-0.01], RAW-RAP (-0.07) [-0.06], WRF-RAP (-0.11) [-0.12], WRF-NAM (-0.14) [-0.16], and RAW-NAM (-0.47) [-0.45]. Forecast skills averaged over all ground stations are in good agreement with those averaged over the marine layer region, suggesting the chosen ground stations are representative of the region. Again, the main contributor to the station-averaged forecast skill is station 4, indicating a lack of simulated inland Sc penetration. CLDDA, WEMPP, and COMBO all show positive skill at stations 1-3. Over the marine layer region, both COMBO and CLDDA exhibit positive skill, with COMBO performing slightly better.

Figure 13: Monthly mean errors in GHI (left) and clear sky index $k_c$ (right) over the coastal land-only marine layer region (defined in Figure 6) computed from satellite data. These errors are spatially averaged over 972 WRF grid points in the inner domain.

Figure 14: Forecast skill against satellite GHI persistence forecast for each configuration as a function of station ID as well as over the marine layer (ML) region. Positive values indicate performance superior to 24-hr persistence forecast.

The simulated evolution of coastal Sc can be observed by computing the monthly mean of hourly $k_c$ bias errors, as shown in Figure 15. The WRF-NAM configuration greatly underpredicts cloud cover over the coast, while WRF-RAP was able to predict some coastal cloud cover in the San Diego region, as well as over the Los Angeles basin. Applying WEMPP to NAM initial conditions slightly improved prediction of cloud cover over most of the coastline, but inland coverage is still lacking. The most significant improvement in spatial Sc coverage was gained by applying CLDDA. Both CLDDA and COMBO configurations show similar cloud cover over land, but all preprocessed configurations overpredicted cloud cover over the ocean, with COMBO showing the most overprediction over ocean. While RAW-NAM is the worst performer overall,
RAW-RAP appears to better capture early morning cloud cover than the WRF-RAP configuration and shows similar inland cloud coverage at 0600 PST as simulations preprocessed by CLDDA, though lifetime of coastal clouds is much shorter. Satellite persistence forecast is not shown in Figure 15 due to its near-zero MBE.

Figure 15: Monthly mean of hourly $k_c$ bias errors, from 0600 to 1000 local time (PST) (sunrise was around 0442 PST) for each configuration (each hourly panel is an average of 30 data points). Magnitudes are in clear sky index $k_c$, with warm (cool) colors representing underprediction (overprediction) of cloud cover.
4. Conclusions and future work

In this study, two initial condition preprocessing schemes WEMPP, CLDDA, and their combination were compared against baseline WRF simulations initialized with NAM and RAP data sets. Comparison against ground station measurements showed CLDDA as the best performer, consistently surpassing persistence forecast at 3 out of 4 ground stations, with the COMBO configuration as a close second.

Supplementary validation against satellite ground irradiance data showed CLDDA to offer the most significant improvement in the prediction of both spatial coverage and lifetime of coastal stratocumulus, while WEMPP led to minor improvements. Their combination, however, provided the most improvement overall over the coast, while also causing the most overprediction of cloud cover over the ocean. All preprocessed schemes overpredicted cloud cover over the ocean.

Quantitative comparisons against persistence forecast were determined by the forecast skill metric, where positive (negative) values indicate superior (inferior) performance to persistence forecast. Against ground stations, all configurations showed negative forecast skill on average, mostly due to poor forecast skill at the most inland station, indicating insufficient simulated inland penetration of Sc. Slightly positive skills ranging from 0.01 to 0.08 were demonstrated by CLDDA (at 3 stations), COMBO (2), and RAW-RAP (1). Against satellite observations, positive forecast skills ranging from 0.06 to 0.07 were demonstrated by CLDDA and COMBO when averaged over the ground station locations. CLDDA and COMBO also demonstrated positive forecast skill ranging from 0.05 to 0.06 when averaged over the coastal marine layer region.

The lack of significant forecast skill indicates that accurate forecasts of marine layer Sc cannot be obtained solely through initial conditions processing and data assimilation, though a large improvement has been made. Pragmatically, however, persistence forecasts for behind-the-meter generation require collocated measurements which is typically not the case for distributed generation. Persistence forecasts also are unable to capture large changes from day to day, while NWP models are designed to predict these changes and (at least with continued improvement of the physics parameterizations) forecast irradiance over large spatial domains. The error magnitudes presented here are representative for days with low cloud cover, which primarily occurs from May through July and intermittently for the rest of the year (Clemesha et al., 2016). Forecast accuracy would be better on other days which are predominantly clear.

Since the amount of liquid water $q_l$ added in CLDDA as 110% RH was “calibrated” using coastal data (Mathiesen et al., 2013), it is not surprising that CLDDA is without bias along the coast. The CLDDA and WEMPP findings point to the need for modification of the addition of water vapor with less water to be added over the ocean and more further inland. The finding of larger WEMPP biases further inland can be explained by Figure 4: the NAM parent model clouds’ inland penetration is insufficient and cannot be corrected relying solely on parent model fields. CLDDA, on the other hand, leverages satellite measurements, reducing both MBE and MAE.

Although CLDDA provided the best forecast improvement, its performance depends on the accuracy of the satellite cloud mask. Nighttime low clouds such as stratocumulus are notoriously difficult to detect as their thermal infrared signature is not significantly different than that of the underlying ocean or land. Regions wrongly detected as clear in observed cloud fields (e.g. around 33.5°N, 117.5°W in Figure 4) sometimes appear over both land and ocean, nullifying any advantage gained from CLDDA. The additional application of WEMPP as in the COMBO configuration can correct such errors at least for ocean and coastal regions, but the increased cloud cover over the ocean is potentially problematic. If satellite data quality is an issue, the combination of both WEMPP and CLDDA may be considered; otherwise, only CLDDA should be applied. If satellite data is unavailable, WEMPP is a computationally cheap method to generate an initial liquid cloud field provided the initial cloud field is predominantly stratus or stratocumulus. All schemes complete on order of a few minutes on a 3.4 GHz Intel i7 machine. CLDDA and COMBO, however, require additional time to retrieve the relevant satellite data which is dependent on the data source.

From this study, the role of initial conditions on the prediction of coastal stratocumulus is shown to be significant. Although improvements in both spatial coverage and cloud lifetime were accomplished, validation against satellite observations still show underprediction of cloud cover, suggesting the need of more accurate parameterization of physical processes such as entrainment of warm, dry air from the free troposphere,
advection, and surface fluxes (Ghonima et al., 2016). The lack of inland cloud cover also suggests possible systematic underestimation of inversion base height (IBH): since the IBH physically constrains Sc cloud tops, clouds cannot be present in regions where the land elevation exceeds the IBH. Future work will thus focus on improving physical parameterization of stratocumulus processes.

**Acknowledgements**

The authors would like to express their gratitude to the Space Science and Engineering Center (SSEC) and Cooperative Institute for Meteorological Satellite Studies (CIMSS) at the University of Wisconsin-Madison for their GOES Sounder cloud product which serves as the input data for CLDDA, Clean Power Research for providing SolarAnywhere® (2013) high resolution satellite-derived surface irradiance data, and San Diego Gas & Electric Company (SDGE) for access to their ground station irradiance measurements. North American Mesoscale Forecast System (NAM) data was obtained from the National Centers for Environmental Prediction (NCEP) at the National Oceanic and Atmospheric Administration (NOAA). Historical Rapid Refresh (RAP) data were obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, Climate and Environmental Sciences Division. Funding was provided by the California Solar Initiative RD&D program.

The authors also extend thanks to Dr. Rauf Gardashov for providing results from his sunglint location algorithm.

**References**


