

Title: Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States

Authors: Patrick Mathiesen¹ and Jan Kleissl¹ (corresponding author)

¹Department of Mechanical and Aerospace Engineering, University of California, San Diego.
9500 Gilman Dr., La Jolla, CA, 92093, USA.
phone: +1 858 534 8087, email: jkleissl@ucsd.edu

Abstract:

Numerical weather prediction (NWP) models are generally the most accurate tools for forecasting solar irradiation several hours in advance. This study validates the North American Model (NAM), Global Forecast System (GFS), and European Centre for Medium-Range Weather Forecasts (ECMWF) global horizontal irradiance (GHI) forecasts for the continental United States (CONUS) using SURFRAD ground measurement data. Persistence and clear sky forecasts are also evaluated. For measured clear conditions all NWP models are biased by less than 50 W m^{-2} . For cloudy conditions near solar noon these biases can exceed 200 W m^{-2} . In general, the NWP models (especially GFS and NAM) are biased towards forecasting clear conditions resulting in large, positive biases.

Mean bias errors (MBE) are obtained for each NWP model as a function of solar zenith angle and *forecast* clear sky index, kt^* , to derive a bias correction function through model output statistics (MOS). For forecast clear sky conditions, the NAM and GFS are found to be positively biased by up to 150 W m^{-2} , while ECMWF MBE is small. Outside of the relatively few clear forecasts that were actually cloudy, the reason for this bias is that the GFS and especially the NAM forecasts can exceed clear sky irradiances by up to 40%, indicating an inaccurate clear sky model. For forecast cloudy conditions ($kt^* < 0.4$) the NAM and GFS models have a negative bias of up to -150 W m^{-2} . ECMWF forecasts are most biased for moderate cloudy conditions ($0.4 < kt^* < 0.9$) with an average over-prediction of 100 W m^{-2} .

MOS-corrected NWP forecasts based on solar zenith angle and kt^* provide an important baseline accuracy to evaluate other forecasting techniques. MOS minimizes MBE for all NWP models. Root mean square errors are also reduced by 50 W m^{-2} , especially for intermediate clear sky indices. The MOS-corrected GFS provides the best solar forecasts for the CONUS with an RMSE of about 85 W m^{-2} . ECMWF is the most accurate forecast in cloudy conditions, while GFS has the best clear sky accuracy.

Keywords: Model output statistics (MOS), Numerical Weather Prediction (NWP), Solar Forecasting

1. INTRODUCTION

Accurate forecasting of solar irradiance is essential for the efficient operation of solar thermal power plants, the management of energy markets, and the widespread implementation of solar photovoltaic (PV) technology (Lew and Piwko, 2010). Early forecasting attempts used model output statistics (MOS, Jensenius and Cotton, 1981) to establish regression models between solar irradiance measurements and available forecast variables. Subsequent forecasts of these variables were used as input to MOS, predicting solar irradiation. Using this method, Jensenius and Cotton (1981) forecasted daily global horizontal solar irradiation (GHI) on a 6 to 30-hour forecast horizon with a root mean square error (RMSE) of 13.2% of extraterrestrial radiation received.

Similarly, artificial neural networks (ANNs, Elizondo et al., 1994, Guarnieri et al., 2008) recognize patterns in data by mimicking the logical processes of a biological neural network and have been successfully applied to solar forecasting. Using training data, typically years of measured ground data (Elizondo et al., 1994), ANNs have been developed to reduce overall relative RMSE (rRMSE) of GHI by as much as 15% when compared to 12 to 18-hour ahead NWP forecasts (Guarnieri et al., 2008).

For physically-based forecasting, cloud cover and cloud optical depth are the most important parameters affecting solar irradiance. Through processing of satellite or ground imagery, clouds can be detected, characterized, and advected to predict GHI accurately up to 6 hours in advance (Hamill and Nehr Korn, 1993; Hammer et al. 1999). Hammer et al. (1999) demonstrated 30 minute GHI forecasts accurate to within 17% rRMSE and 30% rRMSE for 2-hour forecast horizons. For intra-day forecasts, a reduction in rRMSE by 7%-10% compared to persistence forecasts was found.

Forecasts beyond 6 hours, up to several days ahead, are generally most accurate if derived from numerical weather prediction (NWP). NWP models predict GHI using columnar (1D) radiative transfer models (RTM, Heinemann et al. 2006). In most existing NWP models solar irradiation is only used to force the average surface energy balance, making GHI variability less relevant. Consequently, computationally expensive RTMs are only run once every 30 to 60 minutes while model time-steps are on the order of one minute. Heinemann et al. (2006) showed that the MM5 mesoscale model can predict GHI in clear skies without mean bias error (MBE). However, the bias was highly dependent on cloud conditions; in overcast conditions, the MM5 model MBE was 129% (Heinemann et al., 2006). Perez et al. (2007) examined the accuracy of the National Digital Forecast Database (NDFD), a derivative of the operational NWP models published by the National Center for Environmental Prediction (NCEP). After a local correction function was applied, Perez et al. (2007) found that for 8 to 26-hour forecast horizons, the NDFD had a GHI rRMSE of 38%. This was more accurate than satellite derived forecasts for similar forecast horizons (rRMSE = 46%).

Due to complex cloud microphysics and limitations in spatial resolution, clouds and their radiative properties are difficult to predict in numerical models. Consequently, NWP models are expected to show inherent regional or global biases limiting forecast accuracy. Armstrong (2000) compared the Dudhia (1989) radiative model forecast GHI to a single ground measurement site and found a relationship between MBE and cloud cover. Similarly, Remund et al. (2008) examined NWP biases

compared to a single site to find that European Centre for Medium-Range Weather Forecasts (ECMWF) and Global Forecast System (GFS) next day GHI forecasts have an MBE of 19%. This MBE was shown to be approximately constant for intra-day (hour-ahead) to 3 day ahead forecast horizons (Remund et al. 2008). Breitzkreuz et al. (2009) included aerosol effects on irradiance into the ECMWF model to show a reduction in GHI rRMSE for 2 to 3 day ahead forecasts from 11.5% to 7.2% under clear sky conditions. Similarly rMBE was improved from -9.8% to 5.1%.

Consistent error patterns allow for MOS to be used to produce a bias reduction function for future forecasts (Heinemann et al., 2006). Bofinger and Heilscher (2006) used MOS locally with ECMWF GHI forecasts to create daily solar electricity predictions accurate to 24.5% RMSE. Lorenz et al. (2009) related forecasted solar zenith angle (SZA) and clear sky index (kt^* , Eq. 2) to ECMWF MBE for Germany, revealing a consistent over prediction (up to 100 W m^{-2}) for moderately cloudy conditions. Using a MOS correction function eliminated bias and reduced RMSE by 5% for 24 h forecasts.

This paper focuses on the analysis and MOS correction of GHI forecasts from three operational NWP models within the continental United States (North American Model, NAM, GFS, and ECMWF; Section 2). A fourth operational NWP model, the Rapid Update Cycle (RUC), was not evaluated. The approach presented here expands upon previous work. Remund et al. (2008) compared six months of ECMWF, NDFD, and GFS-WRF forecasts against only three ground measurement sites. Furthermore, bias reducing MOS corrections were not examined. Additionally, Perez et al. (2007) evaluated the NDFD at a single site. There, however, the MOS approach was based on cloud cover parameters and not irradiance. For this paper, all available operational NWP models (Section 2) are validated against a nationwide network of meteorological stations for up to 1 year. The MOS correction function is established directly from the NWP model output for solar radiation. Consistent biases are shown as a function of measured (Section 4.1) and forecast clear sky index (Section 4.2) motivating the application of a MOS correction as a function of solar zenith angle and clear sky index. Implementation of the MOS correction mitigates bias errors of the NWP forecasts and significantly reduces RMSE (Section 4.3). Finally, it is concluded (Section 5) that MOS corrections lead to more accurate prediction of solar irradiance, given a location, time, and basis forecast model. Only intra-day (or hour-ahead) forecasts are considered since NWP forecast accuracy has been shown to be approximately constant over a 3 day horizon (Remund et al. 2008) and intra-day forecasts are relevant for energy markets.

2. NUMERICAL WEATHER PREDICTION AND GROUND VALIDATION DATA

2.1 North American Mesoscale Model (NAM)

The NAM forecast is published by the National Oceanic and Atmospheric Administration's (NOAA) NCEP on a 0.113° E-W by 0.111° N-S (approximately 12 km x 12 km) grid spanning the entire continental United States (CONUS, Fig.1, Table 1). Hourly output, available up to 36 hours ahead, is published four times daily at 00, 06, 12, and 18 UTC. Additionally, the NAM forecast is available up to 84 hours ahead for a 3-hour temporal resolution. The NAM forecast is a product of the Weather Research and

Forecasting – North American Mesoscale (WRF-NMM) model. Further details on the numerical model and parameterization methods are available in Skamarock (2007).

Among over 125 output variables, total downward short wave radiation at the surface (Fig. 1) is reported. This is equivalent to global horizontal irradiance (GHI) or the total amount of radiation, direct and diffuse, incident on a horizontal surface at the earth’s surface. Downward shortwave radiation is a non-native forecast parameter that is calculated from several radiative transfer models: A derivative of the Eta Geophysical Fluid Dynamics Laboratory shortwave model (GFDL-SW, Lacis and Hansen, 1974); the Dudhia (Dudhia et al., 1989) shortwave model as used by MM5; or the Goddard (Chou and Suarez, 1994) shortwave model (Skamarock et al., 2007). The operational NAM model analyzed here use the GFDL-SW model. Using the principle of multiple scattering and Mie theory, the GFDL-SW model assumes that the primary source of radiative absorption, reflection, and scattering in the atmosphere is from concentrations of large particles such as water droplets in clouds. Cloud optical depth and albedo are parameterized using a separate cloud model and are based on prognostic numerical variables such as water mixing ratios (liquid and ice), temperature, and pressure (Stephens, 1978; Stephens, 1984). Ozone and carbon dioxide concentrations from climatological tables are used. Radiative transmission is subsequently calculated in an assumed plane-parallel atmosphere with homogeneous layers. The GFDL-SW model is strictly one-dimensional, i.e. GHI is only affected by conditions present in the column of atmosphere directly above the grid point (Skamarock et al., 2007).

Table 1: NWP Forecast Summary

	NAM	GFS	ECMWF
Spatial resolution	0.11°	0.5°	0.25°
Radiative Transfer Model	GFDL-SW	RRTM-SW	McRad
GHI variable name	dswrfsfc	dswrfsfc	surface solar radiation downwards
Output temporal resolution/forecast horizon	1 h / 36 h	3 h / 180 h	3 h / 144 h; 6 h / 240 h
Temporal averaging	Instantaneous	3 h, 6 h alt.	3 h, 6 h, 9 h, etc.
Time period analyzed	3/21/10 – 2/8/11	6/23/10 – 2/8/11	9/1/09-8/31/10

2.2 Global Forecast System (GFS)

Unlike the NAM, the GFS is a global forecasting model. Also published by NOAA through NCEP (publicly available at <http://nomads.ncep.noaa.gov/>, NCEP, 2010), the GFS model has a spatial resolution of 0.5° E-W by 0.5° N-S (Table 1). Temporally, the GFS forecast is published at 00, 06, 12, and 18 UTC, at an average time-step of 3 hours up to 180 hours (7.5 days) ahead. GFS forecasts are reported as alternating 3 (1st, 3rd, etc. forecasts in a series) and 6-hour (2nd, 4th, etc.) averages. Consistent 3-hour resolution forecasts are calculated from the raw data.

The SW radiative model employed by the GFS is the rapid radiative transfer model (RRTM-SW, Mlawer and Clough, 1997; Mlawer, 1997). RRTM uses the correlated-k method to transform the spectral

dependence of radiative absorption coefficients (k) into a continuous cumulative distribution function. The transformed space is subsequently discretized and characteristic k values calculated for each sub-interval. Using a two-stream adding method, radiative transfer calculations are performed independently for each sub-interval across all vertical layers. This is similar to the individual spectral-point line-by-line radiative transfer model (LBLRTM, Clough et al., 1992) calculations.

Absorption effects from water vapor, ozone, oxygen, and methane are considered. Additionally, cloud optical depth, albedo, asymmetry factor, and effective particle radius contribute to layer radiative properties. Similar to the Stephens (1984) parameterization, cloud properties are largely dependent on liquid water path (LWP), ice water path (IWP), temperature, pressure, and location (Hou et al., 2002).

2.3 European Centre for Medium Range Weather Forecasts (ECMWF)

Like the GFS, the ECMWF is a global forecast model. Spatially, ECMWF data is available on a 0.25° E-W by 0.25° N-S grid (Table 1). Temporally, ECMWF forecasts have a time-step size of 3 h and are published twice daily (00 UTC and 12 UTC) up to 10 days in advance. Data is published as progressive averages (i.e. 3 h, 6 h, 9 h, etc.) for each series. From this, 3-hour resolution forecasts are calculated. For this study, only intra-day forecasts were considered (24-hour forecast horizon). Downward surface solar radiation data released at 12 UTC between September 1st, 2009 and August 31, 2010 were purchased.

The current operational ECMWF model contains the McRad radiation model (Morcrette et al., 2007). The shortwave portion of McRad is based on RRTM. Constituents accounted for in McRad are water vapor, carbon dioxide, ozone, methane, nitrous oxide, aerosols, and various chlorofluorocarbons (CFCs). Similar to the GFDL-SW model, important forecast cloud properties are albedo and optical depth and are parameterized primarily from LWP and effective droplet radius (Slingo et al. 1989, Table 3 in Morcrette et al., 2007). From the radiative properties, a two stream adding method is used to solve the radiative model at each level, resulting in surface GHI.

2.4 SURFRAD Ground Measurement Network

Seven Surface Radiation Budget Network (SURFRAD) stations across the CONUS provide an accurate nationwide database to validate NWP forecasts (Fig. 1). SURFRAD stations are equipped with Eppley Precision Spectral Pyranometers (PSP) (SURFRAD, 2009a) capable of measuring GHI to within $\pm 2\%$ (Campbell, 1992). The PSPs are calibrated through NOAA's Solar Radiation Facility. One minute GHI is reported as an average of sixty 1-sec instantaneous measurements. Due to pyranometer thermal offset, night time GHI values are generally negative (SURFRAD, 2009b). For this reason, all negative values of GHI were set to zero. Data flagged by the baseline quality control (QC) assessment function due to historically inconsistent, unphysical, and uncharacteristic data (SURFRAD, 2009b) were removed from this analysis. SURFRAD GHI data were aggregated to one hour averages for comparison with the (hourly or longer) output time-step of NWP forecasts.

3. FORECASTING METHODS

3.1 Forecasting Methods

Fig. 2 depicts a sample time series of the forecast methods that will be evaluated. The most basic “forecast”, C , is the Ineichen clear sky model (Ineichen and Perez, 2002). C primarily depends on location, time, and elevation. Additionally, C is dependent on atmospheric turbidity variations due to aerosols, ozone, and water vapor, which are input as monthly averages of the Linke turbidity from the SoDA database on an 8 km grid (Ineichen and Perez, 2002).

The persistence forecasts, P_1 and P_{24} , assume static weather conditions. The measured clear sky index (kt_m) is defined as the ratio of measured GHI to C ,

$$kt_m = \frac{GHI_{MEAS}}{GHI_{CSK}}. \quad (\text{Eq. 1})$$

P_{24} forecasts GHI by assigning the kt_m at solar noon of the previous day to the entire intra-day forecast (kt_p). P_{24} is calculated by multiplying kt_p with the clear sky model, C . The 1 hour persistence forecast, P_1 is calculated by multiplying kt_p (the measured clear sky index averaged over the previous hour) with C . In this way, P_1 is a ‘best-case’ persistence forecast for hour-ahead forecasting, while P_{24} is a typical persistence forecast for day-ahead forecasting.

Forecasts originating shortly before the CONUS sunrise are the timeliest for planning intra-day power plant load-following. Thus, the only NWP forecasts (Sec. 2.1-2.3) considered for this study originate at 1200 UTC. As NWP forecast errors have been shown to be constant over a forecast horizon of 1 hour to 3 days (Perez et al. 2010), the results of this paper can also be considered applicable to day-ahead (i.e. next day) forecasts.

Interpolation of NWP output is necessary to temporally align NWP output and ground measurement data. Since sun position and solar zenith angle are non-linear functions of time, solar irradiance cannot be a linear process. Therefore, linear interpolation of GHI is inappropriate (N , Fig. 2). Instead, clear sky index kt^* interpolation should be used

$$kt^*(t_1) = \frac{N(t_1)}{GHI_{CSK}(t_1)}, \quad (\text{Eq. 2})$$

Where GHI_{CSK} is calculated from C . kt^* is then interpolated to time t_2 and the NAM forecast becomes $N_K = kt^*(t_2) GHI_{CSK}(t_2)$.

Since the clear sky model, C , is nearly linear over the hourly time-step of the NAM output, N and N_K produce essentially identical forecasts (Fig. 2). N_K is the correct forecast and will henceforth be considered the primary forecasting method for the NAM model.

For the GFS and ECMWF, GHI forecasts are *average* irradiances applicable to the entire forecast interval. 3-hour clear sky indices are calculated through Eq. 2 and multiplied by C to produce the hourly G and E forecasts. Table 2 compares each forecasting method for a single day forecast horizon.

Table 2: Comparison of the forecast timesteps over a single day. Boxes represent averaging periods of constant clear sky index kt^* (kt_m for P_1 and P_{24}). N_K is interpolated from hourly instantaneous values and

assumed to apply over the entire hourly interval. The GFS provides a combination of 3 and 6-hour interval forecast, from which 3-hour forecasts can be back calculated (striped box).

UTC hour	12	13	14	15	16	17	18	19	20	21	22	23	24	1	2	3	4	5
NAM																		
GFS																		
ECMWF																		
P ₁																		
P ₂₄																		

3.2 Spatial Averaging

N_k , G , and E are spatially discrete functions for GHI. Girodo (2006) showed that spatial averaging of ECMWF output over $0.5^\circ \times 0.5^\circ$ reduces the rRMSE of NWP forecasts. Here, forecast GHI for each station is calculated by taking the mean output GHI of all NWP grid points within various distances (10 km, 50 km, 100 km, and 200 km). The mean absolute error (MAE) is not biased by large error events and is the best metric for determining the advantage of spatial averaging (Sec. 4.1).

3.3 Error Metrics

Equations 2-5 define error metrics used in NWP forecast evaluation: the mean absolute error (MAE), mean bias error (MBE), relative mean bias error (rMBE), and relative root mean squared error (rRMSE), respectively;

$$MAE = \frac{1}{N} \sum_{i=1}^N |GHI_{NWP,i} - GHI_{MEAS,i}| \quad (\text{Eq. 3})$$

$$MBE = \frac{1}{N} \sum_{i=1}^N (GHI_{NWP,i} - GHI_{MEAS,i}) \quad (\text{Eq. 4})$$

$$rMBE = \frac{1}{N} \sum_{i=1}^N \frac{GHI_{NWP,i} - GHI_{MEAS,i}}{GHI_{MEAS.(SZA,kt^*)}} \quad (\text{Eq. 5})$$

$$rRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(\frac{GHI_{NWP,i} - GHI_{MEAS,i}}{GHI_{MEAS.(SZA,kt^*)}} \right)^2}. \quad (\text{Eq. 6})$$

Relative errors are obtained by normalizing to the mean measured global horizontal irradiance ($\overline{GHI_{MEAS}}$) as a function of solar zenith angle (SZA) and clear sky index. Through the use of this time and measurement dependent normalization, as opposed to constant as in many other validation studies (e.g. Lorenz et al. 2009), mid-day relative errors are reduced through normalization and early morning/late evening errors are amplified. As errors will be considered as a function of kt^* and SZA, this normalization provides a metric relevant to the period over which the error is calculated.

3.4 MOS Correction

Following Lorenz et al. (2009) a stepwise multivariate fourth-order regression (Rogers, 2007) is applied to derive the MOS correction function.

$$GHI_c(SZA, kt^*) = \alpha \cos^4(SZA) + \beta(kt^*)^4 + \dots \quad (\text{Eq. 7})$$

GHI_c is the model MBE for a given SZA and clear sky index. Weighted regression coefficient calculation ensures accurate representation of unevenly distributed data (see Fig. 5). For all combinations of SZA and kt^* , GHI_c provides the best irradiance correction for any future forecast. As an example, the MOS-corrected NAM forecast, $N_{K,c}$, is computed as

$$N_{K,c} = N_K - GHI_c. \quad (\text{Eq. 8})$$

To simulate operational forecasting, local MOS correction functions are calculated dynamically using only the most recent 56 days of MBE analysis (rather than the entire dataset) separately for each SURFRAD station. This provides a training dataset, independent of the evaluation dataset, and a unique MOS correction for each $N_{K,c}$. This method is similarly applied for G and E and designed to eliminate systematic bias error. Since MOS corrections are not applicable for SZAs larger than 75° , evening and early morning corrected forecasts are not calculated.

4. RESULTS

4.1 Numerical Forecast Evaluation

4.1.1 Effect of Spatial Averaging – Mean Absolute Errors

Spatial averaging of the N_K (Fig. 3) and G (not shown) forecasts results in a reduction in MAE. As the scale of spatial averaging increases the MAE across the majority of measured clear sky indices decreases. Only for small measured clear sky indices ($kt_m < 0.25$) does MAE increase with increased spatial averaging, especially for the 200 km spatial averaging radius. Girodo (2006) and Lorenz et al. (2009) also found reduced errors for averaging over a 100 km x 100 km square grid for the ECMWF forecast. Henceforth, only the 100 km radius spatial average will be considered for the analyses.

4.1.2 Mean Bias Errors of N_K , G and E forecasts

Fig. 4 reveals similar MBE profiles for each NWP model. Given measured clear conditions and SZAs near solar noon, all NWP forecasts are negatively biased by between 0 and 50 W m^{-2} (Fig. 4), indicating the presence of cloudy forecasts that resulted in clear periods. These results may be biased by location and season, since northern stations experience a $\cos(SZA) > 0.9$ only for a short part of the year.

In clear skies, MBE increases with SZA reaching 50 W m^{-2} for $\cos(SZA) = 0.35$ for N_K and approximately 0 W m^{-2} for G and E . At $\cos(SZA) < 0.4$, many N_K forecasts predict up to 40% more GHI than clear sky

models ($kt^* = 1$ to 1.4, see Fig. 5a). This over-prediction of GHI in measured clear conditions is surprising as it potentially indicates an inaccurate NWP clear sky model.

For measured cloudy conditions ($kt_m < 0.8$) and relevant SZA ($SZA < 75^\circ$), N_k , G , and E are positively biased by up to 225 W m^{-2} . Positive MBEs in measured cloudy conditions ($kt_m < 0.8$) suggest that either cloudy conditions were not predicted or that the prescribed cloud optical depths or cloud cover fractions were too small.

Some of these patterns can be explained by the binary nature (either cloudy or clear) of solar forecasts. Examination of the data distribution of predicted and measured conditions (Table 3 with an overcast threshold of $kt = 0.8$) shows that 20.6% of all N_k forecasts are measured cloudy days which were predicted to be clear. False clear forecasts occurred in 19.7% of G forecasts and 12.4% of E forecasts. Forecast cloudy, but measured clear conditions were only observed for 5.7% (N_k), 5.5% (G), and 6.9% (E) of forecasts. These results are indicative that the NWP models are generally biased towards predicting clear conditions and at a larger clear sky index than measured. Overall, the ECMWF model had the smallest likelihood of incorrect forecasts (19.3%) while both the NAM and GFS were incorrect for 25% of forecasts

Table 3: Normalized frequency of clear and cloudy forecasts and observations for each NWP model. The threshold to distinguish between cloudy and clear conditions is a clear sky index of 0.8.

SURFRAD	N_k		G		E	
	Cloudy	Clear	Cloudy	Clear	Cloudy	Clear
Cloudy	0.187	0.206	0.170	0.197	0.282	0.124
Clear	0.057	0.550	0.055	0.579	0.069	0.525

4.2 MOS Correction of Forecast Total Mean Bias Errors

4.2.1 N_k Forecast

Fig. 5a is the histogram for the NAM forecast between March 21st, 2010 and February 8th, 2011, as a function of SZA and kt^* . Consistent with Table 3, more data exist for predicted clear skies than for cloudy conditions. Fig. 5a indicates that the NAM model often forecasts GHI 20% to 40% in excess of clear sky irradiance, especially for large SZAs. While low atmospheric turbidity during clear sky conditions can result in GHI measurements slightly larger than C , the frequency and magnitude of forecasts with $kt^* > 1$ indicates a fundamental problem with either the NAM radiative transfer model or the primary inputs to the model from the NAM forecast. G forecasts only exhibit $kt^* > 1$ for SZAs greater than 75° , which could be caused by a different choice of clear sky model. $kt^* > 1.1$ are not observed for E forecasts.

The MBE profiles in Fig. 4 demonstrate consistent MBE patterns. However, kt_m is unknown prior to measurement and is unsuitable as an input parameter for MOS correction. Instead MBE profiles as a function of the predicted clear sky index, kt^* must be used. Fig. 6a shows the MBE of the N_k forecast as

a function of $\cos(\text{SZA})$ and kt^* . For forecast clear sky conditions ($kt^* > 0.8$), \mathbf{N}_k over-predicts measured data by up to 150 W m^{-2} and on average by 73.9 W m^{-2} . In cloudy conditions ($kt^* < 0.5$), the \mathbf{N}_k model under-predicts measured GHI, on average by -32.3 W m^{-2} . Overall, the magnitude of the negative MBE becomes larger for small SZA, while the positive MBE for clear sky conditions becomes slightly smaller for small SZA.

While \mathbf{N}_k is an overall inaccurate GHI forecast, Fig. 6a shows that NAM bias trends are continuous and therefore, correctable. Using MOS (Sec. 3.4), a correction polynomial was fit to the data (e.g. Fig. 6b, coefficients given in Appendix A). Fig. 6b well represents the typical MBE characteristics of \mathbf{N}_k . Corrections are subsequently applied according to Eq. 8, yielding $\mathbf{N}_{k,c}$. Note that as temporally dynamic local MOS corrections (Sec. 3.4) were applied to simulate operational conditions, Fig. 6b was not directly used in forecast correction. The MBE of $\mathbf{N}_{k,c}$ (Fig. 6c) decreases to less than $\pm 50 \text{ W m}^{-2}$ and overall from 57.5 W m^{-2} to 7.0 W m^{-2} . For a perfect MOS correction, a zero MBE would be expected for $\mathbf{N}_{k,c}$. Here, the MBE is not zero is due to the polynomial fit MOS correction which cannot fully resolve the variability of Fig. 6a.

4.2.2 \mathbf{G} Forecast

Fig. 7a shows the average MBE profile for the GFS forecast, \mathbf{G} . The bias error dependence on kt^* is qualitatively similar to the NAM model (Fig. 6a). For forecasted clear sky conditions ($kt^* > 0.8$), \mathbf{G} generally over-predicts GHI, on average by 41.0 W m^{-2} . Similarly, for overcast conditions with moderate to small optical depths ($kt^* < 0.5$), the GFS forecast is negatively biased by on average -1.8 W m^{-2} . The MBE shows a much stronger dependence on SZA compared to the NAM model, increasing during mid-day ($\text{MBE}_{\text{AVG}} = -67.4 \text{ W m}^{-2}$ for $\cos(\text{SZA}) > 0.7$ and $kt^* < 0.5$) as a result of the larger GHI magnitudes observed then.

Fig. 7b shows the MOS correction function fit to Fig. 7a for the entire data set. From Eq. 7, the MOS function applies a negative correction to \mathbf{G} for predicted clear skies ($kt^* > 0.8$). Correspondingly, a large positive correction is applied for forecast overcast conditions ($kt^* < 0.5$). On average, the MBE of the MOS-corrected forecast, \mathbf{G}_c (Fig. 7c) is reduced from 35.4 W m^{-2} to 5.2 W m^{-2} .

4.2.3 \mathbf{E} Forecast

The MBE profile of \mathbf{E} (Fig. 8a) is significantly different than \mathbf{N}_k or \mathbf{G} . Predicted clear skies ($kt^* > 0.9$) and overcast conditions with $kt^* < 0.2$ are least biased ($\text{MBE} < 50 \text{ W m}^{-2}$). For predicted moderately cloudy conditions \mathbf{E} is most significantly biased, by 42.6 W m^{-2} on average. While there are several areas of negative bias, the MBE is in general positive, with some biases exceeding 150 W m^{-2} . Lorenz et al. (2009) found a similar MBE profile for ECMWF data using ground measurements in Germany.

The MOS correction function for \mathbf{E} (Fig. 8b) targets the large positive bias for moderately cloudy conditions. It should be noted that the negative bias correction for very small predicted clear sky indices ($kt^* < 0.2$) is an artifact of the correction polynomial, which is irrelevant as no such data occurred. This was also observed in the MOS function for \mathbf{G} (Fig. 7b). The MOS correction reduces overall MBE from

31.4 W m⁻² to 0.5 W m⁻² (Fig. 8c). Using this metric, the MOS correction for **E** is the most effective in reducing systematic errors.

4.3 Root Mean Square Errors

For $kt_m = 1$, **C** is expected to have an RMSE = 0 W m⁻². For $kt_m > 0.86$, **C** indeed has the smallest RMSE. However, several instances of $kt_m > 1$ were observed, contributing to the non-zero RMSE for $kt_m = 1$. For overcast conditions, the RMSE of **C** drastically increases.

P₁ is the most accurate forecast for overcast conditions and only slightly less accurate than **G** and **E** for clear conditions. Since hourly cloud cover conditions are relatively static the one-hour persistence forecast, **P**₁, is accurate. For longer persistence forecasting on the order of a day (**P**₂₄), clear conditions are more static than cloudy. Correspondingly **P**₂₄ RMSE increases with decreasing kt_m , approaching a RMSE greater than 400 W m⁻² for $kt_m < 0.15$. Transitional periods between synoptic weather patterns and diurnally driven phenomena cause large errors in 24-hour persistence forecasts. For moderate cloudy conditions, the measured clear sky index can vary significantly, making the persistence forecasts inaccurate (e.g. Fig. 2 from 1800-2400 UTC).

The high spatial variability of irradiance during cloudy conditions is difficult to forecast by NWP. Consequently, NWP RMSE (Fig. 9a) increases as kt_m decreases. Under measured clear sky conditions, all NWP forecasts have low RMSE (RMSE < 125 W m⁻²). In clear conditions ($kt_m = 1$), **G** is the most accurate. Excluding clear conditions, **E** has a lower RMSE than other NWP forecasts with typical improvements of 30 W m⁻² compared to **N**_K and **G**.

After MOS correction (Fig. 9b) the same general trends for NWP RMSE are present, albeit at a smaller RMSE level. The impact of MOS on RMSE is shown in Fig. 9c. Positive changes represent increases in RMSE. The RMSE of the NAM forecast **N**_{K,c} is reduced by up to 50 W m⁻² for $0.15 < kt_m < 0.95$. Similarly, **G**_C is improved over all clear sky indices less than 0.9 by as much as 110 W m⁻². The RMSE of **E** is also significantly lowered for all clear sky indices less than 0.7. For the largest clear sky indices ($kt_m > 0.94$), MOS corrections increased the RMSEs for each NWP forecast. Uncorrected forecasts for measured clear conditions (Fig. 4) were relatively unbiased making it difficult to reduce RMSE. For all other clear sky indices, the MOS correction function significantly reduced RMSE.

5. SUMMARY AND CONCLUSIONS

In this study, five primary intra-day forecasts were analyzed: The Ineichen clear sky model, **C**; the persistence forecasts, **P**₁ and **P**₂₄; the clear sky index (kt^*) interpolated NAM forecast, **N**_K; the 3-hour constant kt^* GFS forecast, **G**; and the 3-hour constant kt^* ECMWF forecast, **E**. It is of interest to compare bias-corrected NWP model forecasts to other more advanced products such as NDFD, satellites, sky imagers, or forecasts by specialized renewable energy forecast providers. While other derived forecast products may be more accurate, the accuracy of bias-corrected NWP forecasts provides a useful reference to establish the value of other (more expensive) products.

Forecast accuracy was evaluated through RMSE and MBE. 24-hour persistence and clear sky forecasts (P_{24} , and C) were inaccurate except for measured clear conditions (Fig. 9) and should be used only as baseline forecasts. The accuracy of 1-hour persistence forecasts is comparable to NWP models for clear conditions, but better for $kt_m < 0.5$.

The method presented by Lorenz et al. (2009) and known as model output statistics (MOS) was employed to reduce MBE in forecasted GHI. To determine forecast accuracy, it was necessary to differentiate between MBE as a function of measured clear sky index (kt_m) and forecasted clear sky index (kt^*). NWP models were validated using kt_m , while MOS necessitated the use of the forecast variable, kt^* . Using fourth order multivariate regression, correction functions were fit to the observed MBE as a function of solar zenith angle (SZA) and clear sky index (kt^*). Table 4 summarizes the results of the MOS correction when averaged across all SURFRAD stations. Errors separated by station are provided in Table 5 of Appendix B and are consistent with the overall errors listed in Table 4. The local MOS functions, used in conjunction with forecast SZA and kt^* , reduced RMSE by 20.1 W m^{-2} (N_k), 17 W m^{-2} (E), and 25.6 W m^{-2} (G) and effectively eliminated MBE. RMSE reductions were largest for intermediate kt^* at about 50 W m^{-2} . MBEs were similar across different sites with the smallest errors at Desert Rock, NV (site with the largest kt_m) and the largest errors at Pennsylvania State University, PA (site with the lowest kt_m). MOS-corrected RMSE were smallest at the most clear site, Desert Rock, and largest at Fort Peck, Montana. For all stations, seasonal changes in bias errors limit MOS-correction accuracy. For several SURFRAD stations, MBE were observed to change significantly from month to month. As the MOS-correction uses 56 days of training data, sudden changes in MBE dependence on SZA and kt^* are not immediately represented increasing the MBE of the MOS-corrected forecasts.

For the original models, ECMWF performed best, closely followed by GFS. Surprisingly, the NAM model performed the worse despite its higher spatial resolution and specialized application to North American Meteorology. Significant problems were found with the clear sky model applied in N_k which over-predicted GHI in clear conditions by up to 40%. After MOS correction, the GFS model was most accurate, indicating that the errors in GFS forecasts are more systematic. A combination of GFS forecasts for clear conditions and ECMWF conditions for cloudy conditions may yield the lowest overall RMSE (Fig. 9b).

Table 4: MBE and RMSE for NWP forecasts before and after MOS correction averaged across all sites. Table 5 shows the errors separated by site.

	N_k	G	E
MBE (W m^{-2})	57.5	35.4	31.4
MBE _C (W m^{-2})	7.0	5.2	0.5
RMSE (W m^{-2})	134.2	110.5	123.2
RMSE _C (W m^{-2})	114.1	84.6	106.2

The NWP models were shown to be significantly biased towards predicting clear skies. Thus, when plotted against kt_m (Fig. 4), negative biases are rare. Overall, N_k , G , and E predicted more false clear days than false cloudy days (Table 3). Of the forecasts that erroneously predicted clear or cloudy conditions (Table 3), 78.3% of N_k (75.6% for G and 64.2% of E) were clear forecasts resulting in cloudy days. This was a primary contributor to the large positive MBE observed in forecast clear conditions (Figs. 6a, 7a, and 8a). However, the many accurate clear sky forecasts (55.0% of all N_k , 57.9% of G , and 52.5% of E) reduced the magnitude of the MBE in clear conditions. Although false cloudy forecast were considerably less frequent (5.7% of all N_k , 5.5% of G and 6.9% of E), far fewer cloudy days were predicted overall (Fig. 5a). These few large negative bias events account for part of the 150 W m^{-2} MBE for $kt^* < 0.6$ for N_k and G (Fig. 6a/7a).

Evaluating ECMWF forecast accuracy in Germany, Lorenz et al. (2009) showed that NWP MBE was largest for cloudy conditions with moderate clear sky indices ($0.3 < kt^* < 0.6$), while forecasted clear conditions were relatively unbiased. This study revealed a similar ECMWF MBE profile for SURFRAD stations (Fig. 8a). Positively biased for moderately overcast conditions, the overall MBE of E was 31.4 W m^{-2} . Consistent with the results of this study, Remund et al. (2008), found E MBEs ranging from 15 to 43 W m^{-2} at 3 different sites.

MOS application to the NWP irradiance output was successful in minimizing bias and reducing RMSE, but did not provide information as to the source of the MBE. Errors originating from the radiative transfer model (RTM), cloud model, and prognostic variable errors can all contribute to the MBE. MOS corrections in the measured clear sky regime ($kt_m > 0.9$) did not reduce RMSE (Fig. 9c). This is because the MOS could not distinguish between RTM errors (over-prediction of GHI even for clear skies, especially for NAM) and cloud model errors (incorrect parameterization of RTM inputs). Consequently, many initially accurate forecasts were unnecessarily corrected. Rather than correcting all forecasts, differentiating between the sources of the error is the first step into understanding which forecasts need to be corrected. Further research will investigate the relationship of MBE to other prognostic variables, most notably those involved in cloud parameterization.

A final consideration of NWP forecasting is resolution. Even the $0.1^\circ \times 0.1^\circ$ NAM spatial resolution is insufficient to resolve most clouds and only an average cloud cover can be forecasted for a given point. For the GFS and ECMWF the resolution is even coarser. However, even if the spatial resolution was finer, the temporal output intervals would not permit the examination of time dependent cloud cover variability, important in predicting ramp rates and bands of variability for solar power plants. While NWP model time-steps are on the order of minutes, the radiative transfer models are run less frequently, and the output is only hourly (NAM) or every 3 hours (GFS and ECMWF). Consequently, any patterns with characteristic time scales less than an hour are unresolved. Linking observed temporal variability in GHI to native NWP forecasts will require further research.

Acknowledgements

We acknowledge funding from California Energy Commission under BOA-248. Masao Kanamitsu (UC San Diego) provided helpful advice on NWP models. Jordan Alpert and Fanglin Yang (NCEP) aided in the interpretation of the GFS forecast. Helpful discussions with Manajit Sengupta (NREL) are acknowledged.

References

- Armstrong, M. 2000. Comparison of MM5 forecast shortwave radiation with data obtained from the atmospheric radiation measurement program. Master of Science Scholarly Paper, University of Maryland, USA.
- Bofinger, S., and Heilscher, G. Solar Electricity Forecast – Approaches and first results. 21st PV Conference. Dresden, Germany, 2006.
- Breitkreuz, H., Schroedter-Homscheidt, M., Holzer-Popp, T., and Dech, S. 2009. Short-range direct and diffuse irradiance forecasts for solar energy applications based on aerosol chemical transport and numerical weather modeling. *J. Applied Meteorology and Climatology*. **48:9**, 1766-1779.
- Campbell Scientific, 1992. Eppley PSP – precision spectral pyranometer instruction manual. *Campbell Scientific Instruction Manual*.
- Chou, M., and Suarez, M., 1994. An efficient thermal infrared radiation parameterization for use in general circulation models. *NASA Tech. Memo*. NASA/TM-1994-104606. 1-85.
- Clough, S., Iacono, M., Moncet, J., 1992. Line-by-line calculations of atmospheric fluxes and cooling rates: Application to water vapor. *J. Geophysical Research*. **97:14**, 15761-15785.
- Dudhia, J., 1989. Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *J. Atmospheric Sciences*. **46:20**, 3077-3107.
- Elizondo, D., Hoogenboom, G., and McClendon, R. 1994. Development of a neural network model to predict daily solar radiation. *Agriculture and Forest Meteorology*. **71:1-2**, 115-132.
- GFS. 2003. The GFS Atmospheric Model. <http://www.emc.ncep.noaa.gov/gmb/moorthi/gam.html>. Accessed 8/23/2010.
- Girodo, M. 2006. Solarstrahlungsvorhersage auf der Basis numerischer Wettermodelle. Ph.D. dissertation, Oldenburg Univ., Germany.
- Guarnieri, R., Martins, F., Oereira, E., and Chuo, S. 2008. Solar radiation forecasting using artificial neural networks. National Institute for Space Research, **1**, 1-34.
- Hamill, T. and Nehr Korn, T. 1993. A short-term cloud forecast scheme using cross correlations. *Weather and Forecasting*, **8:4**, 401-411.

- Hammer, A., Heinemann, D., Lorenz, E., and Lückehe, B. 1999. Short-term forecasting of solar radiation: A statistical approach using satellite data. *Solar Energy* **67:1-3**, 139-150.
- Heinemann, D., Lorenz, E., and Girodo, M. 2006. Forecasting of solar radiation. *Solar Energy Management for Electricity Generation*, E.D. Dunlop, L. Wald, and M. Sári (Eds.). Nova Publishers, New York. Chapter 7, p. 83-94.
- Hess, M., Koepke, P., Schult, I. 1998. Optical Properties of aerosols and clouds: the software package OPAC. *Bulletin American Meteorological Society*. **79**, 831-844.
- Hou, Y., Moorthi, S., and Campana, K. 2002. Parameterization of Solar radiation transfer in the NCEP models. *NCEP Office Note* 441, 1-34.
- Ineichen, P., and Perez, R. 2002. A new airmass independent formulation for the Linke turbidity coefficient. *Solar Energy*. **73:3**, 151-157.
- Janjic, Z., Gall, R., and Pyle, M. 2010. Scientific documentation for the NMM solver. NCAR Technical Note NCAR/TN-477+STR.
- Jensenius, J. and Cotton, G. 1981. The development and testing of automated solar energy forecasts based on the model output statistics (MOS) technique. *Satellites and Forecasting of Solar Radiation: Proceedings of the First Workshop on Terrestrial Solar Resource Forecasting and on Use of Satellites for Terrestrial Solar Resource Assessment*. 22-29
- Kasten, F. and Young, A. 1989. Revised optical air mass tables and approximation formula. *Applied Optics*, **28**, 4735-4738.
- Lacis, A., and Hansen, J. 1974. A parameterization for the absorption of solar radiation in the earth's atmosphere. *J. Atmospheric Sciences*, **31:1**, 118-133.
- Lew, D., and Piwko, R. 2010. Western wind and solar integration study. NREL technical report, NREL/SR-550-47434. Golden, Colorado: National Renewable Energy Laboratory. 536 pgs.
- Lorenz, E., Hurka, J., Heinemann, D., and Beyer, H. 2009. Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **2:1**, 2-10.
- Meek D. and Hatfield, J. 1994. Data quality checking for single station meteorological databases. *Agriculture and Forest Meteorology*, **69:1-2**, 85-109.
- Mlawer, E. and Clough, S. 1997. Shortwave and longwave enhancements in the Rapid Radiative Transfer Model. *Proceedings of the 7th Atmospheric Radiation Measurement (ARM) Science Team Meeting*. U.S. Department of Energy, CONF-9603149.

Mlawer, E., Taubman, S., Brown, P., Iacono, M., and Clough, S. 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophysical Research*. **102:14**, 16663-16682.

Morcrette, J., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Doblus-Reyes, F., Hague, J., Hamrud, M., Haseler, J., Kaiser, J., Leutbecher, M., Mozdzyński, G., Razinger, M., Salmond, D., Serrar S., Suttie, M., Tompkins, A., Untch, A., and Weisheimer, A. 2007. Recent advances in radiation transfer parameterizations. *ECMWF Tech. Memo*. 539, 1-52.

NCEP, 2010. NOAA Operational Model Archive and Distribution System. <http://nomads.ncep.noaa.gov/>. Accessed 8/2/2010.

Perez, R., Ineichen, P., Moore, K., Kmiecik, M., Chain, C., George, R., and Vignola, F. 2002. A new operational model for satellite-derived irradiances: Description and validation. *Solar Energy*. **73:5**, 307-317.

Perez, R., Moore, K., Wilcox, S., Renné, D., Zelenka, A. 2007. Forecasting solar radiation – preliminary evaluation of an approach based upon the national forecast database. *Solar Energy*. **81:6**, 809-812.

Perez, R., Kivalov, S., Schlemmer, J., Hemker, K., Renné, D., and Hoff, T. 2010. Validation of short and medium term operational solar radiation forecasts in the US. *Solar Energy*. **84:12**, 2161-2172.

Remund, J., Perez, R., and Lorenz, E. 2008. Comparison of solar radiation forecasts for the USA. 2008 European PV Conference, Valencia, Spain.

Rogers, S. 2007. 2D weighted polynomial fitting and evaluation. *Matlab Central File Exchange*. <http://www.mathworks.com/matlabcentral/fileexchange/13719-2d-weighted-polynomial-fitting-and-evaluation>. Accessed 8/25/2010.

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Wang, W., and Powers, J. 2007. A description of the advanced research WRF version 2. NCAR Technical Note NCAR/TN-468 +STR.

Slingo, A. 1989. A GCM Parameterization for the shortwave radiative properties of water clouds. American Meteorological Society. **46:10**, 1419-1427.

Stephens, G., 1978 Radiation Profiles in extended water clouds. II: parameterization schemes.. *J. Atmospheric Sciences*. **35:11**, 2123-2132.

Stephens, G. The parameterization of radiation for numerical weather prediction and climate models. 1984. *Monthly Weather Review*. **112**, 826-867.

SURFRAD. 2009a. SURFRAD Network Overview. <http://www.srrb.noaa.gov/surfrad/surfradpage0.html>. Accessed 8/24/2010.

SURFRAD. 2009b. SURFRAD Data README. <ftp://ftp.srrb.noaa.gov/pub/data/surfrad/README>. Accessed 8/24/2010.

Appendix A: MOS Polynomial Values

The correction polynomial (Eq. 7) can be calculated from the coefficients in this table by summing over $\cos(\text{SZA})^{(\text{COLUMN1})} \cdot (kt^*)^{(\text{COLUMN2})} \cdot (\text{COEFFICIENT OF FORECAST})$

POWER(cos(SZA))	POWER(kt*)	N_K	G	E
0	0	-155.3	-36.7	-124.3
1	0	927.9	-13.3	-545.5
0	1	485.7	77.2	1047
2	0	-1924	2242	-2448
1	1	-1563	-2453	-684.2
0	2	-662.8	472.9	-2598
3	0	2111	-4199	-2703
2	1	-425.5	1313	-2442
1	2	3324	3802	470
0	3	217.1	-962.9	2492
4	0	-1108	561.9	889.2
3	1	748.8	3483	999.8
2	2	-447.9	-3433	917.5
1	3	-1531	-1180	-698.9
0	4	41.4	438.2	-816.3

Appendix B: MBE and RMSE by SURFRAD Station

MBE and RMSE before and after MOS correction. Average measured GHI and kt_m are also shown.

STN	NWP		MBE (W m ⁻²)	MBE _c (W m ⁻²)	RMSE (W m ⁻²)	RMSE _c (W m ⁻²)
	Model					
Bondville, IL	N_K		59.1	4.6	137.7	117.0
GHI _{avg} = 356.8 W m ⁻²	G		39.4	3.3	108.6	82.8
$kt_{m,avg}$ = 0.73	E		23.5	0.7	113	100.9
Desert Rock, NV	N_K		41.0	10.6	93.5	81.6
GHI _{avg} = 474.7 W m ⁻²	G		18.2	3.2	83.2	76.9
$kt_{m,avg}$ = 0.87	E		22.3	1.4	96.4	87.8
Fort Peck, MT	N_K		53.1	11.8	124.1	104.5
GHI _{avg} = 319.9 W m ⁻²	G		32.4	10.7	98.8	74.7
$kt_{m,avg}$ = 0.73	E		26.8	4.1	121.4	107.9
Goodwin Creek, MS	N_K		52.9	1.2	144.0	125.4
GHI _{avg} = 391.3 W m ⁻²	G		36.4	7.1	122.7	91.3
$kt_{m,avg}$ = 0.76	E		10.8	-0.9	127.9	109.8
Penn State University, PA	N_K		77.5	6.5	145.8	118.5
GHI _{avg} = 314.1 W m ⁻²	G		49.2	0.4	116.2	83.7

$kt_{m,avg} = 0.64$	<i>E</i>	42	-1.1	117.4	97.7
Sioux Falls, SD	<i>N_K</i>	63.2	7.8	137.1	166.3
$GHI_{avg} = 314.8 \text{ W m}^{-2}$	<i>G</i>	36.0	6.3	107.7	80.1
$kt_{m,avg} = 0.75$	<i>E</i>	30.3	0.1	120.4	106.6
Boulder, CO	<i>N_K</i>	55.1	6.5	149.0	128.6
$GHI_{avg} = 378.9 \text{ W m}^{-2}$	<i>G</i>	35.8	5.9	129.4	99.3
$kt_{m,avg} = 0.75$	<i>E</i>	62.6	0	157.6	128.2

Figure Captions

Figure 1: NAM GHI Forecast [$W m^{-2}$] for April 10th, 2010 at 2000 UTC for the NAM domain over North America. Triangles show SURFRAD stations.

Figure 2: GHI forecasts compared against Desert Rock, NV SURFRAD ground measurements on July, 17th 2010. **N**: linear interpolated NAM; **N_K**: clear sky index interpolated NAM; **G**: GFS three hour constant clear sky index; **E**: ECMWF three hour constant clear sky index; **P**: persistence; **C**: clear sky forecast.

Figure 3: **N_K** MAE ($W m^{-2}$) as a function of measured clear sky index (kt_m) for different averaging radii of the NAM model. 50, 100, and 200 km averaging radii correspond to approximately 80, 300, and 1250 grid points, respectively.

Figure 4: NWP forecast MBE ($W m^{-2}$) as a function of SZA and measured clear sky index for all SURFRAD stations for a) **N_K**; b) **G**; c) **E**.

Figure 5: Histogram as a function of solar zenith angle (SZA) and clear sky index (kt^*) for a) **N_K**; b) **G**; c) **E**. Data is presented in \log_{10} format, i.e. a value of 2 indicates 100 observations. All data to the left of the black line ($SZA > 75^\circ$) were not used in the MOS analysis.

Figs. 6, 7, 8: MBE ($W m^{-2}$) of **N_K** (Fig. 6), **G** (Fig. 7), and **E** (Fig. 8), as a function of solar zenith angle (SZA) and forecasted clear sky index (kt^*) compared to SURFRAD measurements: a) raw data, b) polynomial fit to the data (Eq. 7); c) MBE of the corrected forecasts: **N_{K,c}**, **G_c**, and **E_c**.

Figure 9: RMSE ($W m^{-2}$) as a function of measured clear sky index (kt_m): a) (**N_K**, **G**, **E**, **P₁**, **P₂₄**, and **C**); b) (**N_{K,c}**, **G_c**, **E_c**, **P₁**, **P₂₄**, and **C**); c) Change in RMSE due to MOS correction.

Suggested Placing of Figures

Figure 1: p. 4, immediately prior to Table 1

Figure 2: p. 6, immediately following the first paragraph (line 7)

Figure 3: p. 8, line 19, immediately following heading of 4.1.1

Figure 4: p. 8, line 27, immediately following heading of 4.1.2

Figure 5: p. 9, line 29, following the first paragraph of 4.2.1

Figure 6, 7, 8: p. 10, line 6, following the second paragraph of 4.2.1

Figure 9: p. 11, line 4, following heading of 4.3