P2.7 INTEGRATION OF ASOS WEATHER DATA INTO MODEL-DERIVED SOLAR RADIATION

Brian N. Belcher* and Arthur T. DeGaetano
Northeast Regional Climate Center, Cornell University, Ithaca, NY

1. INTRODUCTION

Due to the scarcity of solar radiation measurements, model estimation of solar radiation is necessary for a variety of engineering applications. These models commonly incorporate different combinations of hourly meteorological surface observations such as diurnal temperature range, precipitation, relative humidity, wind speed or cloud cover to estimate solar radiation.

Over the past decade, observation methods and instrumentation used to report hourly weather conditions have changed at most airports. Automated systems such as the National Weather Service's Automated Surface Observation Systems (ASOS) have provided gradual replacement of human observers since 1992. The implementation of ASOS introduced spatial limitations on cloud detection both vertically and horizontally since ceilometers are unable to detect cloud cover above 3840 meters (12,600 feet) or near the horizon. Along with these spatial limitations, differences between human perception and ASOS instrument sensitivity also result in biases when reporting the amount and height of cloud coverage or the presence of surface weather phenomena such as fog or haze (and the resulting horizontal visibility). These differences contribute to biases between observation methods. For example, analysis of 10 locations across the U.S. reveals that ASOS observations report clear conditions nearly twice as frequently as human observers (Fig. 1). These differences in cloud observation method also result in a higher frequency of manual cloud cover reports for most sky conditions (including reports with multiple cloud layers), with the exception of single layer overcast skies.

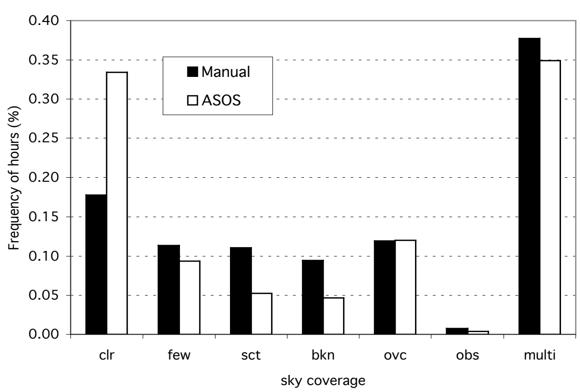

Fig. 1. Frequency of reported sky coverage. Frequencies for 'few', 'sct', 'bkn' and 'ovc' only include hours in which a single cloud layer is reported; 'multi' refers to reports of more than one cloud layer. Frequency of clear ('clr') and obscured ('obs') reports are also given.

* *Corresponding author address:* Brian N. Belcher, Northeast Regional Climate Center, 1123 Bradfield Hall, Cornell University, Ithaca, NY 14853-1901; email: bnb2@cornell.edu

These biases must be compensated for when ASOS data are used as input for solar radiation models. A new model has been developed to estimate solar radiation using ASOS observations as model input. The semi-physical model developed by DeGaetano et al. (1993), which uses human observations as input to estimate global solar radiation for the Northeast U.S., was used as a base model. The new model can be applied to different climate regimes and allows for estimation of hourly global, direct and diffuse solar radiation by using either human or ASOS observations as model input.

2. DATA

Hourly solar radiation data necessary for model development and evaluation were obtained from a variety of different networks across the United States. Some of these networks are regional in scope, while others provide data on a national level. Together, these networks of stations provide hundreds of locations at which solar radiation is measured across the United States.

Hourly surface observations were obtained from stations using ASOS. Meteorological variables necessary for input into the solar radiation model include surface pressure, dew point temperature, cloud layer coverage, cloud layer heights, ceiling heights, visibility and present weather (e.g. fog). These reports are available on an hourly basis at 883 stations in the United States. An additional version of this model takes advantage of satellite estimates of cloud layers with heights greater than 3840 meters (12,600 feet) in order to augment ASOS observations that lack cloud data above this level. These data are derived by using a carbon dioxide slicing (CDS) technique (Menzel et al. 1998) on data from GOES. This technique provided estimates of cloud heights, cloud opacity and total sky cover to supplement cloud observations at each ASOS site used in model development.

Hourly solar radiation and ASOS observations were seldom coincident at one location, however model development required concurrent observations in order to establish empirical relationships between surface hourly reports and solar radiation observations. Stations were therefore used if they were located within 30 km of each other and had at least 3 years of concurrent data available. These requirements resulted in 53 global solar radiation-ASOS station pairs across the U.S. Due to data availability, less than half of these station pairs (22) were used for direct and diffuse solar radiation model development.

3. MODEL DEVELOPMENT

The model of DeGaetano et al. (1993) serves as the basis of the new model that has been developed for the use with ASOS data. This base model, developed at the Northeast Regional Climate Center, is subsequently referred to as the 'NRCC model'.

3.1 Background of the original NRCC model

The semi-physical solar radiation model developed by Meyers and Dale (1983), the basis for the original NRCC model (DeGaetano et al. 1993), expressed the flux density of global solar radiation on a horizontal surface at a given location, I, in the form:

$$I = I_0(\cos Z)T_R T_g T_w T_A T_C \qquad (1)$$

where I_0 is the solar constant; Z is the solar zenith angle; and T_R, T_g, T_w, T_A and T_C are transmission coefficients for Rayleigh scattering, permanent gas absorption, absorption by water vapor, and absorption and scattering by the aerosol and clouds, respectively. Methods of determining T_R, T_g, and T_w are identical to those outlined in Meyers and Dale (1983) and DeGaetano et al. (1993). Since determination of these transmissivities are unaltered during model development, readers are referred to these references for technical details.

Absorption and scattering due to aerosols is computed based on the optical air mass using the equation given by Houghton (1954):

$$T_A = x^m. \qquad (2)$$

The constant '*x*' in equation 2 is determined empirically using observed hourly solar radiation data. The value of x = 0.85 used for the northeastern United States by the original NRCC model is substantially lower than the value of x = 0.935 suggested by Meyers and Dale, that was based on U.S. stations outside of the Northeast.

The original NRCC model also treats cloud transmissivity empirically using the relationship

$$T_c = \prod_{i=1}^{n} \left[1 - c_i(1 - t_i)\right] \Big/ (1 - r_e r_c) \, , \qquad (3)$$

where c_i is the fractional cloud coverage for the *i*th layer, t_i is the transmission coefficient of the *i*th layer, n is the number of cloud layers and r_e and r_c are the albedos of the surface and clouds, respectively. The value of c_i is set to 0.0 for clear skies, 0.3 for scattered conditions, 0.7 for broken conditions, and 1.0 for overcast conditions. Empirical cloud transmission coefficients, t_i, were derived by solving equations 1 and 3 for T_C and t_i, respectively, during hours when a single cloud layer was reported at the development station. These development hours were restricted to 1100-1400 local time, the period of maximum solar radiation, in order to limit the estimation errors that result when clouds are not in the direction of sun. Values of t_i were derived for six different height ranges throughout the atmosphere. Surface albedo, r_e, is assumed to be 0.65 when snow depth exceeds 1 inch and 0.2 otherwise. Cloud albedo, r_c, is assumed to be 0.50 for clouds with bases less than 5486 meters (18,000 feet), otherwise cloud reflection is ignored.

Additional aspects of the original NRCC model improve solar radiation estimates when sky conditions are obscured or when horizontal visibility is restricted. Empirical transmissivities were determined for reports of fog, haze, total and partial obscurations, and for hours when visibility is less than 10 miles when no precipitation or obstructions to visibility are reported. Without these additions, the model tends to overestimate the lowest observed solar radiation values.

3.2 Modifications to the existing NRCC model

Recalculation of aerosol and cloud transmissivities was performed to compensate for changes in observation practices and reporting techniques. Separate transmissivities are used for estimating global horizontal and direct normal solar radiation, and diffuse horizontal radiation is calculated as the residual of these values (global horizontal minus direct horizontal).

Evaluation of hourly surface observations in Fig. 1 revealed that clear sky conditions are reported nearly twice as frequently by ASOS instrumentation compared to human observers. Many of these clear reports by ASOS may be contaminated by cloud cover present outside of the vertical or horizontal ranges of the ceilometer. It is therefore expected that appropriate atmospheric transmissivities should differ when using ASOS data as model input rather than human observations.

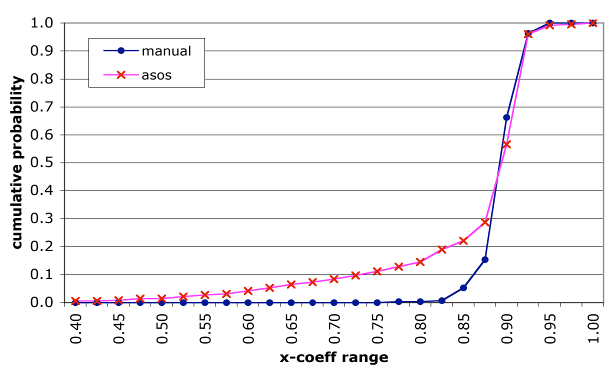

Fig. 2. Empirical cumulative frequency distributions of hourly x-coefficient values using manual or ASOS observations for model development at North Platte, Nebraska.

This is shown when comparing the empirical cumulative frequency distributions of hourly values of '*x*' in equation 2 (subsequently referred to as the x-coefficient) using human and ASOS observations (Fig. 2). A higher frequency of lower empirical x-coefficient values occur with ASOS observations, resulting from the presence of clouds that are not detected by the automated instrumentation. From the shape of these distributions it is apparent that the x-coefficient distribution is approximately normal (mean ≈ median) for human observations compared to the skewed distribution (mean < median) when using ASOS observations. To limit the positive mean error bias that would result by using the median x-coefficient in the modified NRCC model, the mean is used instead to better represent conditions with undetected clouds by ASOS. For consistency, mean x-coefficient values are also used for manual observation input, however this has little effect on solar radiation estimation because of the normal distribution of the x-coefficients. Consistent with Fig. 2, it is found that the mean value of the x-coefficient decreases when using ASOS observations (x=0.89) instead of human observations (x=0.93). When CDS observations are incorporated to supplement ASOS observations (ASOS-CDS), the mean x-coefficient value over all station pairs falls between manual and ASOS values as more information about clouds above the ASOS detection limit is included.

Additional relationships were incorporated into the modified NRCC model in order to more accurately estimate the appropriate x-coefficient values. For instance, the x-coefficients show dependence on the daily solar progression. A linear regression was fit to the relationship between solar zenith angle and x-coefficients to account for some of this variability. X-coefficients for hours with small solar zenith angles (x~0.85) are typically lower than for hours when the sun is near the horizon (x~0.95). This dependence is strongest in more humid locations, and nearly non-existent in more arid locations where x~0.95 independent of zenith angle. This difference is presumed to occur due to a higher frequency of cloud coverage above 3840 meters (12,600 feet) at more humid locations, resulting in a larger number of reported clear conditions by ASOS that are potentially contaminated by upper-level clouds.

In order to account for some of the x-coefficient variability in different climate regimes (humid or arid), seasonal climatologies of average daily relative humidity were determined at each station and their relationship with the slopes and y-intercepts of the linear regression fit to x-coefficients were assessed. Such relationships were strongest during the summer, with coefficients of determination for linear regression fits of about 0.35 for global and 0.63 for direct solar radiation. Because of these relationships, the original NRCC model was modified to calculate the x-coefficient using the following series of equations when using ASOS or ASOS-CDS observations:

$$\begin{aligned} A &= F(RH); \\ B &= F(RH); \\ x &= A + (z \cdot B), \end{aligned} \tag{4}$$

where A is the y-intercept and B is the slope of the linear relationship between solar zenith angle and x-coefficient, RH is the average daily relative humidity for the appropriate season, z is the solar zenith angle, and x is the x-coefficient. Here, A and B are represented as functions of RH. Once the calculation of the x-coefficient was complete, equation 2 was used as in the original NRCC model to obtain the aerosol transmissivity.

Cloud transmissivities (T_i) were first determined using hours with single cloud layer reports without the presence of fog or haze. This was done for each cloud condition: few (FEW), scattered (SCT), broken (BKN) and overcast (OVC). For each hour, a value of T_i was determined based on the following equation:

$$T_i = \left(\frac{I_{OBS}}{I_{CLR}}\right) * (1 - r_e r_c) \quad . \tag{5}$$

Here, T_i represents the cloud transmissivity without ground-cloud-ground reflection effects, I_{OBS} is the measured solar radiation value, I_{CLR} is the calculated clear sky solar radiation value from the modified NRCC clear sky model, r_e is the surface albedo and r_c is the cloud albedo. T_i is also equivalent to the numerator of equation 3 in the original NRCC model with n=1, and can be viewed as an 'intermediate' cloud transmissivity value that represents transmittance after radiation has passed through the atmosphere from space, but before any of the radiation that reaches the ground is reflected back into the atmosphere and further interacts with the clouds. Table 1 provides values of mean T_i assigned to different conditions and cloud layer heights in the modified NRCC model when using ASOS observations for model input.

Table 1. Single layer coud transmissivities used with ASOS observations. Each transmissivity value is assigned to specific cloud layer height ranges (meters) and cloud coverage.

Layer (m)	FEW	SCT	BKN	OVC
0 – 609	0.79	0.73	0.64	0.30
609 – 1219	0.85	0.81	0.70	0.37
1219 – 1828	0.86	0.82	0.69	0.40
1828 – 2438	0.85	0.78	0.64	0.45
2438 – 3048	0.84	0.73	0.59	0.48
3048 – 3840	0.77	0.68	0.57	0.53

When multiple cloud layers are present, the original NRCC model assumes that T_i is equivalent for each layer reported, even though its value was developed during conditions with single cloud layers. Under conditions with multiple cloud layers, T_i was multiplied for each cloud layer reported, as specified in equation 3. Due to a large underestimation of solar radiation during conditions with multiple cloud layers, this assumption was re-evaluated by calculating new cloud transmissivities for each cloud layer present, not just for the first layer under single cloud conditions. For instance, cloud transmissivities for the second of two reported cloud layers is defined as:

$$T_{i2} = \left(\frac{I_{OBS}}{I_{CLR} \times T_{i1}} \right) * (1 - r_e r_c) \ , \qquad (6)$$

and for the third of three layers,

$$T_{i3} = \left(\frac{I_{OBS}}{I_{CLR} \times T_{i1} \times T_{i2}} \right) * (1 - r_e r_c) \ , \qquad (7)$$

where T_{i1} is the cloud transmissivity defined in equation 5 (T_i) for conditions with single cloud layers, and T_{i2} and T_{i3} are cloud transmissivities for the second and third reported layers, respectively. When ASOS-CDS observations are used, up to 5 different reported layers are possible, and equations 5-7 can be applied to compute T_{i4} and T_{i5} for this situation. In the modified NRCC model, these transmissivities are used along with the term to account for ground-cloud-ground reflection effects to obtain the new total cloud transmissivity, T_C :

$$T_C = \prod_{k=1}^{n} T_{i(k)} \Big/ (1 - r_e r_c) \ , \qquad (8)$$

where n is the number of reported cloud layers and all other variables are defined as above.

DeGaetano et al. (1993) also showed that the original NRCC model reduced an overestimation bias on days with very low observed solar radiation by including reports of fog (T_F), haze (T_H), horizontal visibility (T_V) and obscurations (T_O) into the model. These transmissivities were also recomputed during model development.

4. MODEL EVALUATION

Cross validation procedures were performed to evaluate the modified NRCC model on a daily and hourly basis. Daily model statistics are given for each of the three versions of this model (Manual, ASOS and ASOS-CDS) in Table 2 for global horizontal, direct normal and diffuse horizontal solar radiation estimation. The use of human observations produces the most accurate solar radiation estimation overall. This can be attributed to the more detailed information about sky and surface conditions that human reports provide. Less accurate solar radiation estimation occurs with the use of automated observations (ASOS), but is improved by incorporating information about cloud cover above 3840 meters (ASOS-CDS). The typical daily percent error for estimation of global horizontal solar radiation (10%-15%) is comparable to other semi-physical and empirical solar radiation models (DeGaetano et al. 1995; Petersen et al. 1995; Maxwell 1998; Thornton and Running 1999). Model errors are also comparable to, although slightly larger than, errors in satellite-based estimates of daily global horizontal solar radiation that are typically less than 10% (Pinker et al. 1995; Jacobs et al. 2002). Errors resulting from estimation of direct normal and diffuse horizontal solar radiation were compared to numerous decomposition models. Such comparisons revealed comparable errors (20% - 25%) for estimation of direct normal solar radiation and slightly larger errors (27% - 39%) for estimation of diffuse horizontal solar radiation.

Table 2. Daily results from cross validation, using Manual, ASOS or ASOS-CDS observations as input to estimate (a) global horizontal, (b) direct normal and (c) diffuse horizontal solar radiation (MJ/m^2). Data were not available for development of direct and diffuse models using manual observations.

(a)	RMSE	ME	MAE	MAE %
Manual	1.99	-0.01	1.46	10.4
ASOS	2.94	-0.01	2.20	13.9
ASOS-CDS	2.56	0.06	1.87	12.0

(b)	RMSE	ME	MAE	MAE %
Manual	--	--	--	--
ASOS	6.30	-0.40	4.97	24.1
ASOS-CDS	5.35	0.13	4.19	20.4

(c)	RMSE	ME	MAE	MAE %
Manual	--	--	--	--
ASOS	2.26	-0.45	1.69	30.3
ASOS-CDS	2.03	-0.30	1.57	27.9

The sign of the estimation bias for global horizontal solar radiation is evenly distributed across all stations for each model type (ASOS and ASOS-CDS), but an uneven bias distribution is more apparent for direct normal and diffuse horizontal radiation estimation (Table 3). In general, the biases of each radiation component do not show any spatial consistency in any particular region of the country. Other models have shown positive biases for global horizontal (~60% positive, ~40% negative; Thornton and Running 1999), direct normal (~70% positive, ~30% negative; Maxwell 1998)

Table 3. Number of stations with positive and negative biases when estimating global horizontal, direct normal and diffuse horizontal solar radiation using ASOS or ASOS-CDS observations as input.

ASOS	+	-	ASOS-CDS	+	-
Global	27	27	Global	28	26
Direct	8	14	Direct	11	11
Diffuse	5	17	Diffuse	5	17

and diffuse horizontal (~80% positive, ~20% negative; Maxwell 1998) solar radiation.

The typical scatter of points that results from comparison of daily estimated and observed global horizontal solar radiation is given in Fig. 3. The apparent offset of the scatter concentration from the 1:1 line, particularly for days with high amounts of observed solar radiation, is a compensation that the modified NRCC model makes to account for days in which observations were not representative of reported conditions. These patterns are a result of incorporating mean x-coefficient values (instead of median values) into the model in order to minimize the positive mean error bias that results from cloud cover that is not observed during reported clear hours.

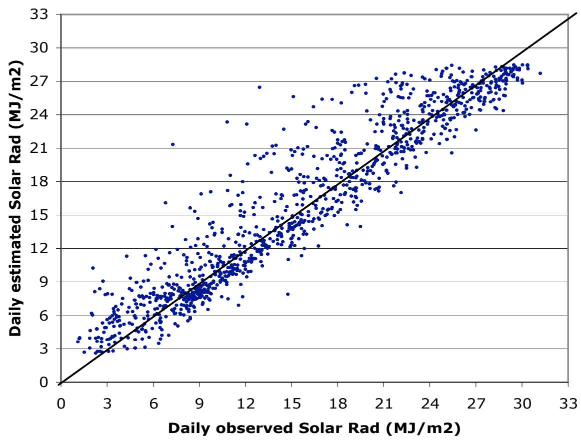

Fig. 3. Comparison of daily estimated and observed global horizontal solar radiation (MJ/m^2) at Scottsbluff, Nebraska.

Consistent with the daily results, hourly solar radiation estimation is most accurate when using human observations and least accurate when using ASOS observations that are not supplemented with CDS cloud estimates. The errors in estimating hourly global horizontal solar radiation from the modified NRCC model (15% - 19% for combined sky conditions) are similar to the 15% - 20% errors typically seen for hourly solar radiation estimates using satellite data (Pinker et al. 1995; Jacobs et al. 2002).

Three statistical discontinuity tests were used to detect inhomogeneities in the modeled solar radiation time series that could be attributed to the shift from manual to ASOS observations (Alexandersson 1986; Easterling and Peterson 1995; Perreault et al. 2000). These tests were performed on 11 stations in the Northwest, Northeast, Midwest and Southwest U.S. at which at least 5 years of data were available both before and after the ASOS commissioning date. Only 4 of the 11 stations show a break in their estimated solar radiation time series within 1 year of the ASOS commissioning date. Discontinuities were also detected in observed monthly solar radiation time series at some locations. While the number of breaks detected in the observed and modeled data time series were comparable within a 5-yr period of ASOS implementation, the step changes occurring in the time series of model bias show a slight tendency to concentrate closer to the ASOS commissioning date.

5. SUMMARY AND CONCLUSIONS

A new model was developed to estimate solar radiation using ASOS observations as model input. Development of such a model was necessary to compensate for biases that have resulted from ASOS replacement of human observers over the past decade. Evaluation of daily model output reveals typical errors of approximately 10% when human observations are used, and approaching 14% when ASOS data are used. When CDS cloud estimates are used in conjunction with ASOS data, typical errors decrease to approximately 12%. These results reflect that while combining ASOS and CDS data provides better sky condition representation, it does not specify actual conditions as frequent as the human observer. This model is also capable of estimating direct normal and diffuse horizontal solar radiation, with typical errors of 20%-24% and 27%-30%, respectively. These results are comparable to other contemporary solar radiation models.

6. REFERENCES

Alexandersson, H. 1986: A homogeneity test applied to precipitation data. *J. Clim.*, **6**, 661-675.

DeGaetano, A.T., K.L. Eggleston and W.W. Knapp, 1993: Daily solar radiation estimates for the northeastern United States. Northeast Regional climate Center Pub. No. RR 93-4, Ithaca, NY.

DeGaetano, A.T., K.L. Eggleston and W.W. Knapp, 1995: A comparison of solar radiation estimates for the northeastern United States using the Northeast Regional Climate Center and National Renewable Energy Laboratory models, *Solar Energy*, **55**, 185-194.

Easterling, D.R, and T.C. Peterson, 1995: A new method for detecting undocumented discontinuities in climatological time series. *Int. J. Climatol.*, **15**, 369-377.

Houghton H.G., 1954: On the annual heat balance of the Northern Hemisphere. *J. Meteorol.*, **11**, 1-9.

Jacobs, J.M., D.A. Myers, M.C. Anderson and G.R. Diak, 2002: GOES surface insolation to estimate wetlands evapotranspiration. *J. Hydrol.*, **266**, 53-65.

Maxwell, E.L., 1998: METSTAT – the solar radiation model used in the production of the NSRDB. *Solar Energy*, **62**, 263-279.

Menzel, W.P., F.C. Holt, T.J. Schmit, R.M. Aune, A.J. Schreiner, G.S. Wade and D.G. Gray, 1998: Application of GOES-8/9 soundings to weather forecasting and nowcasting. *Bull. Amer. Meteor. Soc.*, **79**, 2059-2077.

Meyers, T.P. and R.F. Dale, 1983: Predicting daily insolation with hourly cloud height and coverage. *J. Clim. Appl. Meteorol.*, **22**, 537-545.

Perreault, L., J. Bernier, B. Bobee, and E. Parent, 2000: Bayesian change-point analysis in hydrometeorological time series Part 1: the normal model revisited. *J. Hydrol.*, **235**, 221-241.

Petersen, M.S., P.J. Lamb and K.E. Kunkel, 1995: Implementation of a semi-physical model for examining solar radiation in the Midwest. *J. Appl. Meteorol.*, **34**, 1905-1914.

Pinker, R.T., R. Frouin and Z. Li, 1995: A review of satellite methods to derive surface shortwave irradiance. *Remote Sens. Environ.*, **51**, 105-124.

Thornton, P.E. and S.W. Running, 1999: An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. Forest Meteorol.*, **93**, 211-228.