Development of a semi-parametric PAR partitioning model for the United States

James C. Kathilankal[1*], Thomas L. O'Halloran[2], Andres Schmidt[3], Chad V. Hanson[3], and Beverly E. Law[3].

[1] LI-COR Inc. 4647 Superior Street, Lincoln, NE, 68504, United States

[2] Sweet Briar College, Department of Environmental Sciences, Sweet Briar, VA, 24595, United States

[3] Department of Forest Ecosystems and Society
College of Forestry, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331, United States

*Corresponding author: James Kathilankal

Abstract

A semi-parametric PAR diffuse radiation model was developed using commonly measured climatic variables from 114 site-years of data from 19 AmeriFlux sites. The model has a logistic form and improves upon previous efforts, using a larger data set and physically viable climate variables as predictors, including relative humidity, clearness index, surface albedo, and solar elevation angle. Model performance was evaluated by comparison with a simple cubic polynomial model developed for the PAR spectral range. The logistic model outperformed the polynomial model with an improved coefficient of determination and slope relative to measured data (logistic: R^2 = 0.76; slope=0.76; cubic: R^2 = 0.72; slope=0.73), making this the most robust PAR-partitioning model for the US subcontinent currently available.

1. Introduction

Photosynthetically Active Radiation (PAR) is the 0.4-0.7µm spectral range that is absorbed by plants and drives the process of photosynthesis (McCree, 1972). PAR at the ground surface has two primary incoming streams, diffuse and direct; which are significantly affected by the amount of clouds and aerosols in the atmosphere. These two radiant components differ in the way they transfer energy through plant canopies thus affecting canopy photosynthesis processes differently than what would occur at the leaf scale (Misson *et al*., 2005). Increased diffuse PAR fraction (the ratio of diffuse or isotropic PAR to total PAR (diffuse + direct beam)) in the atmosphere has been correlated with higher light use efficiency and increased CO_2 assimilation (e.g. Weiss and Norman., 1985, Gu. *et al*., 1999, 2002 and 2003, Knohl *et al*., 2008., Mercardo *et al*., 2009 and Still *et al*., 2009). Many of these studies utilize models of diffuse radiation (usually in the 0.15 to 4.0 µm shortwave range) to estimate the diffuse fraction rather than direct measurements.

Diffuse PAR can be estimated from models that range in complexity from spectral parameterization schemes like SPCTRAL2 (Bird and Riodan, 1986) and SMARTS2 (Gueymard, 1995) to simple linear regression models relating diffuse radiation fraction to extra terrestrial PAR (Hassika and Berbigier, 1998 and Tsubo and Walker., 2005). Jacovides *et al*. (2009) developed a third order polynomial model after applying 25 point moving average on clearness index (k_{tp}) (the ratio of global irradiance to extraterrestrial irradiance) data collected over a three year period over Athens, Greece. Butt *et al*. (2010) used a proxy cloud fraction (ratio of calculated total solar irradiance at a surface to the measured) to estimate diffuse PAR fraction.

Most diffuse fraction models are developed for global solar irradiation and very few models are developed from PAR data sets. The models developed for global solar radiation have been used in studies to convert the diffuse global solar irradiance to diffuse PAR fractions (e.g. Gu *et al*., 2002). Regression type models of diffuse shortwave radiation usually employ linear (e.g. Orgil and Hollands, 1977; Reindl *et al*., 1990), logistic (Boland *et al*., 2001; Ridley, 2010) or higher order polynomial type (e.g. Erbs *et al*., 1982., Spitters *et al*., 1986; Chandrasekaran and Kumar 1994, Miguel *et al*., 2001; Oliveria *et al*., 2002;; and Jacovides *et al*., 2006) equations relating clearness index (k_{tp}) to estimate diffuse fraction (k_{dp}). Reindl *et al*. (1990) used multiple regression analysis and identified air temperature, dew point and sine of the solar elevation angle as important parameters determining the partitioning of total irradiance into diffuse and direct components. Solar elevation angle and clearness index were used as inputs in models developed by Maxwell (1987) and Skartveit and Olseth (1987). Other parameters used in modeling diffuse fraction include dew point temperature, albedo and hourly variability index (root mean square difference between clearness index of an hour in question with respect to its preceding and succeeding hour) e.g. Perez *et al*., (1992) and Skartveit *et al*., (1998). The BRL model (Ridley *et al*., 2010) uses hourly clearness index, apparent solar time, solar elevation angle, daily clearness index and a persistence index similar to the variability index to calculate the diffuse fraction. Muneer and Munawwar (2006) used sunshine fraction, cloud fraction and air mass along with clearness index in predicting the diffuse fraction of global irradiance.

The objective of our study is to develop a simple semi-parametric diffuse PAR model applicable for the US, employing the AmeriFlux (Hargrove, *et al*., 2003) data set of above-canopy observations that have high spatio-temporal resolution. Development of such a model will aid future investigations of the effect of diffuse radiation on photosynthesis and light-use efficiency in response to climate. Although diffuse radiation is not regularly measured at all AmeriFlux sites, multiple year records from 19 sites are available for model development. The model presented here is developed with a dataset that is larger and more temporally and spatially diverse than any previous efforts, making it the most robust and broadly applicable diffuse PAR model developed to date. The model development is based on the BRL model as the logistic relationship used in this shortwave diffuse radiation model can be adopted for PAR diffuse fraction but with more pertinent drivers. The model is primarily intended for aiding researchers in understanding ecosystem response in terms of carbon and energy exchange in relation to the diffuse PAR fraction with data recorded at the site.

2. Methodology and Data analysis

The dataset used for model development and testing consists of multiple year records of PAR and diffuse fraction obtained from the AmeriFlux network. A detailed description of the sites utilized in this study is presented in Table 1. The sites selected consist of forested ecosystems, shrublands and croplands covering a wide latitude range (35-70°N). The geographical location of the sites is presented in Figure 1 in the form of a map. Sites which are close to one other may appear as single points on the map due to resolution of the map. The diffuse fraction data are mostly obtained using the BF3 sensor (Delta-T devices, Cambridge, UK). The BF3 sensor uses an array of photodiodes with a shading pattern that provides shade to some of the photodiodes while others remain exposed. This instrument has a resolution of 1 μmol m^{-2} s^{-1} and an accuracy of 15%. The data from BF3 sunshine recorders have been used in other studies relating cloud fraction to diffuse fraction (Butt *et al*., 2001).

For our study, data collected when solar elevation angles were <10° were removed to avoid cosine response issues. Although the data set contained records in hourly and half hourly formats, we averaged data to obtain hourly values for consistency. The hourly radiation values were checked against the quality controls proposed by the European Commission Daylight. This quality control eliminates data points based on the following criteria: R_d>1.1; R_S, R_S >1.2R_E; R_d>0.8R_E; R_S< 5 Wm^{-2} and R_b>R_E, where R_d is the total diffuse radiation R_S is the total incoming solar irradiance, R_E is the extra terrestrial irradiance and R_b is the direct normal irradiance.

Data points were eliminated when hourly rainfall values were greater than 5 mm, relative humidity values were 100%, or when dew point exceeded air temperature, as under these conditions, the measurement accuracy might be affected by water droplets formed on the sensor. Outliers were removed visually after the initial quality check so as to remove bad data which could occur due to electronic noise or instrument malfunction that could produce physically impossible values. After implementing the quality control check, the dataset consisted of 302926 hourly records from 114 site years.

Extraterrestrial PAR (R_{EP}) was calculated with solar elevation angle at a location according to

$$R_{EP} = R_C\left[1+0.033\cos(360t_d/365)\sin\beta\right] \tag{1}$$

where, R_C is the solar constant (2776.4 μmol m^{-2} s^{-1}., Spitters *et al.*, 1986); sin β is the sine of the solar elevation angle and t_d is the day number since 1st January.

3. Model development

The model developed here is similar in structure to the multi-predictor logistic model (BRL) developed by Ridley *et al*. (2010) for global solar irradiance, except we use additional predictors that directly affect the diffuse fraction and we also use a considerably larger data set. The predictors in the BRL model include daily clearness index (K_t), sine of the solar elevation angle (sin β), persistence index (ψ) and apparent solar time (AST).

$$k_d = \frac{1}{1+\exp(-5.38+6.63k_t+0.006AST-0.007\sin\beta+1.75K_t+1.31\psi} \tag{2}$$

The logistical form of the model has been identified as more robust than previously published piecewise linear or other non-linear forms (Boland *et al*., 2001; 2008). The goal of our work is to develop a model that is constrained by more commonly measured micrometeorological variables, rather than estimated variables like persistence index. The important factors considered in this study are PAR clearness index (k_{tp}), relative humidity (RH), albedo (α) and sine of solar elevation angle (sin β). Clearness index is widely used in one predictor models for PAR partitioning (Jacovides *et al*., 2009) as it is directly related to cloud fraction. Relative humidity is positively related with cloud cover (Walcek, 1994) and a greater diffuse fraction is often associated with higher humidity values. The effect of relative humidity on the relationship between k_{tp} and k_{dp} observed in our data set is presented in Figure 1a. The data are binned into linearly space bins of relative humidity classes and they indicate increased diffuse PAR fractions associated with higher relative humidity classes. Increased surface albedo resulting from changes in canopy reflectance or presence of snow can alter the diffuse fraction estimates. Skartveit *et al*. (1998) proposed a correction for clearness index estimation to account for the multiple reflections occurring between the surface and instrument dome when albedo is over 0.15. However in this study we consider albedo as a contributing factor to diffuse fraction as multiple reflections between the surface and clouds can enhance the diffuse fraction available for photosynthesis (Campbell and Norman, 2008; Knohl and Baldocchi, 2008; and Winton, 2005). Albedo of most vegetated surfaces can reach up to 0.25 and can vary widely as a function of leaf area index, disturbance history and snow cover. The effect of surface albedo on the relationship between k_{tp} and k_{dp} is presented in Figure 2b. The diffuse PAR fraction in general shows an increasing trend with increased albedo, but the trend shows some variations, probably due to the confounding effects of other factors. Increased albedo can result in increased diffuse fraction for the same clearness index compared to lower albedo values. The PAR diffuse fraction model developed in this study takes the logistic form

$$k_{dp} = \frac{1}{1+e^{-z}}$$

, where z is given as

$$z = a + b\mathrm{k_{tp}} + c\mathrm{RH} + d\alpha + e\sin\beta \qquad (3)$$

and a, b, c, d and e are fitted empirical coefficients determined in our analysis. The empirical coefficients were obtained by fitting the model to the data set. The relationship presented in equation 3 tends to underestimate diffuse fraction under clear sky conditions (Figure 3a). As a correction, a second logistic fit is applied to the data for $k_{tp}>0.78$. The values of the coefficients for the logistic model along with their 95% confidence intervals are presented in Table 2. The model performance is compared with a one predictor model developed by Jacovides *et al*. (2009). This model was selected for comparison as it was developed using data in the PAR spectral range and used a simple predictor (k_{tp}) that could be estimated for a large data set from multiple locations. This cubic polynomial model which relates diffuse PAR fraction as a function of smoothed PAR clearness index (moving average window size of 25) takes the following form after fitting to this data set:

$$k_{dp} = 0.8637 + 1.2699k_{tp} - 5.6676k_{tp}^{\ 2} + 3.8088k_{tp}^{\ 3} \quad (4)$$

The original cubic polynomial model had prescribed limits within which the model operated and constant values were assigned to k_{dp} values for k_{tp} values above and below a particular range. The modified cubic polynomial model presented in equation 4 is valid for $0.13<k_{tp}<0.865$, whereas for $k_{tp} \leq 0.13$, $k_{dp} = 0.9413$ and $k_{tp} \geq 0.865$, $k_{dp} = 0.18655$. These set points were chosen to provide a smooth transition from the inflection points in the model output. The model coefficients were estimated using a robust nonlinear regression method in MATLAB (Mathworks, Inc). The fit of data to the adjusted logistic model and the cubic model for the data set is presented in Figure 3b and Figure 3c. The percentage differences between measured diffuse fraction k_{dp} and modeled diffuse fraction k_{dpm} is plotted in Figure 4 as a function of each of the predictor variables in unequally spaced bins with an equal number of data points.

The model fits were assessed by randomly selecting one third of the data as an evaluation data set for statistical analysis. The comparison between measured and modeled diffuse PAR for the logistic and cubical model for the evaluation data set is provided in Figure 5. The performance of both models was further compared by using a bootstrap regression between the measured and modeled diffuse fractions with a data re-sampling of 10000 times to account for the errors in measuring the independent variable (measured diffuse fraction) from the evaluation data set. The results of the bootstrap regression comparison for the two models are presented in Table 3. The root mean square error percentage (RMSE %) (Jacovides, 2006) of the model fits to the evaluation data set is also presented in Table 3. The influence of seasonality on the logistic model accuracy was examined by plotting the RMSE (%) and R^2 of the regression between measured and modeled values as a function of the various months (Figure 6) for the entire data set. Since seasonality can influence the model fit, the logistic model was fit to the entire data set, by classifying the data into the four different seasons. The seasons were classified as summer (June 20 to September 21), fall (September 22 to December 20), winter (December 21 to March 19) and spring (March 20 to June 19). This enabled the development of seasonal model empirical coefficients, which are presented in Table 4. The model fit for the different sites is also presented by plotting the RMSE (%) and R^2 of the regression between the measured and modeled values for the various sites (Figure 7). The sites are arranged on the x-axis on an increasing latitudinal gradient and the figure illustrates the model fit across the sites.

3. Discussion

The multi-parameter logistic model predicts different diffuse fractions for the same clearness index for different combinations of albedo, solar elevation angle and relative humidity. The percentage difference between measured and modeled diffuse fraction generally indicate an underestimation by the model. The largest differences are associated with clearness index values around 0.67, albedo values of 0.24, moderate relative humidity (between 50-60%) and solar elevation angles of 46˚ (Figure 3). The logistic model thus produces the largest errors under moderately clear sky conditions, during the late morning and afternoon periods and when the atmosphere has moderate humidity. The PAR clearness index values close to 0.67 represents a clear sky condition above which the diffuse PAR fraction stays constant with increasing total PAR. The inability of the model to accurately capture this behavior results in large errors around this clearness index threshold. Further higher PAR clearness index values indicate low diffuse PAR fraction levels, which along with the above mentioned PAR clearness index threshold can lead to uncertainties in the measurement of the diffuse PAR fraction by the sensor. Albedo value of 0.24 produced the large errors as this is in the range of most vegetated surfaces and hence other confounding factors contributes to model errors around this albedo range. The cubic polynomial model evaluated in this study produces the largest errors during periods of high solar elevation angle, in contrast to the original model, which exhibited maximum error during the early morning/late evening hours (Jacovides *et al*., 2010). The cubic polynomial model percentage errors showed a similar behavior in relation with clearness index and albedo as the logistic model, but produced the largest errors under low humidity in contrast with the logistic model. The regression analysis indicates better performance of the logistic model over the cubic model, with a higher slope, lower intercept, and a larger coefficient of determination (R^2) (Table 3 and Figure 5). The RMSE (%) values also indicate a comparatively lower error for the logistic model (30.59 %) compared to the cubic polynomial model (32.68 %). The errors in the developed model could be attributed to other confounding factors such as seasonal effects, changes in atmospheric turbidity caused by air pollution or aerosol loading, and location differences. The fact that a combined data set from different locations was used in this study can lead to minimization of the dependence of the k_{dp}–k_{tp} correlation on local conditions (Jacovides *et al*., 2006). The model coefficients developed over the various seasons are similar in nature and the fit of the seasonal models to the data indicate similar R^2 and RMSE (%) values. This indicates the robustness of the logistic model developed in this study as only a marginal improvement was obtained for certain seasons by determining seasonal coefficients. The largest RMSE (%) values and the lowest R^2 values were observed for the summer months. The model performance stays constant throughout the year except for the period from September to December when the RMSE (%) decrease and the R^2 value increases. The largest RMSE (%) values were observed during the summer months, as in Jacovides *et al*. (2006) (Figure 6). The model fit done over the individual sites indicate larger errors (higher RMSE (%)) values as latitude increases. The upper latitude experience lower solar elevation angles which does impact the model accuracy. The lowest R^2 for the model fit was observed for sites in the middle of the country.

4. Concluding remarks

A logistic diffuse radiation model was developed using a large hourly radiation dataset obtained from the AmeriFlux network. The model performance was evaluated against a cubic polynomial model and its strengths and weaknesses were assessed. The goal was to develop a diffuse PAR model that employs commonly measured climatic/weather variables as predictors and is applicable for sites in the contiguous United States. The logistic model improves upon other PAR diffuse fraction models as it was developed using a large data set comprising of multi-year records from multiple sites. Future work includes application of this model to estimate diffuse radiation effects and contributions to annual net ecosystem exchange over various biomes represented by the AmeriFlux data.

Acknowledgements

This research was supported by the Office of Science (BER), U.S. Department of Energy (DOE, Grant nos. DE-FG02-06ER64307 and DE-FG02-06ER64316). The authors also would like to thank the AmeriFlux Network for providing the data and making it publically available. We would like to thank the principal investigators of the participating sites and other technical staff who put together this data set and the AmeriFlux QA/QC laboratory at Oregon State University for helping to ensure the quality of data of the AmeriFlux database.

Code Availability

The model is a very simple logistic model and it can be implemented very easily in any programming software or spread sheet based software like MS excel. A Matlab based function is provided. This function requires inputs of incoming PAR, relative humidity, albedo and sine of the solar elevation angle.

References

Baldocchi, D., Falge, E., Gu, L. H., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X. H., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K. T. P., Pilegaard, K., Schmid, H. P., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, Bulletin of the American Meteorological Society, 82, 2415-2434, 2001.

Bird, R. E., and Riordan, C.: Simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmospheres., Journal of Climate and Applied Meteorology, 25, 87-97, 1986.

Boland, J., Scott, L., and Luther, M.: Modelling the diffuse fraction of global solar radiation on a horizontal surface, Environmetrics, 12, 103-116, 2001.

Boland, J., Ridley, B., and Brown, B.: Models of diffuse solar radiation, Renewable Energy, 33, 575-584, 10.1016/j.renene.2007.04.012, 2008.

Butt, N., New, M., Malhi, Y., da Costa, A. C. L., Oliveira, P., and Silva-Espejo, J. E.: Diffuse radiation and cloud fraction relationships in two contrasting Amazonian rainforest sites, Agricultural and Forest Meteorology, 150, 361-368, 2010.

Campbell, G. S., and Norman, J. M.: An Introduction to Environmental Biophysics, Springer-Verlag, New York, 1998.

Chandrasekaran, J., and Kumar, S.: Hourly diffuse fraction correlation at a tropical location, Solar Energy, 53, 505-510, 1994.

De Miguel, A., Bilbao, J., Aguiar, R., Kambezidis, H., and Negro, E.: Diffuse solar irradiation model evaluation in the North Mediterranean belt area, Solar Energy, 70, 143-153, 2001.

Gu, L., Baldocchi, D., Verma, S. B., Black, T. A., Vesala, T., Falge, E. M., and Dowty, P. R.: Advantages of diffuse radiation for terrestrial ecosystem productivity, J. Geophys. Res., 107, 4050, 2002.

Gu, L. H., Fuentes, J. D., Shugart, H. H., Staebler, R. M., and Black, T. A.: Responses of net ecosystem exchanges of carbon dioxide to changes in cloudiness: Results from two North

American deciduous forests, Journal of Geophysical Research-Atmospheres, 104, 31421-31434, 1999.

Gu, L. H., Baldocchi, D. D., Wofsy, S. C., Munger, J. W., Michalsky, J. J., Urbanski, S. P., and Boden, T. A.: Response of a deciduous forest to the Mount Pinatubo eruption: Enhanced photosynthesis, Science, 299, 2035-2038, 2003.

Gueymard, C. A.: Parameterized transmittance model for direct beam and circumsolar spectral irradiance, Solar Energy, 71, 325-346, 2001.

Hargrove, W., Hoffman, F., and Law, B.: New analysis reveals representativeness of the AmeriFlux network, EOS, Transactions American Geophysical Union, 84, 529, 10.1029/2003EO480001, 2003, 2003.

Hassika, P., and Berbigier, P.: Annual cycle of photosynthetically active radiation in maritime pine forest, Agricultural and Forest Meteorology, 90, 157-171, 1998.

Jacovides, C. P., Tymvios, F. S., Assimakopoulos, V. D., and Kaltsounides, N. A.: Comparative study of various correlations in estimating hourly diffuse fraction of global solar radiation, Renewable Energy, 31, 2492-2504, 10.1016/j.renene.2005.11.009, 2006.

Jacovides, C. P., Boland, J., Asimakopoulos, D. N., and Kaltsounides, N. A.: Comparing diffuse radiation models with one predictor for partitioning incident PAR radiation into its diffuse component in the eastern Mediterranean basin, Renewable Energy, 35, 1820-1827, 2009.

Knohl, A., and Baldocchi, D. D.: Effects of diffuse radiation on canopy gas exchange processes in a forest ecosystem, Journal of Geophysical Research-Biogeosciences, 113, -, Artn G02023

Doi 10.1029/2007jg000663, 2008a.

Knohl, A., and Baldocchi, D. D.: Effects of diffuse radiation on canopy gas exchange processes in a forest ecosystem, J. Geophys. Res., 113, G02023, 2008b.

Mercado, L. M., Bellouin, N., Sitch, S., Boucher, O., Huntingford, C., Wild, M., and Cox, P. M.: Impact of changes in diffuse radiation on the global land carbon sink, Nature, 458, 1014-U1087, Doi 10.1038/Nature07949, 2009.

Muneer, T., and Munawwar, S.: Improved accuracy models for hourly diffuse solar radiation, Journal of Solar Energy Engineering-Transactions of the Asme, 128, 104-117, 10.1115/1.2148972, 2006.

Oliveira, A. P., Escobedo, J. F., Machado, A. J., and Soares, J.: Correlation models of diffuse solar-radiation applied to the city of Sao Paulo, Brazil, Applied Energy, 71, 59-73, 2002.

Orgill, J. F., and Hollands, K. G. T.: Correlation equation for hourly diffuse radiation on a horizontal surface, Solar Energy, 19, 357-359, 1977.

Perez, R., Ineichen, P., Seals, R., and Zelenkal, A.: Dynamic global to direct irradiance conversion models, ASHRAE Transations, 98, 354-369, 1992.

Reindl, D. T., Beckman, W. A., and Duffie, J. A.: Diffuse fraction correlation, Solar Energy, 45, 1-7, 1990.

Ridley, B., Boland, J., and Lauret, P.: Modelling of diffuse solar fraction with multiple predictors, Renewable Energy, 35, 478-483, 10.1016/j.renene.2009.07.018, 2010.

Skartveit, A., and Olseth, J. A.: A model for the diffuse fraction of hourly global radiation, Solar Energy, 38, 271-274, 1987.

Skartveit, A., Olseth, J. A., and Tuft, M. E.: An hourly diffuse fraction model with correction for variability and surface albedo, Solar Energy, 63, 173-183, 1998.

Spitters, C. J. T., Toussaint, H., and Goudriaan, J.: Seperating the diffuse and direct component of global radiation and its implications for modeling canopy photosynthesis. 1. Component of incoming radiation, Agricultural and Forest Meteorology, 38, 217-229, 1986.

Still, C. J., Riley, W. J., Biraud, S. C., Noone, D. C., Buenning, N. H., Randerson, J. T., Torn, M. S., Welker, J., White, J. W. C., Vachon, R., Farquhar, G. D., and Berry, J. A.: Influence of clouds and diffuse radiation on ecosystem-atmosphere CO2 and (COO)-O-18 exchanges, Journal of Geophysical Research-Biogeosciences, 114, 10.1029/2007jg000675, 2009a.

Still, C. J., Riley, W. J., Biraud, S. C., Noone, D. C., Buenning, N. H., Randerson, J. T., Torn, M. S., Welker, J., White, J. W. C., Vachon, R., Farquhar, G. D., and Berry, J. A.: Influence of clouds and diffuse radiation on ecosystem-atmosphere CO2 and CO18O exchanges, J. Geophys. Res., 114, G01018, 2009b.

Tsubo, M., and Walker, S.: Relationships between photosynthetically active radiation and clearness index at Bloemfontein, South Africa, Theoretical and Applied Climatology, 80, 17-25, 10.1007/s00704-004-0080-5, 2005.

Walcek, C. J.: Cloud Cover and Its Relationship to Relative Humidity during a Springtime Midlatitude Cyclone, Monthly Weather Review, 122, 1021-1035, doi:10.1175/1520-0493(1994)122<1021:CCAIRT>2.0.CO;2, 1994.

Winton, M.: Simple optical models for diagnosing surface-atmosphere shortwave interactions, Journal of Climate, 18, 3796-3805, 2005.

Table 1: Site location and ecosystem type information.

Sl. No	Site Code	Site Name	Vegetation	Latitude	Longitude	% Data
1	USFuf	Flagstaff Unmanaged Forest	Evergreen needle Forest	35.09	-111.76	5.19
2	USFmf	Flagstaff Managed Forest	Evergreen needle Forest	35.14	-111.73	5.44
3	USFwf	Flagstaff Wildfire	Grasslands	35.45	-111.77	5.74
4	USVar	Vaira Ranch	Grasslands	38.41	-120.95	10.30
5	USMMS	Morgan Monroe State Forest	Deciduous Broadleaf Forest	39.32	-86.41	7.02
6	USNe1	Mead Irrigated	Croplands	41.17	-96.48	13.11
7	USNe2	Mead Irrigation Rotation	Croplands	41.17	-96.47	12.64
8	USNe3	Mead Rainfed	Croplands	41.18	-96.44	13.10
9	USBar	Bartlett Experimental Forest	Deciduous Broadleaf Forest	44.07	-71.29	7.51
10	USMe2	Metolius Intermediate Pine	Evergreen needle Forest	44.45	-121.56	6.35
11	USKut	KUOM Turf Grass Field	Grasslands	45.00	-93.19	1.50
12	USHo1	Howland Forest Main	Evergreen needle Forest	45.20	-68.74	2.31
13	USHo3	Howland Forest East	Evergreen needle Forest	45.21	-68.73	2.31
14	USHo2	Howland Forest West	Evergreen needle Forest	45.21	-68.75	2.31
15	USUmd	UMBD Disturbance	Deciduous Broadleaf Forest	45.56	-84.70	0.46
16	USWCr	Willow Creek	Deciduous Broadleaf Forest	45.81	-90.08	0.35
17	USAn3	Anaktuvuk River Unburned	Open Shrub lands	68.93	-150.27	1.70
18	USAn2	Anaktuvuk River Moderate Burn	Open Shrub lands	68.95	-150.21	1.33
19	USAn1	Anaktuvuk River Severe Burn	Open Shrub lands	68.99	-150.28	1.31

Table 2: Logistic model coefficients for clearness index classes. The value given in the brackets is 95% confidence interval

Coefficients	k_{tp}(≤0.78)	k_{tp}(>0.78)
a	2.0394 (2.021,2.058)	1.2450 (1.163,1.325)
b	-5.7165 (-5.739,5.695)	-2.3404 (-2.427,-2.254)
c	1.3600 (1.344,1.376)	0.7100 (0.685,0.735)
d	0.8638 (0.838,0.890)	0.4228 (0.395,0.451)
e	0.3032 (0.287,0.320)	-1.9463 (-1.973,-1.920)

Table 3: Model performance comparison using regression analysis. The values given in the brackets are the standard error of the estimates obtained by resampling evaluaton data 10000 times. The root mean square error estimate from the measured and modeled values is also presented.

Model statistics	Logistic model	Cubic model
Slope	0.76 (±6.0x10^{-6})	0.73 (±7.0x10^{-6})
Intercept	0.12 (±4.0x10^{-6})	0.14 (±4.0x10^{-6})
R^2	0.76 (±8.0x10^{-6})	0.72 (±9.0x10^{-6})
RMSE (%)	30.59	32.68

Table 4: Logistic model coefficients for clearness index classes for the various seasons. The value given in the brackets is 95% confidence interval. The R^2 and the RMSE obtained by comparing the model output to the observed data is also provided.

Model Params	Summer		Fall		Winter		Spring	
	k_{tp}(≤0.78)	k_{tp}(>0.78)	k_{tp}(≤0.78)	k_{tp}(>0.78)	k_{tp}(≤0.78)	k_{tp}(>0.78)	k_{tp}(≤0.78)	k_{tp}(>0.78)
a	2.571 (2.531,2.612)	1.990 (1.767,2.212)	2.046 (2.003,2.089)	1.472 (1.338,1.606)	1.949 (1.911,1.987)	0.912 (0.765,1.060)	2.111 (2.077,2.146)	2.131 (1.964,2.297)
b	-5.586 (-5.622 -5.546)	-2.834 (-3.061,-2.606)	-5.671 (-5.720,-5.623)	-2.315 (-2.450,-2.180)	-5.470 (-5.513,5.427)	-2.188 (-2.339,-2.038)	-6.173 (-6.218,-6.127)	-3.106 (-3.284,-2.928)
c	1.432 (1.403,1.461)	1.121 (1.069,1.173)	1.259 (1.222,1.294)	0.277 (0.232,0.322)	1.476 (1.440,1.512)	0.931 (0.886,0.977)	1.241 (1.211,1.271)	0.473 (0.427,0.519)
d	-2.244 (-2.346,-2.142)	-2.071 (-2.272,-1.869)	0.578 (0.516,0.639)	0.656 (0.591,0.721)	1.158 (1.121,1.194)	0.497 (0.461,0.533)	0.787 (0.724,0.849)	0.822 (0.733,0.910)
e	-0.077 (-0.106,-0.048)	-2.090 (-2.144,-2.036)	0.460 (0.406,0.515)	-2.535 (-2.594,-2.475)	0.058 (0.017,0.099)	-1.867 (-1.924,-1.811)	0.822 (0.790,0.854)	-2.041 (-2.082,-2.000)
R^2	0.75		0.77		0.75		0.76	
RMSE (%)	31.00		30.81		29.61		29.57	

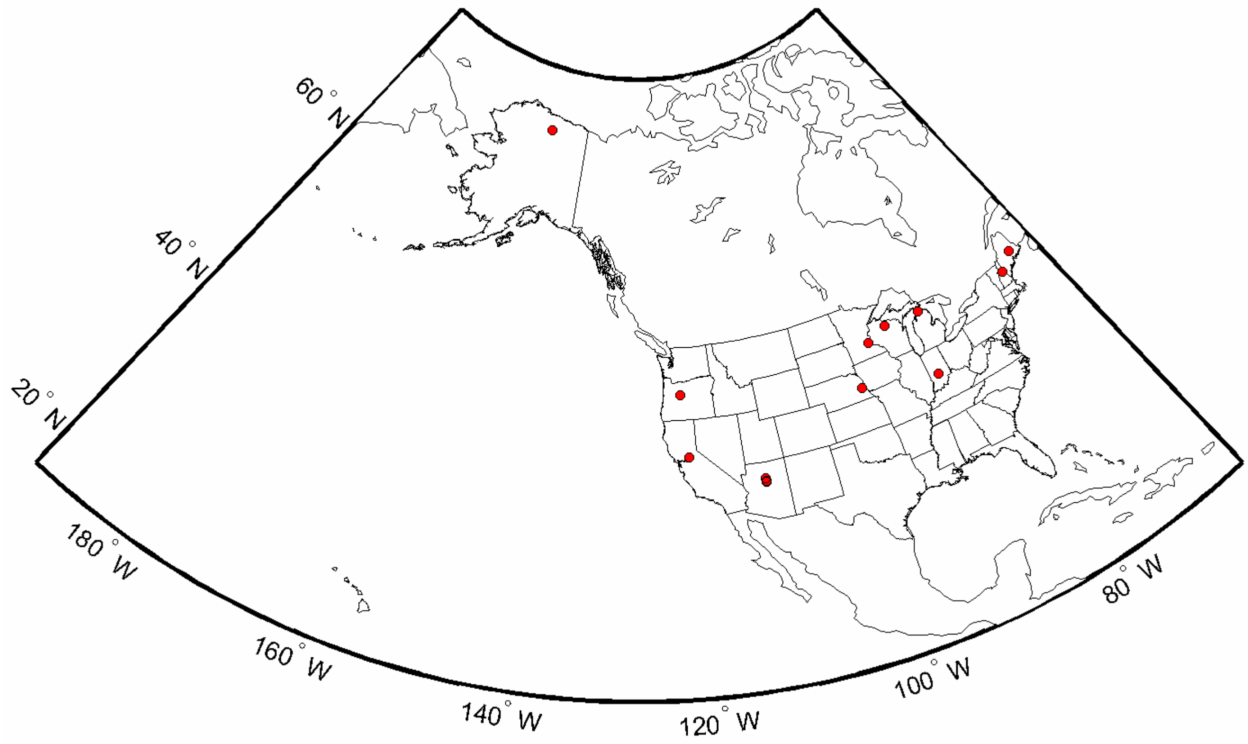

Figure 1: Location of sites presented on the USA map. Many sites which are closer together can appear as a single point on the map.

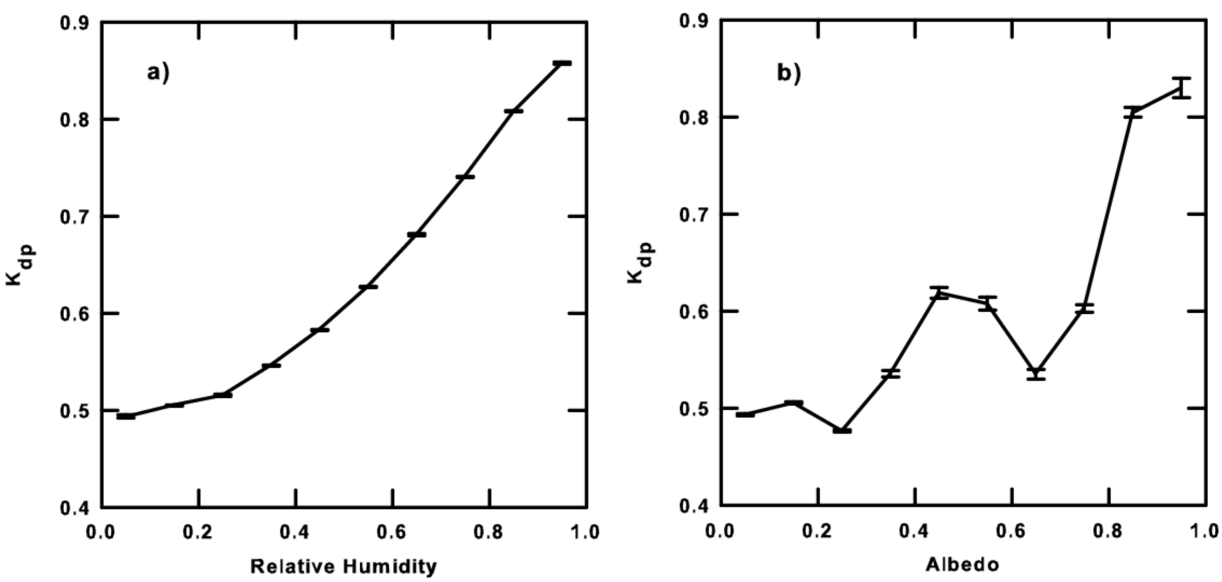

Figure 2: Relative humidity and albedo effects on k_{tp}-k_{dp} relationship

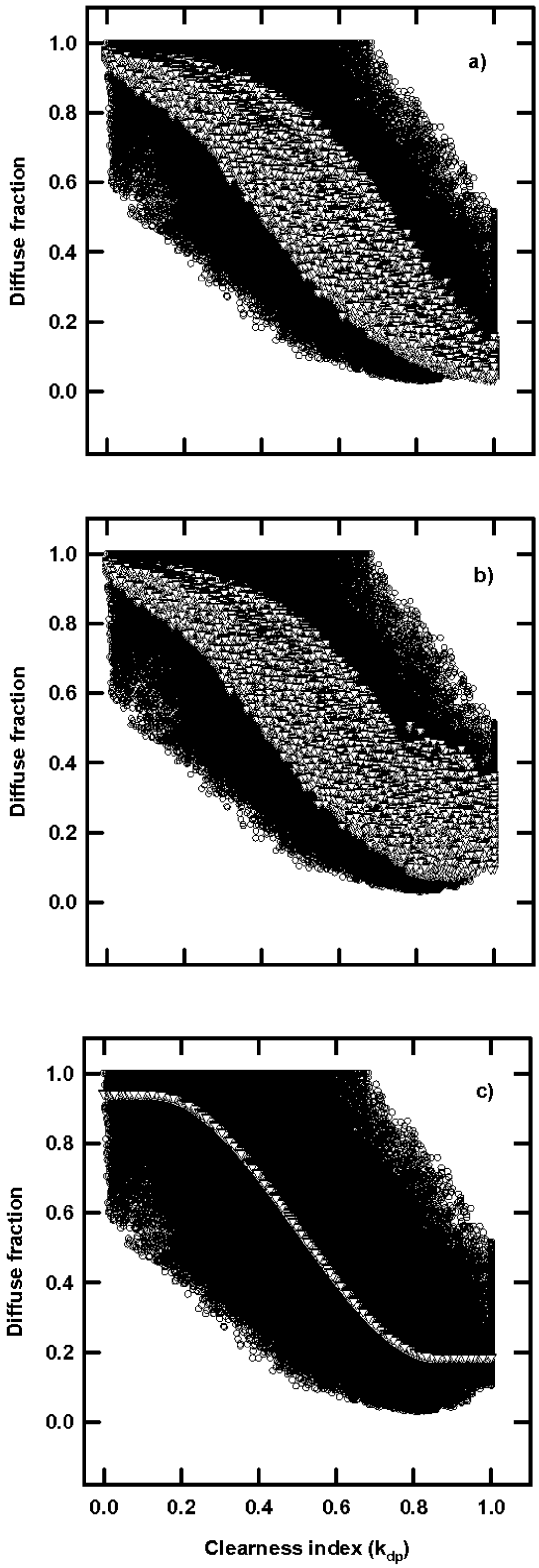

Figure 3: Model fit for the proposed multi-parameter logistic model (a and b) and cubic model (c). Panel (a) represents the initial fit to the logistic form and panel (b) indicates the modification to the initial logistic fit with a second logistic fit

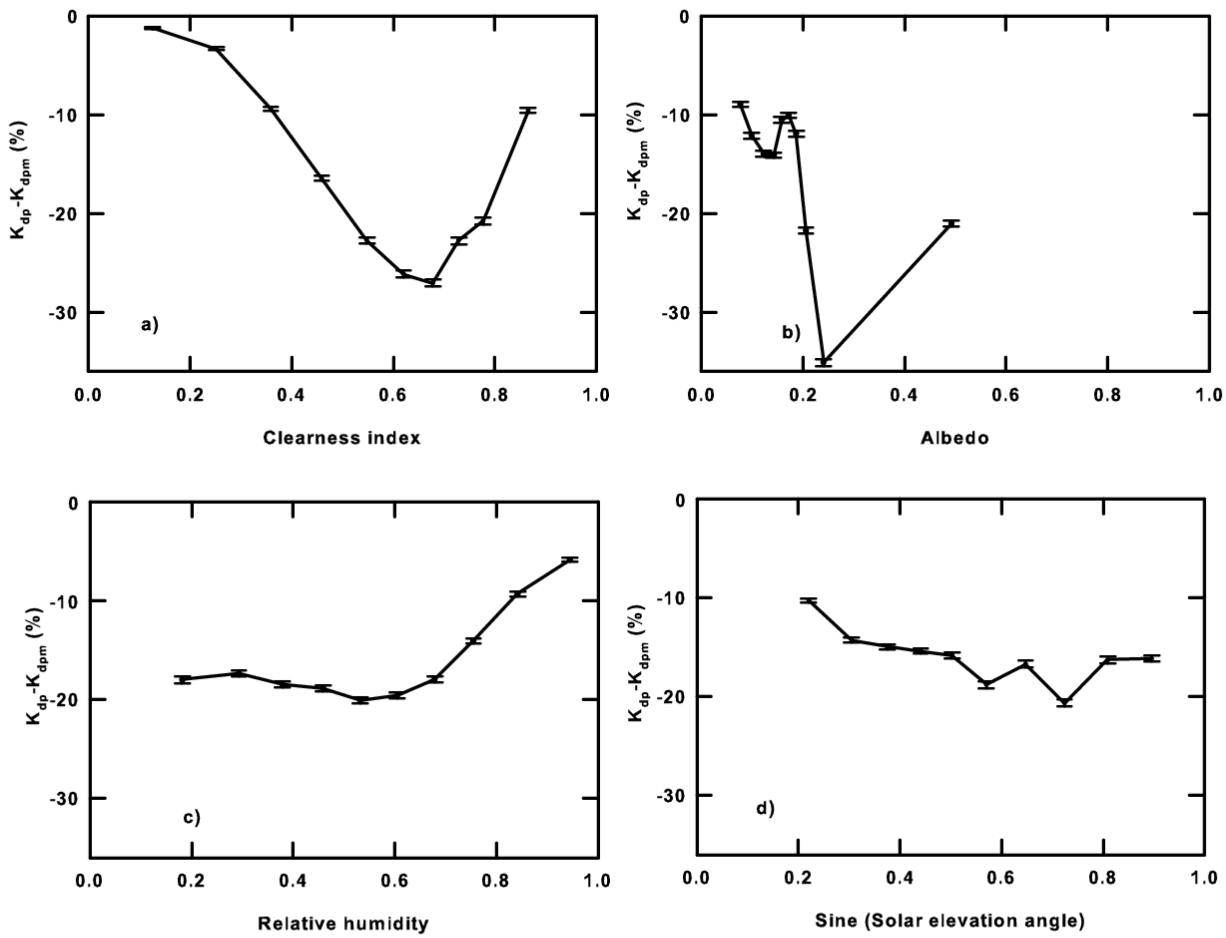

Figure 4: Percentage differences between measured and modeled diffuse radiation as a function of predictor variables

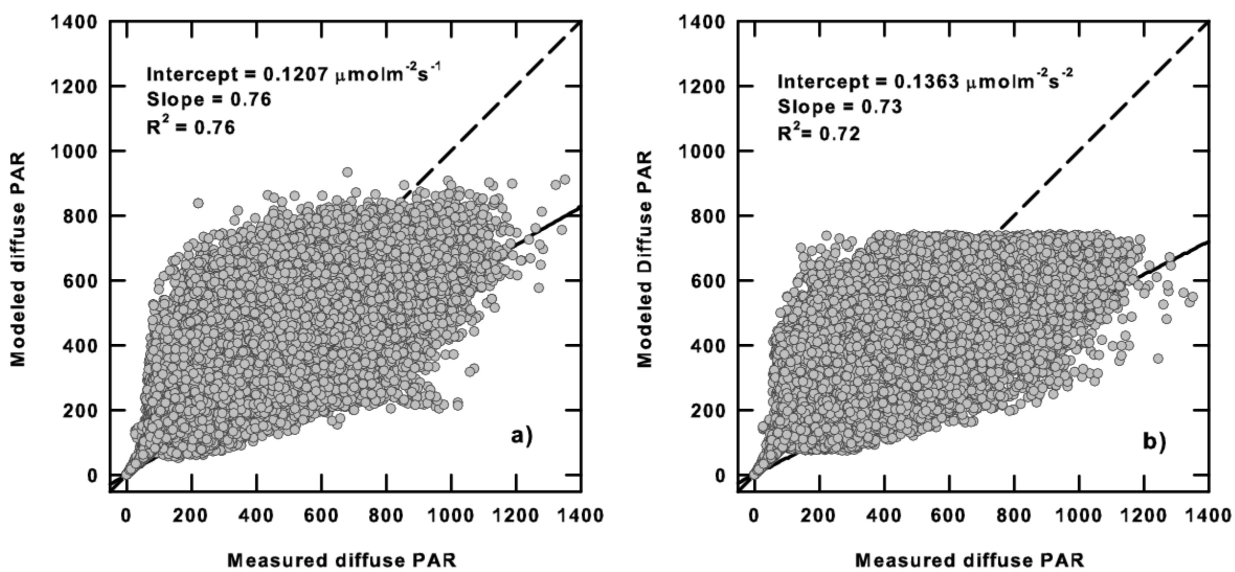

Figure 5: Comparison between measured and modeled diffused PAR a) logistic model b) cubic polynomial model. The regression statistics presented are for the bootstrap regression between the measure and modeled variables. All units are in µmol m^{-2} s^{-1}

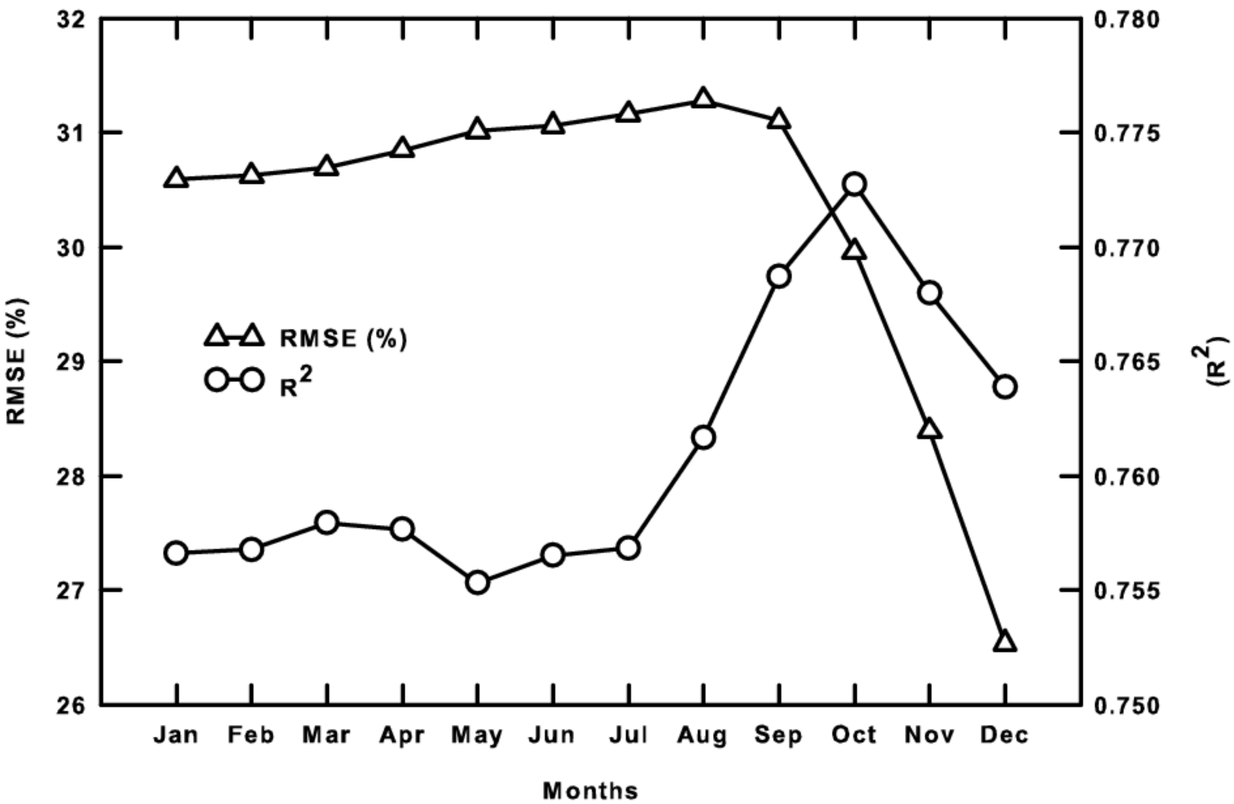

Figure 6: Model performance in terms of RMSE (%) and R^2 over various months of the year

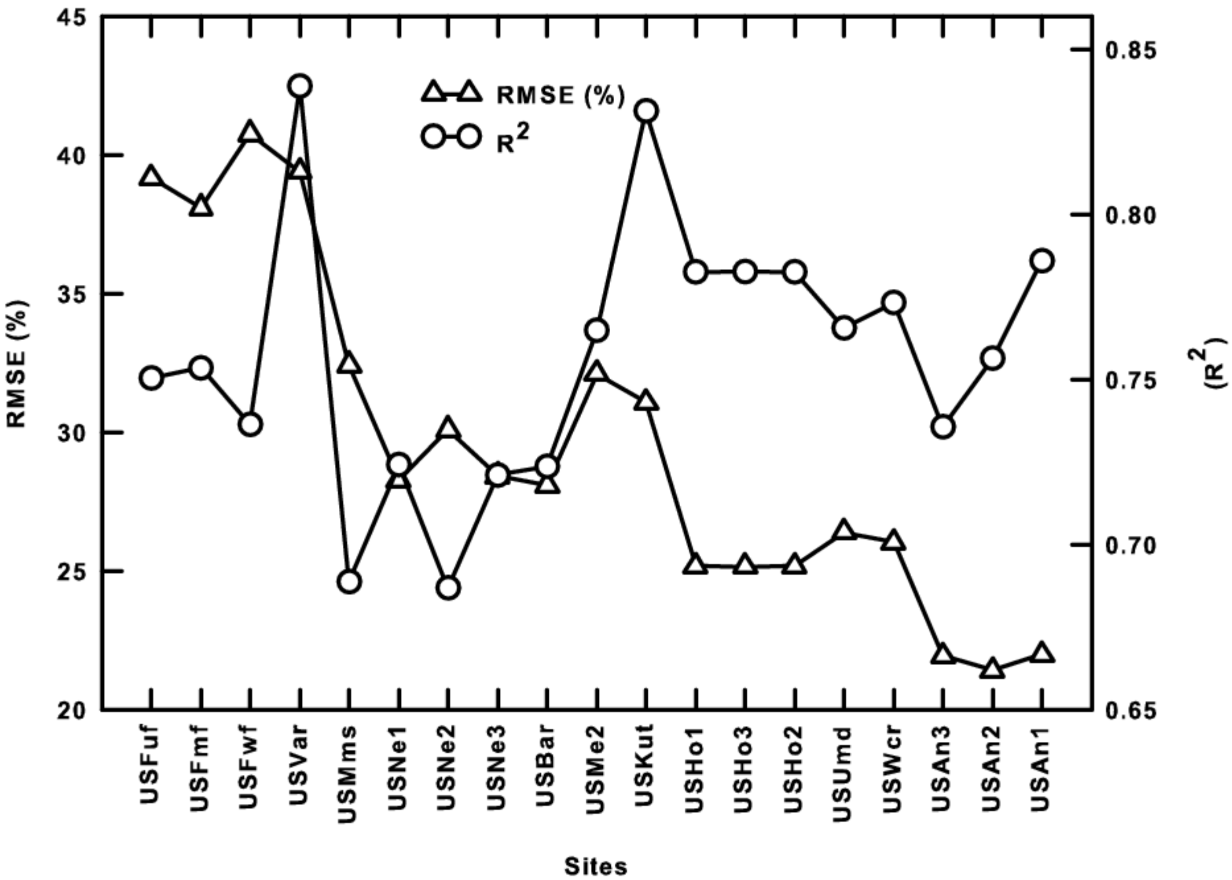

Figure 7: Model performance in terms of RMSE (%) and R^2 over the various sites. The sites are arranged on the x axis following an increasing latitudinal gradient from left to right.