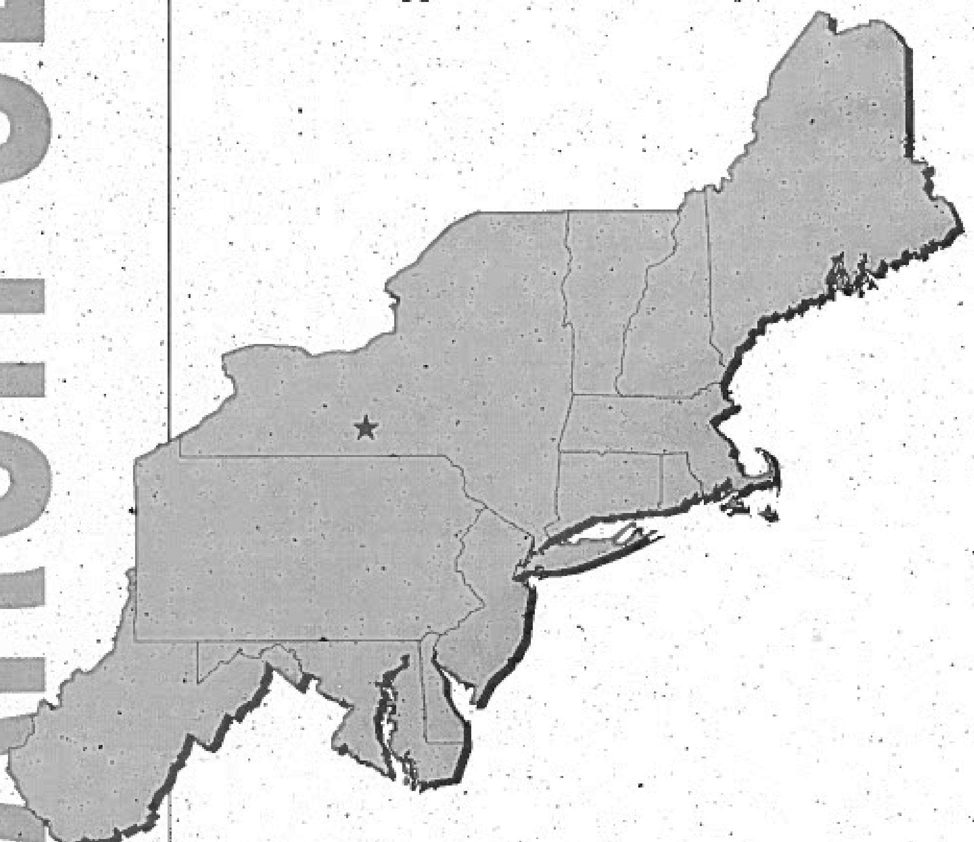
RESEARCH SERIES

NORTHEAST REGIONAL CLIMATE CENTER

Daily Solar Radiation Estimates for the Northeastern United States

Arthur T. DeGaetano
Keith L. Eggleston
Warren W. Knapp



Cornell University
Ithaca, New York
Publication No. RR 93-4
May 1993

The mission of the Northeast Regional Climate Center (NRCC) is to facilitate and enhance the collection, dissemination and use of climate data as well as to monitor and assess climatic conditions and impacts in the twelve-state, northeastern region of the United States. Implementing this mission involves three programmatic objectives: 1) the development and management of regional climate data bases, 2) the dissemination of information and educational services regarding climate and its impacts, and 3) the performance and support of applied climate research.

Established in 1983, the Northeast Regional Climate Center (NRCC) is one of six regional climate centers now operating throughout the nation. These regional centers serve as sources of climate data and information to public and private institutions and individuals as well as expertise on local and regional climate problems. The Center's staff cooperate with State Climatologists and research scientists in disseminating climate data and information, analyzing environmental and economic impacts of climate variability, and developing new applications of weather and climate data for agriculture, business, industry, and government operations.

The NRCC Research Report series is intended to make available to interested users the full results of climate research that has been supported by the NRCC. This report series supplements the normal reporting of research results in professional journals and provides an outlet for more complete and comprehensive accounts of work performed than is generally possible in journals.

For further information please write or call:

Northeast Regional Climate Center
1123 Bradfield Hall
Cornell University
Ithaca, New York 14853-1901
(607) 255-1751

The Northeast Regional Climate Center is supported by a Grant from the National Oceanic and Atmospheric Administration.

Daily Solar Radiation Estimates for the Northeastern United States

Arthur T. DeGaetano
Keith L. Eggleston
Warren W. Knapp

Northeast Regional Climate Center
Research Series
Publication No. RR 93-4
May 1993

INTRODUCTION

Solar radiation provides the energy that drives processes ranging from crop and insect development to active and passive solar energy systems. Despite its importance, solar radiation remains among the least measured climatological variables. A network of 26 stations operated by the National Weather Service provided the only source of nationwide solar radiation observations prior to 1977. The quality of the hourly radiation data collected by these SOLMET (SOLar METeorological) stations deteriorated over time, eventually reaching a point at which errors of as large as 30% existed (Thekarhara, 1976). Because of this problem, publication of SOLMET data was discontinued in 1972. More recently, the solar radiation observations collected prior to 1975 have been restored using an algorithm which accounted for the slow deterioration of the pyranometers used to collect the data. These adjusted values are now available for the period 1952-1975. Derived estimates of solar radiation at 222 additional "ERSATZ" sites have also been developed from the restored radiation estimates at the SOLMET sites. Radiation values at the ERSATZ sites are based on a correlation with observed cloud amount and the number of minutes of possible sunshine.

From 1977-1980, hourly solar radiation measurements were made at a network of 38 National Weather Service SOLRAD sites across the country. Upgrades to the instrumentation and instrument maintenance procedures were implemented in this network to prevent the data from deteriorating with time. However observations from this network were discontinued after only three years. In 1988, hourly solar radiation measurements began again at 31 sites nationwide. These data are archived at the National Climatic Data Center as part of TD-3281 Airways Solar Radiation.

In 1992, the National Renewable Energy Laboratory completed the National Solar Radiation Data Base (NSRDB). This data set provides hourly values of solar radiation at 239 stations for the period 1961-1990 based on observations from the SOLMET sites and several research facilities, together with model estimates based on total and opaque sky cover. Although the modeled solar radiation values were quality controlled, documentation provided with the NSRDB does not provide estimates of the accuracy of individual modeled daily solar radiation values. Instead, the annual average daily radiation values given in the NSRDB were compared to the corresponding SOLMET and ERSATZ values. These comparisons indicate that 40% of the NSRDB estimates exceed the solar radiation values documented in the combined SOLMET and ERSATZ data base (Marion and Meyers, 1992). Since the NSRDB is intended to provide high quality long-term solar radiation averages, the individual hourly and daily solar radiation values given in the data base must be used with caution.

Because the inventory of solar radiation measurements is so limited, numerous models and methods to estimate solar radiation from more commonly observed meteorological parameters have been proposed and implemented. For example, a model by Schmetz and Raschlke (1978) estimates total solar radiation based on observations of cloud amount and type. A model developed by Rangarajan et al. (1984) uses cloud height and coverage. Observed sunshine hours have been used by several investigators as a basis for solar radiation estimates (e.g.. Revfeim, 1981; Andretta et al., 1982). Other methods using daily maximum and minimum temperature (Bristow and Campbell, 1984) and latitude and precipitation (Reddy, 1987) have also been proposed. At best, these methods can estimate solar radiation with an average relative error of approximately 10%.

Geostationary satellite brightness measurements have also been used to estimate daily solar radiation values at the earth's surface. Gautier et al. (1980) achieved estimates of daily solar radiation which averaged within 10% of measured mean values. Duchon et al. (1992) used GOES estimated solar radiation data to drive a water budget model. In comparison to pyranometer measurements, they found that the GOES values overestimated surface solar radiation under cloudy conditions and underestimated the measured data as radiation values increased. Wetzel et al. (1993) showed that GOES derived estimates of solar radiation tracked pyranometer measured values fairly well under cloud-free conditions.

More sophisticated semi-physical models are also available to estimate total (direct beam and diffuse) solar radiation on a horizontal surface, hereafter referred to as simply solar radiation. Meyers and Dale (1983) modified earlier semi-physical models described by Atwater and Brown (1974) and Atwater and Ball (1981). The Meyers and Dale model uses standard meteorological observations as input and accounts for the attenuating effects of Rayleigh scattering, absorption by water vapor and permanent atmospheric gases, absorption and scattering due to aerosols, and absorption, scattering and reflection from clouds. Comparisons of calculated solar radiation estimates with observed values by Meyers and Dale showed mean absolute errors ranging from 1.05 MJ m^{-2} day^{-1} at Los Angeles to 1.80 MJ m^{-2} day^{-1} at Miami. In both cases these errors were approximately 10% of the mean daily solar radiation values.

In the midwestern United States, Meyers and Dale's model is currently used operationally by the Midwest Climate Center to drive crop yield and soil moisture prediction models and provide estimates of total solar radiation. Petersen (1990) conducted additional evaluations of the accuracy of solar radiation estimates produced by the Meyers and Dale model. Using observations from three midwestern sites, mean absolute errors were found to range from 8.4% to 17% of the mean daily solar radiation.

Due to the paucity of direct measurements and the receipt of frequent requests for solar radiation data, NRCC has adapted and implemented the Meyers and Dale model to provide daily solar radiations estimates for sites in the northeastern United States. Comparisons of the radiation estimates produced by the modified model and observations at northeastern sites show mean absolute errors of less than 13% and a mean error very close to zero.

MODEL DESCRIPTION

To adapt the Meyers and Dale model for use in the Northeast, several adjustments and modifications were incorporated. These changes primarily involved the modeling of absorption due to aerosols in the atmosphere and by clouds. In their model, Meyers and Dale express the flux density of total solar radiation on a horizontal surface at a given location, I, in the form:

$$I = I_o(\cos Z)T_R T_g T_w T_a T_c \quad (1)$$

where I_o 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 aerosols and clouds, respectively. Values of I are calculated for each hour and summed from sunrise to sunset to produce a daily total. The mean value of the solar constant was increased from that used by Meyers and Dale to 1367 Wm^{-2} to reflect the currently accepted value and adjusted to account for interannual differences in the earth-sun distance. Using latitude, ϕ, the solar declination angle, δ, and the hour angle, H, the solar zenith angle is computed using the relationship:

$$Z = \cos^{-1}(\sin\phi \sin\delta + \cos\phi \cos\delta \cos H). \quad (2)$$

The transmission coefficients for Rayleigh scattering and permanent gas absorption are combined and calculated using an empirical relationship described by Kondratyev (1969). The relationship, which considers forward as well as back scattered radiation, is

$$T_R T_g = 1.021 - 0.084[m(0.00949p + 0.051)]^{-0.5} \quad (3)$$

where p is the station pressure (kPa) and m is the optical air mass at mean sea-level. The optical air mass is given by

$$m = 35(1224 \cos^2 Z + 1)^{-0.5}. \quad (4)$$

The transmission coefficient for water vapor absorption,

$$T_w = 1 - 0.077(um)^{0.3}, \quad (5)$$

is a function of optical air mass, m, as well as the precipitable water vapor in the atmospheric column, u. Surface dew point is used to estimate u using a relationship given by Smith (1966):

$$u = \exp[0.1133 - \ln(\lambda + 1) + 0.0393T_d] \quad (6)$$

where T_d is the surface dew point temperature and λ is an empirically derived constant based on latitude and season.

Absorption and scattering due to aerosols is computed using the equation

$$T_a = x^m. \quad (7)$$

Meyers and Dale obtained a value of 0.935 for the empirical constant, x, using data from several U.S. stations, all of which were located outside the Northeast. Considerably better results for northeastern stations are obtained using a value of x=0.85. This value represents the median value obtained by solving equations 1 and 7 for T_a and x, respectively, under clear sky conditions (T_c=1) using surface and solar radiation observations at Ithaca, NY.

Cloud attenuation is also expressed empirically using the relationship

$$T_c = \prod_{i=1}^{3} [1 - c_i(1 - t_i)] \times (1 - r_e r_c) - 1. \quad (8)$$

Surface albedo (r_e) is assumed to be 0.65 when snow depth exceeds one inch and 0.2 otherwise. For clouds with bases less than 5486 m (18,000 ft.), the cloud albedo, r_c, is assumed to be 0.50. Cloud reflection is ignored for clouds with bases above 5486 meters as are multiple cloud reflections. Although up to five cloud layers can be recorded, reports of more than three layers are infrequent. Thus, only three layers were considered in estimating cloud attenuation. The fractional sky coverage of the ith layer is represented by c_i. This value is set to 0.0 for clear skies and 1.0 for overcast conditions. For broken and scattered conditions, c_i is given values of 0.7 and 0.3, respectively. Empirical transmission coefficients, t_i, are assigned to each cloud layer. These values were derived by solving Equations 1 and 8 for T_c and t_i, respectively, during hours with a single cloud layer at Ithaca, NY over the period April 1990 through September 1992. Median values are given in Table 1. Hours used to determine cloud transmission coefficients were limited to 1100–1400 local time to minimize observational errors which result when clouds are not in the direction of the sun.

If sky conditions are reported as obscured, then $t_i = 0.12$ and $c_i = 1.0$. Values of $t_i = 0.72$ and $c_i = 0.5$ are used for the lowest layer when the sky is partially obscured (-X). If not reported, a partial obscuration is assumed during hours when visibility is less than 1 mile. If clear skies, fog and a visibility greater than one mile are simultaneously reported, then $t_i = 0.93$. During hours in which clear skies are reported along with haze, the value of t_i is also set equal to 0.93. If no clouds or obstructions to visibility are reported, but the visibility is less than 10 miles, $t_i = 0.98$. These transmissivity values were empirically derived using Ithaca observations.

Table 1. Transmission coefficients, t_i, for indicated cloud heights and coverage.

Height (m)	Scattered	Broken	Overcast
≤ 609	0.84	0.61	0.16
610 - 1219	0.84	0.61	0.30
1220 - 3048	0.86	0.62	0.35
3049 - 5486	0.84	0.61	0.44
> 5486	0.96	0.77	0.66
> 5486 (thin)	0.98	0.91	0.86

TREATMENT OF MISSING DATA

During certain time periods within the historical hourly data record, only every third cloud height observation is archived for some stations. In such cases, cloud base heights are inferred from the reported sky coverage and ceiling height, which continued to be archived hourly, using a method devised by Petersen (1990). In her procedure, the reported ceiling height is assumed to be equal to the height of the lowest broken or overcast layer. Scattered layers below the ceiling are assigned heights 1524 or 3048 m below the ceiling, depending on its height. Overcast or broken layers which were visible and reported above the ceiling are similarly assigned heights 1524 or 3048 m greater than the ceiling height. Scattered layers, which occurred in the absence of a ceiling, are uniformly assigned random heights between 0 and 7620 m.

If sky coverage or ceiling data is also missing for a given hour, the average of the preceding and subsequent hour's transmission factors, t_i, is used. If either of these t_i values are missing, the remaining value is assigned to the missing hour. A similar procedure is used to handle missing dew point and station pressure observations. The daily average value, however, is used as a surrogate during periods when three or more consecutive hours are unavailable.

VALIDATION

To assess the accuracy of the radiation values estimated by the model, output values were compared with daily solar radiation measurements at seven stations in the northeastern United States. Coincident meteorological and solar radiation observations are available only at National Weather Service offices in Burlington, VT, and Pittsburgh, PA. At the five other comparison sites, radiation measurements from research facilities located near meteorological observing sites were used. Mean errors (ME), mean absolute errors (MAE), and root mean square errors (RMSE) were computed for the period of record at each site. The results appear in Table 2. Errors were also calculated seasonally and tended to be largest in summer and smallest during winter. However, when expressed as a percentage of the seasonal average

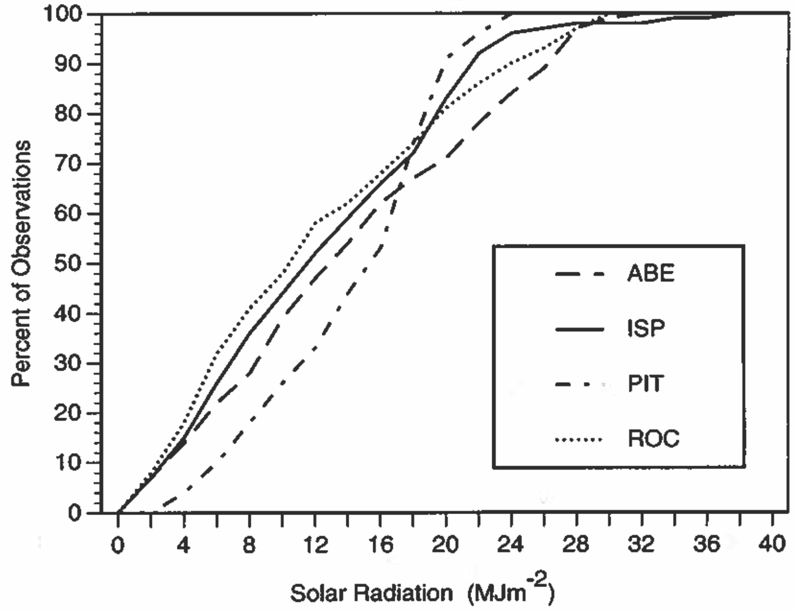


Figure 1. Cumulative frequency distribution of solar radiation observations at Pittsburgh, PA (PIT) as well as those near Allentown, PA (ABE), Islip, NY (ISP) and Rochester, NY (ROC). Radiation measurements at Ithaca, NY closely follow the distribution of ROC data. Observations at Burlington, VT and near Columbus, OH are essentially coincident with the ABE distribution.

observed solar radiation, errors similar to those given in Table 2 resulted during all seasons.

The cumulative frequency distributions of available observed solar radiation at the seven sites are illustrated in Figure 1. These daily observations are uniformly distributed throughout the year in all cases. Compared to the five other sites, the distributions of observed solar radiation at Pittsburgh and Islip show an unusual increase in the frequency of solar radiation observations at the upper end of the distributions. In addition, the maximum observed solar radiation recorded at these sites is anomalously low at Pittsburgh and high at Islip compared with the other stations which are located at roughly the same latitude. Based on these comparisons, it appears that instrumentation problems may have biased the data at

Table 2. Mean error (ME), mean absolute error (MAE), and root mean square error (RMSE) associated with modeled solar radiation at the indicated sites. Errors have units of MJm^{-2} and are also given as a percentage of the average observed solar radiation (OSR). OBS indicates the number of daily observations.

STATION	ME	%	MAE	%	RMSE	%	OSR	OBS
Allentown, PA	0.1	0.6	1.5	11.0	2.0	14.5	13.6	2523
Burlington, VT	-0.1	1.1	2.2	16.5	3.1	23.3	13.3	142
Columbus, OH	0.6	4.3	1.6	12.2	2.2	16.9	13.1	361
Islip, NY	-1.4	-9.0	2.0	13.1	2.5	16.3	15.3	2010
Ithaca, NY	-0.2	-2.0	1.5	12.0	2.0	16.3	12.3	400
Pittsburgh, PA	1.3	13.1	2.1	21.2	2.8	28.2	9.9	158
Rochester, NY	-0.2	-2.0	1.6	13.6	2.2	18.5	11.9	2460
All stations	-0.3	2.3	1.7	12.7	2.2	16.8	13.3	8054
w/o Islip and Pitt.	0.0	0.0	1.6	12.5	2.1	16.4	12.8	5886
Original model *	1.8	14.0	2.3	18.0	2.9	22.6	12.8	5886

* Original version of the Meyers and Dale model using data from all sites, excluding Islip and Pittsburgh.

Pittsburgh and Islip. For this reason, Table 2 contains a summary of the errors associated with all seven sites as well as other comparisons with Pittsburgh and Islip omitted. With the omission of these two stations, the mean error is reduced to zero and the MAE and RMSE are slightly lower.

The results in Table 2 compare favorably with a similar error analysis based on four midwestern stations carried out by Petersen (1990) and that based on a network of 12 stations used by Meyers and Dale (1983) in their original assessment of the model's performance. The stations used by Meyers and Dale were widely distributed across the U.S., but did not include any in the northeastern states. A total of 913 observations were available at the four midwestern sites used by Petersen. In her analysis, Petersen found mean errors which averaged 0.83 MJm^{-2}, while MAE and RMSE values averaged 1.75 and 2.3 MJm^{-2}, respectively. The MAE was 11.9% of the average observed solar radiation. The comparisons conducted by Meyers and Dale used 732 observations and gave ME, MAE and RMSE values of -0.1, 1.3 and 1.7 MJm^{-2}, respectively.

For a further comparison, the results produced by running the original version of the Meyers and Dale model for the five northeastern sites are also listed in Table 2. It is evident that our modifications and adjustments to the original Meyers and Dale model improve the accuracy of solar radiation estimates for northeastern sites.

CONCLUSIONS

Solar radiation measurements in the northeastern United States are relatively few in number and have limited periods of record. Since this semi-physical model requires only standard hourly surface observations as input, current solar radiation estimates can be generated at several dozen hourly observing sites within the region. Historical solar radiation estimates dating back to 1948 can be computed at approximately 30 sites in the Northeast which have digitized hourly data. Such a data base of estimated solar radiation values will provide essential data for applications ranging from crop modeling to the design of solar energy systems. In addition, the model estimates provide a comparative standard for the quality control of solar radiation measurements in the region.

ACKNOWLEDGMENTS

We would like to thank Dr. Kim Kroll of the Rodale Institute and Mr. Victor Cassella of Brookhaven National Laboratory for providing solar radiation measurements near Allentown and Islip, respectively. We are also indebted to Dr. John Hubbard of the State University of New York at Brockport for the solar radiation observations used in the Rochester comparisons. Mary Schoen Petersen of the Midwest Climate Center furnished the radiation data for Columbus, OH. Her suggestions were also helpful in writing several of the model's subroutines.

REFERENCES

Andretta, A., B. Bartoli, B. Coluzzi, V. Cuomo, M. Francesca and C. Serio, 1982: Global radiation estimation from relative sunshine hours in Italy. *J. Appl. Meteor.*, 21, 1377-1384.

Atwater, M.A. and J.T. Ball, 1981: A surface solar radiation model for cloudy atmospheres. *Mon. Wea Rev.*, 109, 878-888.

_________ and P.S. Brown, 1974: Numerical computation of the latitudinal variations of solar radiation for an atmosphere of varying opacity. *J. Appl. Meteor.*, 13, 289-297.

Bristow, C.L. and G.S. Campbell, 1984: On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. For. Meteorol.*, 31, 159-166.

Duchon, C.E., J.M. Salisbury, T.H. Lee Williams and A.D. Nicks, 1992: An example of using Landsat and GOES data in a water budget model. *Water Resour. Res.*, 28, 527-538.

Gautier, C., G. Diak and S. Masse, 1980: A simple physical model to estimate incident solar radiation at the surface from GOES satellite data. *J. Appl. Meteor.*, 19, 1105-1112.

Kondratyev, K., 1969: *Radiation in the Atmosphere*, Academic Press, 912 pp.

Marion, W. and D. Myers, 1992: *A comparison of data from SOLMET/ERSATZ and the National Solar Radiation Data Base.* National Renewable Energy Laboratory Technical Paper NREL/TP-463-5118, 13 pp.

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

Petersen, M.S., 1990: *Implementation of a semi-physical model for examining solar radiation in the midwest.* Midwest Climate Center Research Report 90-01.

Rangarajan, S., M.S. Swaminathan and A. Mani, 1984: Computation of solar radiation from observations of cloud cover. *Solar Energy*, 32, 553-556.

Reddy, S.J., 1987: An estimation of global solar radiation and evaporation through precipitation - a note. *Solar Energy*, 38, 97-104.

Revfeim, K.J.A., 1981: Estimating solar radiation income from bright sunshine records. *Quart. J. Roy. Meteor. Soc.*, 107, 427-435.

Schmetz, J. and E. Raschke, 1978: A method to parameterize the downward solar radiation. *Arch. Met. Geoph. Biokl. Ser. B.*, 26, 143-151.

Smith, W.L., 1966: Note on the relationship between total precipitable water and surface dewpoint. *J. Appl. Meteor.*, 5, 726-727.

Thekarhara, M.P., 1976: Solar radiation measurement: techniques and instrumentation. *Solar Energy*, 18, 309-325.

Wetzel, M.A., Y. Qiao, J.A. Warburton, D.L. Reinke and C.L. Combs, 1993: *Use of satellite cloud climatologies for climate studies at WRCC.* Preprint, AMS Eighth Conference on Applied Climatology, Anaheim CA., 218-223.

NOTES

NOTES

NRCC RESEARCH SERIES

Knapp, W.W. and K.L. Eggleston, *Some Impacts of Recent Climate Variability on the Northeast,* NRCC Research Publication RR 91-1.

Wilks, D.S., *Gamma Distribution Probability Tables for Use in Climatology,* NRCC Research Publication RR 91-2.

Samelson, D., *A Simple Method for Predicting Snowpack Water Equivalent in the Northeastern United States,* NRCC Research Publication RR 92-1.

Wilks, D.S., *Spline Interpolated Parameters for Adjusting Climatological Precipitation Distributions using the 30- and 90-Day Outlooks,* NRCC Research Publication RR 92-2.

Cember, R.P. and D.S. Wilks, *Climatological Atlas of Snowfall and Snow Depth for the Northeastern United States and Southeastern Canada,* NRCC Research Publication RR 93-1.

DeGaetano, A.T., K.L. Eggleston, and W.W. Knapp, *A Method to Produce Serially Complete Daily Maximum and Minimum Temperature Data for the Northeast,* NRCC Research Publication RR 93-2.

DeGaetano, A.T., W.W. Knapp, and K.L. Eggleston, *Standardizing Growing Degree Day Totals for Differences in Temperature Observing Schedules,* NRCC Research Publication RR 93-3.

NRCC DIGITAL DATA SETS

Eggleston, K.L. and D.S. Wilks, *Gridded Monthly Precipitation Distribution Parameters for the Continental United States,* NRCC Data Set DS 92-1.

Cember, R.P., K.L. Eggleston, and D.S. Wilks, *Digital Snowfall and Snow Depth Probabilities for the Northeastern United States and Southeastern Canada,* NRCC Data Set DS 93-1.

CORNELL
UNIVERSITY
Department of Soil, Crop and Atmospheric Sciences
Ithaca, New York 14853

Produced by
Northeast Regional Climate Center
9/93 1C 125-8435