Spatial Interpolation of Daily Maximum and Minimum Air Temperature Based on Meteorological Model Analyses and Independent Observations

Arthur T. DeGaetano and Brian N. Belcher

Northeast Regional Climate Center

Department of Earth and Atmospheric Sciences

Cornell University

Ithaca, NY

Submitted to Journal of Applied Meteorology and Climatology

July, 2006

Corresponding author address:

Dr. Art DeGaetano, 1119 Bradfield Hall,

Cornell University , Ithaca, NY 14850

Email: atd2@cornell.edu

Abstract

Hourly meteorological forecast model initializations are used to guide the spatial interpolation of daily Cooperative Network station data in the Northeastern United States. The hourly model data are transformed to daily maximum and minimum temperature values and interpolated to the station points after standardization to station elevation based on the model temperature lapse rate. The resulting bias (interpolation – observation) is computed and then interpolated back to the model grids allowing daily adjustment of the temperature fields based on independent observations. These adjusted data can then be interpolated to the resolution of interest. For testing, the data are interpolated to stations that were withheld during the construction of the bias field.

The use of the model initializations as a basis for interpolation improves upon the conventional interpolation of elevation-adjusted station data alone. When inverse distance weighted interpolation is used in conjunction with data from a 40 km model grid, mean annual absolute errors averaged 5% smaller than those from interpolation of station data alone for maximum and minimum temperature, a significant decrease. Using data from a 20km model grid, reduces mean absolute error during June by 10% for maximum temperature and 16% for minimum temperature. Adjustment for elevation based on the model temperature lapse rate improved the interpolation of maximum temperature, but had little effect on minimum temperature. Winter minimum temperature errors were related to snow depth, a feature that likely contributed to the relatively high autocorrelation exhibited by the daily errors.

1. Introduction

The uneven and relatively sparse distribution of stations that report daily climatological variables has led to a rich literature describing and comparing techniques to interpolate these point measurements. In most cases, the impetus for interpolation has been the use of these climatological observations for various agricultural, ecological or hydrological applications. Meyers (1994) reviews the basic statistical methodologies that form the foundation for most climatological interpolations. Traditional approaches have ranged from relative simple inverse-distance weighting to more complex techniques like kriging, splines (Hutchinson, 1991), and artificial neural networks (e.g. Rigol et al. 2001).

Jarvis and Stuart (2001a, 2001b) compare techniques for interpolating daily maximum and minimum air temperature. Their comparisons cover both interpolation methods (Javis and Stuart 2001b) and the selection of interpolation variables (Javis and Stuart 2001a). In terms of methods, they find only subtle differences in the performance of partial thin-plate splines (Hutchinson, 1991) with elevation as a linear covariate, detrended ordinary kriging (Deutsch and Journel, 1992) and detrended optimal inverse distance weighting. Cross-validated root mean square errors (rmse) for these methods over a single year averaged 0.83°C for daily maximum temperature and 1.15°C for daily minimum temperature.

The spatial interpolation of temperature is invariably influenced by elevation. Elevation emerged as the strongest covariate for estimating both daily maximum and minimum temperature Jarvis and Stuart (2001a). This relationship tended to be stronger for maximum than minimum temperature. Elevation has guided the spatial interpolation of temperature in numerous other studies as well (e.g. Price et al., 2000; Johnson et al., 2000; Kurtzman and Kadmon, 1999). In addition to the principal influence of elevation, Jarvis and Stuart (2001a) show that northing (north map coordinate), directional distances to the coast, and urbanization emerge as among the most important covariates. Typically more variables were required to interpolate daily minimum temperature than maximum temperature, however even in the case of minimum temperature the addition of more than three covariates has minimal effect on interpolation accuracy (Jarvis and Stuart 2001b). The degree of urbanization was particularly important covariate for minimum temperature (Jarvis and Stuart 2001a). After addressing elevation, Choi et al. (2003) showed a 30% reduction in minimum temperature interpolation errors by accounting for urbanization. Nonetheless, elevation exerted the primary influence.

Quantitatively, rmse for maximum temperature interpolation tend to be less than those for minimum temperature. For daily data, Jarvis and Stuart (2001b) show rmse in the range of 0.8 to 0.9°C for maximum temperature and between 1.1 and 1.2°C for minima. Courault and Monestiez (1999) use a kriging approach to interpolate daily temperature, but stratify their interpolations by circulation pattern. The accuracy of these interpolations for southeast France was improved only modestly by the stratification, as elevation had a much greater effect. In this region, root mean square errors were near 1.2°C for maximum temperature and 1.6°C for minimum temperature, higher values than those reported Jarvis and Stuart (2001b) for England and Wales. Thornton et al. (1997) use a Gaussian weighting filter to interpolate several meteorological variables over complex terrain. Mean absolute errors are typical of other approaches that account for elevation, 1.76°C and 1.95°C for daily maximum and minimum temperature, respectively.

Eischeid et al. (2000) evaluate several interpolation methods for their ability to estimate missing daily maximum and minimum temperature. For daily maximum temperature the best estimation procedure gave median monthly rmse values (for stations west of the Mississippi River) ranging from 3.32°C (January) to 2.44°C (July). For minimum temperature these values increased to 3.62°C (January) and 2.68°C (July). Elevation was not directly considered in either case. Johnson et al. (2000) demonstrate the use of the parameter-elevation regressions on independent slopes (PRISM) model to interpolate mean annual temperature. Mean absolute errors for annual maximum temperatures were between 0.86°C and 1.00°C, with a value closer to 1.50°C for minimum temperature.

Rigol et al. (2001) interpolated minimum temperature across the United Kingdom using an artificial neutral network (ANN). When trained on neighboring observations and terrain variables such as elevation, distance to the nearest river and a measure of terrain roughness, estimates were associated with a rmse of 1.15°C. Snell et al. (2000) used an ANN to simulate the downscaling of general circulation model (GCM) output. Sets of four and sixteen stations were chosen to simulate GCM grids and used to estimate daily maximum temperature at eleven interior stations. Based only on data from the simulated GCM grids they show the ANN to be superior to more conventional spatial averaging and inverse distance weighting approaches.

Willmott and Matsuura (1995) discuss two “smart” interpolation procedures. Topographically informed interpolation begins with an annual temperature field at a network of stations. Each temperature is reduced to sea-level based on the average environmental lapse rate. Temperatures are then interpolated to digital elevation model

(DEM) grid points on this constant elevation field with the final temperature at each DEM grid estimated based on an average lapse rate and the elevation of the point.

Topographically and climatologically informed interpolation (TCII) uses data from a second more spatially dense temperature network, but having observations over a different time period. Data from the high-density network are interpolated to DEM grids and the lower-resolution network station locations using topographically informed interpolation. Differences between the interpolated and observed values at the lower resolution network stations (δT) are then computed and reflect the influence of the different observation period. The δT values are then interpolated to the DEM grid and used to adjust the temperatures that were interpolated from the high-resolution station network.

In this study an alternative interpolation methodology is described. It uses the principal of TCII as a guide, but relies on gridded meteorological forecast model initialization data (as the high density network) and independent daily temperature observations (as the low density network). Unlike TCII, both the model data and observations reflect the same time period. The premise is that the enhanced spatial resolution of the initializations and the physical treatment of variables such as land use and elevation by the model allow for more accurate interpolations, while the available observations can be used to adjust any biases inherent to the model initializations. The method is evaluated on data from the northeastern United States for 2005 and compared to two conventional distance-based interpolations that use only the station observations.

2. Methods

a. Data

Real time hourly Rapid Update Cycle model (RUC) initializations for 2005 at 40 km x 40 km horizontal resolution were archived from an operational data feed. Benjamin et al. (1998) provide a description of the RUC model, with Smirnova et al. (2000) providing additional information about the land use scheme. Although RUC initializations at finer horizontal and vertical resolutions are available from archived sources, the lower-resolution operational values that were available provide a stringent evaluation of the interpolation methods relative to other approaches and allow the use of the interpolations in conjunction with an array of real-time agricultural models. This latter attribute was the impetus for this study. In addition, the 40 km x 40 km resolution of the available real-time RUC initializations is similar to that of the North American Regional Reanalysis (NARR) dataset (Mesinger et al., 2006). This facilitates the use of this dataset as input to the interpolation procedure described in this study and allows for the potential development of interpolated climatological data fields.

A limited subset of data for each RUC grid point was extracted. These included metadata describing grid latitude, longitude, elevation and land cover type. In addition, hourly temperature data at 2 m and 1000, 950, 900, and 850 hPa were extracted, as were the heights of the pressure levels. These temperature and height data, which are constrained by the vertical resolution of the model in the lower atmosphere, provided a means of computing hourly temperature lapse rates at each RUC grid. RUC grid points encompassed the spatial domain shown in Figure 1, which is hereafter referred to as the Northeast.

Observed temperature data were from a subset of Cooperative Observer Network stations that report daily maximum and minimum temperature data in real time. The set was restricted to those sites with 24-hour observation schedules ending in the morning (6 –9 a.m. local time). This is the most common observation time in the network and alleviated the problems of interpolating data between sites with different observation schedules. This restriction also eliminated the small number of sites that may have been incorporated into the RUC initializations. Between 350 and 400 Cooperative Network observations are typically available on a given day. Figure 1 shows a representative distribution of these stations. Given the real time application of the interpolation, the station data were not subject to final quality control that typically occurs several months after observation. Nonetheless, the data did undergo preliminary quality control that screens for same-station data inconsistencies, but does not make spatial between-station comparisons.

To be comparable to the daily Cooperative Network maximum and minimum temperatures, a spline curve (Press et al., 1992) was fit to the hourly RUC temperatures within the 8 a.m. to 8 a.m. daily observation period and the preceding 7 a.m and subsequent 9 a.m. values. Including these boundary values eliminated a few uncommon inconsistencies associated with starting or ending the spline at the potentially lowest value within the 24-hour window. Based on the spline, the maximum and minimum temperatures within the 24-hour daily interval are obtained.

b. Elevation Adjustment

On day *n*, a minimum (maximum) temperature existed at each RUC grid point based on the spline interpolation. This value was also assigned a time of occurrence based on the initialization time corresponding to the lowest (highest) RUC hourly temperature. Before interpolating these minimum (maximum) temperatures horizontally to station *j*, the temperature at each RUC grid (T_r) was adjusted to the elevation of station *j*, (z_j) using the equation:

$$T_j = T_R - \Gamma_R z_j . \quad (1)$$

T_j is the adjusted RUC temperature and Γ_R is the linear change in temperature for the hour of minimum (maximum) temperature occurrence. Γ_R is calculated between the RUC pressure levels that encompass the station elevation. Thus, instead of characterizing the ambient topography, the RUC temperatures are projected to a plane with an elevation corresponding to that of the station. During summer, T_R was typically –8.2°C m^{-1} for maximum temperature and –3.7°C m^{-1} for minimum temperature. Typical winter T_R values were 7.3°C m^{-1} and 6.5°C m^{-1} for maximum and minimum temperature, respectively.

c. Horizontal Interpolation

Horizontal interpolation of the adjusted RUC temperatures to the coordinates of the station was then accomplished on this equal-elevation surface via multiquadric interpolation (Nuss and Titley, 1994). This method is of the same class of interpolation routines as the thin-plate spline procedures evaluated in previous studies (e.g. Jarvis and Stuart, 2001b). Multiquadric interpolation is a widely used means of meteorological data interpolation that yields results that are similar or superior (e.g. Sokolov and Rintoul, 1999) to those based on thin plate splines.

The multiquadric interpolation procedure includes three tunable variables 1) r, the radius beyond which grids do not influence the interpolation 2) λ, a smoothing parameter and 3) c, the multiquadric parameter. Based on Nuss and Titley (1994), λ was fixed at a relatively low value of 0.0025. This provides minimal smoothing, and based on the earlier work low, relatively stationary cross-validated rmse. The multiquadric parameter essentially ensures that the multiquadric function has continuous derivatives. The analyses were not sensitive to the choice of this small positive value, a result shared by Nuss and Titley (1994). Hence, this parameter was set to 0.06.

The radius of influence was also fixed based on a set of sensitivity analyses. In these analyses, (not shown) interpolation errors were found to decrease quickly as r increased from 0.4° to 0.5°. Including stations beyond a 0.8° radius contributed little to reducing the interpolation error. A relatively large 2.0° radius of influence was used in the present evaluations to limit areas in which insufficient data precluded interpolation and also to allow a more selective choice of stations based on topographic considerations in addition to horizontal distance. This radius is near the middle of the range of radii evaluated in Nuss and Titley (1994).

For comparison purposes, separate horizontal interpolations were computed using inverse distance squared weighting. The 2.0° radius of influence was maintained for this analysis.

Given this background, the overall interpolation procedure progresses as shown in Figure 2. Multiquadric (or inverse distance weighting) interpolation is used in step 3 to interpolate the RUC temperature to the location of the target station. Likewise in step 5, the computed biases are interpolated via the multiquadric (or inverse distance weighting) procedure back to each RUC grid point. The biases, defined at interpolated RUC – station observation, at several stations are used in the interpolation in step 5.

Once biases have been interpolated to all grid points, the interpolated bias is subtracted from the original RUC temperature at each grid, producing an interpolated temperature field at the resolution of the RUC that has been ground-truthed based on the set of independent station observations.

If interpolation to a finer scale is desired, a similar procedure is followed. In addition to interpolation to the station locations in Steps 1 and 2, the RUC temperatures are also elevation-adjusted and interpolated to points on a finer grid. For example, the finer grid might be represented by a 5 km DEM. For computational efficiency, a smaller radius of influence is used in this interpolation, such that it assures the presence of at least one grid point in each of the four compass quadrants about the finer scale (DEM) grid point. Following these interpolations and the computation of biases in Step 3, the station biases are interpolated to the finer scale grid points (with $r = 2°$) and the original RUC interpolations adjusted to account for the observed biases.

c. Validation

An analogous procedure is used to validate the RUC-based interpolations in the subsequent section based on cross-validation. Each station serves as the finer scale grid point and its bias is withheld during the interpolation of biases from the remaining stations (Step 5). This procedure is applied to each station iteratively resulting in an independent interpolated temperature and observation at each station location from which error statistics can be computed.

These cross-validation trials were repeated using three modifications of the above interpolation procedure for comparison purposes. To quantify the benefits of assimilating the RUC analyses, the interpolation of elevation detrended observations from stations within the radius of influence using multiquadric interpolation (MQ_{obs}) and separately inverse distance weighting (IDW_{obs}) was evaluated. Likewise, the RUC-based multiquadric interpolation (MQ_{RUC}) was applied without the preliminary adjustment for elevation differences (MQ_{noelev}). Table 1 summarizes these different analyses, as well as method discussed in the next section.

3. Results

Table 2 summarizes the cross-validation results for these methods based on the 134,741 station-days evaluated. A RUC-based method is consistently associated with the smallest error. For minimum temperature, regardless of interpolation method, biases are typically small (< 0.02), positive (interpolation warmer than observations) and not significantly different from zero. In terms mean absolute error (mae) and root mean square error (rmse), the IDW_{RUC} errors are 4-10% lower than the MQ_{obs} and IDW_{obs} errors. These differences are significant at the $\alpha = 0.01$ level.

For maximum temperature, the biases are larger and negative. The errors associated with the IDW methods are significantly different from zero ($\alpha = 0.01$). The bias for the IDW_{obs} method is notably larger than that of the other methods. In terms of mae and rmse, however the IDW_{RUC} interpolation gives values that are about 7% lower than those given by the other procedures. These mae differences are statistically significant ($\alpha = 0.01$).

For minimum temperature the effect of adjusting the RUC data for elevation (prior to MQ interpolaton) is trivial. The elevation adjustment plays a larger role for maximum temperature, as the mae is reduced by 10% when an adjustment for elevation is made prior to interpolation. This is to be expected given the relatively coarse 50 hPa vertical resolution of the RUC initializations. Lapse rates computed at this resolution are a better representation of a well-mixed boundary layer profile, such as might be expected near the time of maximum temperature occurrence, rather than a more stratified morning sounding.

a. Seasonal biases

On a monthly basis, a strong season cycle in maximum temperature interpolation errors is absent in MQ_{RUC}, IDW_{RUC} and MQ_{obs} (Fig. 3a). In terms of rmse the methods behave similarly. Only a subtle seasonal cycle exists, with minimum rmse during the summer. IDW_{RUC} gives the lowest rmse during all months.

For minimum temperatures (Fig. 3b), the largest (most positive) biases generally occur during the winter months, albeit the seasonal cycle is again weak. In terms of variance, the seasonal cycle is amplified, with larger rmse during the winter months for all interpolation methods. The IDW_{RUC} interpolations are consistently the least variable. However MQ_{RUC} is associated with the smallest mean biases in almost every month. Median errors tended to be larger (more negative in the case of maximum temperatures) indicating some non-normality in the errors. Based on the median errors, those from the MQ_{RUC} interpolations were consistently closer to zero.

b. Geographical biases

There is not a consistent spatial pattern to the errors associated with the maximum and minimum temperatures from the RUC-based interpolations (Fig. 4). Figure 4 is representative of the spatial pattern of both maximum and minimum temperature errors produced by the IDW_{RUC} and MQ_{RUC} methods during the summer and winter months.

The largest errors appear as isolated anomalies interspersed across the region. Presumably these errors are artifacts of unique microclimates affecting specific stations or non-meteorological peculiarities in the station observations. There are several pockets of relatively high errors along the Atlantic Coast. However, there is not a consistent relationship between distance from the coast and interpolation error. Exceptions are six stations that occupy RUC grids with land use classified as water. In these grids, interpolation bias for minimum temperature is consistently negative during both seasons, indicating warmer RUC values. Presumably the water surface moderates the diurnal cycle in the RUC analysis. A similar bias, however, is absent for maximum temperature. Analyses based on MQ_{obs}, show more spatial homogeneity in the bias field, with more widespread areas of high bias along the Atlantic coast and the Appalachian Mountains.

In terms of individual stations, several of the stations with the largest mae during summer also experience among the largest mae in winter. Correspondence between stations with the largest maximum and minimum temperature mae during the same season is limited. Exceptions include three stations in Maine that exhibit large summer errors (two experience large winter errors as well). The Maine stations exemplify the influence that erroneous station observations exert on the interpolation. Figure 5 shows daily maximum temperature series at Jonesboro and Machias, Maine. These sites are 17.4 km apart. From Figure 5, it appears that the observer at Machias “shifts” the daily observations, recording them on the most likely day of occurrence (which is typically the calendar day before observation) rather than the day of observation, as specified by the observation procedures. It is also possible based on Figure 5 that Machias’ morning observation time may be incorrectly specified in the metadata, particularly since the minimum temperatures also appear to be shifted. Machias is also problematic in winter. However, its observational inconsistencies do not exert as strong of an influence on neighboring stations.

Other than sharing large mae, the stations with largest interpolation errors have little in common. The sites encompass a range of elevations and are dispersed across the Northeast. There is not an indication that distance to the Great Lakes influences the errors, although there are few morning observation stations along the Lakes' shorelines (Fig. 1). Station density differences across the region appear to have minimal effect. Aside from the erroneous Maine station, a number of sites are located in northern New England. This may be an indication that the varied topography of the region influences the interpolation. There is a tendency for biases to become more positive with increasing elevation. However, a linear regression of bias versus elevation explains less than 1% of the variation in all cases (not shown).

In an effort to isolate and adjust for the effects of local topographic variations, a 5-arc second DEM was used to characterize the slope of the terrain along an east-west and separately north-south transect centered on each station. Using the north-south transect as an example, the topographic half-angle, A_{south}, between the station and the closest DEM grid to the south of the station is computed by:

$$A_{south} = \frac{180.0b}{\pi}\left(\tan^{-1}\left[\frac{|Z_s - Z_D|}{f(Y_s - Y_D)}\right]\right), \quad (2)$$

where Y and Z are the latitude and elevation of the station (S) and DEM grid (D). The variable f is a function of latitude and is used to convert latitude to distance (m). When Z_s

$< Z_D$, b equals 1.0, , otherwise b equals -1.0. An analogous function is used to compute the topographic half-angle to the north of the station, A_{north}.

The topographic angle along the transect is then given by:

$$Ta = 180 - A_{south} - A_{north} \tag{3}$$

Using this convention, stations located within depressions have topographic angles < 180, while those located on domes have topographic angles > 180. The topographic angle is set to 180 in cases with sloped terrain such that Z_D south $< Z_s < Z_D$ north. There was no indication that topographic angle, defined in this manner, influenced the magnitude or sign of the interpolation bias. An analogous procedure can be used for east-west transects, with longitude substituted for latitude. Combining Ta for both transects allows the three-dimensional topography to be characterized.

c. Temporal and Meteorological biases

Daily interpolation errors within three subregions, central New York, southern New England and northern Virginia (Figure 1) were assessed in terms of temporal biases and ambient weather conditions. Figure 6 shows that within each subregion the lag-1 autocorrelation of mae exceeds 0.30 and in some cases approaches 0.60. There are also geographical and seasonal differences in autocorrelation behavior and magnitude. During winter (January and February, since December data for 2005 are not sequential), the mae for minimum temperature exhibits high lag-1 autocorrelation in all subregions (Fig 6 a and b). Assuming normality, these correlations are significant at the 95% level. Lag-1 autocorrelation for minimum temperature bias, however, is lower, particularly in the southern New England subregion, where this value becomes negative. This is an indication that large daily interpolation errors in southern New England have a tendency to be followed by large errors of the opposite sign. Maximum temperature bias exhibits marginally significant lag-1 auto correlation in the central New York and northern Virginia (not shown) subregions. This is not the case in southern New England.

In summer, high positive correlation between the interpolation errors tends to persist for several days. This is particularly true for minimum temperature (both mae and bias) in the southern New England and northern Virginia (not shown) areas (Fig. 6d). In these regions, high autocorrelation for maximum temperatures errors is limited to a one-day lag. In contrast, maximum temperature mae and bias exhibit high correlations to lags of 6 days in central New York (Fig. 6c). In this subregion autocorrelation in the summer minimum temperature mae and bias series is generally low.

To investigate the factors that contributed to the largest daily interpolation errors, the five largest daily average mae values for each subregion (values averaged over all stations within a subregion) for maximum and minimum temperature were isolated for summer and winter months. Surface weather maps based on 0000 and 1200 UTC observations were examined to determine if there was a specific synoptic feature that was associated with the largest (and separately smallest mae). During winter, there was a tendency for large interpolation errors to occur under high pressure for both maximum and minimum temperature. Of the 30 cases examined (five from each subregion for maximum and minimum temperature) 27 occurred with high pressure centered over the subregion.

In each subregion, several of the largest error values (for both maximum and minimum temperature occurred during the period from January 30 – February 6.

Otherwise, there was little correspondence in the dates of largest (or smallest) mae among the subregions or between maximum and minimum temperature.

During summer, ten of the 15 largest maximum temperatures examined occurred on days on which a stationary front affected the subregion. For minimum temperature, nine of the 15 instances of the largest errors occurred under high pressure. The June 15-20 period was associated with large maximum and minimum temperature errors in all subregions. In terms of the lowest errors, high pressure was again the predominant synoptic feature for minimum temperature. The lowest errors, however, were consistently found in highs of maritime tropical origins, while polar highs were typically associated with the large error cases. Except for the period from July 15-17, there was little correspondence between the dates of the smallest errors among the subregions or between maximum and minimum temperature.

A closer examination of the days with the largest errors revealed four general casual mechanisms. Errors in the observed data values continued to be problematic as discussed previously. Likewise, large interpolation errors resulted from synoptic patterns such as stationary and cold fronts that produced a sharp gradient of temperature across the subregion of interest. These cases were often confounded by the passage of frontal boundaries near the start or end of the 8 a.m. observation period.

Microclimatic factors also appear to influence the accuracy of the interpolations. On August 9, 2005, a 1023 hPa surface high pressure system was centered over Pennsylvania. The RUC 2 m minimum temperature analysis showed a weak temperature gradient across the subregion, in agreement with the observations (Fig. 7a). There is also good agreement between the RUC analysis temperature values and observations, with the exception of the 12°C observation (Fig. 7a). Local experience indicates that cold air drainage typically affects the cooler site. The 6°C temperature difference between this site and the station with the 18°C reading to the southwest is typical of the temperature variation on nights with strong radiational cooling conditions.

Attempts to identify and account for such microclimatic effects were not successful. Using topographic angles computed with Eq. 2 and 3, sites prone to cold air drainage were characterized by $Ta < 180$. After computing biases at each station (step 4 of the interpolation procedure), the bias field was examined to determine if biases of a specific sign preferentially occurred at sites with $Ta < 180$. The binomial probability distribution was used for this assessment. If a preferred bias was indicated, the interpolation of biases in step 5 was limited to stations with $Ta < 180$. Thus, in interpolating to a station with $Ta < 180$ in the cross-validation analysis, only those sites with $Ta < 180$ were used when a preferential bias was present.

Although this procedure (denoted MQ_{Ta}, hereafter) reduced the interpolation errors in some cases, the overall effect was minimal. Across all stations (not just those with the largest mae), this procedure had little effect on bias or mae (Table 3). It should be noted that restricting stations to those with $Ta < 180$ limited the number available for interpolation. As a result, stations more distant from the target station were weighted more heavily in the interpolation. This increase in distance presumably counteracts any benefit derived from assuring similar topographic character.

The final source of large interpolation errors results from poor representations of the observed temperature field by the RUC analysis. Such cases, which were rare, are analogous to the erroneous station observation discussed earlier. An example of this case is shown in Figure 7b. Across the northern Virginia region, the RUC initialization averages more than 5°C warmer than the observations. In addition, the RUC temperatures decrease by more than 5°C from the southern to northern portions of the subregion. The observed temperatures show little variation. If the RUC temperature gradient was absent, interpolation error would likely be small, since the station observations would have adjusted the warm bias inherent in to the RUC analysis. In this case, the gradient introduces an erroneous gradient to the bias field that negatively affects the subsequent adjustment and interpolation.

4. Discussion

The analyses presented in the previous section indicate that mesoscale meteorological model analyses provide a robust foundation for climatological station data interpolation. The RUC-based interpolation procedure gives results that are comparable or superior to those presented in the literature for other regions (e.g. Jarvis and Stuart, 2001; Courault and Monestiez, 1999; Thornton et al., 1997). When elevation-detrended station-based interpolation procedures, such as those used in previous studies, are applied to the present study domain, the model-based interpolations are superior, particularly in terms of mae. Both methods yield similar biases. Having demonstrated the initial benefit of the RUC-based interpolation scheme, future research efforts should investigate means by which interpolation accuracy might be improved. Such efforts should focus on three primary areas.

The first deals with interpolation methodology. The literature suggests similar interpolation accuracy among partial thin plate smoothing splines, elevation-detrended ordinary kriging and detrended inverse distance weighting methods (Jarvis and Stuart, 2001b). The current results, confirm these results given the similarities between the MQ_{obs} and IDW_{obs} error statistics. However, in previous evaluations the inverse distance method is outperformed by the spline-based approach.

There are also differences in methodological approach. Most notably Jarvis and Stuart (2001b) optimized the power function in their IDW approach and the tunable spline parameters on a daily basis. The current work adopted a static set of interpolation method parameters, as the intent was to compare model and station-based results. It should noted that for a given method, the model-based interpolations are consistently associated with lower errors.

Further improvements in the RUC-based interpolations are likely through real-time optimization of the interpolation parameters. Tuning of the parameters can be accomplished through a series of cross-validation trials. For a given day, cross-validation trials based on different combinations of parameters could be compared and a set of parameters selected such that error across the entire interpolation domain or subregions are minimized. In this case, the stations withheld for the computation of cross-validated errors would also need to be withheld during parameter optimization to obtain independent error statistics The use of artificial neural network approaches in conjunction with the model data should also be evaluated.

The autocorrelation of the interpolation biases might also be exploited to improve the accuracy of the interpolations. In a pilot investigation, runs of same-signed cross-validation errors were tracked at each station. When the likelihood of realizing a run of the observed length by chance was sufficiently small ($\alpha = 0.05$, based on a binomial probability distribution), an adjustment was made based on the median bias exhibited on days within the run. Thus, the interpolated temperature at the station on day *n* was adjusted by the median bias on days *n-k* through *n-1*, where *k* is the length of the run. At the small sample of stations for which this procedure was tested, the results were encouraging.

Despite this, it is unclear how this procedure could be adapted to the interpolation of data to points other than existing stations. Potentially there is a physical cause for the persistent error magnitude that could be exploited to improve the interpolations. Snowcover is a plausible mechanism. Figure8 shows that, in the central New York subregion, interpolation error tends to increase with snowcover. Over 25% of the variance in minimum temperature mae is explained by snow depth in this figure. Modifications to the interpolation procedure that account for such features, although beyond the scope of this work, warrant future investigation.

Finally, interpolation accuracy has the potential to be improved via the use of higher resolution RUC analyses. To examine this possibility 20 x 20 km RUC initializations were obtained for June, 2005. The 20km RUC also has a finer vertical resolution, with temperature and height information available at 25 hPa intervals. Maximum temperature interpolations benefited from the use of the higher resolution RUC initializations (Table 3). The mae associated with MQ_{RUC} was 88% of the 40 km value, a similar reduction in mae resulted for IDW_{RUC}. With the exception of IDW_{RUC}, maximum temperature interpolation biases were higher based on the 20km data. Smaller reductions (about 7%) in mae for minimum temperature were also realized using the 20 km RUC. Most of the improvement (for both maximum and minimum temperature) appears to come from the increase in horizontal resolution, rather than the increase in the number of vertical layers, since a similar reduction in mae occurs for MQ_{noelev}. Likewise, the MQ_{Ta} errors are not affected by the increase in vertical resolution.

Across much of the domain, the 40km grids encompass multiple observation sites. Thus, elevation, land use and meteorological features, that are captured by the model grid, are also represented by the station field. The finer resolution grid presumably captures influences that are not detectable in the station data and incorporates these into the interpolations. Operationally, the increase in accuracy provided by the finer grid needs to be weighed against computation time requirements, particularly if cross-validation is used to optimize interpolation parameters. Likewise future research examining the influence of even finer resolution model initializations is warranted, but beyond the scope of this work. The RUC model is currently operational at 13 km horizontal resolution.

5. Conclusions

Mesoscale meteorological model initializations such as those available from the RUC model show promise as a means to guide the interpolation of independent daily maximum and minimum temperatures. In cross-validation trials, the mean annual absolute errors associated with the 40-km-RUC-based interpolations averaged 5% smaller than those from elevation-detrended multiquadric interpolation of station data alone for maximum temperature and minimum temperature. These improvements are statistically significant at the 99% level, based on a two-sample t-test. During June data from a 20km RUC grid, reduced mae by 10% for maximum temperature and 16% for minimum temperature compared to two conventional station-based techniques. Additional evaluations for other regions and time periods using the outlined methods would strengthen these conclusions.

Adjustment for elevation based on the RUC lapse rates had little effect on the MQ_{RUC} interpolation of minimum temperature. The interpolation of maximum temperature was improved by the adjustment for elevation. Mean absolute errors were approximately 10% higher when elevation was ignored. Presumably, the vertical resolution of the RUC initializations used was too coarse to resolve the stratified vertical temperature profile that characterizes early morning soundings. However, under better mixed afternoon conditions, the model lapse rates provided a more robust means of incorporating the influence of elevation than empirical detrending.

The RUC-based interpolation errors displayed only subtle seasonality. Likewise large errors did not preferentially occur in association with specific topographic or static surface cover types, nor did they favor a particular synoptic pattern. However, there is some evidence that winter minimum temperature errors are related to snow depth. High autocorrelation is a feature of the daily errors in summer and to a lesser degree winter.

The results demonstrate that the accuracy of real time spatial interpolations of daily maximum and minimum temperature can be improved through the use of meteorological model analyses. It is likely that these results could be extended to produce longer-term climatological data interpolations. This application of the interpolation procedure would require the use of long-term gridded model data sets such as the North American Regional Reanalysis (NARR). Spatially, the 40 km RUC model and NARR are of a similar resolution. However, the three-hour temporal resolution of the NARR might be problematic as it will complicate the computation of instantaneous daily maximum and minimum temperature. The RUC-based approach also requires testing in regions with more diverse topography, such as the western United States before the outlined procedure can be applied outside of the Northeast.

Finally, the current results point to the utility of the interpolation procedure as a means by which station data can be quality controlled. Erroneous station data can be identified both as outliers in the bias field and through a unique pattern of adjacent bullseyes of large errors of opposite signs. Identification of such observation errors in real time would improve the accuracy of the interpolations. It may also offer improvements to existing spatial quality control methods that are applied prior to final archival of the climate data.

6. Acknowledgments

This work was supported by USDA Special Project NYC-125584. Partial support was also obtained from NOAA contract EA133E-02-CN-0033. The programming expertise of Laura Joseph facilitated our efforts. A subset of the RUC data was obtained from the Atmospheric Radiation Measurement (ARM) Program sponsored by the U.S. Dept. of Energy, Office of Science, Office of Biological and Environmental Research, Environmental Sciences Division.

7. References

Benjamin, S. G., J. M. Brown, K. J. Brundage, B. E. Schwartz, T. G. Smirnova, T. L. Smith and L. L. Morone,1998: *RUC-2 - The Rapid Update Cycle Version 2 NWS Technical Procedures Bulletin 448,* 18 pp. [available from NOAA/ERL Forecast Systems Laboratory, Boulder, CO or online http://maps.fsl.noaa.gov/ruc2.tpb.html]

Choi, J, U. Chung and J.I. Yun, 2003: Urban-effect correction to improve accuracy of spatially interpolated temperature estimates in Korea, *J. Appl. Meteorol.* **42**, 1711-1719.

Courault, D. and P. Monestiez, 1999: Spatial interpolation of air temperature according to atmospheric circulation patterns in southeast France, *Int. J. Climatol.*, **19**, 365-378.

Deutsch, C.V. and A.G. Journel, 1992: *GSLIB Geospatial Software Users Guide*. Oxford University Press, 340 pp.

Eischeid, J.K., P.A. Pasteris, H.F. Diaz, M.S. Plantico and N.J. Lott, 2000: Creating a serially complete, national daily time series of temperature and precipitation for the western United States, *J. Appl. Meteorol.*, **39**, 1580-1591.

Hutchinson, M.F., 1991: The application of thin plate smoothing splines to continent-wide data assimilation. *Data Assimilation Systems*, J.D. Jasper Ed., BMRC Research Report No. 27, Bureau of Meteorology, Melbourne, Australia, 104-113. [Available from Bureau of Meteorology, G.P.O. Box 1289K, Victoria 3001, Australia].

Jarvis, C.H. and N. Stuart, 2001a: A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part I: The selection of "guiding topographic and land cover variables, *J. Appl. Meteorol.* **40**, 1060-1074.

Jarvis, C.H. and N. Stuart, 2001b: A comparison among strategies for interpolating maximum and minimum daily air temperatures. Part II: The interaction between number of guiding variables and the type of interpolation method, *J. Appl. Meteorol.* **40**, 1075-1084.

Johnson, G.L., C. Daly, G.H. Taylor and C.L. Hanson, 2000: Spatial variability and interpolation of stochastic weather simulation model parameters, *J. Appl. Meteorol.*, **39**, 778-796.

Kurtzman, D. and R. Kadmon, 1999: Mapping of temperature variables in Israel: a comparison of different interpolation methods, *Climate Research*, **13**, 33-43.

Mesinger, F, G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jovic J. Woollen, E. Rogers, E. H. Berbery, M. B. Ek, Y. Fan, R. Grumbine, W. Higgins, H. Li, Y. Lin, G. Manikin, D. Parrish, and W. Shi, 2006: North American Regional Reanalysis. *Bull. Amer. Meteorol. Soc.*, 87, 343–360.

Meyers, D.E. 1994: Spatial interpolation: An overview. *Geoderma*, **62**, 17-28.

Press, W.H, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, 1992: *Numerical Recipes in FORTRAN*, Cambridge, 963 pp.

Price, D.T., D.W. McKenney, I.A. Nalder, M.F. Hutchinson and J.L. Kesteven, 2000: A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data, *Agric. For. Meteorol.*, **101**, 81-94.

Nuss, W.A. and D.W. Titley, 1994: Use of multiquadric interpolationfor meterological objective analysis*, Mon. Wea. Rev.*, **122**, 1611-1631.

Rigol, J.P., C.H. Jarvis and N. Stuart, 2001: Artificial neural networks as a tool for spatial interpolation, *Int. J. Geographical Information Science*, **15**, 323-343.

Smirnova, T.G., J.M. Brown, S.G. Benjamin, and D. Kim, 2000: Parameterization of cold-season processes in the MAPS land-surface scheme. *J. Geophys*. Res., 105, D3, 4077-4086.

Snell, S.E., S. Gopal and R.K. Kaufmann, 2000: Spatial interpolation of surface air temperatures using artificial neural networks: Evaluating their use for downscaling GCMs, *J. Climate*, **13**, 886-895.

Sokolov, S. and S.R. Rintoul, 1999: Some remarks on interpolation of nonstationary oceanographic fields, *J. Atmos. Ocean Technol*., **16**, 1434-1449.

Thornton, P.E.. S.W. Running and M.A. White, 1997: Generating surfaces of daily meteorological variables over large regions of complex terrain, *J. Hydrol*., **190**, 214-251.

Willmott, C.J. and K. Matsuura, 1995: Smart interpolation of annually averaged air temperature in the United States, *J. Appl. Meteorol.,* **34**, 2577-2586.

Table Legends

Table 1. List of interpolation analyses conducted.

Table 2. Average daily bias , mean absolute error (mae), and root mean square error (rmse) by interpolation method for maximum and minimum temperature. Bold values are the smallest among the methods.

Table 3. Comparison of average daily bias, mean absolute error (mae), and root mean square error (rmse) during June, 2005 based on 20km and 40 km RUC analyses.

Figure Captions

Figure 1. Coordinates of 40 km RUC grid points (black dots) and location of morning observation time Cooperative Network Stations (gray circles). The gray boxes highlight subregions in which an analysis of the causes of large station errors was conducted.

Figure 2. Flow chart of RUC temperature interpolation procedure. The numbers correspond above each box refer to the sequence of steps referred to in the text. The panels to the left illustrate the interpolation steps. The letter A denotes the a station location. Likewise the word bias is associated with the difference between the interpolated RUC temperature and the observed temperature. Thus, the word BIAS also denotes station locations. In the final pattern the letter T denotes the adjusted RUC temperatures at each grid.

Figure 3. Monthly mean error (solid) and root mean square error (dashed) associated with MQ_{RUC} (squares), IDW_{RUC} (circles) and IDW_{obs} (triangles) interpolation of daily a) maximum and b) minimum temperature.

Figure 4. Average seasonal MQ_{RUC} interpolation errors for June, July and August minimum temperature. Areas with absolute errors ≥ 1°C are shaded. Contours are at 0.5°C intervals.

Figure 5. Observed daily maximum and minimum temperatures at Jonesboro, (solid) and Machias, (dotted) Maine during July 2005.

Figure 6. Lag correlation for daily maximum (solid) and minimum (open) temperature mae (squares) and bias (circles) for a) Central New York (JF), b) southern New England (JF), c) Central New York (JJA) and southern New England (JJA). Correlations above the dashed horizontal line are significant ($\alpha = 0.05$) assuming Gaussian data.

Figure 7. RUC 2m temperature analysis (contours) superimposed with a) minimum temperatures at Cooperative Network stations in central New York (bold values) and b) maximum temperatures at Cooperative Network stations in northern Virginia (bold values).

Figure 8. Minimum temperature mae at Cooperative Network stations in central New York (using IDW_{RUC}) as a function of snow depth.

Table 1. List of interpolation analyses conducted.

Acronym	Interpolation Method	Data	Remarks
MQ_{RUC}	Multiquadric	RUC initialization	
IDW_{RUC}	Inverse distance*	RUC initialization	
MQ_{obs}	Multiquadric	Station observations	Adjusted for elevation
IDW_{obs}	Inverse distance*	Station observations	Adjusted for elevation
MQ_{noelev}	Multiquadric	RUC initialization	Lapse rate adjustment omitted
MQ_{TA}	Multiquadric	RUC initialization	Terrain-based bias adjustment

* Distance is squared in the IDW procedures

Table 2. Average daily bias, mean absolute error (mae), and root mean square error (rmse) by interpolation method for maximum and minimum temperature. Bold values are the smallest among the methods.

	Maximum Temperature			Minimum Temperature		
Method	Bias (°C)	Mae (°C)	Rmse (°C)	Bias (°C)	Mae(°C)	Rmse (°C)
MQ_{RUC}	**-0.014**	1.16	1.69	0.007	1.47	2.05
IDW_{RUC}	-0.018	**1.08**	**1.57**	0.014	**1.37**	**1.90**
MQ_{noelev}	-0.016	1.29	1.82	**0.006**	1.48	2.04
MQ_{obs}	-0.016	1.15	1.68	0.010	1.49	2.09
IDW_{obs}	-0.062	1.14	1.64	-0.015	1.43	1.98

Table 3. Comparison of average daily bias , mean absolute error (mae), and root mean square error (rmse) during June, 2005 based on 20km and 40 km RUC analyses.

	Maximum Temperature			Minimum Temperature		
Method	Bias (°C)	Mae (°C)	Rmse (°C)	Bias(°C)	Mae(°C)	Rmse (°C)
20 km MQ_{RUC}	-0.047	1.16	1.72	-0.001	1.32	1.79
40 km MQ_{RUC}	-0.025	1.33	1.87	0.007	1.41	1.90
20 km IDW_{RUC}	-0.030	1.03	1.50	0.024	1.20	1.64
40 km IDW_{RUC}	-0.045	1.21	1.72	0.005	1.28	1.73
20 km MQ_{noelev}	-0.049	1.30	1.85	-0.003	1.32	1.77
40 km MQ_{noelev}	-0.027	1.45	1.99	0.005	1.40	1.88
20 km MQ_{Ta}	–0.068	1.20	1.77	0.001	1.32	1.80
40 km MQ_{Ta}	–0.041	1.36	1.92	0.008	1.41	1.90

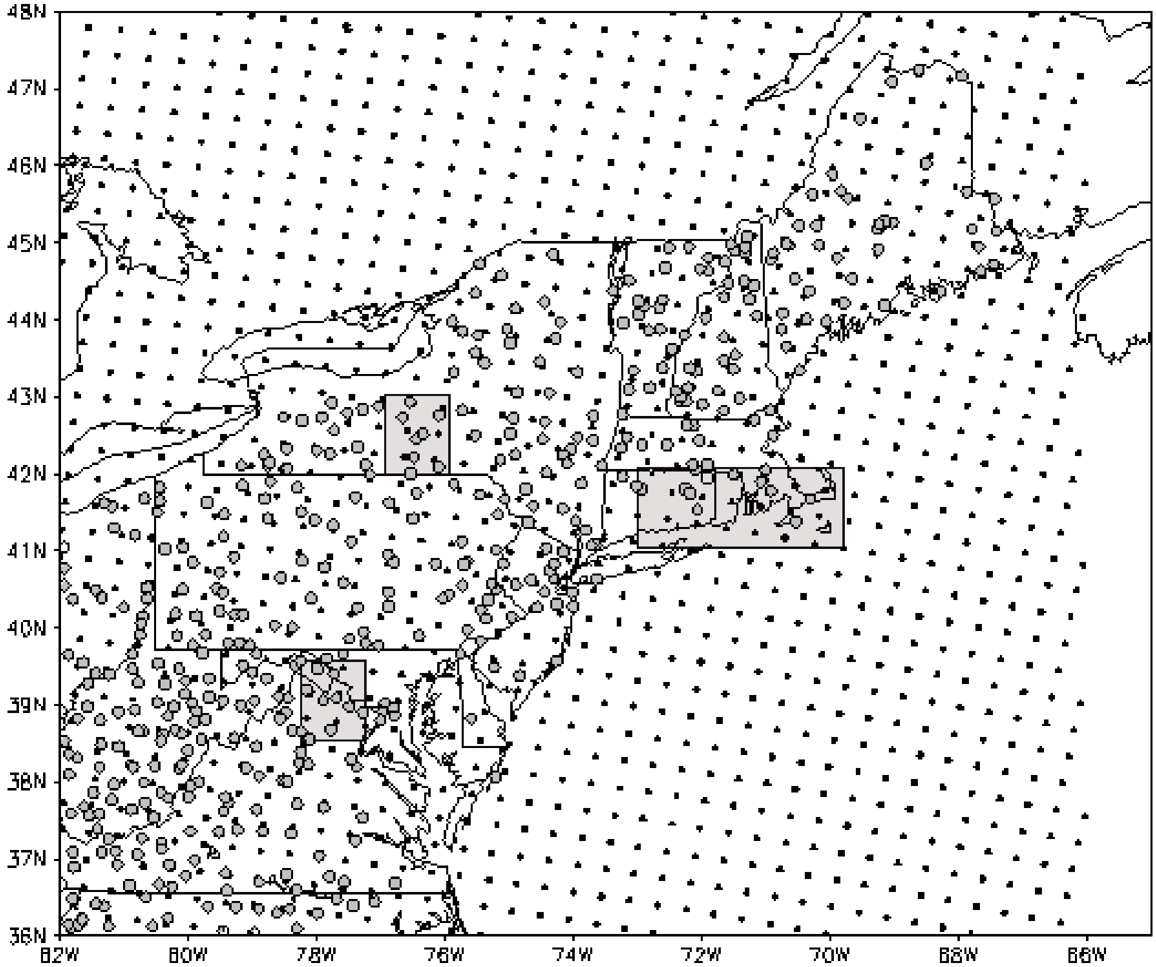

Figure 1. Coordinates of 40 km RUC grid points (black dots) and location of morning observation time Cooperative Network Stations (gray circles). The gray boxes highlight subregions in which an analysis of the causes of large station errors was conducted.

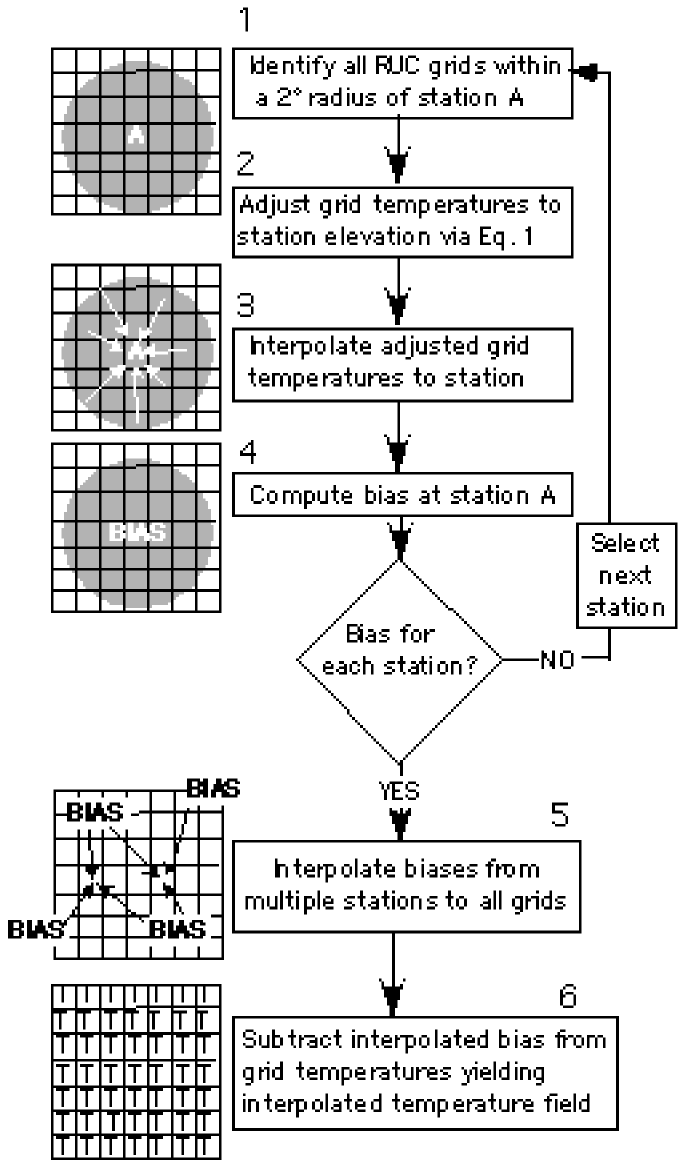

Figure 2. Flow chart of RUC temperature interpolation procedure. The numbers correspond above each box refer to the sequence of steps referred to in the text. The panels to the left illustrate the interpolation steps. The letter A denotes the a station location. Likewise the word bias is associated with the difference between the interpolated RUC temperature and the observed temperature. Thus, the word BIAS also denotes station locations. In the final pattern the letter T denotes the adjusted RUC temperatures at each grid.

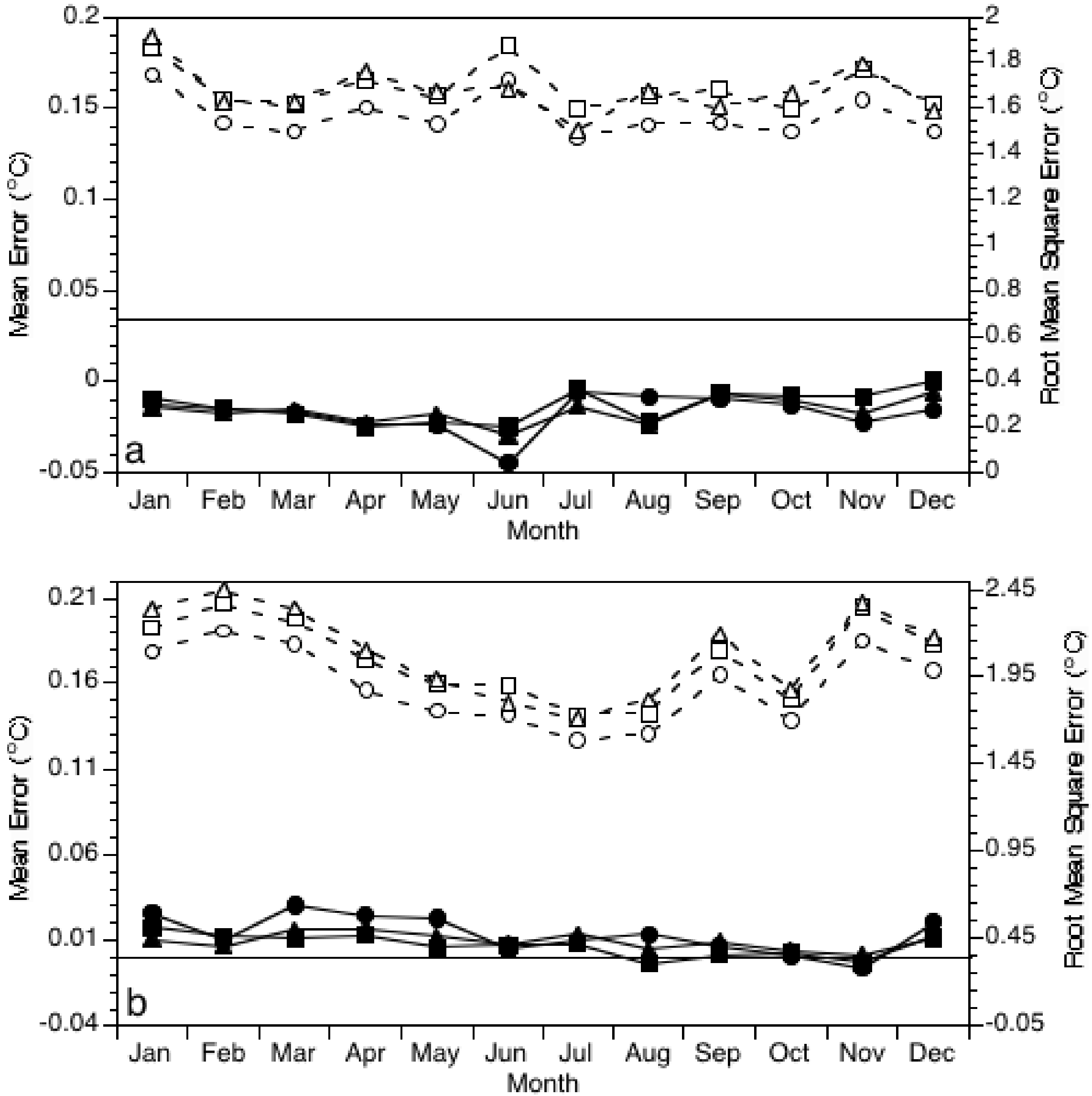

Figure 3. Monthly mean error (solid) and root mean square error (dashed) associated with MQ_{RUC} (squares), IDW_{RUC} (circles) and MQ_{obs} (triangles) interpolation of daily a) maximum and b) minimum temperature.

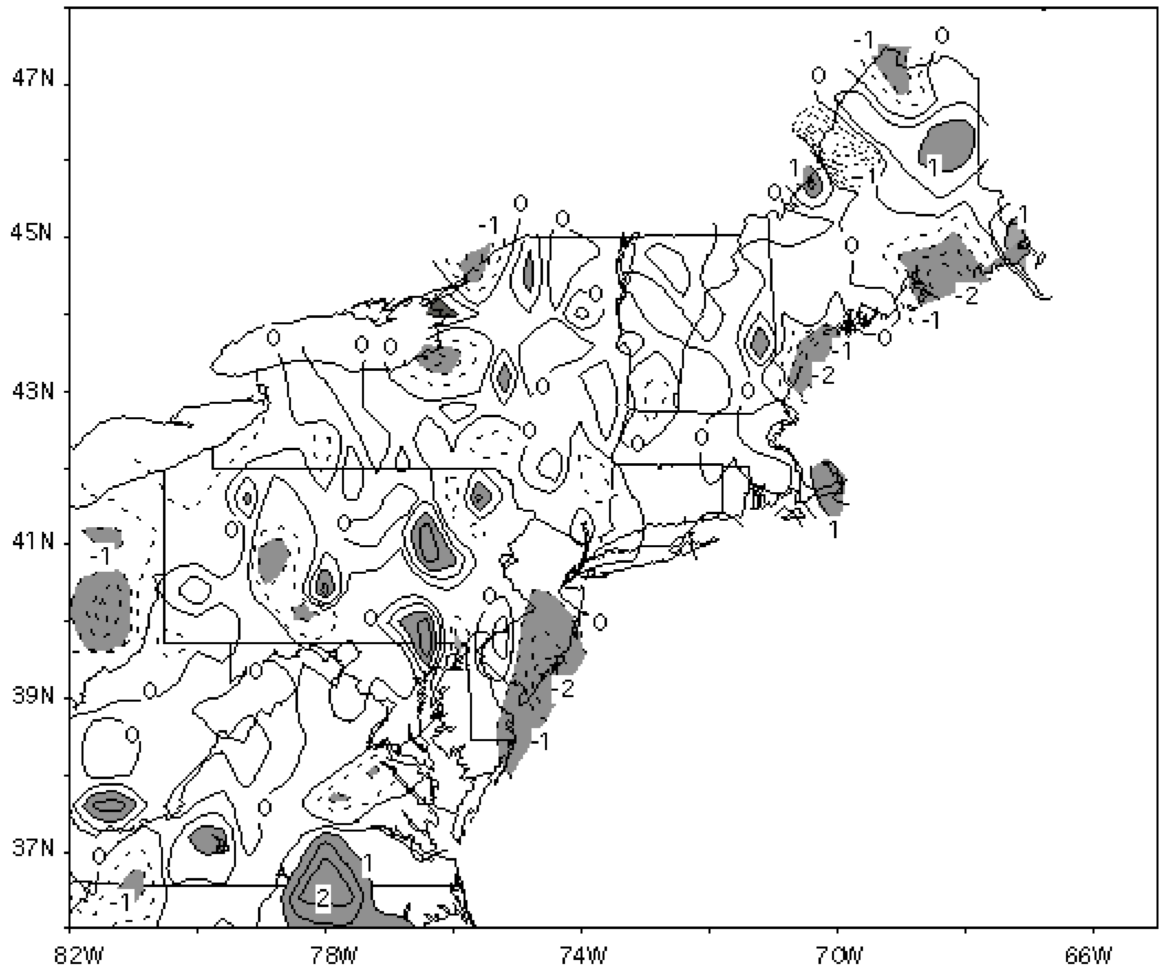

Figure 4. Average seasonal MQ_{RUC} interpolation errors for June, July and August (JJA) minimum temperature. Areas with absolute errors ≥ 1°C are shaded. Contours are at 0.5°C intervals.

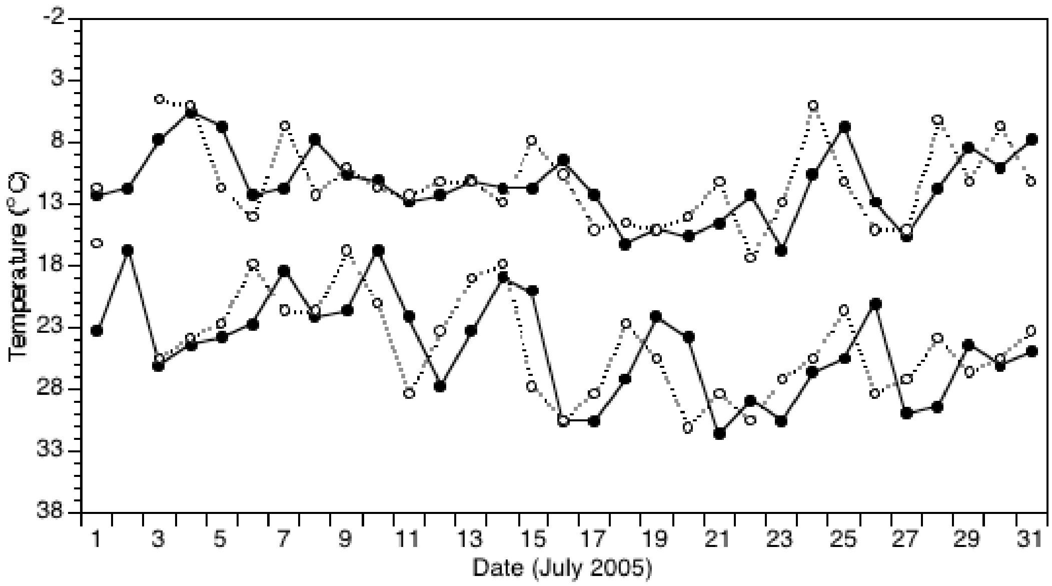

Figure 5. Observed daily maximum and minimum temperatures at Jonesboro, (solid) and Machias, (dotted) Maine during July 2005.

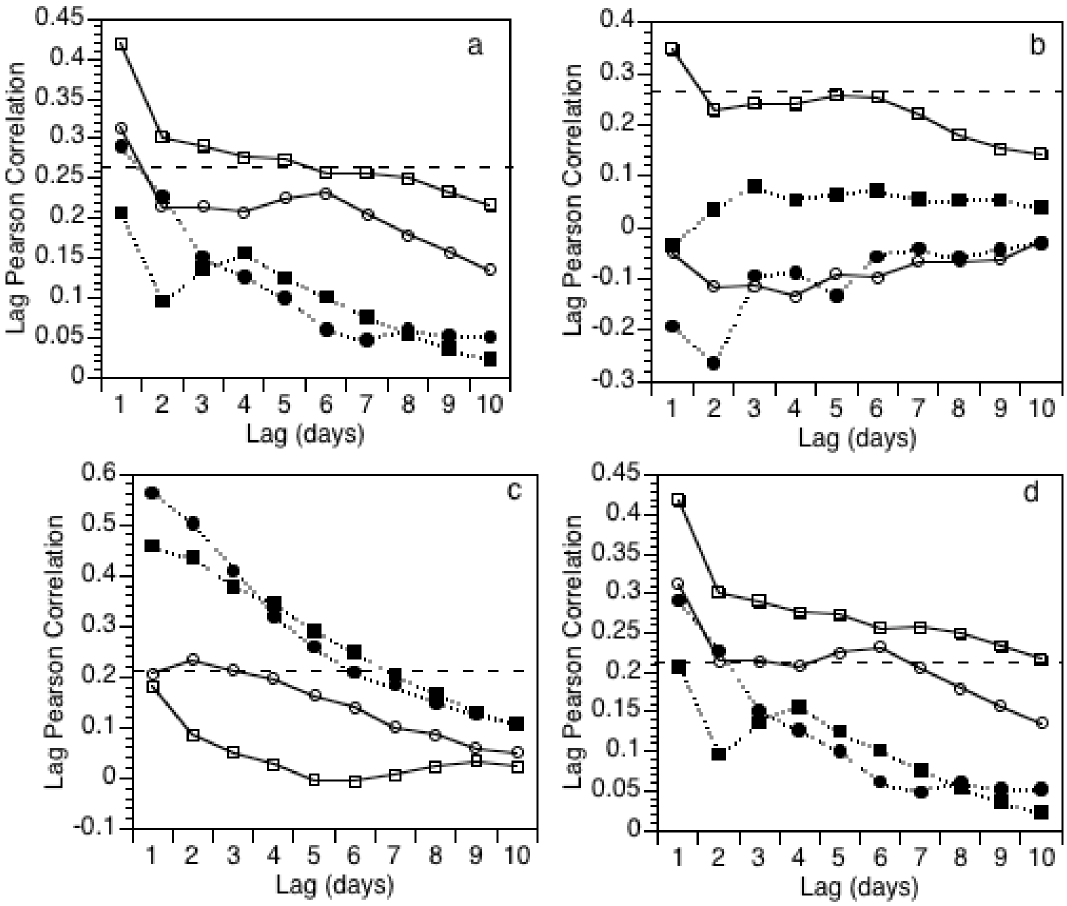

Figure 6. Lag correlation for daily maximum (solid) and minimum (open) temperature mae (squares) and bias (circles) for a) Central New York (JF), b) southern New England (JF), c) Central New York (JJA) and southern New England (JJA). Correlations above the dashed horizontal line are significant ($\alpha = 0.05$) assuming Gaussian data.

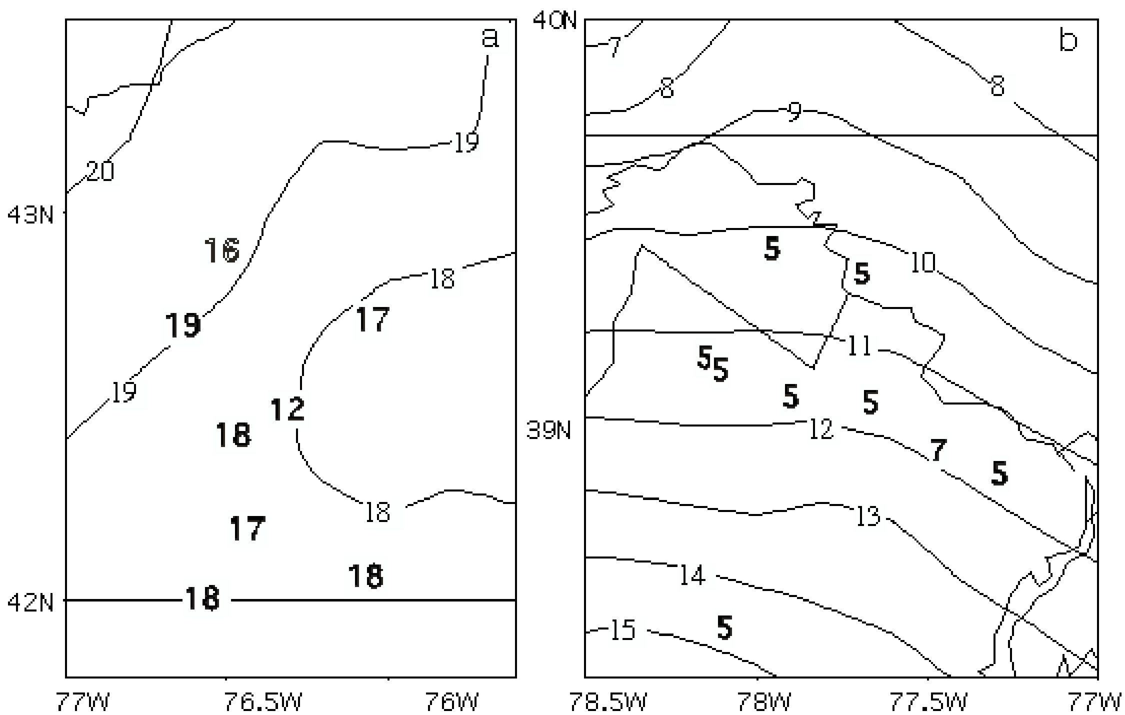

Figure 7. RUC 2m temperature analysis (contours) superimposed with a) minimum temperatures at Cooperative Network stations in central New York (bold values) and b) maximum temperatures at Cooperative Network stations in northern Virginia (bold values).

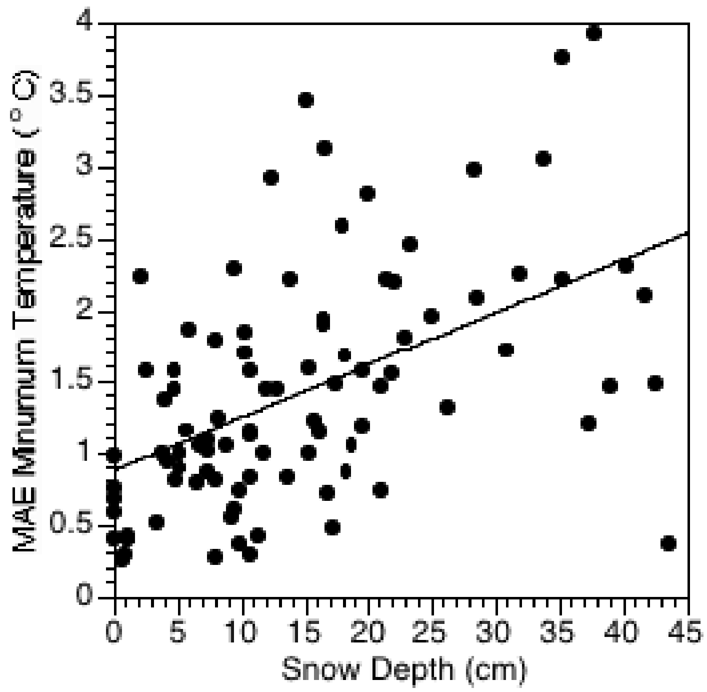

Figure 8. Minimum temperature mae at Cooperative Network stations in central New York (using MQ_{RUC}) as a function of snow depth.