Generating Scenarios of Local Surface Temperature Using Time Series Methods

ROBERT S. CHEN

Alan Shawn Feinstein World Hunger Program, Brown University, Providence, Rhode Island

PETER J. ROBINSON

Department of Geography, University of North Carolina, Chapel Hill, North Carolina

(Manuscript received 27 January 1990, in final form 5 February 1991)

ABSTRACT

A method for creating scenarios of time series of monthly mean surface temperature at a specific site is developed. It is postulated that surface temperature can be specified as a linear combination of regional and local temperature components, where the regional component is an areal-average free tropospheric temperature such as may be produced by general circulation model climate simulations, and the local component is a response to purely local surface effects. Three models were tested for the North Carolina area, using weighted least squares linear regression techniques. The first related free tropospheric temperature directly to the site-specific local temperature, while the second added specifically local parameters as independent variables. For the four stations tested, this second model increased the explained variance from about 60% to near 70%. The third model incorporated a local energy and moisture balance process model to specify interactions between local variables, but yielded no significant increase in explanation. Scenarios of time series of future temperatures at a specific site can be obtained by using the statistical regression models and incorporating a random error term based on the regression residuals. For any postulated areal mean tropospheric temperature change this procedure yields a time series that retains the variability characteristics appropriate to the site-specific series.

1. Introduction

There is increasing concern with the potential impacts of climatic change. Most impacts occur and must be assessed on relatively small space and time scales. It is necessary, therefore, to develop and test methods of creating scenarios of altered local climate. A scenario is a set of plausible future climatic conditions, developed from sound scientific principles, but with no suggestion of a forecast attached (National Research Council 1983). Most information about future climates comes from general circulation models (GCM), but their present resolution is too coarse to provide the required detail, and this is likely to remain the case into the foreseeable future. A potentially fruitful approach is to combine current GCM outputs with the observational record of surface and upper atmosphere conditions to provide the required scenarios (Lamb 1989).

Climate scenarios developed to date have used various approaches to the above problem (Robinson and Finkelstein 1989). It has been most directly faced by Kim et al. (1984), who essentially used area averages of monthly mean surface observations to develop statistical relationships for individual stations within the area. Wilks (1989) modified and extended this approach for daily values of three surface weather variables and utilized a Monte Carlo procedure to improve the statistical realism of precipitation patterns. Other workers have taken GCM grid-cell outputs, assumed that they affect the whole grid-cell area uniformly, and simply "added" the changes to the present local climates (Cohen and Allsop 1988; Smith and Tirpak 1988). This approach automatically assumes that the variability about the mean does not change as the mean climate changes, although the validity of this assumption is not clear (Mearns et al. 1984). Other workers (e.g., Gleick 1987) have assessed potential impacts by postulating hypothetical scenarios, while others (e.g., Webb and Wigley 1985) have used analogues with paleoclimatic conditions. A similar type of problem, but in the context of long-range weather forecasting, has been addressed by Klein (1983) and Klein and Walsh (1983). Their approach has been to use the broad-scale forecasts of upper air conditions and attempt to relate these to local surface conditions using statistical techniques. Some recent climate scenario development efforts (Wigley et al. 1990) have used a similar conceptual framework.

The present work also builds upon the concepts introduced by Klein (1983). It develops scenarios of time series of monthly mean local air temperature from re-

Corresponding author address: Dr. Peter Robinson, Dept. of Geography, 203 Saunders Hall, University of North Carolina, Chapel Hill, NC 27599-3220.

gionally averaged free air temperatures, such as might be developed from information generated by GCMs. The underlying hypothesis is that the surface temperature at a specific site is composed of a regional free-air component and a local component controlled by surface characteristics. These components can be linked using suitable statistical and process model techniques. A hierarchy of methods of increasing complexity are here developed to address this linkage problem. Since an important consideration in scenario development for impact assessment is the need for reliable results without burdensome data or computational requirements, consideration must be given to the degree of explanation obtained at each stage. No attempt is made here to develop a formal "parsimonious model," but the degree of explanation at each stage is explicitly considered. Each of the methods is tested using data from a region centered on the state of North Carolina.

2. Theory

a. General

It is generally accepted that most climate elements, and temperature in particular, are spatially more conservative once the influence of the surface is removed. Thus, it is reasonable to assume that GCM grid cell output, although most often presented in terms of surface temperatures, is more akin to conditions above the planetary boundary layer. It can also be postulated that the actual local surface temperature is a function of both the regional temperatures created by these upper air conditions and the local influences of surface characteristics. The way in which these scales are linked depends on the synoptic situation. When the advection of air through the local site is weak, the atmospheric boundary layer comes into thermal equilibrium with the energy fluxes and boundary conditions that prevail at the site. Local conditions then dominate. When advection is strong, air from the region outside the site is transported through it and may be modified only slightly by local climatological conditions. Regional influences then prevail. With moderate advection or an unstable atmospheric boundary layer, there may be significant mixing of local and regional air, and intermediate conditions prevail. A site may experience all three situations during a month, and the monthly mean temperature is a combination of them all.

Our basic hypothesis is that monthly mean temperatures are a *linear* combination of local and regional influences. Thus, we may write:

$$T_s = \Theta_r T_r + \Theta_l T_l + \mu_t \qquad (1)$$

where T_s is the site-specific monthly average surface temperature, T_r and T_l are the monthly average regional and local temperature components, the Θ's are weighting factors to be estimated and μ_t is an error term. Here T_s may be regarded as a combination of the temperature of the air dominated by regional processes and the temperature dominated by local processes, the relative contributions being specified by the Θ, while any short-term nonlinearities are incorporated into the error term.

The regional temperature component can be ascertained directly as the tropospheric free-air temperature T_a. This can be averaged over an area similar to that of a GCM grid cell and used as a surrogate for the grid-cell averaged surface temperatures currently produced by GCMs. It can be postulated that GCMs produce estimates of T_a that are more reliable than those of T_s, since the former is controlled primarily by the hemispheric and global processes that are most explicitly modeled in GCMs. The T_a is not as strongly influenced by the subgrid scale processes that may be inadequately parameterized in GCMs. This present approach differs from that of Kim et al. (1984), who used a regional temperature obtained by averaging surface, not upper-air, temperatures from a group of stations.

There is no comparable approach to the determination of the local temperature component. Although it is possible to treat T_l as a residual between a site-specific surface temperature and the areal-averaged value from a group of neighboring stations, this provides little useful information in the context of a changed climate. A more useful approach is to incorporate explicitly those local variables, such as incoming solar radiation or surface wind speed, which must have an effect on local temperatures. The methods of incorporation explored here produce the three statistical models considered below. Throughout this work T_s is treated as being represented by a single surface observing station.

b. Model I

The simplest relationship between the regional temperature and a site-specific surface temperature which can be formulated assumes that $\Theta_l = 0$. Thus our, first model, Model I, is a linear relationship of the form:

$$T_s = \Theta_a T_a + \mu_1 \qquad (2)$$

where Θ_a is a transfer function relating the tropospheric free-air temperature to the local surface temperature and μ_1 is an error function. For the T_a and T_s time series, μ_1 can be treated either as a random variable implying a time-independent relationship between T_a and T_s, or as a time-varying stochastic function that allows for serial autocorrelation in the relationship between the two time series. In the latter form the error term acts as a surrogate for the local temperature component of (1). Thus, while Model I does not explicitly incorporate both regional and local effects, it provides the baseline model reflecting primarily the contribution of the regional component. This approach has been used in statistical forecasts of monthly T_s from 70 kPa geopotential height fields on spatial scales similar to those used in GCMs (Klein 1983).

c. Model II

Extension of Model I requires the explicit incorporation of the local temperature component. This must involve the local variables that have an influence on local temperatures. In Model II this is accomplished using a multiple regression approach such that

$$T_l = \sum_{i=1}^{n} \Theta_i L_i + \mu_l \tag{3}$$

where the L_i represent the individual variables incorporated and the Θ_i their individual weights.

The local variables chosen for inclusion must either directly contribute to the surface energy balance or be indicative of the synoptic conditions. Five variables were selected for investigation, based on their relevance and data availability. These were the solar radiation flux at the earth's surface, the amount of sky cover, the precipitation amount, the surface wind speed, and the surface pressure. Solar radiation is the key input to the surface energy balance and therefore has a great influence on surface temperature. Sky cover directly influences the incoming solar radiation, but also affects outgoing longwave radiation and thus has a major influence on nighttime temperatures. Precipitation influences soil moisture, which in turn affects the surface albedo, soil heat flux and heat capacity, and the latent and sensible heat fluxes. It is also likely to be related to cloud amount. Surface wind speed directly affects near-surface turbulence and thus the sensible and latent heat fluxes and the surface temperature. It is also likely to be related to the synoptic situation, thus also being related to the cloud amount, solar radiation, and precipitation. Surface pressure is included as a variable indicative of the synoptic situation since, in summer for example, high pressure is usually associated with a westward shift of the Bermuda high which creates hot, dry and clear conditions in the study area.

Inclusion of these five local variables, therefore, leads to Model II:

$$T_s = \Theta_a T_a + \sum_{i=1}^{5} \Theta_i L_i + \mu_2. \tag{4}$$

d. Model III

The major drawback of the Model II approach is that it assumes linear statistical relationships between the L_i and T_s. It cannot account for likely nonlinear interaction effects among the climate elements. This can be overcome, at least partially, by explicitly developing and incorporating a local process-based climate model. This process model was constrained to have output in a form similar to the local climate variables already introduced, so that it could be incorporated into the statistical analysis. Thus the resultant hybrid model, Model III, can be formulated as

$$T_s = \Theta_a T_a + \sum_{i=1}^{5} \Theta_i L_i + \Theta_m T_m + \mu_3 \tag{5}$$

where T_m is the model-generated temperature and Θ_m is the model weighting factor to be estimated.

The local climate model (LCM) used here was developed specifically to provide the term $\Theta_m T_m$ in (5). It was not intended to provide an independent prognostic model of surface temperatures. Further, it was constrained to use relatively readily accessible input data. Full details of the LCM are given in Chen (1987), and only a brief outline is given here. The LCM uses energy and moisture budget calculations in two linked "boxes," one for the near-surface air layer, the other for the surface soil layer, to determine surface temperatures from a prescribed regional temperature. The top of the atmospheric box is prescribed at 85 kPa, allowing direct connection with the layer for which T_a is determined. The soil layer is 2 m thick, with an assumed water content at field capacity of 0.2 m. The horizontal dimensions are 1 km × 1 km. Vegetation and soil conditions are assumed to be uniform within the box, keeping the model and its data requirements relatively modest. For similar reasons, the effects of snow on surface albedo and on the energy and moisture fluxes are ignored, not an unreasonable assumption in the test area when monthly mean conditions are of prime concern.

The general form of the energy and moisture budget equations used, patterned after those of Sellers (1965), are

$$\delta B_s = K_s + L_{as} - L_{sa} - \mathrm{LE}_{sa} - H_{sa} + G_{gs} \tag{6a}$$

$$\delta B_a = L_{sa} - L_{as} + L_{ra} - L_{ar} + \mathrm{LE}_{sa} - \mathrm{LE}_{ar} + \mathrm{LC}_{as} + H_{sa} - H_{ar} \tag{6b}$$

$$\delta S_s = P_{rs} + C_{as} - E_{sa} - O \tag{6c}$$

$$\delta S_a = E_{sa} - E_{ar} - C_{as} \tag{6d}$$

where B and S are the energy storage and moisture content of the box, K is absorbed solar radiation, L is longwave radiation, LE, H, and G are the latent, sensible and ground heat fluxes, LC is the latent heat released by condensation of excess water vapor, E is evaporation, P is precipitation, C is condensation, and O is runoff. Subscript "s" indicates the soil surface box, "a" the atmospheric box, "g" refers to conditions below the surface box, "r" to conditions above the atmospheric box. Double subscripts indicate the flux directions. Each component of (6) is derived from one or more of the five variables already used in Model II, using techniques readily available in the literature (Chen 1987). Precipitation, incoming solar radiation, cloudiness, surface pressure, and wind speed are prescribed exogenous variables. All except the latter are regarded as functions of the regional climate. Wind speed for near-surface conditions is used only to calculate the sensible and latent heat fluxes, using a method similar to that of Deardorff (1978). No advection into or out of the box is allowed.

Equations (6) are solved simultaneously for both boxes, with the soil and air temperatures and moisture contents as the prognostic variables. Since the model must simulate behavior on a monthly time scale, seasonal variations in the exogenous variables are explicitly incorporated. However, the response of the model to these variables is not linear on a daily time scale, since they tend to be correlated and affect both energy and moisture fluxes simultaneously. Consequently, daily time steps are used, and monthly statistics computed from them.

The LCM in essence produces estimates of the difference between the free tropospheric temperature and the local surface temperature. As such, it can be used directly as another climatic "parameter" in the formulation of (5).

3. Data

Model performance was tested using upper air and surface data centered on North Carolina (Fig. 1). The hypsometric equation was used to determine the daily mean temperature of the 85–30 kPa layer for the five upper air stations in the region. Data were obtained from the TD-6200 series Upper Air Digital Files (National Climatic Data Center 1986). The 85 kPa surface, corresponding to a geopotential altitude of approximately 1.5 km (National Oceanic and Atmospheric Administration et al. 1986), is above the normal planetary boundary layer, whereas the 30 kPa level, at about 9.2 km, is below the tropopause. Thus, the mean temperature of the 85–30 kPa layer approximates the free air temperature. The arithmetic mean for the 5 stations was used to determine the monthly mean T_a, representing the free air temperature for a region covering roughly 5° lat by 10° long.

For surface conditions, four stations were used. Three, Sterling, Virginia, and Greensboro and Cape Hatteras, North Carolina, were colocated with the radiosonde sites, while the fourth, Raleigh-Durham, North Carolina, was selected to allow analysis of a site without radiosonde data but with a full range of surface meteorological observations. Daily temperature data were obtained from a local quality controlled dataset (Wiser 1983) derived from the National Climatic Data

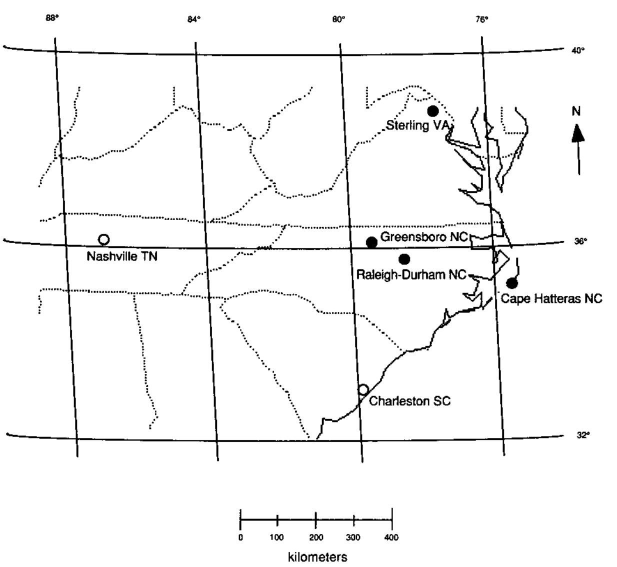

FIG. 1. Study area and stations used. Hollow circles, radiosonde data only; shaded circles, surface data only, solid circles, both radiosonde and surface data.

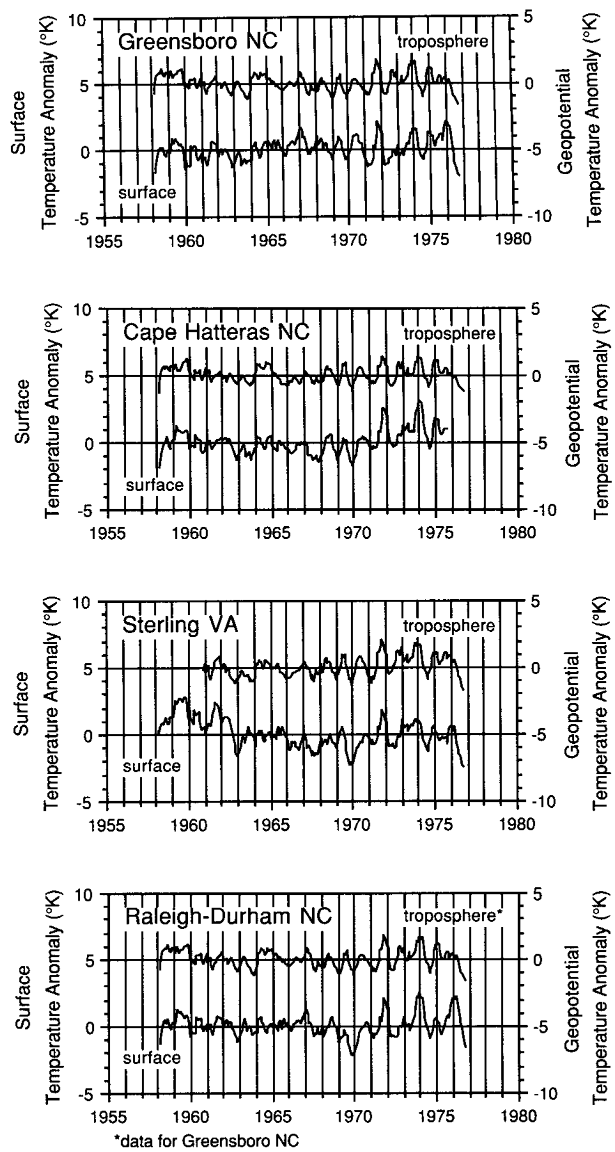

FIG. 2. Comparison of smoothed (5-month running mean) surface and tropospheric temperature anomalies at four stations, 1958–76.

TABLE 1. Correlation coefficients between monthly mean tropospheric temperature at the five radiosonde stations, 1958–76. Values above the diagonal are for actual temperatures, those below are for anomalies. All values are significant at the 99% confidence level.

Station	Greensboro	Cape Hatteras	Sterling	Charleston	Nashville
Greensboro	—	.9947	.9948	.9936	.9936
Cape Hatteras	.9414	—	.9916	.9946	.9821
Sterling	.9263	.8916	—	.9856	.9857
Charleston	.9278	.9360	.8173	—	.9859
Nashville	.9212	.7962	.8391	.8453	—

Center's Climatological Data archive. Other surface meteorological parameters were obtained from the SOLMET data (National Climatic Data Center 1978, 1979). Monthly averages were calculated from the daily values.

For all observations, monthly anomalies were obtained using 1958–76 monthly means. The use of anomalies removes the influence of the strong seasonal cycle but does not remove the short-term temporal autocorrelation which is a feature of temperature time series. The 1977–86 data were reserved and used for testing the methods.

The tropospheric and surface temperature anomalies at the 4 surface stations (Fig. 2) display no regionwide long-term trends in either series. For the tropospheric series, comparisons among the five radiosonde stations (Table 1) indicate that the individual series differ in detail but display a high degree of interstation correlation. This strongly suggests that an arithmetic average provides an adequate regional-tropospheric temperature.

4. Results

a. Model I

The most straightforward way of estimating the coefficients of Model I is through an ordinary least squares (OLS) linear regression. Here the error term is treated simply as a random variable. The result, using Greensboro as an example, is shown in Fig. 3. For each station, using the Pearson correlation coefficient, the variance explained by Model 1 exceeds 40% (Table 2). This result, however, must be treated with caution because the OLS regression approach treats all observations as independent, and this is unlikely to be the case with temperature. Monthly mean climate data, even when dealing with anomalies, frequently have strong serial autocorrelations.

An alternative approach, therefore, is to use weighted least squares (WLS) regression, which produces statistically more efficient estimates of the coefficients when autocorrelation is present (Priestley 1981). Thus the regression residuals, the error term in (2), are treated as time-dependent processes with "autoregressive" (AR) terms, i.e.,

$$\mu_t = \alpha_1\mu_{t-1} + \cdots + \alpha_p\mu_{t-p} + Z_t, \qquad (7)$$

where $\mu_{t-1} \ldots \mu_{t-p}$ represent previous error terms with their corresponding coefficients $\alpha_1 \ldots \alpha_p$, and Z_t is a random variable with mean zero and constant variance. The squared multiple correlation coefficient includes the contribution of the AR model of the residuals in explaining the variance in T_s.

WLS regressions were computed using standard procedures, including backward stepwise elimination, to find the most parsimonious AR model (SAS Institute 1984). The WLS method increases the value of the squared multiple correlation coefficients by 5% or more at the four stations (Table 2). In addition, the estimated values of Θ_a change by up to 10% (Table 3). For Greensboro, for example, this means that the WLS prediction of T_s would be 4% higher than the OLS prediction of T_s, given the same time series of T_a as the independent variable.

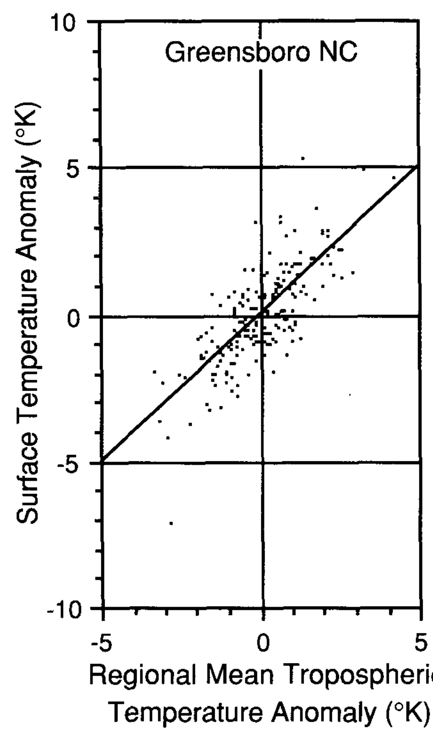

FIG. 3. Monthly anomaly of surface temperature at Greensboro versus the monthly anomaly of regional mean tropospheric temperature. The regression line is based on ordinary least squares estimates.

TABLE 2. The squared multiple correlation coefficient (r^2), Akaike's Information Criterion (AIC), and Schwarz's Bayesian Information Criterion (BIC) for each of the models relating T_s to T_a for each station tested.

	Greensboro	Cape Hatteras	Sterling	Raleigh-Durham
		r^2		
Direct regression (OLS)	.572	.602	.402	.600
Direct regression (WLS)	.626	.675	.522	.659
With local variables	.712	.706	.570	.704
With process model	.727	.708	.575	.714
		AIC		
Direct regression (OLS)	671	589	780	648
Direct regression (WLS)	647	568	735	620
With local variables	595	539	717	594
With process model	584	540	716	588
		BIC		
Direct regression (OLS)	678	595	787	655
Direct regression (WLS)	664	568	752	641
With local variables	625	573	744	624
With process model	619	577	746	622

It is possible to compare the results of the two methods using Akaike's Information Criterion (AIC) (Priestley 1981) or Schwarz's Bayesian Information Criterion (BIC) (Katz and Skaggs 1981). These indicate the information content of a model by incorporating both the goodness of fit using the mean squared error and the model parsimony using a penalty function for the number of parameters included. They differ only in the form of the penalty function. Although the former is more generally used, Katz and Skaggs (1981) indicate that the latter may be more appropriate for some meteorological variables. According to both criteria, the WLS method, with its lower mean square error, yields lower (more significant) AIC and BIC values than the OLS method (Table 2).

The WLS method provides a means of predicting T_s from T_a, using (2) and (7) in combination with the parameter estimates in Table 3. Again using Greensboro as a typical example, the results for the 1977–84 temperature anomalies indicate that the main characteristics of the series are retained (Fig. 4a). A linear regression between the actual and predicted anomalies (Fig. 4b) yields

$$P = -0.150 + 0.545A \quad (r^2 = 0.563),$$

where P is the predicted and A the actual temperature anomaly. Here P has lower variability than A because of the omission of the random error term Z_t in (7). However, P represents the best (unbiased) predictors

TABLE 3. Comparison of regression equation results of observed monthly anomalies of surface-air temperature as a function of regional tropospheric monthly-mean temperature, using the ordinary least squares and weighted least squares methods. Standard errors are given in parentheses.

	Greensboro, NC	Cape Hatteras, NC	Sterling, VA	Raleigh-Durham, NC
		Ordinary least squares		
Intercept	−0.003(0.070)	−0.044(−0.064)	−0.003(0.088)	−0.003(0.066)
Θ_a	1.006(0.058)	0.969(0.054)	0.906(0.074)	1.014(0.055)
		Weighted least squares		
Intercept	−0.016(0.144)	−0.043(0.156)	0.002(0.203)	−0.012(0.098)
Θ_a	1.045(0.053)	0.878(0.047)	0.872(0.064)	1.031(0.049)
Θ_1	0.232(0.064)	0.317(0.063)	0.296(0.063)	0.265(0.063)
Θ_2	—	0.159(0.063)	—	—
Θ_3	—	—	0.142(0.063)	—
Θ_4	0.152(0.063)	—	—	0.141(0.063)
Θ_6	0.169(0.063)	—	0.179(0.063)	0.107(0.064)
Θ_{10}	—	0.161(0.063)	—	—
Θ_{12}	—	—	—	−0.141(0.063)

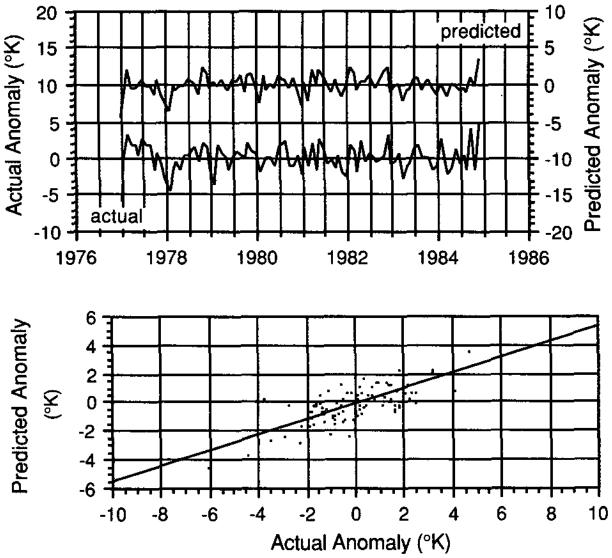

FIG. 4. Comparison of actual and predicted monthly anomalies of surface air temperature for 1977–84 at Greensboro: (a) time series of actual and predicted anomalies; (b) scatter plot of actual versus predicted anomalies.

of T_s. If a random error with variance corresponding to that of the WLS regression residuals is included in the prediction equation, the resulting time series of T_s is no longer an unbiased prediction, but rather a scenario, a distinction considered in more detail in Section 5.

b. Model II

Inclusion of the five local climate variables at the four stations in Model II WLS regressions indicates an improvement in r^2, AIC, and BIC over those of the simple direct regression model (Table 2). However, only T_a and the solar radiation flux are consistently significant above the 99% confidence level (Table 4). The closely related parameter, sky cover, is highly significant at Greensboro and Raleigh-Durham, the two sites in piedmont North Carolina that subjectively have a similar climate, and is significant at the 95% level at the other two sites. These two piedmont North Carolina sites also have a significant wind speed effect, with surface temperatures decreasing as surface wind speed increases. This is understandable given the original postulate that wind speed is a prime determinant of the respective roles of regional and local conditions. Both Cape Hatteras and Sterling have higher average wind speeds, suggesting that the regional climate effects, rather than local ones, should be more important. This suggestion is not completely supported by the values of the T_a coefficients, which are all significant at the 99% level, but are marginally closer to a one-to-one relationship with T_s in the piedmont. There may, however, be limiting values for the wind speed at which the underlying assumptions begin to break down. Certainly it is only at Sterling that surface pressure plays a significant role. In no case is precipitation significant, suggesting that the role of surface moisture is everywhere minor.

One possible confounding factor in these estimates is the effect of seasons. For example, precipitation may be associated with warmer-than-normal conditions in winter but colder-than-normal conditions in summer. Such seasonal effects may be tested by adding to (4) a set of 18 seasonal interaction terms:

$$\sum_{i=1}^{5} \sum_{j=1}^{3} \Theta_{ij} L_i S_j + \sum_{j=1}^{3} \Theta_{aj} T_a S_j$$

where S_1 is a "dummy" variable with value 1 during December–February and 0 otherwise, S_2 is 1 during March–May and 0 otherwise, and S_3 is 1 during June–August and 0 otherwise. The months September–November are indicated by $S_1 = S_2 = S_3 = 0$. At Greens-

TABLE 4. Summary of WLS regression results for four stations for T_s as a function of T_a and local variables (a) for Model II [Eq. (4)] (without process model) and (b) Model III [Eq. (5)] (with process model—only those weighting coefficients significant at the 5% level being shown).

	Greensboro		Cape Hatteras		Sterling		Raleigh-Durham	
Parameter	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
T_a (K)	0.980*	0.985*	0.916*	0.912*	0.831*	0.838*	0.979*	0.976*
Solar flux (W m^{-2})	0.00051*	0.00046*	0.00034*	0.00033*	0.00034*	0.00031*	0.00049*	0.00048*
Precip. (mm)	−1.603	−1.533	2.217	2.186	−0.423	−0.485	−0.919	−1.097
Wind speed (m s^{-2})	−0.678*	−0.672*	−0.012	−0.007	0.079	0.033	−0.412*	−0.423*
Sky cover (tenths)	0.510*	0.588*	0.292	0.327*	0.348	0.412	0.526*	0.631*
Pressure (kPa)	0.367	0.459	−0.472	−0.438	1.166*	1.123*	0.027	0.044
Local model (−)	—	0.126*	—	0.043*	—	0.073*	—	0.099*
Θ_1	—	0.171	—	0.299	—	0.247	—	0.274
Θ_4	—	—	—	0.185	—	—	—	0.134
Θ_5	—	—	—	0.128	—	—	—	—
Θ_6	—	0.164	—	—	—	—	—	—

*: 99% confidence level.

boro, only the terms S_1T_a and S_2T_a are statistically significant at the 95% confidence level, implying a larger response of T_s in winter and spring than in summer or fall. At Cape Hatteras a 10 cm change in rainfall corresponds to about a 0.8 K change in T_s in spring but to negligible changes in other seasons. However, overall significance tests for this expanded model indicate that the hypothesis that there are no seasonal effects cannot be rejected.

The results in Table 4 may be used to indicate the first-order effects of changing the input variable, i.e. T_a. A change in T_a of 1 K, holding local effects constant, implies a mean increase in T_s of 0.8–1.0 K. A change in T_a, however, will almost inevitably be associated with a change in one or more of the other variables. The inclusion of the process model, therefore, is intended to explore such interactions.

c. *Model III*

The Model III results, which include the effects of the local process model, show relatively small changes from those for the model with local variables alone (Table 4). As a parameter, it is significant at the 99% level for all stations. The values of the other parameters, however, are changed by rather small amounts, and their significance is not altered. The amount of variance explained (Table 2) increases slightly at all stations. The change in AIC and BIC is also small. At Cape Hatteras and Sterling BIC increases when the process model is added, emphasizing that the increase in explanation provided by the model inclusion is more than offset by the penalty extracted by adding another parameter. In practical terms this penalty is probably underestimated since both criteria treat T_m as the same type of parameter as any of the others, whereas in reality a considerable number of elements and much computation is needed prior to its establishment.

Thus, despite the theoretical strength of including a process model to account for the parameter interaction terms, the results suggest that the modeling approach, at least in this statistical context, is not an efficient approach to adopt. Virtually identical results can be obtained using a purely statistical approach without explicit interactions.

Without these interactions, however, the analysis is confined to assessments where only a single parameter may be varied, all the others being treated as invariant. The results in Table 4 for Greensboro, for example, indicate that increasing the sky cover by one tenth of the celestial dome implies a 0.5–0.6 K increase in T_s on average. Physically, increased sky cover is likely to create a decrease in solar flux. While this will itself influence T_s, the magnitude of the decrease is unknown within the statistical model. Some insight is possible by using the process model in a stand-alone mode. The relationship between sky cover and solar radiation, obtained from the observational data, indicates that such an increase in sky cover would be expected to lead to a 0.4 K decrease in T_m. This implies, from Table 4, a decrease in T_s of about 0.05 K. Hence the net change as a result of a cloud cover increase and its interaction with solar radiation is an increase in T_s amounting to about 0.55 K.

This analysis, using the process model off-line from the main model, suggests that efforts to develop a more explicit incorporation of the interaction within the statistical model may prove fruitful. However, such efforts can only be justified if the interacting relationships between T_a, solar radiation, sky cover, and any other pertinent parameters, are known. In the context of scenario generation, this implies that they are useful when internally consistent sets of data on T_a and regional scale parameters are available, such as might be obtained, for example, from the same GCM simulation.

5. Simulation of surface air temperature

The methods described in Section 4 can be used to create simulated time series of local temperature to use in the production of scenarios of future local climates. An example of the procedure can be given here. For illustrative purposes only, it is assumed that T_a increases by 2 K, a value which might, for example, be the average change in a grid cell in a particular GCM run. The simulation is for Greensboro (Fig. 5) and was begun for January 1959 using the actual regression residuals for July–December 1958. For each month, the actual T_a was increased by 2 K and T_s determined from the parameter estimates from Table 4. In keeping with (7), an error value Z_t was added each month, randomly drawn from a normal distribution with zero mean and a variance equal to that of the original regression residuals. The resulting time series has a mean of 2.007 K and a standard deviation of 1.588 K. The observed situation during this time period (1959–76) is a mean of 0.049 K and a standard deviation of 1.593 K. Thus, the two time series differ by 1.96 K, essentially as expected from the T_a coefficient of Table 4.

This time series is a *simulation* not a *prediction.* Each monthly value in the simulation is not the best unbiased predictor of that month's T_s because of the addition of the randomly selected errors. The advantage of this is that the time series as a whole more nearly retains the statistical character of the actual time series and, in particular, retains the relative contributions of the regional and local components. If the series is recomputed without the random errors, for example, the standard deviation drops to 1.334 K, considerably less than that of the actual series.

By recomputing the time series with different sets of error values drawn from the same normal distribution, it is possible to generate alternative scenarios of surface air temperature. Such scenarios should be useful as inputs into climatic impact assessment models that re-

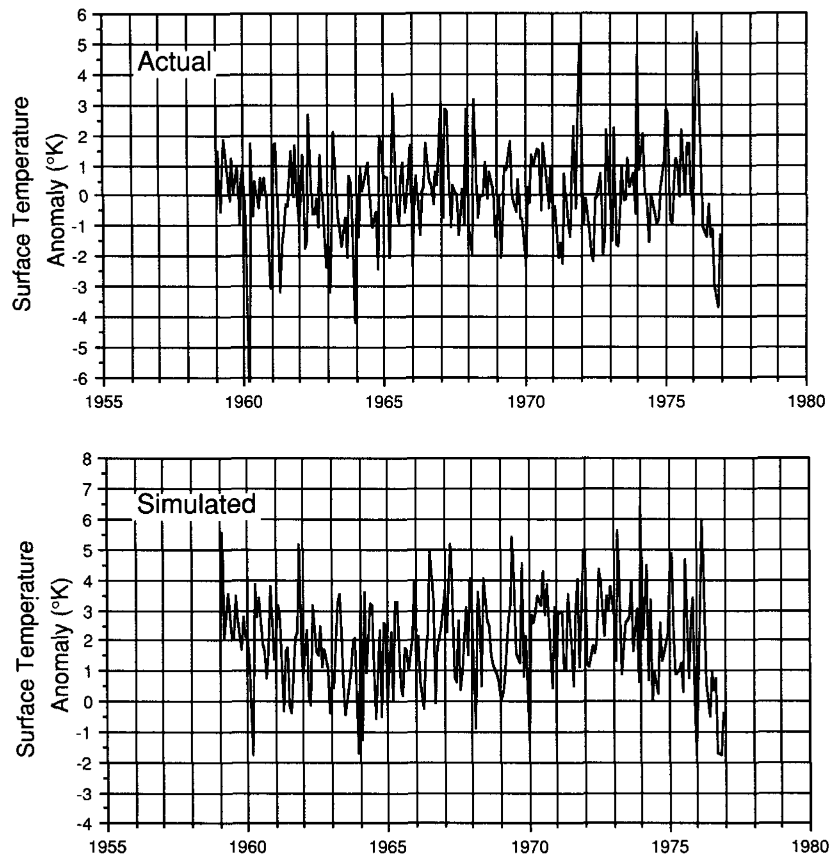

FIG. 5. Comparison of actual and simulated monthly anomalies of surface air temperature for 1959–76 at Greensboro.

quire realistic time series of temperature reflecting not only the new mean but also the statistical properties such as the variance and the likely recurrence and persistence of extreme values (e.g., Chen and Parry 1987).

6. Conclusions

It has been postulated that free tropospheric temperature, such as may be produced by GCM climate simulations, represents the regional component of temperature which has a direct effect on the surface temperature at a point or over a small area. The other temperature component is that resulting from purely local effects. Regression techniques have been used to construct models for the development of time series of temperature to test this postulate and suggest a method of producing temperature scenarios. The methods were tested using free troposphere data for the area of North Carolina and selected stations in the region as local sites.

Three regression models were developed using weighted least squares regression techniques. The first used only the free tropospheric temperature, whereas the second added specifically local parameters as independent variables. The third method added a local energy and moisture balance process model to specify the interactions between the local variables. There was a marked improvement in performance between the first and second models, but the third contributed little additional explanation. The major potential contribution of the process model appears to be in specifying interactions when GCM simulations produce internally consistent datasets containing free tropospheric temperature, sky cover, and solar radiation values.

The statistical regression models are capable of producing simulated time series of future temperature which retain the appropriate variability and lag characteristics, for any postulated mean tropospheric temperature change produced by GCM simulations.

Acknowledgments. This research was supported in part by the Atmospheric Research and Exposure Assessment Laboratory, U.S. Environmental Protection Agency, under cooperative agreement CR-815217-01-0. Support for one of us (RSC) was provided by the John Motley Morehead Foundation, University of North Carolina at Chapel Hill.

REFERENCES

Chen, R. S., 1987: The climate inversion problem: A statistical/modeling approach. Ph.D. Dissertation, University of North Carolina. 124 pp.

——, and M. L. Parry, 1987: Policy-oriented impact assessment of climatic variations. IIASA Research Report RR-87-7, International Institute for Applied Systems Analysis, Laxenburg, Austria. 54 pp.

Cohen, S. J., and T. R. Allsopp, 1988: The potential impacts of a scenario of CO2-induced climatic change on Ontario, Canada. *J. Climate,* **1,** 669–681.

Deardorff, J. W., 1978: Efficient prediction of ground surface temperature and moisture, with inclusion of a layer of vegetation. *J. Geophys. Res.,* **83**(C4), 1889–1903.

Gleick, P. H., 1987: Regional hydrologic consequences of increases in atmospheric CO2 and other trace gases. *Climatic Change,* **10,** 137–161.

Katz, R. W., and R. H. Skaggs, 1981: On the use of autoregressive-moving average processes to model meteorological time series. *Mon. Wea. Rev.,* **109,** 479–484.

Kim, J.-W., J.-T. Chang, N. L. Baker, D. S. Wilks and W. L. Gates, 1984: The statistical problem of climate inversion: Determination of the relationship between local and large-scale climate. *Mon. Wea. Rev.,* **112,** 2069–2077.

Klein, W. H., 1983: Objective specification of monthly mean surface temperature from mean 700 mb heights in winter. *Mon. Wea. Rev.,* **111,** 674–691.

——, and J. E. Walsh, 1983: A comparison of pointwise screening and empirical orthogonal functions in specifying monthly surface temperature from 700 mb data. *Mon. Wea. Rev.,* **111,** 669–673.

Lamb, P. J., 1989: On the development of regional climatic scenarios for policy-oriented climatic impact assessment. *Bull. Amer. Meteor. Soc.,* **68,** 1116–1123.

Mearns, L. O., R. W. Katz and S. H. Schneider, 1984: Extreme high-temperature events: Changes in their probabilities with changes in mean temperature. *J. Climate Appl. Meteor.,* **23,** 1601–1613.

National Climatic Data Center, 1978: *SOLMET Volume 1—Users Manual, TD-9724, Hourly solar radiation-surface meteorological observations.* National Climatic Data Center, NOAA, U.S. Department of Commerce. 98 pp.

National Climatic Data Center, 1979: *SOLMET Volume 2—Final Report, TD-9724, Hourly solar radiation-surface meteorological observations.* National Climatic Data Center, NOAA, U.S. Department of Commerce. 184 pp.

National Climatic Data Center, 1986: *TD-6200 Series NCDC Upper air digital files.* National Climatic Data Center, NOAA, U.S. Department of Commerce. 15 pp.

National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration and United States Air Force, 1986: *U.S. Standard Atmosphere, 1976.* Washington DC, U.S. Govt. Printing Office. 227 pp.

National Research Council, 1983: *Changing Climate.* Washington DC, National Academy Press. 496 pp.

Priestley, M. B., 1981: *Spectral Analysis and Time Series, Vol. 1: Univariate Series; Vol. 2: Multivariate Series, Prediction and Control.* Academic Press. 890 pp.

Robinson, P. J., and P. L. Finkelstein, 1989: *Strategies for the development of climate scenarios,* Final Report, Cooperative Agreement CR-815217, Research Triangle Park, NC. U.S. Environmental Protection Agency, Atmospheric Research and Exposure Assessment Laboratory. 73 pp.

SAS Institute Inc., 1984: *SAS/ETS Users Guide: Verson 5 Edition.* SAS Institute, 738 pp.

Sellers, W. D., 1965: *Physical Climatology.* University of Chicago Press. 272 pp.

Smith, J. B., and D. Tirpak, Eds., 1988: *The Potential Effects of Global Climate Change on the United States: Report to Congress.* U.S. Environmental Protectional Agency, Office of Policy, Planning and Evaluation, Office of Research and Development.

Webb, T. I., and T. M. L. Wigley, 1985: What past climates can indicate about a warmer world. *Projecting the Climatic Effects of Increasing Carbon Dioxide,* MacCracken, M. C. and F. M. Luther, Eds., U.S. Department of Energy, Office of Energy Research, 237–257.

Wigley, T. M. L., P. D. Jones, K. R. Briffa and G. Smith, 1990: Obtaining sub-grid-scale information from coarse resolution general circulation model output. *J. Geophys. Res.,* **95**(D2), 1943–1953.

Wilks, D. S., 1989: Statistical specification of local surface weather elements from large-scale information. *Theor. Appl. Clim.,* **40,** 119–134.

Wiser, E. H., 1983: *HISARS, Hydrologic Information Storage and Retrieval System Users Guide.* Rep. UNC-WRRI-83-201, Water Resources Research Institute, University of North Carolina. 117 pp.