A Comparison among Strategies for Interpolating Maximum and Minimum Daily Air Temperatures. Part I: The Selection of ''Guiding'' Topographic and Land Cover Variables

CLAIRE H. JARVIS AND NEIL STUART

Department of Geography, University of Edinburgh, Edinburgh, Scotland

(Manuscript received 19 April 2000, in final form 11 October 2000)

ABSTRACT

This paper explores the derivation and selection of a comprehensive set of continuous topographic and land cover–related variables to guide the interpolation of daily maximum and minimum temperatures over England and Wales, for an entire annual cycle to a resolution of 1 km. The work draws on and updates historical topoclimatic modeling through use of digital elevation data and land cover data, using the modeling capabilities of geographical information systems. The influential guiding variables under a variety of dominant weather patterns were identified and used to assist with the interpolation of an annual sequence of daily maxima and minima for 1976. North map coordinate (''northing''), elevation, and coastal and urban effects were found to be particularly significant variables in explaining the variation in U.K. daily minimum temperature. Urban factors have not previously been thoroughly investigated, despite the high density of population in England and Wales. Analysis of the residuals from data withheld from the partial thin plate spline interpolation suggests that the incorporation of coastal shape and situation, land cover, and soils data might further improve the modeling of local-scale influences on maximum and minimum temperature. They also suggest that the results achieved (rms errors of 0.8°C for maxima and 1.14°C for minima) may be close to the limits of accuracies achievable at 1-km resolution given the density of temperature observation data and standard exposure of the observing network used.

1. Introduction

Climate variables play a central role in many geographical models of natural and agricultural systems (e.g. Running et al. 1987). Over extensive areas, deriving spatially distributed data using process-based approaches (e.g., Running and Thorton 1996) may be impractical for reasons of computational and theoretical complexity or the unavailability of data. Under such circumstances, interpolation is a popular means of extending point meteorological data observations over geographic areas to form spatially distributed inputs to such models (e.g., Willmott et al. 1985; Landau et al. 1998; Supit 1997). This paper focuses on the task of gridding *daily* data, a task arguably more complex than constructing monthly gridded data (e.g., Goodale et al. 1998; Hutchinson 1991; Lennon and Turner 1995; Price et al. 2000). Methods to improve the accuracy of daily temperature interpolation are important when placed within an ecological modeling context, because there is a considerable sensitivity of the underlying ecological system to these temperature variables. Comprehensive comparisons among more than two interpolation methods when working with daily data remain unusual within the published literature, particularly for both maximum and minimum temperatures over a nationwide extent as in this case. Moreover, several previous comparisons have been undermined because underlying geographical trends were not accounted for equally across all techniques (e.g., Collins and Bolstad 1996; Bolstad et al. 1998).

The approach taken in this paper and in part 2 (Jarvis and Stuart 2001, this issue) is to adopt a ''middle ground'' between data-based geostatistical interpolation and process-based meteorological modeling. This approach is achieved by choosing to guide the process of interpolation using indices developed from topographical and land cover data that are hypothesized to influence the climatic conditions. A similar convergence is evident in published research, with an increasing use of interpolation by those workers more commonly associated with process-based models [e.g., Running et al. (1987) vs Running and Thornton (1996), and Laughlin and Kalma (1990) vs Laughlin et al. (1993)] and, from the interpolation perspective, the view that the next generation of interpolators will be those better able to incorporate knowledge of the underlying geographical processes (Mitás and Mitásovà 1999).

Corresponding author address: Dr. C. H. Jarvis, Dept. of Geography, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, Scotland.
E-mail: chj@geo.ed.ac.uk

The objective of this first paper is to address the derivation and selection of topoclimatic parameters to guide interpolations over the full annual sequence for both maximum and minimum temperature at a daily time step, to 1-km resolution. A number of authors have suggested that there is a need for greater consideration of factors such as the improved consideration of cold-air drainage effects, water bodies, or urban areas (Hutchinson 1991; Landau and Barnett 1996; Bolstad et al. 1998; Price et al. 2000). This work draws on both historical topoclimatic modeling (e.g., Manley 1944; Tabony 1985) and more recent geographical information system (GIS)–based work (Lennon and Turner 1995; Cornford 1997) to derive a rich set of variables for both maximum and minimum daily temperatures over England and Wales. In addition to incorporating proxy variables to account for known physical processes, our approach also effectively provides a ''level playing field'' for comparing the estimation accuracy of different interpolation algorithms. This factor is important because geographical trends must be accounted for before proceeding with most geostatistical methods of interpolation. Given the acknowledged variability of the underlying meteorological processes from day to day (e.g., Bolstad et al. 1998; Cornford 1997), a set of guiding variables is required that perform under all main U.K. weather types.

In addition to the selection of a guiding variable set, this work also assesses its performance as part of the interpolation process. Partial thin plate spline interpolation is used as a ''representative'' method for this purpose (Hutchinson 1991). The accuracy of interpolation results are most commonly interpreted on the basis of a spatially aggregated error computation such as the root-mean-square (rms) error or the correlation coefficient r^2. In addition to these quantitative measures, the paper will consider the geographical variability in the accuracy of the interpolation, given that this may be masked when residual data are aggregated to provide the rms error.

2. Methods

First, methods are presented for deriving topoclimatic and land cover–related variables at continuous positions over the landscape of England and Wales. All topoclimatic variables were computed to a grid spacing of 1 km; example variables are illustrated in Fig. 1.

Using an objective method of selecting influential covariates to guide the local estimation was an important criterion. The process of deriving and selecting appropriate guiding variables was carried out using daily temperature data from 1986, using stepwise linear regression. Temperature data for comparable stations from 1976 were then interpolated using second-order, two-dimensional partial thin plate splines as a representative interpolation tool. Data from 1976 were used for the interpolation experiments to maintain independence from data used to construct the interpolation function.

a. Data

Archived U.K. Meteorological Office (UKMO, now known as the Met Office) data from 174 distributed stations, provided under a quality assured agreement, were used in this study (Fig. 2). These data are collected at ''standard exposure'' for synoptic purposes.

The principal raster data used in the study were the Ordnance Survey Panorama product, a digital elevation model (DEM) at 50-m resolution for England and Wales. These raw data were used to derive an approximate drainage network using the default hydrological commands from the proprietary GIS package ARC-INFO, following which all further processing was carried out using grids at resolutions of 500 m $\times$ 500 m or 1 km. Hutchinson and Gallant (1999) concur that''Mesoscale DEMs, with spatial resolutions from 200 m to 5 km, are appropriate for topographically dependent representations of surface temperature and rainfall, key determinants of biological activity.''

To begin to accommodate known alterations of temperature that arise as a result of urbanization effects, the elevation data were supplemented by ''urban'' and ''suburban'' class data from the Landsat-derived Institute of Terrestrial Ecology (ITE) land cover data set of England and Wales (Fuller et al. 1994) available at a spacing of 25 m $\times$ 25 m. The urban class incorporates all urban developments (without significant quantities of permanent vegetation), and the suburban category includes mixed built/vegetative land covers and small villages or rural industrial estates.

b. Deriving land cover and topoclimatic variables to guide the interpolation of temperature

1) SURFACE TYPE

Processes affecting daily maximum and minimum temperatures include the effects of absorbed solar radiation, internal boundary layers, surface roughness, and urbanization. In this study, particular emphasis is given to a new treatment of urban areas as a guiding variable.

A wide range of empirical measures has been used within the literature to account for surface roughness. First, the standard deviation of height inside local rectangular windows of 5 and 25 km width were computed (stdev5 and stdev25, Fig. 1h). An additional measure of local surface roughness as ''height above the local minimum'' was derived by subtracting the target cell height from the minimum elevation found over similar rectangular areas (localrough5, localrough25, Fig. 1g). These measures follow from the findings of Lennon and Turner (1995) and Cornford (1997).

Several studies indicate that the urban heating effect

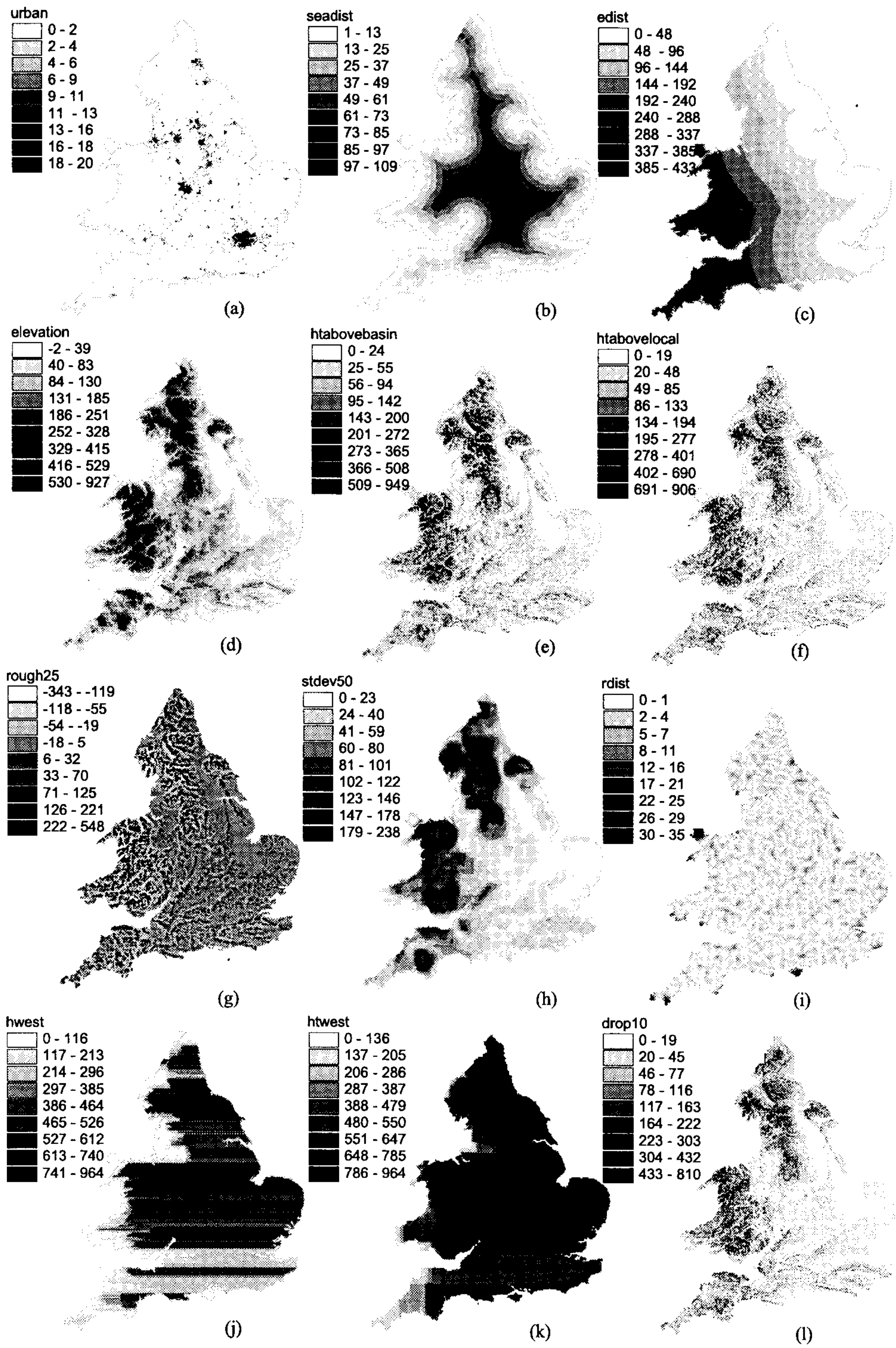


FIG. 1. Examples of the guiding variables derived for the selection process.

declines steeply at the edge of suburban areas (e.g., Oke 1976). Previous work incorporating the heat island effect as a guiding variable for interpolation has used raw distance measures, ignoring the findings that the heat effect is linearly related to the log of the size of the urban population (Oke 1973). An alternative index was therefore modeled using land cover data from ITE regarding urban and suburban areas. Individual cells were assigned a value represented by the natural log of the size and density of the urban area within which they were located (Fig. 1a).

The influence of local boundary layers and surface

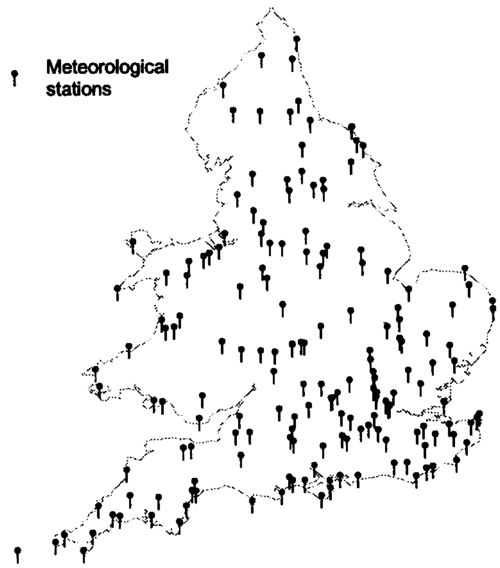


FIG. 2. Locations of the 174 distributed UKMO meteorological stations from which continuous daily maximum and minimum temperature records for 1976 were extracted for this study.

albedo, although important influences at the microscale, were not modeled in this work, because a more regional estimate was sought and because digital land cover and soils data that affect such processes are not available at such fine scales in the United Kingdom. Sea breezes are discussed below as a component of the discussion of topography.

2) TOPOGRAPHY

As Bolstad et al. (1998) have shown, even lapse rates that are defined locally and by season are no substitute for more adaptive (daily adjusted) techniques unless a wider range of parameters is also included. In this study, regression equations between maximum and minimum temperatures and elevation were computed separately and on a daily basis to ensure flexibility according to the prevailing environmental conditions.

Cold-air drainage and evaporation on slopes may affect temperatures in a number of ways. The notion that damp soils, with their cooling effects, are more likely at valley bottoms was encapsulated by representing distance to the nearest rivers of various ‘‘Strahler’’ orders (in which the labeling of a stream reach does not change until a stream of equal or higher order is joined to it). Changes of slope also affect the depth of the surface-cooled atmospheric layer, which is greater on concave rather than convex surfaces, leading to corresponding lower minima on the concave slopes. However, this effect was expected to prove unresolvable in this study given that the average slope of the landscape is very shallow at the 1-km resolution used.

With the exception of work by Cornford (1997), there are few reports that proxy measures to represent the locations prone to katabatic winds have been derived or used to improve interpolated estimates of local minimum daily temperatures. Given the 1-km resolution of modeling in this study, Tabony’s (1985) simpler notion of large-scale ‘‘shelter’’ was adopted. Tabony suggested that the height of a point relative to the valley floor provided a measure of susceptibility to frost. Cornford translated this to the difference between height and the minimum elevation over a 10-km radius in his comparative study. In this work the more process-oriented local drainage basin replaces the geometric radius, acknowledging that indices computed using a set radius could, for example, straddle two drainage basins, leading to a poor representation of windflow over the underlying land surface. Basins of several sizes were generated according to stream order (htabovelarge, htabovemid, . . . , htabovebasin, Figs. 1e,f). Cornford’s simpler algorithm was also computed for comparative purposes (drop4, drop10, Fig. 1i).

For this work, given the variety of prevailing wind systems throughout the year, distance to the south, east, and west coasts is modeled individually (Figs. 1b,c). Gridded variables are also constructed that compute distance inland to a maximum of 100 km (e.g., east100; Barry and Chorley 1982; Linacre 1992). To account for the complexity of the British coastline, the local land: sea ratio within radii of 2 and 5 km (pcoast4, pcoast25) was also computed (Lennon and Turner 1995).

Last, adiabatic warming as a result of the descent of air from mountains and plateaus by föhn winds was considered. In this study, only east–west and north–south barriers are considered, following reports of the föhn effect in relation to North Wales and the Pennines (Mayes and Wheeler 1997, p. 25) and in connection with southerly air flows, in particular (e.g., Mayes and Wheeler 1997, p. 35). After the method of Lennon and Turner (1995), the maximum height was computed in the relevant direction for an unrestricted distance (hsouth, hwest) and that within a 50 km $\times$ 1 km rectangle 25-km north and south of the point in question was computed (htsouth, htwest, Fig. 1h).

c. Selecting gridded variables to guide the interpolation process

To estimate daily air temperatures at locations between meteorological stations using the grids derived from topographic and land cover data to guide the interpolation, the next step was to develop a parsimonious model from the ‘‘best’’ of the 35 candidate gridded variables. Given the need to perform interpolations for several hundred days in succession, a decision was made to select one single set of variables whose coefficients in the derived regression model could be adjusted on a daily basis to improve the accuracy of the interpolation. The size of the selection task was such that, as one component of a broader applied study, only automatic means of selecting variables could be warranted. Others (e.g., Blennow and Persson 1998; Lennon and Turner 1995) have adopted a similar approach in preference to the more subjective choice of variables through the visual inspection of correlation coefficients (e.g., Cornford 1997).

Backward linear regression was used to establish significant relationships separately for the maximum and minimum temperature data and the preprocessed gridded variables. Daily temperature data from 1986 were used, with selected values at least 5 days apart from any other to avoid problems of temporal correlation in the dataset and to maximize the information gained. Only variables significant at the 95% level were allowed to remain a component of the regression equations. To investigate the influence of the general synoptic situation upon daily air temperature estimates, 21 days were used in the regressions for each of the high pressure, low pressure, and low-vorticity classifications as defined by the British ''Lamb'' classification (available at the time of writing online at http://www.cru.uea.asc.uk/~mikeh/datasets/uk/lamb.htm).

The results were compared first by counting the number of times a variable was selected as making a ''significant'' contribution to the regression model (i.e., remained as part of the regression model) and second on the basis of the strength of the partial correlation coefficient. The final selection of variables from the two separate rankings accounting for consistency and strength of contribution was made using a ''Pareto'' selection algorithm. The top 10 covariates were then selected after eliminating those with Pearson partial correlations of more than ± 0.5 with any variables of higher rank.

d. Interpolation using the best variable set: Geographical distribution of residuals

Using the best-performing 10 covariates from data in 1986 to guide the interpolation process, two-dimensional partial thin plate spline interpolation (Hutchinson 1991) was then used to estimate daily maximum and minimum temperatures over England and Wales throughout the annual cycle for the independent 1976 dataset. Within this interpolation procedure, the most influential guiding variables were incorporated as partial linear covariates.

Assessing the degree to which interpolations capture the predominant spatial patterns in temperature is an important component of a geographical study. This paper seeks to provide a broader view of the interpolator performance than that provided by an aggregate rms statistic by computing the geographic distribution of residuals constructed using data withheld from the interpolation process using a jackknife cross-validation technique across England and Wales. The geographical distribution of the daily jackknifed residuals from partial thin plate spline interpolation averaged over the annual period, together with the variance in these daily residual values at the validation locations over the year (Figs. 5–8; see section 3b), was then used to identify any recurrent geographical patterns that might assist in improving the combined guiding variable/interpolation method.

3. Results and discussion

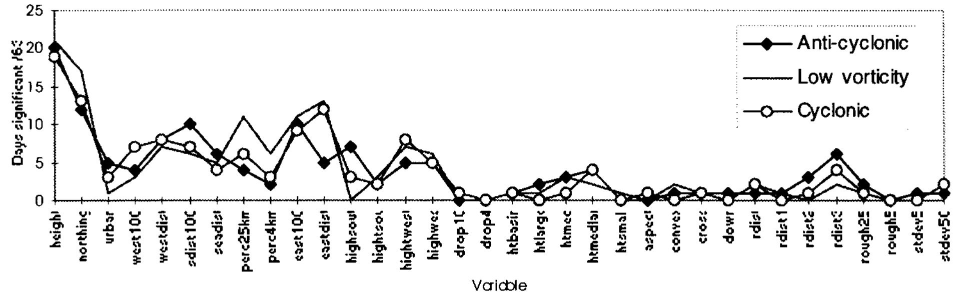
a. Selection of guiding variables

1) Consistency of selection

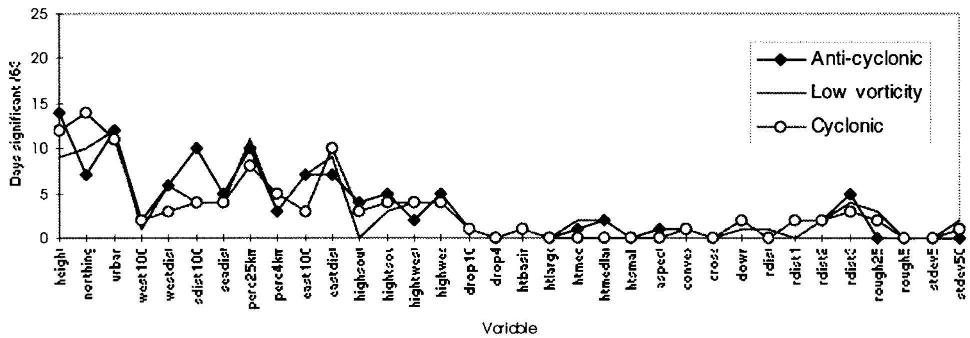
The consistency (number of days/total of 63 days) with which the guiding variables provide a significant contribution to a regression model explaining temperature is graphed by weather type for (a) maximum and (b) minimum daily temperature in Fig. 3. Fewer variables were selected more consistently over the 63 days from 1986 as contributing to the estimates for daily maximum than for daily minimum temperatures.

Elevation, through the lapse-rate effect, was an important predictor variable for maximum temperature (Fig. 3a). ''Northing'' (north mapping coordinate) is considerably more influential for estimating maximum temperatures than for minimum temperatures under all weather conditions, but particularly under conditions of low vorticity. A further cross correlation is evident between northing and the distance from the south coast (sdist100, limited to 100 km). Directional indices of coastal effect proved more influential upon temperature than did nondirectional indices, such as the land:sea ratio, with distances to the east coast most commonly selected as significant. Also, in the case of temperature maxima, measures relating to the maximum height to the west were more consistently chosen, with the broader north–south banded index (hightwest, Fig. 1k) of greater significance than the more restricted hightwes (Fig. 1j). Distance to the nearest river and height above the drainage basin within which a site lies were less often chosen, and slope-related variables were of little influence at this 1-km resolution.

In the case of minimum temperatures, three main variables dominate Fig. 3b. These are elevation, northing, and the urban index, which are all significant (95% confidence) on approximately 12 days of the 63 modeled. The overall number of days on which the elevation variable was found to be significant at the 95% level was, however, lower than anticipated (Fig. 3). The influence of urbanization more strongly influences minimum temperature estimates than maximum temperature estimates. This result is anticipated from the literature (Oke 1987, p. 290). Coastal variables also provide significant contributions to the regression equation for minimum temperatures. Directional variables were consistently selected more frequently than the simpler ''distance from the sea'' measure, with distance from the east coast (unlimited or constrained to 100 km: edist, edist100) of greatest influence. The combined influence of these coastal effects regularly provides a greater contribution



(a) Maximum temperature



(b) Minimum temperature

FIG. 3. Consistency of selection of topoclimatic and land cover–related variables: (a) daily maximum and (b) daily minimum temperatures, 1986.

to explaining minimum temperatures than does elevation. Of the variables measuring the local ratio of land to sea, effects up to 25 km (pcoast25) were found to have influence on minimum temperatures more often than did the restricted 4-km variable (pcoast4). Other indices that were less consistently significant (only 5 days) were maximum elevation to the west and south, height above the base of a ''medium to large'' valley, distance to a river, and slope- and aspect-related parameters. Variables relating to surface roughness (rough25 and stdev50) were selected only occasionally, under conditions of low vorticity.

2) STRENGTH OF RELATIONSHIP

The strength and direction of influential ''guiding variables'' for each of the 63 days modeled were found to be relatively weaker for estimating daily minima than for estimating daily maximum, in accordance with the results for consistency of selection (Fig. 3). Height exerted a negative influence over minimum temperatures, with standardized partial correlation coefficients between -0.6 and -0.4. This correlation coefficient falls slightly below the ''standard'' lapse rate of $-6.5°C$ $(1000\ m)^{-1}$. This result suggests that the manner in which cold airflows are accounted for in the regression model could be further improved. Coastal influences also exert strong effects in the regression model, especially under anticyclonic conditions, with the direction of the relationship altering according to season. Under anticyclonic weather patterns, temperatures increase quickly with increasing distance from the east coast, especially during the summer season. Both the urban index and land:sea ratio are consistently selected for minimum temperature prediction, with partial correlation coefficients between 0.2 and 0.4 and -0.3 and 0.6 depending on weather types.

Variables designed to reflect the degree to which large-scale topography provides a shelter from prevailing weather systems (htwest, hwest, hsouth, htsouth) showed relatively weak relationships with minimum temperature. Considering how basin shape may affect estimates of minimum temperature, both height above the minimum elevation over a 10 km $\times$ 10 km area and height above the most local basin exerted positive influences on minimum temperature observed under all main weather classifications. Of interest, it was the variable drop10 rather than heights above the various sizes of watershed minima (e.g., htlarge, htmed) that provided the anticipated positive relationship between minimum temperature and height above the area minimum that is expected from an understanding of katabatic processes and cold-air ponding within hollows and basin bottoms.

For maximum temperatures, the significance of individual variables showed a similar but often inverted pattern to those for minimum temperatures. The relationship between maximum temperature and elevation was more consistent and slightly more strongly negative

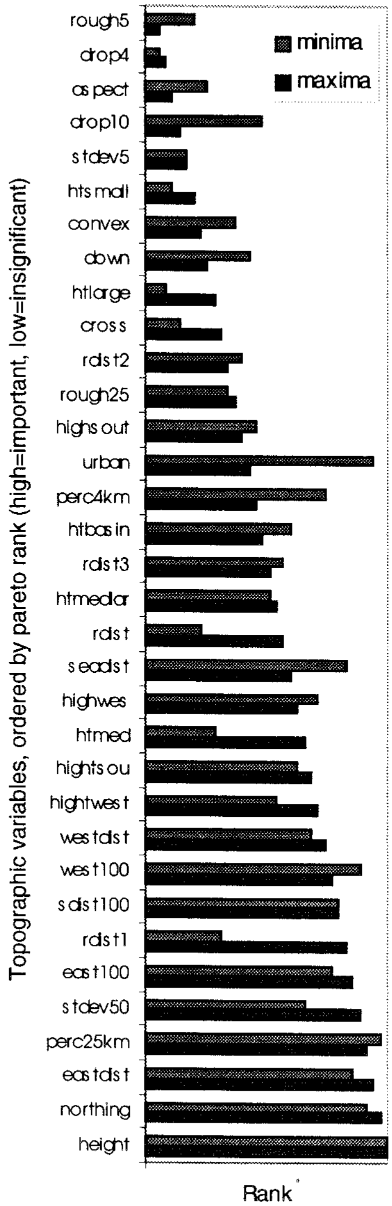


FIG. 4. Summary Pareto rank (significance and consistency) association of topoclimatic variables with maximum and minimum temperatures, 1986.

than that for minimum temperatures, consistent with the literature for monthly temperatures (e.g., Hutchinson 1991). Distance from the west coast showed a strong but bipolar relationship with temperature maxima. The cause of this bidirectionality is attributed to seasonal differences between land–sea heat accumulation. Temperatures decreased strongly with distance eastward under conditions of low vorticity with the reverse situation occurring under anticyclonic and cyclonic conditions.

Terrain shape is known to have a considerable bearing on minimum temperatures in particular (Tabony 1985). The increase in minimum temperatures with increased local roughness may be attributed to improved mixing in the atmosphere, which prevents cold-air ponding in local pockets, especially under mixed and low pressure weather systems.

TABLE 1. Top covariates, maximum and minimum temperature.

Max	Min
Height	Height
Northing	Pcoast25
Edist	Urb
Pcoast25	Northing
Stdev50	Wdist100
East100	Edist
Rdist1	Seadist
Sdist100	Sdist100
West100	Edist100
Westdist	Pcoast40
Htwest	Highwest1
Hsouth	Wdist
Htmed	Stdev50
Hwest	Htsouth
Seadist	Htbasin

The presence of rivers and measures of concavity were included as part of the analysis, but the manner in which these factors interrelated was complex and potentially beyond the capabilities of the linear modeling used in this and similar studies.

3) VARIABLE SELECTION: COMBINING THE EVIDENCE

The variables found to influence maximum and minimum temperature are summarized in Fig. 4 in the order of their combined ranking (strength plus consistency) in the regression model. Urban area is found to be an important predictor for daily minima but not for daily maxima. The effect of directional coastal influences is less pronounced for minima than for maxima. The Pareto technique also allows the influence of the standard deviation in height (stdev 50) to be identified. This influence was otherwise unclear from the individual rankings, owing to its consistent but weak strength of selection.

Several of these predictor variables are known from the regression modeling to be cross-correlated. Restricting the variable set further was therefore an important next step of the analysis. By limiting the major variables listed (Table 1) to those with partial correlations of absolute value 0.5 or less (Table 2), a subset of guiding variables was obtained for estimating both maximum and minimum temperatures.

Perhaps unsurprising, given the strong maritime influence over British climate, multiple measures of distance to the coast in all major directions appear in Table 2, in addition to the nondirectional land:sea ratio (pcoast25). For minimum temperatures, the urban index is highly ranked, owing to its consistent selection, albeit weak strength of relationship. Northing is much less important for estimating minimum temperatures.

TABLE 2. Covariates to be used in the interpolation of maximum and minimum temperatures ($-0.5 <$ partial correlation coefficient > 0.5).

Max	Min
Height	Height
Northing	Pcoast25
Pcoast25	Urb
Stdev50	Northing
Edist100	Wdist100
Rdist1	Seadist
Sdist100	Sdist100
Wdist100	Edist100
Wdist	Hwest
Htwest	Wdist
Hsouth	Stdev5
Htmed	Htsouth

b. Geographical variation in residuals from interpolation over England and Wales

Figure 5 illustrates the annual average daily residual maximum temperature, and Fig. 6 shows the variance of the residuals at the meteorological sites over the year. Locations where the interpolation model performs most poorly (both under- and overpredictions) are summarized in Table 3. Particularly striking is the coastal nature of all of the sites with greatest negative bias (overprediction by the interpolator), the majority of which are exposed southerly and westerly coastal clifftop stations. Similar, if less pronounced, effects may be seen on the Yorkshire coast at Scarborough and Whitby. Because the effect of the coastal warming is computed by covariates on a national basis, and at the majority of sites the estimation of temperature is successful, local coastal situation may explain the overprediction at Bognor.

For the locations where observed maximum temperatures are strongly under predicted, both Valley and Scilly (St. Mary's) are coastal sites, but they are also located at small civilian airports. Airport sites are commonly warmer than nearby (otherwise similar) locations, owing to the reduction in wind speed close to temperature recorders in close proximity to buildings and owing to the blackbody effects of tarmac. The scale of these smaller airports means that they are not identified using the ''urban index'' which was designed to account for these factors in large residential areas.

Some stations with high annual average daily residual error (Fig. 5) also exhibit a large day-to-day variability in their residuals (Fig. 6) (e.g., Bardsey Island, Valley, Scilly). Locations of different topographic characters are also included in Table 3 as a result of the differences in magnitude and direction of residual values at different times of the year. Through their easterly facing locations, for example, both Folkestone and Dungeness are likely to encounter a relatively high frequency of northeasterly winds cooled by the North Sea under anticyclonic conditions. During the summer of 1976, such weather patterns were unusually common. These effects are also seen to a lesser degree at other locations without easterly shelter (e.g., Margate, Shoeburyness). Covariates based upon nationally aggregated data might not be expected to pick up such differences at locations whose characteristic topography is unusual in relation to the dataset as a whole. For sites on sandy soil, such as at Dungeness and in the Breckland, high fluctuations in maximum temperature as a result of the relatively fast ground heating are to be expected. This effect may be the cause of underpredictions in maximum temperature, given that the regression model used does not include soil type.

Poorest performance has so far been associated with sites at exposed coastal situations and airports. A further factor affecting interpolator performance identified from Fig. 5 and Table 3b is high land. Widdybank Fell, for example, is one of the highest meteorological stations in England and Wales, and the sparseness of data at such heights (a function of the meteorological network rather than sampling method) makes accurate estimation difficult whatever the interpolation method. In addition, both Widdybank Fell (Stirling 1997, p. 120) and Newton Rigg (Tufnell 1997) are on record as sites at which extreme low temperatures have been recorded, so that high variances in residuals at these sites are not surprising.

Predictions of minimum temperatures (Figs. 7 and 8) are generally less accurate than for maximum temperatures, resulting in a longer list of ''problem'' stations (Table 4). St. Mary's Airport and Bardsey Island are problem stations for both maximum and minimum temperatures. Several other sites (e.g., Mansfield, London Weather Centre, Margate, and Valley) have modestly strong residuals in common. To check the overall consistency of the temperature interpolation, it was affirmed that, for all points over the landscape and for nine test dates, the values interpolated for minimum temperature were always lower than the maxima estimated at the same sites.

For minimum temperatures, a difficulty of estimating lapse rates owing to a relative lack of high-elevation data again leads to high residuals at upland areas such as Widdybank Fell. Hartburn Grange is another station known to receive cold air from the Pennines under high pressure conditions, and this is likely to be the cause of the high overprediction observed in Table 4. In the contrasting case of Malvern, however, the site is actually warmer than predicted. This may be a result of Malvern's position overlooking the Severn Valley, making it frost free in comparison with stations at lower altitudes on the valley floor (Manley 1994). Nonlinearities in the ''height above basin floor'' predictor variables may therefore be the cause of model underpredictions at this site.

As Table 4 indicates, interpolated estimates at both Manchester and London Weather Centers show strong underprediction relative to that observed. However, in the case of the London site, in particular, the site of the recording station at rooftop elevation (as opposed to the

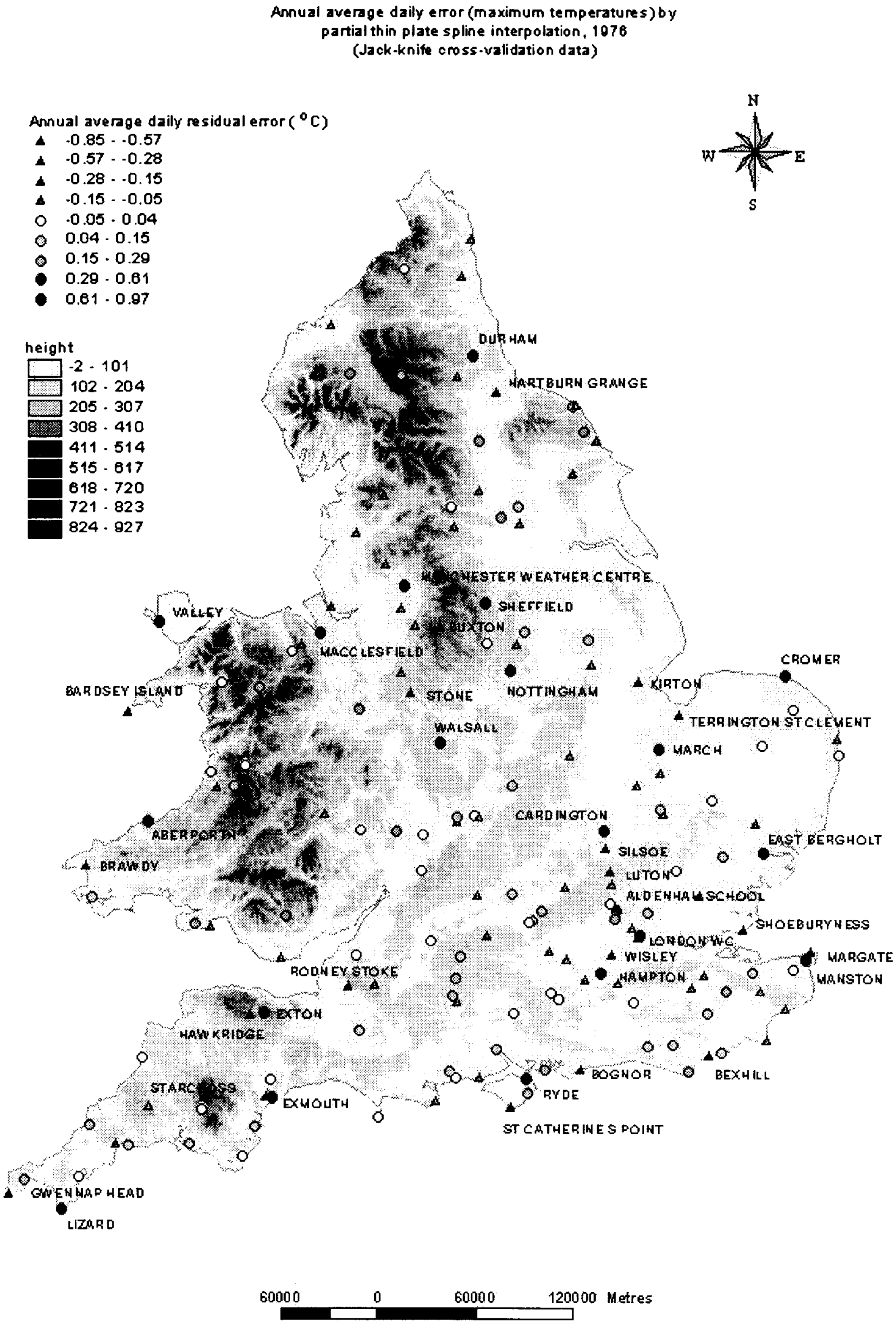


FIG. 5. Average daily residual for maximum temperatures, partial thin plate spline interpolation, 1976.

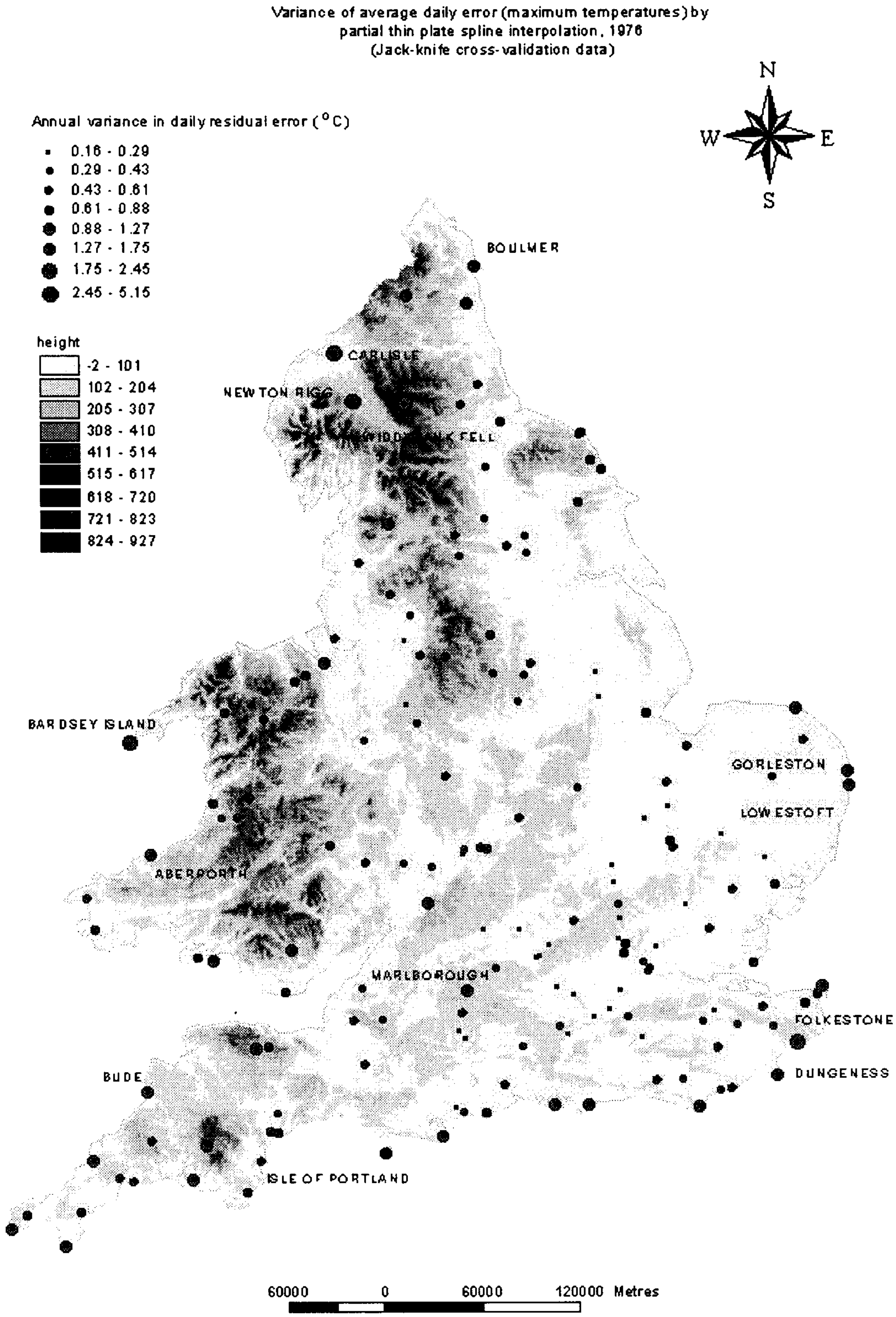


FIG. 6. Variance in daily residual of maximum temperature, partial thin plate spline interpolation, 1976.

TABLE 3. Poorest performing locations for maximum temperature predictions, by (a) average bias and (b) annual variance in bias. In (a), positive figures imply underprediction by the interpolator and negative figures relate to overpredictions.

(a) Meteorological station	Bias	(b) Meteorological station	Variance
Bardsey Island	−0.85	Dungeness	1.75
Bognor Regis	−0.74	Folkestone	2.13
Gwennap Head	−0.59	Widdybank Fell	2.2
St. Catherine's Point	−0.57	Bardsey Island	2.35
Manchester Weather Centre	0.61	Valley	2.45
Cardington	0.67	Scilly (St. Marys)	3.64
Scilly (St. Marys)	0.9	Carlisle	4.29
Valley	0.97	Newton Rigg	5.15

standard flat, grassy exposure) makes these anomalies unsurprising. In future work, these are data that might be better omitted from analyses unless the data were adjusted prior to interpolation. Residuals for Aigburth, Nottingham, and Enfield reveal an average underprediction of minima in these suburban locations. This result suggests that the ''heat island'' effect in the model may need further exaggeration.

The effects of local situation and land cover are more likely to influence minima than maxima, as seen in the cases of Moel Cynnedd and Wisley. For both sites, recording equipment lies in or near to woodland, and forest clearings are well known for their tendency to enhance frost. A number of stations with particular, ''nonstandard'' exposures also appear within Table 4. Plumpton, for example, is west facing, and Trawsfynydd is north facing. Managing such local-scale variations is beyond the scale and national coverage of interpolation within this study, and such residuals serve to highlight the many variations in sub-1-km-resolution climate processes.

4. Discussion

The relative strengths of the relationships between the derived topographic and land cover variables and the observed maximum and minimum temperatures were highly variable. This fact highlights the importance of adaptability in any interpolation system that models daily temperatures for all days in the year, as opposed to the use of a fixed-trend model (e.g., Landau and Barnett 1996). In Fig. 2 this variation could be interpreted to reflect different processes occurring by weather type. When restricted to using a fixed set of variables as in this case, this variability confirms the need to include in the estimating equations both the most consistently selected variables and those that are only rarely selected but are of considerable influence on the days when they are chosen. This requirement does introduce potential redundancies of information, but with the benefit of added flexibility in predicting temperature conditions occurring under unusual but potentially significant weather conditions (e.g., during strong anticyclones).

There were a number of cases in which a relationship was hypothesized, but for which no marked relationship, either in strength or consistency of selection, was found. Measures relating to aspect and gradient fall inside this category. This is attributed to two main factors. First, the resolution at which the variables are derived (1 km) means that only large-scale changes in slope will be identifiable. Slope and aspect *are* known to affect temperature through differences in solar radiation received (Linacre 1992, p. 193), but the relatively crude resolution of modeling is likely to be masking many of the more local effects that result from terrain shape in small valleys. Second, the standard exposure of U.K. meteorological stations means that data are preferentially collected from flat, open areas.

Of the various significant covariates, the incorporation of an urban index to reflect the heat island concept enabled the reduction of residuals for minimum temperatures in central urban areas relative to those reported by others such as Landau and Barnett (1996) and Len-

TABLE 4. Poorest performing locations for minimum temperature predictions, by (a) average bias and (b) annual variance in bias. In (a), positive figures imply underprediction by the interpolator and negative figures relate to overpredictions.

(a) Meteorological station	Bias	(b) Meteorological station	Variance
Hartburn Grange	−1.67	Sandown	2.4
Bracknell Beaufort Park	−1.6	Bardsey Island	2.55
Scilly (St. Marys)	−1.37	Trawsfynydd	2.56
Bastreet	−1.37	Marlborough	2.72
East Hoathly	−1.17	Bracknell (Beaufort Park)	2.85
Moel Cynnedd	−1.15	Bude	2.87
Malvern	1.07	Moel Cynnydd	2.88
Plumpton	1.14	Elmstone	3.59
Manchester Weather Centre	1.16	Bastreet	3.94
London Weather Centre	1.72	Scilly (St. Marys)	5.24

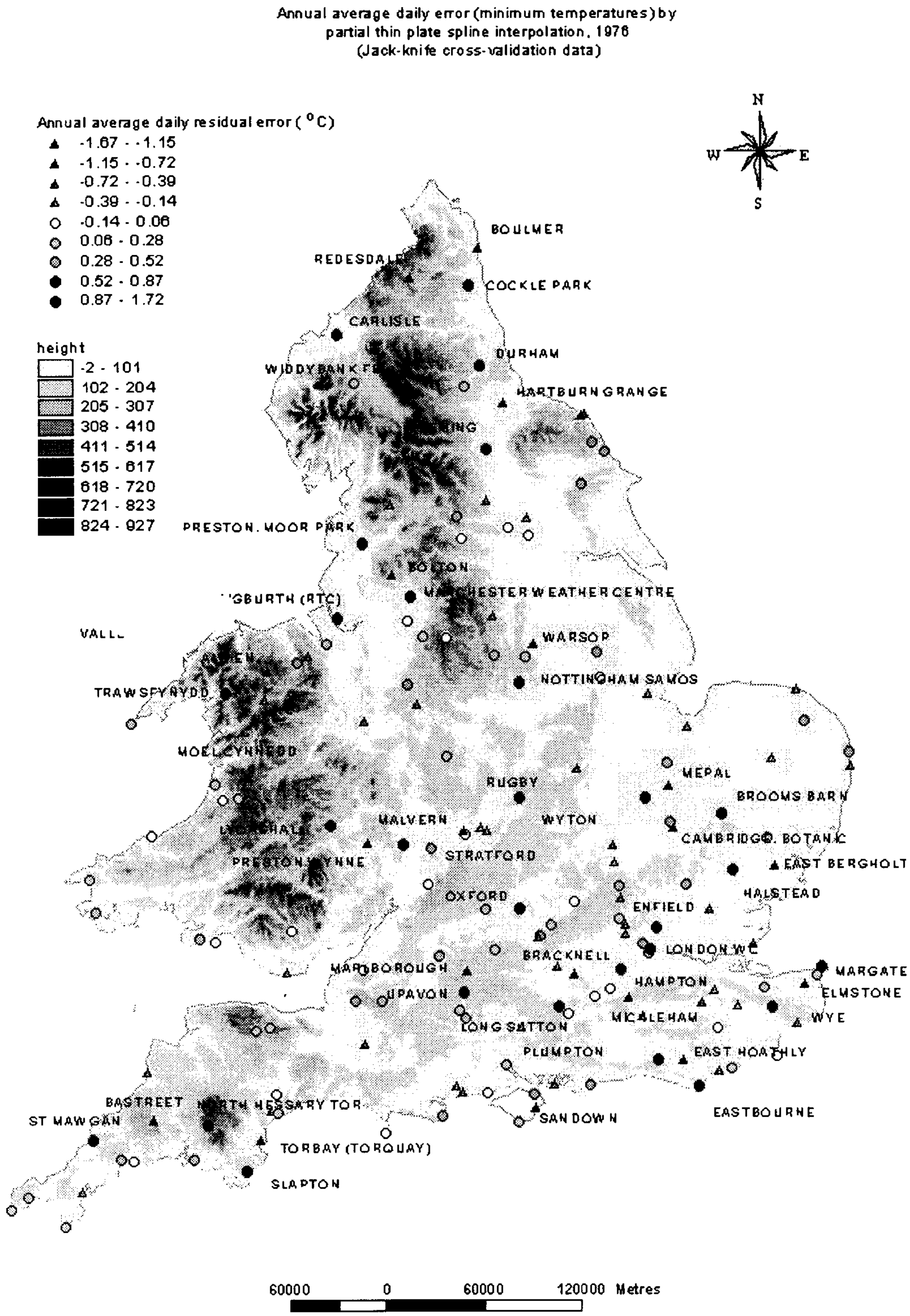


FIG. 7. Average daily residual of minimum temperature, partial thin plate spline interpolation, 1976.

non and Turner (1995). Temperatures remain slightly underpredicted on the outskirts of large cities, suggesting an amendment to the manner in which the index tails off in proportion to suburban land cover. As with previous studies for Britain, residual error was highest on the coastal margins, even with the incorporation of multiple directional variables. Investigation of the most problematic locations suggests that incorporating coast-

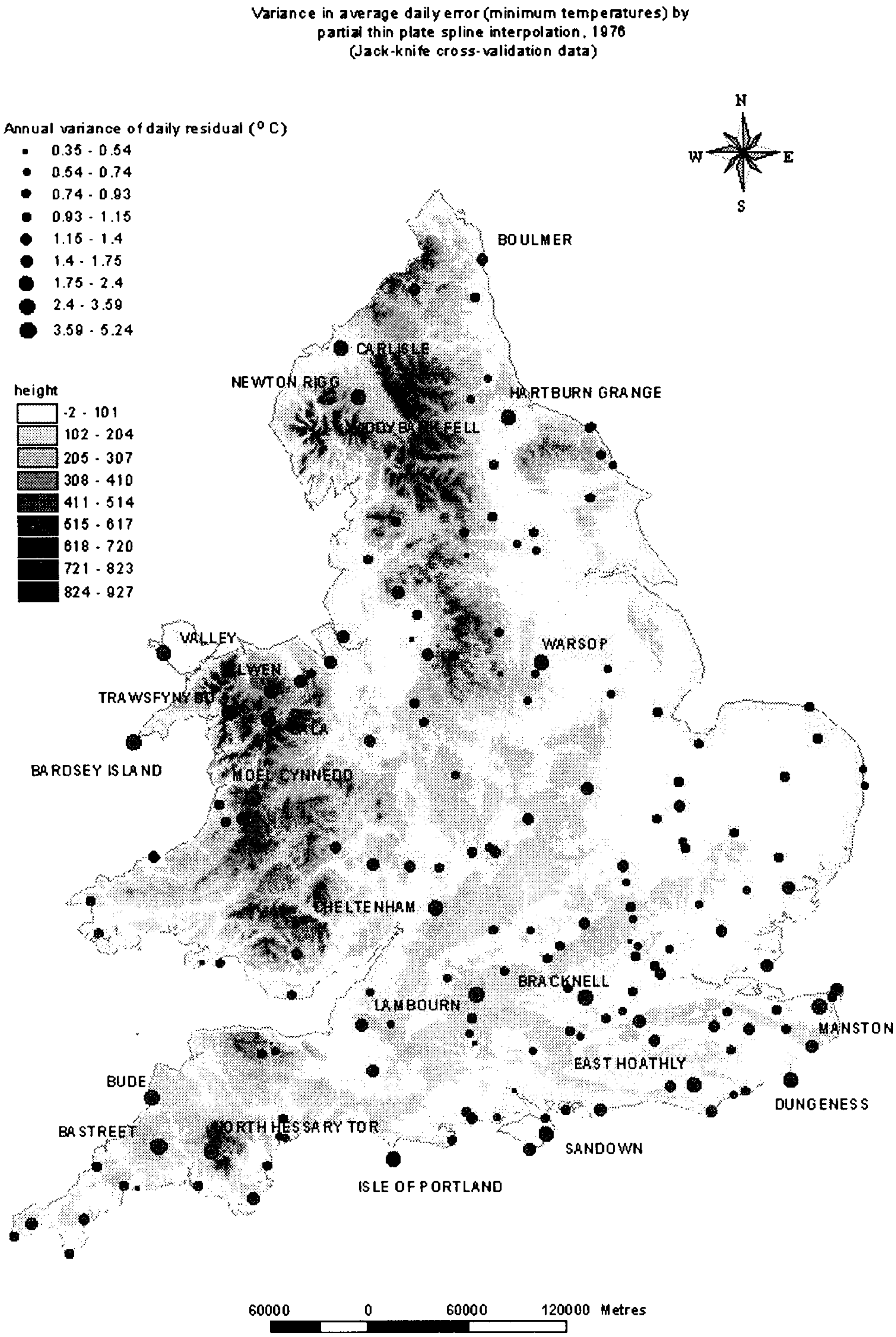


FIG. 8. Variance of daily residual in minimum temperature, partial thin plate spline interpolation, 1976.

al situation and landscape form (e.g., clifftop and local aspect) would be of benefit to future work. Large residuals at the highest elevations reflect extrapolation at underrepresented sites in the meteorological network as a whole, but data scarcity (in geographical or covariate space) was otherwise not associated strongly with the highest estimation errors. The variability of the land surface at large and small scales proved influential, with the incorporation of watershed-based rather than arbitrary geometrical functions (e.g., Lennon and Turner 1995) found to be beneficial for representing the likelihood of cold-air ponding at the basin scale. There is considerable scope for further improving the modeling of the large-scale process of cold-air ponding, for which land cover data would be an important component.

The mapping of residuals for data points withheld from the interpolation process provided a means to explore the geographical component of error of the interpolation process in contrast to the more usual rms statistic. As such they describe errors that arise both from the limits of the guiding variables used to improve the estimation and equation itself in addition to the distribution of stations that can lead to sparse sampling of the temperature surface in certain areas.

Last, it has been suggested that the improvements in accuracy that are gained by incorporating additional appropriate guiding variables are greater than those obtained by using a more complex mathematical interpolation algorithm instead (Cornford 1997). The balance between interpolator choice relative to the inclusion of guiding variables will be explored further in the following paper (Jarvis and Stuart 2001).

5. Conclusions

In this study, variables were chosen for inclusion in the interpolation process on the basis of *both* their strength of relationship and according to their consistency of selection to account for both average and extreme weather patterns. Multivariate linear modeling identified elevation, urban land use, coastal processes, and topographic variability as particularly dominant factors affecting British daily maximum and minimum temperatures at a grid resolution of 1 km. Previous work interpolating daily and monthly temperatures has often ignored the potential impact of urban development on minimum temperatures (e.g., Lennon and Turner 1995; Landau and Barnett 1996). Analysis of the residuals from partial thin plate spline interpolation suggests that the incorporation of coastal shape and situation, land cover, and soils data might improve the modeling of local-scale climate processes.

A larger set of terrain variables was needed to guide the interpolation number of minimum temperature than of maximum temperatures, but stronger influences were identified between terrain and temperature for maxima. The strength and direction of importance for these variables varied by season, and by day according to weather type, suggesting that adapting the manner in which covariates are used to guide interpolation on a daily basis is needed to ensure accurate estimations yearround. Operational constraints meant that a single set of guiding covariates was selected for further use despite this variability, but the work nevertheless advances upon the static and more limited trend model used by Landau and Barnett (1996) in an applied modeling context. The combined guiding variable/partial thin plate spline interpolation method achieved rms accuracies of 0.80° and 1.14°C for daily maximum and minimum temperatures aggregated over the year, which is comparable to other studies (e.g., Landau and Barnett 1996; Bolstad et al. 1998). The successful interpolation of a nationwide sequence of daily maximum and minimum temperatures through a full annual cycle to a resolution of 1 km using a partial thin plate spline method more commonly used for smoother monthly temperature surfaces has not previously been reported in the literature.

Acknowledgments. This work was initiated under a collaborative Ph.D. studentship for Claire Jarvis at the Department of Geography, University of Edinburgh, funded by Central Science Laboratory (CSL). Thanks are owed to Dr. R. Baker for initial work on selecting UKMO stations for analysis, to Dr. M. Hims for his ongoing interest in the work, and to Dr. R. Harker for providing UKMO meteorological data. The partial thin plate spline results were computed using the ANUSPLIN software, made available to the project by Dr. M. Hutchinson, whose comments on an earlier draft of this paper were appreciated. Neil Stuart is grateful to the Royal Society of Edinburgh for the grant of a Research Fellowship for 2000, during which this paper was completed. The opinions expressed in this paper are those of the authors personally.

REFERENCES

Barry, R. G., and R. J. Chorley, 1982: *Atmosphere, Weather and Climate,* Methuen, 407 pp.

Blennow, K., and P. Persson, 1998: Modelling local-scale frost variations using mobile temperature measurements with a GIS. *Agric. For. Meteor.,* **89,** 59–71.

Bolstad, P. V., L. Swift, F. Collins, and J. Régnière, 1998: Measured and predicted air temperatures at basin to regional scales in the southern Appalachian Mountains. *Agric. For. Meteor.,* **91,** 161–176.

Collins, F., and P. V. Bolstad, 1996: A comparison of spatial interpolation techniques in temperature estimation. *Proc. Third Int. Conf./Workshop on Integrating GIS and Environmental Modelling,* Santa Fe, NM, National Centre for Geographical Information and Analysis, CD-ROM.

Cornford, D., 1997: The development and application of techniques for mapping daily minimum air temperatures. Ph.D. thesis, The University of Birmingham, 161 pp. [Available from The University of Birmingham, Edgbaston, B15 2TT, United Kingdom.]

Fuller, R. M., G. B. Groom, and A. R. Jones, 1994: The land-cover map of Great Britain—an automated classification of Landsat Thematic Mapper data. *Photogramm. Eng. Remote Sens.,* **60,** 553–562.

Goodale, C. L., J. D. Aber, and S. V. Ollinger, 1998: Mapping precipitation, temperature and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Res.,* **10,** 35–49.

Hutchinson, M. F., 1991: The application of thin plate smoothing splines to continent-wide data assimilation. *Data Assimilation Systems,* J. D. Jasper, Ed., Bureau of Meteorology, 104–113. [Available from Bureau of Meteorology, G.P.O. Box 1289K, Melbourne, Victoria 3001, Australia.]

——, and G. C. Gallant, 1999: Representation of terrain. *Geographical Information Systems,* P. A. Longley et al., Eds., 2d ed., John Wiley and Sons, 105–124.

Jarvis, C. H., and N. Stuart, 2001: 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. Meteor.,* **40,** 1075–1084.

Landau, S., and V. Barnett, 1996: A comparison of methods for climate data interpolation, in the context of yield predictions from winter wheat simulation models. *Aspects Biol.,* **46,** 13–22.

——, R. A. C. Mitchell, V. Barnett, J. J. Colls, J. Craigon, K. L. Moore, and R. W. Payne, 1998: Testing winter wheat simulation models' prediction against observed UK grain yields. *Agric. For. Meteor.,* **89,** 85–99.

Laughlin, G. P., and J. D. Kalma, 1990: Frost risk mapping for landscape planning: A methodology. *Theor. Appl. Climatol.,* **42,** 41–51.

——, M. F. Hutchinson, and B. G. Mackey, 1993: An intuitive approach to analysing small point-source spatial data sets. *Int. J. GIS,* **7,** 21–38.

Lennon, J. J., and J. R. G. Turner, 1995: Predicting the spatial distribution of climate: Temperature in Great Britain. *J. Anim. Ecol.,* **64,** 370–392.

Linacre, E., 1992: *Climate Data and Resources: A Reference and Guide,* Routledge, 366 pp.

Manley, G., 1944: Topographical features and the climate of Britain: A review of some outstanding effects. *Geog. J.,* **103,** 241–262.

Mayes, J., and D. Wheeler, 1997: The anatomy of regional climates in the British Isles. *Regional Climates of the British Isles,* D. Wheeler and J. Mayes, Eds., Routledge, 9–45.

Mitás, L., and H. Mitásová, 1999: Spatial interpolation. *Geographical Information Systems,* P. A. Longley et al., Eds., 2d ed. John Wiley and Sons, 481–492.

Oke, T. R. 1973: City size and the urban heat island. *Atmos. Environ.,* **7,** 769–779.

——, 1976: Inadvertent modification of the city atmosphere and the prospects for planned urban climates. *Proc. Symp. on Meteorology Related to Urban and Regional Land-Use Planning,* Asheville, NC, WMO, 151–175.

——, 1987: *Boundary Layer Climates.* Methuen, 435 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. Meteor.,* **101,** 81–94.

Running, S. W., and P. E. Thornton, 1996: Generating daily surfaces of temperature and precipitation over complex topography. *Proc. Second Int. Conf. on Environmental Modelling and GIS,* Breckenridge, CO, National Centre for Geographical Information and Analysis, 93–98.

——, R. R. Nemani, and R. D. Hungerford, 1987: Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evapotranspiration and photosynthesis. *Can. J. For. Res.,* **17,** 472–483.

Stirling, R., 1997: *The Weather of Britain.* Giles de la Mare, 306 pp.

Supit, I., 1997: Predicting national wheat yields using a crop simulation and trend models. *Agric. For. Meteor.,* **88,** 199–214.

Tabony, R. C., 1985: Relations between minimum temperature and topography in Great Britain. *J. Climatol.,* **5,** 503–520.

Tufnell, L., 1997: North-west England and the Isle of Mann. *Regional Climates of the British Isles,* D. Wheeler and J. Mayes, Eds., Routledge, 158–180.

Willmott, C. J., C. M. Rowe, and W. D. Philpot, 1985: Small scale climate maps: A sensitivity analysis on some common assumptions associated with gridpoint interpolation and contouring. *Amer. Cartographer,* **12,** 5–16.