INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **25**: 1023–1039 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/joc.1160

THE DEVELOPMENT OF A NEW SET OF LONG-TERM CLIMATE AVERAGES FOR THE UK

MATTHEW PERRY* and DANIEL HOLLIS
Met Office, FitzRoy Road, Exeter, Devon EX1 3PB, UK

Received 1 April 2004
Revised 10 January 2005
Accepted 10 January 2005

ABSTRACT

Monthly and annual long-term average datasets of 13 climate variables are generated for the periods 1961–90 and 1971–2000 using a consistent analysis method. Values are produced for each station in the Met Office's observing network and for a rectangular grid of points covering the UK at a horizontal spacing of 1 km. The variables covered are mean, maximum, minimum, grass minimum and soil temperature, days of air and ground frost, precipitation, days with rain exceeding 0.2 and 1 mm, sunshine, and days with thunder and snow cover.

Gaps in the monthly station data are filled with estimates obtained via regression relationships with a number of well-correlated neighbours, and long-term averages are then calculated for each site. Gridded datasets are created by inverse-distance-weighted interpolation of regression residuals obtained from the station averages. This method does not work well for days of frost, thunder and snow, so an alternative approach is used. This involves first producing a grid of values for each month from the available station data. The gridded long-term average datasets are then obtained by averaging the monthly grids.

The errors associated with each stage in the process are assessed, including verification of the gridding stage by leaving out a set of stations. The estimation of missing values allows a dense network of stations to be used, and this, along with the range of independent variables used in the regression, allows detailed and accurate climate datasets and maps to be produced. The datasets have a range of applications, and the maps are freely available through the Met Office Website.

KEY WORDS: long-term average; UK; climate normal; gridded data; missing data estimation; spatial interpolation; regression

1. INTRODUCTION

1.1. Aims

Long-term averages (LTAs) are a simple but effective means for describing the state of the climate. They have a number of important uses, including placing values for an individual day, month or year into context, evaluating global or regional climate models, hydrological modelling, and monitoring climate change through the comparison of LTAs for different periods. For maximum usefulness, the averages need to be provided not only as values for individual observing sites, but also as gridded datasets.

This paper describes the production of new sets of UK monthly and annual averages for the periods 1961–90 and 1971–2000. These averages have been calculated for each observing site in the Met Office's climate data archive and for a high resolution 1 km × 1 km grid of points covering the UK.

Averages have been produced for each of the 13 variables listed in Table I. These variables have various characteristics. Seven relate to temperature, three to rainfall, and one relates to each of sunshine, thunder and snow. In addition, six are averages of daily values, one is a summation and six are counts of 'days of'. These characteristics determine the statistical properties of the data, which in turn influence the choice of analysis techniques for deriving the LTAs.

* Correspondence to: Matthew Perry, Met Office, FitzRoy Road, Exeter, Devon EX1 3PB, UK; e-mail: matthew.perry@metoffice.gov.uk

Table I. List of the climate variables for which long-term averages have been produced

Climate variable	Statistic	Unit
Maximum temperature	Average of daily (09–09) maxima	°C
Minimum temperature	Average of daily (09–09) minima	°C
Mean temperature	Average of daily maxima and minima	°C
Grass minimum temperature	Average of daily minima	°C
30 cm soil temperature	Average of daily measurements at 0900	°C
Sunshine duration	Average of daily durations	h/day
Precipitation amount	Sum	mm
Days of rain $\geqslant$0.2 mm	Count	Days
Days of rain $\geqslant$1 mm	Count	Days
Days of air frost	Count of days where min temp <0 °C	Days
Days of ground frost	Count of days where grass min temp <0 °C	Days
Days of snow lying	Count of occurrences at 0900	Days
Days of thunder	Count of days when heard	Days

The variables chosen are those used by the Met Office in the production of UK climate summaries, for which gridded climate normals are required in order to put monthly statistics into context.

It is recognized that there are some notable omissions, such as wind speed, wind direction, humidity, visibility, solar radiation, snow depth and days of snow falling. It is expected that some or all of these variables will be addressed in future projects.

1.2. Background

We have chosen to follow a two-stage process in the calculation of LTA statistics. First, station values are calculated for all stations with at least 4 years of data during the period 1961–2000. This is done by estimating missing values to fill in gaps in the monthly climate statistics, ensuring that station LTAs are calculated from data for the same complete 30 year period and so are not biased by the presence of missing years. Second, the station values are interpolated to a 1 km × 1 km grid using geographical information system (GIS) capabilities.

1.2.1. Estimation of missing data. There are several possible approaches to estimating missing monthly climate data. Spatial methods (e.g. kriging, inverse-distance-weighted interpolation, interpolation of regression residuals) use all data available for a single month to create a model of the climate from which a value for any location can be estimated. Methods based on time series (e.g. constant difference, multiple regression, normal ratio, principal component analysis) use data from a smaller number of stations over a longer time period. These methods are generally preferred because they are usually able to model more localized effects better by taking into account the available data history of each station.

Comparative studies of different methods for estimating missing monthly climate data have been made by Tabony (1983) for UK temperature and sunshine, Tang *et al*. (1996) for Malaysian rainfall, and Xia *et al.* (1999) for temperature, vapour pressure, wind speed and precipitation at German forest sites. They have all found that the best method is a time-series approach. For Malaysian rainfall the best results were achieved using a modified normal ratio method, whereby estimates are generated from a distance-weighted average from a number of neighbouring sites of the monthly anomaly (ratio) from the climate normal. Xia *et al.* (1999) found the best method to be multiple regression with data records from neighbour stations. We use a method based on that proposed by Tabony (1983), who had similar data and purpose to ourselves. In this method, neighbour stations are selected, and their contributions weighted, based on their correlation with the target station. The correlation coefficient is adjusted based on the length of overlapping record between the stations. Linear regression is used to generate the estimates, and a number of neighbours are used to reduce random errors.

This method is unsuitable for 'days of' variables due to their limited range of possible values; so, for these variables, we use a spatial interpolation approach.

1.2.2. Interpolation. Various methods for interpolating irregularly distributed station data to produce gridded LTAs have been tried. These include polynomial regression (Goodale *et al.*, 1998), thin-plate splines (New *et al.*, 1999), kriging of regression residuals (Agnew and Palutikof, 2000), weighted local regression (Daly *et al.*, 2002), and inverse-distance weighting (Nalder and Wein, 1998).

Splines are probably most useful in covering large areas, e.g. the global interpolation of New *et al.* (1999), or in data-sparse areas, e.g. the Canadian study of Price *et al.* (2000). It tends to produce surfaces that are more smoothly rounded compared with the natural spatial variability of the UK climate. Kriging should produce optimum results, but it is rather complex and also tends to give over-smooth surfaces.

The method we have used for the production of gridded datasets of climate normals is inverse-distance-weighted interpolation of residuals from a multiple regression model. We have found that this method gives good results that look realistic according to our knowledge of the UK climate, and it is also easy to implement for the production of maps for a wide range of climate variables. Much of the accuracy of the method is due to the availability of a dense and fairly regularly spaced network of observing stations.

Other workers have used similar methods with good results. Nalder and Wein (1998) tested a method (GIDS) that combines multiple linear regression with distance weighting and found that it gave the lowest errors for both temperature and precipitation normals in Canada, compared with methods such as co-kriging and universal kriging. Brown and Comrie (2002) also used inverse-distance weighting to interpolate regression residuals of precipitation in the southwest USA after finding that ordinary kriging gave poor results. The use of multiple regression to model variations in the climate variable with respect to geographic and topographic factors has been widely demonstrated to add considerable value when used as a prior stage to spatial interpolation (Agnew and Palutikof, 2000; Vicente-Serrano *et al.*, 2003).

2. ANALYSIS METHODS

2.1. Data

The starting point for the current analysis is an array of monthly climate statistics for the period 1961 to 2001. This is extracted from the Met Office's database of climate statistics, which is derived from a database of daily observations from the Met Office network of observing sites across the UK. These values have been subjected to a thorough quality control procedure, with suspect and some missing values being replaced by estimates.

Some gaps in the daily record do remain, however, and monthly values are only extracted subject to a maximum number of missing daily values in the month. Owing to the large volume of data involved, it was not possible to fill in gaps on a daily basis. For precipitation and all 'days of' variables, no missing days are permitted because this would introduce bias into the results, as these are cumulative variables and because of the higher daily variability of these variables. No check was made as to whether the precipitation gaps occurred during dry periods, for example, because this would be a large amount of effort that would only add a small amount of extra data to an already dense network. For the other variables, a test was done to compare the average error introduced by having increasing numbers of missing days in the month with the average error associated with estimating the monthly value (see Section 3.1). For temperature, the root-mean-square (RMS) estimation error was between the RMS error for two and three missing days, so a maximum of two missing days was allowed. For sunshine, the results indicated that up to five missing days should be allowed.

Initially the data array is relatively empty, a consequence of the rate at which sites open and close, with around 5% of the observing network changing each year. For air temperature, 1490 stations reported at some point between 1961 and 2000, but only an average of 560 of these were open at any one time. This gives an array that is 38% complete. The density of the network varies between climate elements, but the array is always around 40% complete. For rainfall there are a total of 12 100 stations, for grass temperature 1190, for snow and thunder 1020, for sunshine 730 and for soil temperature 570.

Relatively few stations have data for every year of a 30 year averaging period from which LTAs can be calculated. In order to avoid bias caused by averages calculated from different periods, the solution is to fill in the gaps using an appropriate estimation technique.

2.2. *Estimation of missing monthly values*

A test was carried out to compare a spatial interpolation estimation method with a basic time-series based method (the constant difference approach). It was found that the constant difference approach performed significantly better overall for maximum temperature. It was decided to follow closely the approach of Tabony (1983), which improves on the constant difference method by estimating missing monthly values using linear regression against data from neighbouring stations in periods where the records overlap. The six best neighbours are chosen based on their correlation with the target station, and a weighted average of the estimates from the six neighbours is used as the final estimate. The method was programmed in Fortran for automatic use.

For each station with gaps in the record, the correlation coefficient r with each other station is calculated. The length of the overlapping record is taken into account by converting to Fisher's z statistic, and subtracting two standard errors from z to get the lower confidence limit z_{low}. This is used to rank the neighbours, and for each missing value the six highest ranked neighbour stations with data available in that month are used (although less than six will be used if there are not enough stations with a strong enough relationship with the target station). A linear regression equation is calculated for each neighbour, using data for the years of overlapping data with the target station, and this equation is used to calculate the estimate from that neighbour. A weighted average (based on the correlation coefficient) of the estimates from the six neighbours is taken to arrive at the final estimate for the missing value. This ensures that the stations with the strongest and most reliable relationships with the target station have the greatest input to the final estimate. This procedure is followed separately for each month (i.e. all Januarys are analysed separately from all Februarys, etc.), unlike Tabony, who smoothed regression parameters over the 12 months. It is only attempted for stations with at least 4 years of original data.

The infilling method only makes use of neighbours with a value of z_{low} greater than zero. This is equivalent to setting a lower limit on the correlation coefficient that is dependent on sample size (i.e. the length of the overlapping record). Clearly, altering the number of standard errors that are subtracted from z (i.e. the level of confidence required that there is a true relationship between the stations) will alter the minimum correlation coefficient, which may change the number of available neighbours. It will also affect the ranking and relative weights of the neighbours. The method was tested with different levels of confidence to see what impact this had on the quality of the estimates. It was found that setting the number of standard errors at either two or three gave the best results. Results with more rigorous thresholds (standard errors greater than three) are sometimes better, but are more variable, and increasing numbers of estimates are unable to be produced due to a lack of qualifying neighbours. It was decided to choose two standard errors in order to minimize the error whilst maintaining a very low level of missing estimates. Table II shows the implicit lower limit on r when two standard errors are subtracted from z for various lengths of overlapping record.

This method does not perform well for the non-rainfall 'days of' variables (air frost, ground frost, thunder and snow cover), which have highly skewed distributions and often have very few occurrences. In fact, stations often have a series of identical values (e.g. zero) for the period when they were open, making it impossible to form regression relationships with other stations. Thus, a different approach is required for these variables, and the spatial regression and interpolation method (see Section 2.4.1) was used. This involved producing a grid of values for each month of the 1961–2000 period and then generating an estimate for each gap in the data array by interpolating a value from the relevant grid. The monthly grids are produced using all available original data for each month, a network of stations which is changing gradually during the period. In particular, the network for snow and thunder from 1961 to 1970 was very sparse, which will affect the quality of the results in this period. Unlike the 'Tabony' method, the spatial method is not able to model individual station characteristics or local effects such as frost hollows, but does produce estimates that are representative of the area.

Table II. The minimum correlation coefficient, for different sample sizes, below which neighbouring stations are not used in the infilling process

Length of overlapping record (years)	Minimum r
<4	Not used
4	0.96
6	0.82
10	0.64
20	0.45
30	0.37
40	0.32

2.3. *Calculation of station LTAs*

Once the gaps in the array have been filled, LTAs for the periods 1961–90, 1971–2000 and 1991–2000 can be calculated for each station from the complete array.

There may still be some gaps in the array for variables infilled using the Tabony method. These will mostly be for stations with less than 4 years of data, which are not used; but other stations, especially those with a fairly short record and when the network of stations is sparse, may not have any well-correlated neighbours from which to generate an estimate.

Only 0.2–0.4% of monthly values remain missing for most variables, except for grass minimum temperature and sunshine, which have slightly more gaps up to 0.8%. In these situations, the station LTA is calculated from the available years of data as long as the number of years exceeds a given threshold. This threshold is set by comparing the error introduced by having missing years in the averaging period with the error associated with creating an estimate for the location by interpolation from a grid (see Table VII). The threshold varies from 7 years of data required in a 30 year period for grass minimum temperature, 10 years for minimum temperature, 12 years for sunshine, and 14 years for maximum temperature, to 16 years for precipitation. An example, for minimum temperature, of the crossover graph used to set these thresholds is given in Figure 1. This shows the RMS error (RMSE) from tests made on complete stations for the 1961–90 period. Periods of consecutive years are excluded, either starting from 1961, ending in 1990, or starting in any year.

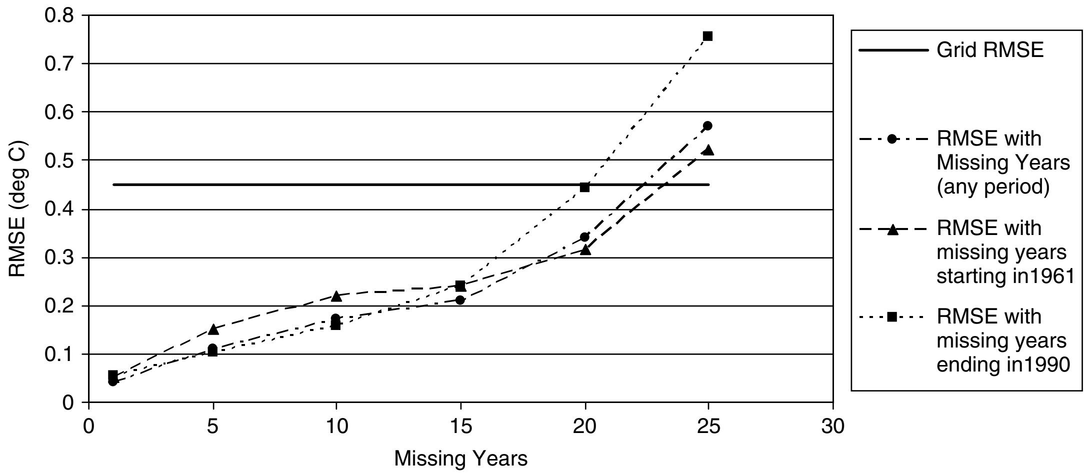

Figure 1. Comparison of grid errors with errors due to missing years for minimum temperature, showing crossover point

2.4. *Generation of gridded datasets*

For most variables, the station LTAs are interpolated to a regular 1 km × 1 km grid of values covering the UK (the 'average then grid' approach). All available monthly station LTAs are used, and the network of stations will vary slightly between months due to each month having been considered separately for the estimation of missing values and thresholds of data availability. However, for the non-rainfall 'days of' variables, the long-term average grids are created by averaging the monthly grids from each year of the 30 year period ('grid then average' approach).

Although 'average then grid' is generally the preferred method, 'grid then average' is used for days of frost, snow and thunder because the estimates have been made using the same method as is used for the gridding, and so they are not adding any extra information.

2.4.1. Description of the methodology. The method used to create the LTA grids from an input of irregularly spaced station data is described in this section. The gridding was carried out using functionality programmed into an ESRI ArcView 3.2 GIS.

The heart of this process is the interpolation of station data onto a regular 1 km × 1 km grid using inverse-distance-weighted averaging (Burrough, 1986). The value at each grid point is calculated as a weighted average of surrounding station values (Perry and Hollis, 2005). The value of the power parameter needs to be chosen, together with the radius within which points will be used in the weighted average. The inverse-distance-weighted method uses all data values within a specified search radius, but the radius may optionally be expanded if fewer than 12 stations are found. It also has an option to select a modified weighting function that prevents the weight going to infinity when station and grid point coincide, and an optional adjustment for spatial variations in station density.

Many climate variables exhibit dependencies on the geographic characteristics of the surrounding area. Geographic effects are removed from the data prior to interpolation by creating a model of those effects using multiple regression analysis (with the station data as the dependent variable). Table III shows the factors that are available for inclusion as independent variables in the regression analysis.

The regression estimate for each station is subtracted from the corresponding LTA to obtain a set of regression residuals that should be largely free of geographic effects. These residuals are then interpolated onto a regular grid. Finally, the regression model is evaluated at each grid point and the result is added to the interpolated residual, thus producing a grid of the original climatological variable.

2.4.2. Exclusion of poorly fitting stations. After an initial run of the gridding analysis is made, an analysis of the regression residuals at stations combined with a visual inspection of the grids is used to assess whether there are any poorly fitting stations that should be excluded from the gridding analysis process.

Stations were ranked by the mean absolute residual over the 12 months for both 1961–90 and 1971–2000. The maximum residual error in any month was also considered. Suitable thresholds for the exclusion of stations were set in order to capture outliers. For some variables, especially rainfall, it was necessary to be more flexible with the exclusion of stations because the regression had been unable to model all

Table III. Topographic and geographic variables used in the regression model

Easting and northing	To capture spatial trends. They are incorporated as second- or third-order cross-product polynomials
Altitude	The elevation above mean sea level, from a 500 m digital elevation model interpolated to 100 m
Terrain shape ('north', 'south', 'east' and 'west')	Mean altitude over a circle of radius 5 km offset by 10 km to the north, south, east and west
Sea	The proportion of sea within a 5 km radius
Urban	The proportion of urban land use within a 5 km radius, calculated from a mid-1970s land-use dataset that does not cover Northern Ireland

significant spatial variations in the data. Each LTA grid was inspected for bull's-eyes or other peculiar features. Expert climatological knowledge was used throughout this process, and stations were removed from the analysis where necessary. In a few cases the analysis method itself needed to be fine tuned, typically by adjusting the smoothness of the interpolation scheme. The grids were then recalculated and rechecked. Depending on the climate variable, between two and five passes were required to produce satisfactory grids.

Stations may fit poorly when they are affected by local micro-climates, such as frost hollows. This is a weakness in the gridding techniques, which are unable to model these very localized effects. They may also fit poorly if the estimation of missing values has been inaccurate, which may occur if the station had a short original record. Some of the sunshine stations excluded were poorly exposed, with the sun being blocked at certain times, e.g. by surrounding mountains or trees. The largest proportion of stations was excluded for grass minimum temperature, which suffers from variations in measurement period between different types of station (18–09, sunset–09, or 09–09), as well as being highly affected by local factors, such as soil type. For raindays $\geqslant$0.2 mm, some stations are affected by dew, causing values of 0.2 mm to be returned.

One of the main applications of the LTA grids is in the production of climatological maps, where they are used in the calculation of climate anomalies. The use of the LTA grid also allows anomalies to be calculated for any station, including recently opened sites (for which there are no station LTAs). Problems arise, however, if the grids are used to generate anomalies for a station with a long-term average that is poorly fitted by the grid, i.e. sites where the LTA obtained by interpolation from the grid is significantly different from the station LTA. To avoid generating inaccurate anomaly values, it was decided that stations with poorly fitted LTAs would be dropped from the production of climatological maps and, for consistency, that they be excluded from the production of the LTA grids.

If a station is excluded then it is removed from the gridding process for all months, even though for some months the fit may not be particularly poor. The criteria used to identify poorly fitting stations are summarised in Table IV, and the number of stations excluded for each climate variable is given in Table V.

3. DEVELOPMENT AND VERIFICATION OF METHODS

3.1. Assessment of estimation error

The accuracy of the estimation method was tested by excluding values for 6 years (evenly spaced over the period) at each of 20 stations (selected randomly from sites with at least 30 years of data between 1961 and 2001). This was done for 1 month of each season, for maximum temperature, sunshine, rainfall, and days of rain $\geqslant$1 mm. The estimates produced by the infilling method were then compared with the actual measured values. The results, averaged across all the test months, are summarized in Table VI. Rainfall is the hardest variable to estimate, because it has high variability between years.

The overall error for station LTA values will depend on the mean station absolute bias, which is given in Table VI. For all variables, the level of uncertainty associated with these final values will generally be very low, but it will increase as the length of the original data record decreases, both because more years are missing and need to be estimated, and because individual estimates are likely to be less accurate. Stations with few well-correlated neighbours, such as those in data-sparse areas or environments, will also be susceptible to higher errors.

3.2. Choice of gridding models

To identify the best 'analysis profile' (i.e. the best combination of regression model and interpolation method) for each climate variable, error statistics were calculated using a 10% random sample of stations that had been excluded from the gridding process. This was done for 1 month of each season, or for 'grid then average' variables 1 month for each season in each of three sample years. An initial profile was chosen based on prior knowledge of the geographic factors affecting the climate variable, the density of the network, and the statistical properties of the data. An estimate of the value at the location of each excluded station

Table IV. Criteria for determining whether to use a station to produce the LTA grid. Unless specified, the test statistic is the mean absolute regression residual. Where two variables are given in the first column, stations are excluded from the production of the LTA grids for both variables if any of the criteria in the second column are met

Climate variable(s)	Criteria for exclusion
Maximum temperature	$\geq 1\,^{\circ}$C
Minimum temperature/days of air frost	Min temp $\geq 1\,^{\circ}$C
	or min temp $\geq 0.9\,^{\circ}$C and air frost ≥ 3 days
	or min temp $\geq 0.8\,^{\circ}$C and air frost ≥ 3.5 days
Mean temperature	Max temp $\geq 1\,^{\circ}$C
	or min temp $\geq 1\,^{\circ}$C
	or mean temp $\geq 0.8\,^{\circ}$C
Grass minimum temperature/days of ground frost	Grass min temp $\geq 1.2\,^{\circ}$C
	or grass min temp $\geq 1.1\,^{\circ}$C and ground frost ≥ 3 days
	or grass min temp $\geq 1\,^{\circ}$C and ground frost ≥ 3.5 days
30 cm soil temperature	$\geq 0.9\,^{\circ}$C, plus visual inspection of grids for stations with the highest residuals in each month
Sunshine duration[a]	≥ 0.5 h per day, plus visual inspection of grids for stations with residual ≥ 0.9 h per day in any month
Precipitation amount	Visual inspection of grids for stations with the highest residuals in each month
Days of rain ≥ 0.2 mm	Visual inspection of grids for stations with the highest residuals in each month
Days of rain ≥ 1 mm	Visual inspection of grids for stations with the highest residuals in each month
Days of snow lying	Inspection of all stations ≥ 3, and all station months with residual >10
Days of thunder[b]	For NCM stations: mean absolute residual >0.72 days
	For DLY3208 stations: Average residual outside the range -0.3 to 0.3
	or mean absolute residual ≥ 0.5
	or max residual ≥ 3, or min residual ≤ -3

[a] For sunshine duration, it was decided that a number of stations in data-sparse areas would be retained, despite exceeding the threshold. The remoteness of these stations from other observing sites meant both that there was no clear evidence for the data being in error and that the stations could still contribute valuable information to the analysis.
[b] For days of thunder, NCM stations operate a genuine 24 h observing regime and should be unbiased. DLY3208 stations do not always observe throughout the whole day and, as such, may have a tendency to underreport the frequency of thunder. Both types of station were assessed against LTA grids produced from just the NCM stations. The final LTA grids were produced after excluding any poorly fitting NCM stations and including any well-fitting DLY3208 stations.

was made by bilinear interpolation from the grid produced from the remaining stations. These estimates were compared with the actual values, then the resultant verification statistics, especially the RMSE, were used to assess whether subsequent adjustments made to the profiles improved the analysis.

Table VII shows the RMSE from the best profile for each climate variable, which was selected as the profile for the final analysis. This gives an idea, for each season, of the average level of error that can be expected in the LTA grids (for 'average then grid' variables) at locations between stations. The accuracy of the grids will vary spatially, however, with areas that have a sparser coverage of stations, and areas with more complex terrain, often having greater than average errors. For rainfall, for example, the RMSE of the test stations in northern Scotland is 24 mm compared with the overall RMSE of 11 mm. This region has the lowest density of stations, as well as the highest rainfall totals. In southeast England and East Anglia, a flat region with comparatively low rainfall and dense station coverage, the RMSE is just 5 mm. Similarly, for mean temperature, northern Scotland has the highest RMSE of 0.44 °C, due to its sparser station coverage and complex terrain, whereas East Anglia has the lowest RMSE of 0.22 °C. For minimum temperature, central England has the highest RMSE of 0.55 °C, probably due to difficulties with modelling frost-prone locations. There is little spatial variation in errors for sunshine.

In 30 of the 36 tests, the RMSE from the final grid (after interpolation of residuals) was lower than that from the regression surface, indicating that the interpolation of residuals is a valuable second stage to the process. This was especially the case for rainfall variables, and for minimum and grass minimum

Table V. Numbers of stations excluded from the production of the LTA grids

Climate variable	Number of excluded stations	Excluded stations (%)
Maximum temperature	12	0.9
Minimum temperature	65	5.2
Mean temperature	72	5.7
Grass minimum temperature	103	10.1
30 cm soil temperature	33	7.5
Sunshine duration	21	3.3
Precipitation amount	92	0.9
Days of rain $\geq$0.2 mm	275	2.8
Days of rain $\geq$1 mm	174	1.8
Days of air frost	65	4.5
Days of ground frost	114	9.8
Days of snow lying	13	1.1
Days of thunder	706	59.9

Table VI. Summary of the average performance of the infilling method

Climate variable	Mean station absolute bias	Mean absolute error	RMSE
Maximum temperature (°C)	0.08	0.19	0.25
Sunshine (%)	4	8	9
Rainfall (%)	8	15	21
Days of rain $\geq$1 mm (days)	0.59	1.19	1.61

Table VII. RMS gridding errors from excluded verification stations (10% sample)

Climate variable	January	April	July	October	Average
Sunshine (h/day)	0.21	0.31	0.35	0.20	0.27
Maximum temperature (°C)	0.24	0.39	0.49	0.29	0.35
Mean temperature (°C)	0.26	0.26	0.30	0.25	0.27
Minimum temperature (°C)	0.39	0.43	0.51	0.47	0.45
Grass minimum temperature (°C)	0.52	0.54	0.58	0.60	0.56
30 cm soil temperature (°C)	0.36	0.44	0.76	0.35	0.48
Precipitation (mm)	13.8	7.4	8.8	12.7	10.7
Days of rain $\geq$0.2 mm	1.22	1.10	1.07	1.17	1.14
Days of rain $\geq$1 mm	0.93	0.81	0.80	0.88	0.85
Days of thunder	0.27	0.19	0.54	0.20	0.30
		March	November		
Days of air frost	3.37	3.28	2.40		3.02
Days of ground frost	5.10	4.97	3.40		4.49
Days of snow cover	3.09	3.00	1.72		2.61

temperature. However, the interpolation made little difference for sunshine, and mean and maximum temperature.

The selection process involved the testing of geographic variables other than those listed above. These included other measures of terrain shape (e.g. slope and aspect), values of *sea* and *urban* calculated using different radii, and a smoothed grid of altitude. The relatively limited set of variables that were finally used reflects the limitations of fitting a single UK-wide regression model to monthly climate statistics. For example, it is recognized that on individual days and in restricted geographical areas it is possible for the sea to have a measurable impact at distances far greater than 5 km from the coast. However, such effects never occur equally around the whole of the UK coastline, nor, in general, do they persist throughout an entire month; consequently, they become difficult to isolate using statistical techniques. It follows that only the most spatially and temporally consistent effects are captured by the modelling process.

The power parameter of the inverse-distance weighting factor was either set to two or three; a power of two is preferred for those variables that are subject to more uncertainty due to modelling difficulties or a sparse network. A radius of 100 km was used for most variables, but for precipitation the radius was 50 km; and for days of rain and snow, 75 km was chosen. Only for days of snow lying did the radius need to be expanded to incorporate 12 stations in some areas. The easting and northing polynomial order is two for those variables with the fewest stations, to avoid spurious extrapolations in the extremes of the UK.

3.3. Gridding results and errors

Having chosen the best analysis model, and refined the analysis by excluding any poorly fitting stations and possibly adjusting the interpolation settings, the chosen profile was used to create the final LTA grids from all remaining data.

Table VIII shows the average number of stations used in the final gridding. It also gives the RMS errors for each climate variable; this measure of how well a grid fits the observations from which it was created is obtained by comparing the observations with co-located values obtained from the grid by bilinear interpolation. For the climate variables analysed by the 'average then grid' method, the comparison was between the LTA grids and the station LTAs. There are 24 grids in all (12 months for each of 1961–90 and 1971–2000). For variables analysed by the 'grid then average' method the quality is assessed in a similar way, except that monthly station statistics are compared with monthly grids (a total of 480 grids, January 1961–December 2000).

Tables IX and X show, for winter and summer half-years, the average value of each coefficient in the regression equation for each climate variable and also the average value of r^2, the proportion of the variance

Table VIII. Summary of the closeness of fit of the LTA grids and monthly grids to the data from which they were derived

Climate variable	Stations	Grids	Mean RMS error
Maximum temperature	1213	24	0.12
Minimum temperature	1165	24	0.07
Mean temperature	1156	24	0.09
Grass minimum temperature	893	24	0.08
30 cm soil temperature	392	24	0.08
Sunshine (h/day)	604	24	0.16
Precipitation (mm)	9669	24	2.7
Days of rain $\geq$0.2 mm	9029	24	1.07
Days of rain $\geq$1.0 mm	9106	24	0.81
Days of air frost	536	480	0.20
Days of ground frost	388	480	0.24
Days of snow lying	370	480	0.08
Days of thunder	149	480	0.35

Table IX. Average winter regression parameters for each climate variable

Climate variable	r^2	Constant	Geographic variables — regression coefficients							
			Sea	Altitude	Alt^2	Urban	North	South	East	West
Sunshine	0.64	2.32	0.05	−0.01	n/a	−0.04	−0.02	−0.03	0.02	−0.09
Maximum temperature	0.94	8.66	−0.04	−0.81	n/a	0.05	n/a	n/a	n/a	n/a
Mean temperature	0.94	5.72	0.27	−0.56	n/a	0.11	0.06	−0.06	−0.02	−0.02
Minimum temperature	0.89	2.77	0.56	−0.38	n/a	0.16	0.03	−0.07	−0.08	−0.02
Grass minimum temperature	0.79	0.55	0.74	−0.05	n/a	0.09	−0.01	−0.09	−0.12	0.01
Soil temperature 30 cm	0.88	6.20	0.15	−0.38	n/a	0.05	n/a	n/a	n/a	n/a
Precipitation	0.67	94	n/a	12	n/a	n/a	−6	10	4	10
Raindays ≥1 mm	0.78	13.0	n/a	0.6	n/a	n/a	−0.2	0.4	0.0	0.5
Raindays ≥0.2 mm	0.73	17.0	n/a	0.6	n/a	n/a	−0.1	0.3	0.0	0.3
Days of air frost	0.66	7.55	−1.58	0.55	0.58	−0.48	−0.31 (mean over all directions)			
Days of ground frost	0.58	13.64	−2.23	0.40	n/a	−0.35	n/a	n/a	n/a	n/a
Days of snow lying	0.53	1.97	n/a	0.78	n/a	n/a	0.18	0.34	0.07	0.20
Days of thunder	0.20	0.26	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table X. Average summer regression parameters for each climate variable

Climate variable	r^2	Constant	Geographic variables — regression coefficients							
			Sea	Altitude	Alt^2	Urban	North	South	East	West
Sunshine	0.79	5.50	0.20	−0.01	n/a	−0.04	0.01	0.01	−0.03	−0.12
Maximum temperature	0.92	16.40	−0.58	−0.89	n/a	0.06	n/a	n/a	n/a	n/a
Mean temperature	0.93	12.35	−0.05	−0.66	n/a	0.16	0.09	0.00	−0.01	−0.02
Minimum temperature	0.86	8.31	0.43	−0.44	n/a	0.24	0.04	−0.03	−0.12	0.01
Grass minimum temperature	0.70	5.99	0.60	−0.23	n/a	0.16	0.06	−0.03	−0.16	0.01
Soil temperature 30 cm	0.85	13.15	−0.05	−0.71	n/a	0.07	n/a	n/a	n/a	n/a
Precipitation	0.62	70	n/a	9	n/a	n/a	−4	5	−2	6
Raindays ≥1 mm	0.75	10.4	n/a	0.5	n/a	n/a	−0.1	0.3	−0.2	0.4
Raindays ≥0.2 mm	0.74	13.7	n/a	0.5	n/a	n/a	0.0	0.2	−0.1	0.3
Days of air frost	0.43	0.43	−0.11	0.19	n/a	−0.06	−0.14 (mean over all directions)			
Days of ground frost	0.29	3.20	−0.77	0.28	n/a	−0.24	n/a	n/a	n/a	n/a
Days of snow lying	0.34	0.20	n/a	0.33	n/a	n/a	0.06	0.00	−0.09	0.07
Days of thunder	0.38	0.90	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

explained by the regression model. Values of the polynomial relationships for easting and northing are given in Tables XI and XII, averaged over all grids for the winter and summer half-years. Note that only *easting* and *northing* were used in the regression model for days of thunder, and that for air frost the effects of terrain shape are captured via the mean altitude within a 5 km radius, rather than by the four offset mean altitudes.

The values of r^2 are clearly much higher for the 'average then grid' variables than for the 'grid then average' variables. Although this may be partly due to differences in the characteristics of the variables, the main reason is likely to be the much longer averaging period of the 'average then grid' variables; 30-year averages are more likely to reflect the climatological impact of geography than monthly means.

The regression coefficients are broadly in line with what might be expected, e.g. temperate conditions near coasts, lower temperatures and higher rainfall at high altitudes, less frost in urban areas, etc. The validity of the terrain shape parameters is more difficult to confirm without a more detailed assessment (they suggest,

Table XI. Average winter polynomial relationship for easting and northing

Climate variable	X	X^2	X^3	Y	Y^2	Y^3	XY	XY^2	X^2Y	X^2Y^2
Sunshine	0.18	−0.12	n/a	−0.16	−0.01	n/a	−0.01	n/a	n/a	n/a
Maximum temperature	−0.01	−0.63	0.36	−1.09	0.37	0.06	−0.18	0.04	0.15	−0.07
Mean temperature	−0.41	−0.97	1.00	−1.98	1.27	0.17	2.12	−2.72	−1.31	1.60
Minimum temperature	−0.70	−1.49	1.72	−2.90	2.28	0.19	4.34	−5.36	−2.70	3.18
Grass minimum temperature	−0.89	−1.32	1.49	−2.30	1.24	0.58	3.28	−3.98	−1.52	1.99
Soil temperature 30 cm	−0.85	0.45	n/a	−1.51	0.46	n/a	0.16	n/a	n/a	n/a
Precipitation	135	−384	257	−112	227	−18	206	−376	−157	211
Raindays ≥1 mm	2.3	−13.2	10.5	−6.5	10.2	1.0	13.9	−22.0	−9.8	12.4
Raindays ≥0.2 mm	0.1	−7.9	6.9	−4.5	6.6	1.1	10.0	−15.2	−6.7	8.6
Days of air frost	−3.07	10.40	−6.16	1.17	−0.27	−1.43	−1.41	5.95	0.30	−3.46
Days of ground frost	0.82	4.98	−3.86	2.56	1.77	−3.61	−4.63	7.43	1.17	−3.29
Days of snow lying	0.02	0.41	n/a	0.09	0.46	n/a	0.32	n/a	n/a	n/a
Days of thunder	−0.11	0.14	n/a	−0.40	0.37	n/a	−0.05	n/a	n/a	n/a

Table XII. Average summer polynomial relationship for easting and northing

Climate variable	X	X^2	X^3	Y	Y^2	Y^3	XY	XY^2	X^2Y	X^2Y^2
Sunshine	0.42	−0.28	n/a	−0.55	0.19	n/a	−0.11	n/a	n/a	n/a
Maximum temperature	1.40	0.00	−0.88	1.56	−2.74	0.41	−4.07	4.68	2.39	−2.89
Mean temperature	0.42	−0.13	−0.07	0.02	−1.21	0.44	−0.79	1.05	0.43	−0.73
Minimum temperature	−0.08	−1.07	1.13	−1.53	0.78	0.12	2.20	−2.45	−1.45	1.47
Grass Minimum temperature	−0.75	−0.12	0.63	−1.31	0.60	−0.03	1.49	−1.33	−0.69	0.66
Soil temperature 30 cm	0.31	−0.25	n/a	−0.55	−0.17	n/a	−0.29	n/a	n/a	n/a
Precipitation	57	−164	111	−39	89	−1	106	−186	−80	103
Raindays ≥ 1 mm	0.0	−4.2	3.8	−1.6	3.6	1.2	6.8	−10.9	−4.6	5.7
Raindays ≥ 0.2 mm	−1.6	−1.1	2.0	−1.0	2.3	1.3	5.3	−7.9	−3.3	4.1
Days of air frost	−0.92	1.92	−0.96	−0.11	−0.02	−0.33	0.00	1.32	−0.03	−0.78
Days of ground frost	0.38	0.52	−0.41	0.09	1.17	−0.63	0.48	−0.12	−0.90	0.24
Days of snow lying	−0.08	0.13	n/a	−0.10	0.24	n/a	0.08	n/a	n/a	n/a
Days of thunder	0.36	0.20	n/a	0.29	−0.34	n/a	−0.23	n/a	n/a	n/a

for example, that sites with high ground to the south and west are wetter than those with high ground to the north).

4. RESULTS

4.1. Output

Results were checked by comparison with existing 1961–90 averages calculated using similar methods in the 1990s. This was done for station averages, and for areal averages based on the gridded datasets. These comparisons showed a good general agreement in the values, with a very small proportion of large differences in station averages. The areal averages show a slight bias in the grass minimum temperature (0.15 °C higher for the new version) and days of ground frost values (0.2 days lower), which may be caused by the fact that most of the excluded stations were colder than expected. There are also marked differences in the precipitation pattern over Scotland. The new methods provide a denser coverage of stations, and much more detail in the gridded datasets.

The station LTAs have been loaded into an Oracle database, together with information on the number of original, estimated, and missing values from which they were derived. The 1 km gridded datasets are stored in ArcView format for use in the production of monthly grids, and are available for other enquiries and investigations. The LTA grids have also been converted to a series of colour-shaded maps, which are freely available via the Internet from http://www.metoffice.gov.uk/climate/uk/averages/19 712 000/mapped.html. Figure 2 shows a sample of four of these maps. The grids have also been used to calculate a set of areal averages for the UK as a whole, and its composite countries, districts, and counties, by averaging all grid points within each area.

4.2. *Comparison between periods*

This paper does not attempt to present a full study of the results produced for trends, patterns, or changes between periods, and it is hoped to analyse the output further in the future. However, some results that have been obtained are presented here, showing that there is potential for interesting patterns to be discovered. Figure 3 shows the change in winter and summer mean temperature and total precipitation between 1961–90 and 1971–2000. It shows areal averages for county and administrative areas. Winter precipitation has increased by about 10% in northern and western Scotland and northwest England. Figure 4 gives more detail on the changes in Scotland, showing that winter precipitation has increased by over 15% in parts of the western Scottish highlands, a significant change given the high level of precipitation in this area. In contrast, the eastern side of the UK has experienced little change between the two periods. There has been a general decrease in summer precipitation between the two periods, especially in northern England and eastern Scotland. The greatest increase in mean temperature occurred in the winter season, and in the southeast of England, where the change was about 0.5 °C. There was also a general increase in the summer, with East Anglia having the highest rise. Autumn temperatures increased the least.

5. CONCLUSIONS

The production of station and gridded long-term averages for 13 climate variables for the periods 1961–90 and 1971–2000, covering the whole of the UK, was successfully achieved using the methods described.

A number of issues were addressed in the development of the methods. The number of missing days to allow per station month was determined objectively. The method used to fill in gaps in the array of station data was chosen, and the selection of neighbour stations was optimized, striking a balance between the strength of the correlation with the target station and the length of the overlapping record. The estimation of missing values for non-rainfall 'days of' variables was problematic, and alternative spatial methods were developed. Further work is planned to investigate alternative statistical techniques for dealing with these types of variable. For the few remaining stations where gaps could not be filled, it was determined objectively how many years of data were required for a long-term average to be calculated from the data rather than estimated.

A two-stage process of multiple regression on geographic and topographic factors followed by inverse-distance-weighted interpolation of residuals was used to generate the LTA grids. Verification statistics give an indication of the level of accuracy of the grids between stations, and indicate that the interpolation stage adds valuable information to the regression surface, especially for rainfall variables. The regression model parameters provide an estimation of the spatial structure of the UK climate, explaining between 29 and 94% of the variance in the data, depending on the climate variable.

The estimation of missing monthly values prior to gridding provides a dense network of observation data that, together with careful quality control and exclusion of unrepresentative stations, enables detailed and accurate high-resolution (1 km × 1 km) gridded datasets to be produced. The accuracy of the estimates was tested and found to be good, especially for temperature and sunshine.

Micro-climatological effects, such as frost hollows, are difficult to capture using the modelling techniques employed in this study. This led to the failure of the LTA grids to fit sufficiently closely to some of the station values, as a consequence of which a number of stations had to be omitted from the analysis.

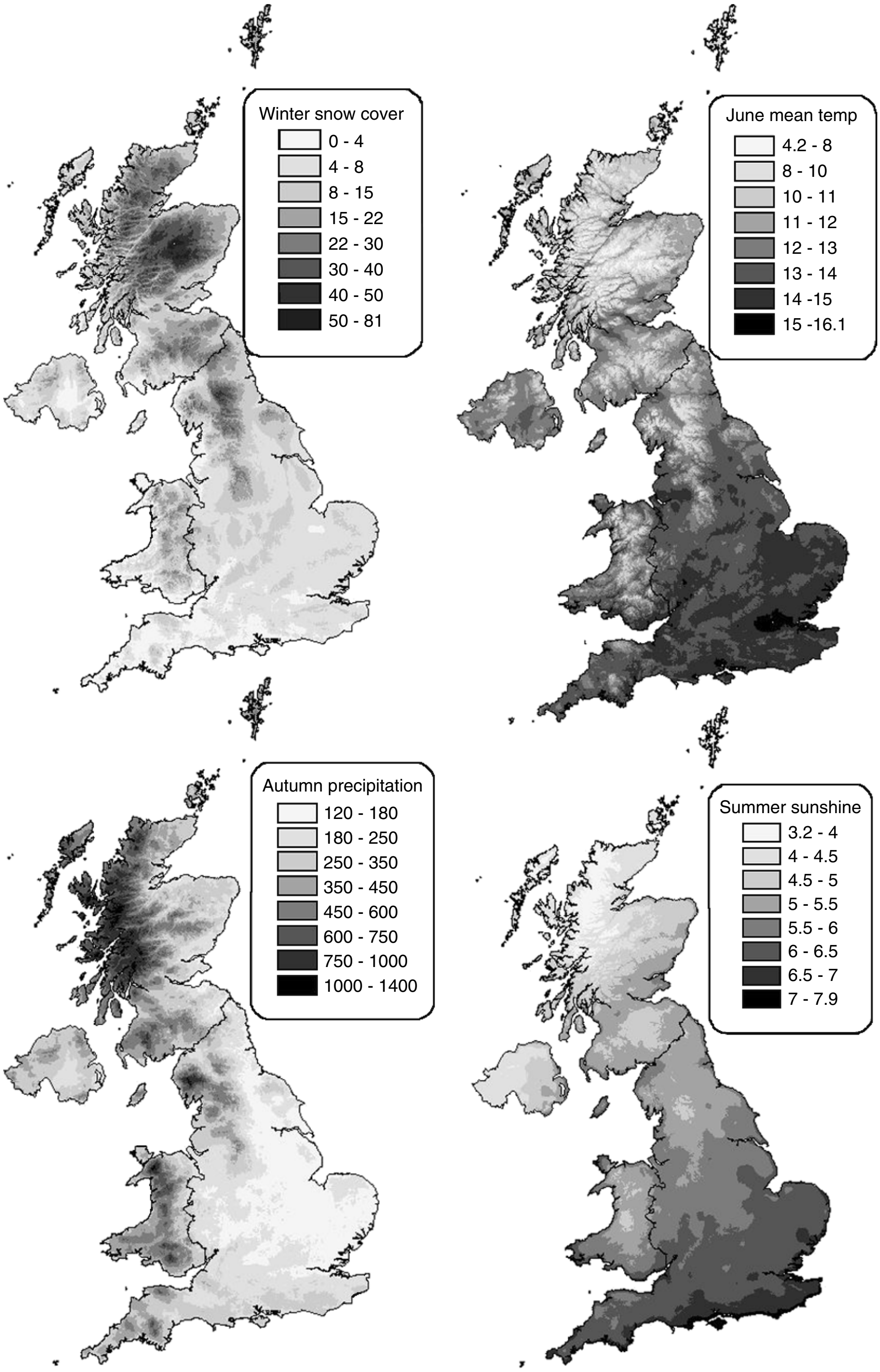

Figure 2. Example 1971–2000 LTA grid maps: (a) winter days of snow cover; (b) June mean temperature (°C), (c) autumn total precipitation (mm); (d) summer sunshine (h/day)

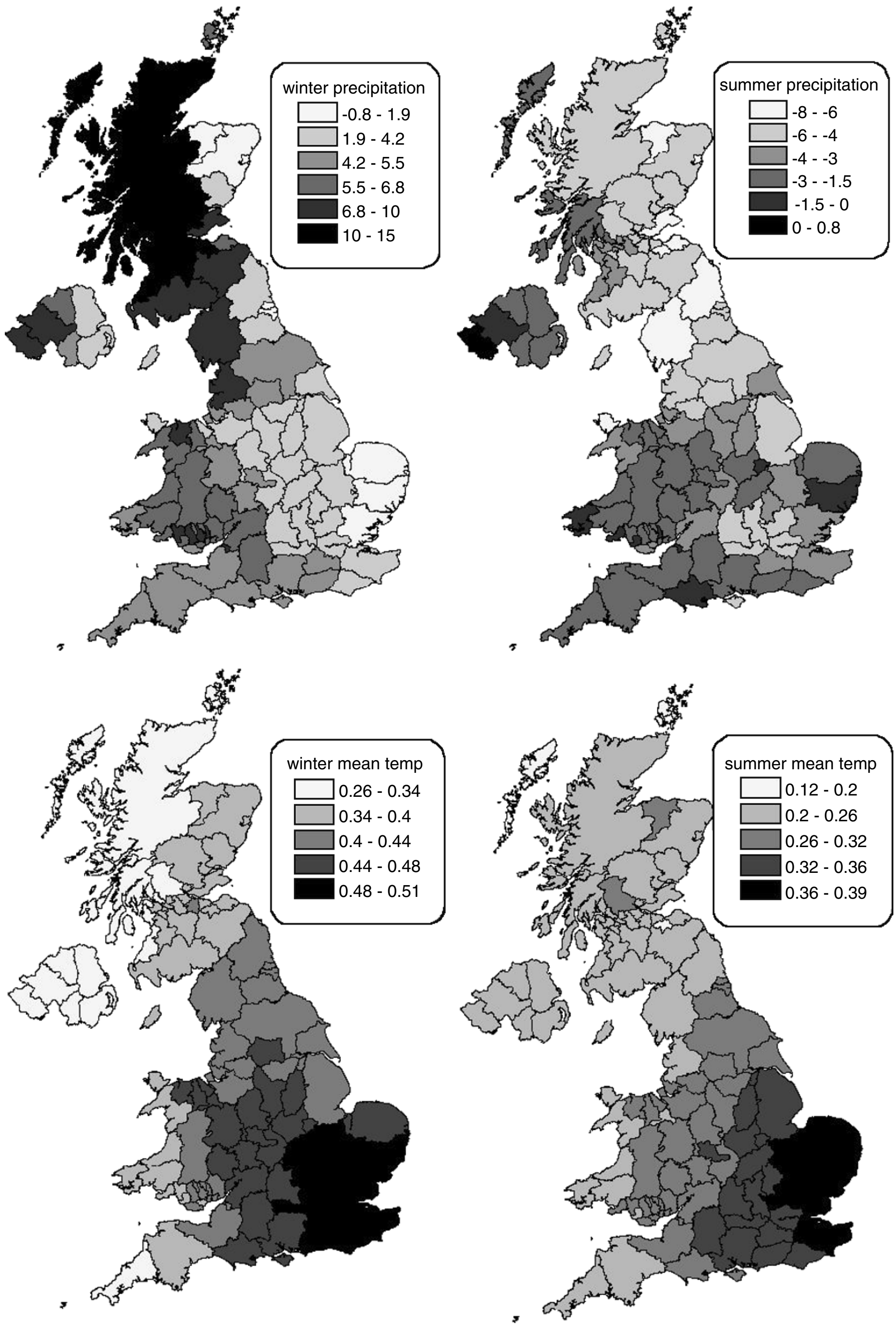

Figure 3. Change between 1961–90 and 1971–2000 periods of summer and winter precipitation and mean temperature

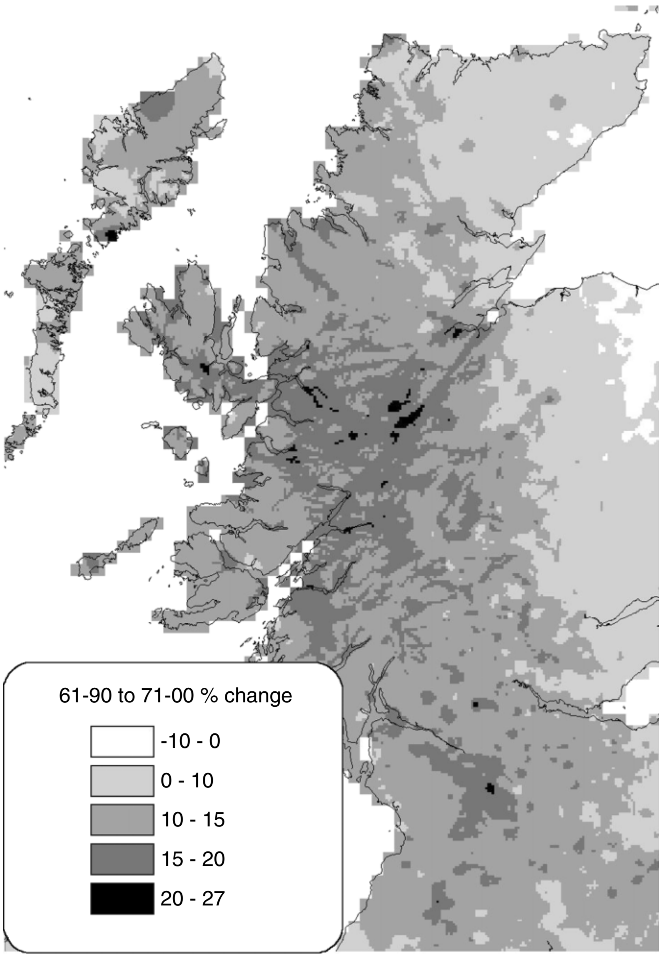

Figure 4. Percentage change in winter precipitation for Scotland, 1961–90 to 1971–2000

It is hoped that further work looking at local regression models and new independent variables, such as soil type and different ways of describing the terrain shape, may help to improve this aspect of the analysis.

This paper shows how missing-data estimation techniques and GIS methods can be combined, and applied to the production of a wide range of climate statistics. The results are proving useful in applications such as hydrological modelling, climate-change research, and putting current weather into context. The 1971–2000 LTA grids show significant and spatially varying changes from those for the 1961–90 period, especially an increase in winter precipitation in the western Scottish Highlands, and increases in temperature. This highlights the importance of ensuring that LTAs are kept up to date.

ACKNOWLEDGEMENTS

We would like to thank the reviewers for their helpful comments, and Malcolm Lee, Richard Tabony, and Simon Tett for their assistance in formulating the analysis methodology and their comments during the writing of this paper.

REFERENCES

Agnew MD, Palutikof JP. 2000. GIS-based construction of baseline climatologies for the Mediterranean using terrain variables. *Climate Research* **14**: 115–127.

Brown DP, Comrie AC. 2002. Spatial modelling of winter temperature and precipitation in Arizona and New Mexico, USA. *Climate Research* **22**: 115–128.

Burrough PA. 1986. *Principles of Geographical Information Systems for Land Resources Assessment*. Oxford University Press.

Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P. 2002. A knowledge-based approach to the statistical mapping of climate. *Climate Research* **22**: 99–113.

Goodale CL, Aber JD, Ollinger SV. 1998. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. *Climate Research* **10**: 35–49.

Nalder IA, Wein RW. 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. *Agricultural and Forest Meteorology* **92**: 211–225.

New M, Hulme M, Jones P. 1999. Representing twentieth century space–time variability. Part I: development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate* **12**: 829–856.

Perry MC, Hollis DM. 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology* **25**: 1041–1054; this issue.

Price DT, McKenney DW, Nalder IA, Hutchinson MF, Kesteven JT. 2000. A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agricultural and Forest Meteorology* **101**: 81–94.

Tabony RC. 1983. The estimation of missing climatological data. *Journal of Climatology* **3**: 297–314.

Tang WY, Kassim AHM, Abubakar SH. 1996. Comparative studies of various missing data treatment methods — Malaysian experience. *Atmospheric Research* **42**: 247–262.

Vicente-Serrano SM, Saz-Sanchez MA, Cuadrat JM. 2003. Comparative analysis of interpolation methods in the middle Ebro Valley (Spain): application to annual precipitation and temperature. *Climate Research* **24**: 161–180.

Xia Y, Fabian P, Stohl A, Winterhalter M. 1999. Forest climatology: estimation of missing values for Bavaria, Germany. *Agricultural and Forest Meteorology* **96**: 131–144.