



A new set of long-term averages for the UK

Daniel Hollis and Matthew Perry

Met Office
FitzRoy Road
Exeter
Devon
EX1 3PB

Version 2.0
5 March 2004

© Crown Copyright 2004

This paper has not been published. Permission to quote from it must be obtained from Group Head, Development Resourcing and Technology



Development of a new set of long-term climate averages for the UK

Daniel Hollis and Matthew Perry
Met Office, FitzRoy Road, Exeter, Devon, EX1 3PB, United Kingdom

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 1km. 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 firstly 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 averages; climate normals; 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 averages for an individual 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 data sets.

This paper describes the production of new sets of UK monthly and annual averages for the periods 1961-1990 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 1km x 1km grid of points covering the UK.

This project is notable for the range of climate variables tackled, and averages have been produced for each of the variables listed in Table 1. These variables have various characteristics. Seven relate to temperature, three to rainfall, plus one each relating to sunshine, thunder and snow lying. 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.

The choice of variables was determined largely by the requirements of the Met Office to produce climate summaries for the UK (in which monthly statistics are placed into historical context through a comparison with climate normals). It is recognised that there are some notable omissions, such as wind speed, wind direction, humidity, visibility, snow depth and days of snow falling. It is hoped that some or all of these variables will be addressed in future projects.

1.2 Background

In order to calculate station LTAs, any gaps in the record of monthly statistics need to be filled in. This ensures that the LTAs are calculated from data for the same complete 30-year period, in order to remove any bias associated with missing years which may be anomalous. It also provides a relatively dense network of stations from which gridded datasets can be created using GIS technology.

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) generally use data from a smaller number of stations over a longer time period. These methods are usually better able to model more localised effects by taking into account the available data history of each station.

Tang et al. (1996) compared twelve different methods for filling in gaps in monthly and annual rainfall records in Malaysia, concluding that the best results are achieved using a modified normal ratio method. This involves calculating, for each site, the ratio of each available data value to the corresponding long-period normal. Where a value is missing, an

estimate is generated by scaling the long-period normal for that site by the distance-weighted average of the ratios for a number of neighbouring sites. Xia et al. (1999) looked at six methods for estimating missing temperature, vapour pressure, wind speed and precipitation values at sites in Germany. The best estimates overall were produced by multiple regression with values from neighbour stations. Tabony (1983) compared five methods and concluded that, though the best general technique will be based on principal component analysis, comparable results can be obtained using an approach based on regression with well-correlated neighbours. We base our method on this latter approach, but use a spatial interpolation method for the more problematic 'days of' variables.

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) and weighted local regression (Daly et al., 2002).

Nalder and Wein (1998) compared seven different approaches to interpolating climate normals from a sparse network of sites in Canada. Their preferred method involved taking a weighted average of corrected values from neighbouring sites, the corrections being based on empirically-derived horizontal and vertical trends. This method, termed GIDS ('gradient-plus-inverse-distance-squared'), produced the lowest cross-validation mean absolute errors for both temperature and precipitation.

Vicente-Serrano et al. (2003) compared 23 methods for interpolating annual mean temperature and precipitation over the Ebro Valley in Spain. They concluded that multiple-regression techniques, using several geographic and topographic variables, produce the most realistic results.

In the current study we develop the approach of Lee et al. (2000), in which inverse-distance weighted averaging is used to interpolate de-trended station data. The de-trending process uses multiple regression analysis to construct a model of the impact of topographic and geographic factors on the local climate. This approach is formally very similar to the GIDS method of Nalder and Wein and has been employed successfully by the Met Office for several years for the production of gridded averages and monthly climatological maps.

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. 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. 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 3.4). For temperature, the RMS estimation error was between the RMS error for 2 and 3 missing days, so a maximum of two missing days were allowed. For sunshine, the results indicated that 5 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 which 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 12100 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-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). Missing monthly values are estimated 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 a number of standard errors from z to get a lower confidence limit (z_{low}). This is used to rank the neighbours, and for each missing value the six highest-ranked neighbours with data available in that month are used (although less than 6 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 procedure is followed separately for each month (i.e. all January's are analysed separately from all February's etc.), unlike Tabony who smoothed regression parameters over the 12 months. It is only attempted for stations with at least four 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 numbers of standard errors to see what impact this had on the quality of the estimates. It was found that setting the number of standard errors at either 2 or 3 gives very similar results. Results with more rigorous thresholds 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 minimise the error whilst maintaining a very low level of missing estimates. Table 2 shows, for various sample sizes, the implicit lower limit on r when two standard errors are subtracted from z .

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) 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 1961-1970 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, and local effects such as frost hollows, but does produce estimates which are representative of the area.

2.3 Calculation of Station LTAs

Once the gaps in the array have been filled, long term averages for the periods 1961-1990, 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 four 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 Section 3.3). 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.

2.4 Generation of Gridded Datasets

For most variables, the station LTAs are interpolated to a regular 1 km x 1 km grid of values covering the UK (the ‘average then grid’ approach). 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.

These methods are used to produce a monthly LTA grid from all available monthly station LTAs. 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.

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 Geographical Information System.

The heart of this process is the interpolation of station data onto a regular 1 km x 1 km grid using inverse-distance-weighted (IDW) averaging. The value at each grid point is calculated as a weighted average of surrounding station values, the weighting function being:

$1 / d^p$ where d is the distance and p is the power parameter.

The value of the power parameter needs to be chosen, together with the radius within which points will be used in the weighted average. Two versions of IDW are available: the standard IDW version uses all data values within a specified search radius, but expands the search if fewer than 12 stations are found. Custom IDW uses all data values within a specified search radius, and 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 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 3 shows the factors which 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 which should be excluded from the gridding analysis process.

Stations were ranked by the mean absolute residual over the 12 months for both 1961-1990 and 1971-2000. The maximum residual error in any month was also considered. Suitable thresholds for the exclusion of stations were set, based on inspection of the grids. 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 significant spatial variations in the data. Each LTA grid was inspected for bull's-eyes or other peculiar features. Where necessary, stations were removed from the analysis. 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 re-calculated and re-checked. 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 (e.g. Santon Downham). This is a weakness in the gridding techniques, which are unable to model these very localised 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, for example by surrounding mountains (e.g. Camps Reservoir) or surrounding trees (e.g. Westonbirt). 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 $\geq 0.2\text{mm}$, some stations are affected by dew causing values of 0.2mm 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 4 and the number of stations excluded for each climate variable is given in Table 5.

3. Development and Verification of Methods

3.1 Assessment of Estimation Error

The accuracy of the estimation method was tested by excluding values for six 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 one month of each season, for maximum temperature, sunshine, rainfall, and days of rain ≥ 1 mm. The estimates produced by the infilling method were then compared to the actual measured values. The results, averaged across all the test months, are summarised in Table 6. Rainfall is the hardest variable to estimate because it has high variability between years. For all variables, the level of uncertainty associated with the LTA value will increase as the length of the original data record decreases, both because the more years are missing, the more years need to be estimated, and because individual estimates are likely to be less accurate.

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 one month of each season, or for 'grid then average' variables one month for each season in each of 3 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 was made by bi-linear interpolation from the grid produced from the remaining stations. These estimates were compared with the actual values, and the resultant verification statistics, especially the RMS error, were used to assess whether subsequent adjustments made to the profiles improved the analysis.

Table 7 shows the RMS error 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 which 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 which have a sparser coverage of stations, and areas with more complex terrain, often having greater than average errors. In 30 of the 36 tests, the RMS error 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 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 recognised 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 5km 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, so 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.

Some aspects of the gridding methods chosen are shown in Table 8, while further details of the regression model for each climate variable can be found in Table 10. The power parameter of the IDW weighting factor was either set to two or three; a power of two is preferred for those variables which are subject to more uncertainty due to modelling difficulties or a sparse network. 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 9 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 bi-linear 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 1961-1990 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, Jan 1961 – Dec 2000).

Table 10 shows the average value of each coefficient in the regression equation for each climate variable (but excluding those for *easting* and *northing*) and also the average value of r^2 , the proportion of the variance explained by the regression model. Again, averages are over all grids (i.e. 24 or 480).

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 5km 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, 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 1961-1990 averages calculated using similar methods in the 1990's. This was done for stations 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 stations 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 onto an Oracle database, together with information on the number of original, estimated, and missing values from which they were derived. The 1km 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 Met Office internet site. Figure 1 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 which have been obtained are presented here, showing that there is potential for interesting patterns to be discovered. Figure 2 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 north and west Scotland and north-west England, but the increase is up to 25% in parts of the western Scottish Highlands. 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 south-east 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 1961-1990 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 optimised, 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 variables. 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 IDW 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 which, together with careful quality control and exclusion of unrepresentative stations, enables detailed and accurate high resolution (1km x 1km) 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. 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 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.

Acknowledgements

The authors would like to thank Malcolm Lee, Richard Tabony, Simon Tett, and George Anderson 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.

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.

Lee MJ, Hollis DM, Spackman E. 2000. From raw data to the internet – producing quality climatological services. Proceedings of the 3rd European Conference on Applied Climatology, Pisa 2000.

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.

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**: 131-14

Table 1: List of the climate variables for which long-term averages have been produced.

Climate Variable	Statistic	Units
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
30cm Soil Temperature	Average of daily measurements at 0900	°C
Sunshine Duration	Average of daily durations	Hours/day
Precipitation Amount	Sum	Millimetres
Days of Rain ≥ 0.2 mm	Count	Days
Days of Rain ≥ 1 mm	Count	Days
Days of Air Frost	Count of days where min temp < 0	Days
Days of Ground Frost	Count of days where grass min temp < 0	Days
Days of Snow Lying	Count of occurrences at 0900	Days
Days of Thunder	Count of days when heard	Days

Table 2: The minimum correlation coefficient, for different sample sizes, below which neighbouring stations are not used in the infilling process.

No. of data points	Minimum r
<4	Not used
4	0.96
6	0.82
10	0.64
20	0.45
30	0.37
40	0.32

Table 3: Topographic and geographic variables used in the regression model

Easting and Northing	To capture spatial trends. They are incorporated as 2nd or 3rd order cross-product polynomials
Altitude	The elevation above mean sea level, from a 500m DEM interpolated to 100m
Terrain Shape	Mean altitude over a circle of radius 5km offset by 10km to the north, south, east and west
Sea	The proportion of sea within a 5km radius
Urban	The proportion of urban land use within a 5km radius, calculated from a 1970's land use dataset which does not cover Northern Ireland

Table 4: 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	$\geq 1^{\circ}\text{C}$.
Minimum Temperature/Days of Air Frost	Min Temp $\geq 1^{\circ}\text{C}$, or Min Temp $\geq 0.9^{\circ}\text{C}$ and Air Frost ≥ 3 days, or Min Temp $\geq 0.8^{\circ}\text{C}$ and Air Frost ≥ 3.5 days.
Mean Temperature	Max Temp $\geq 1^{\circ}\text{C}$, or Min Temp $\geq 1^{\circ}\text{C}$, or Mean Temp $\geq 0.8^{\circ}\text{C}$.
Grass Minimum Temperature/Days of Ground Frost	Grass Min Temp $\geq 1.2^{\circ}\text{C}$, or Grass Min Temp $\geq 1.1^{\circ}\text{C}$ and Ground Frost ≥ 3 days, or Grass Min Temp $\geq 1^{\circ}\text{C}$ and Ground Frost ≥ 3.5 days.
30cm Soil Temperature	$\geq 0.9^{\circ}\text{C}$, plus visual inspection of grids for stations with the highest residuals in each month.
Sunshine Duration†	≥ 0.5 hours per day, plus visual inspection of grids for stations with residual ≥ 0.9 hrs per day in any month.
Precipitation Amount	Visual inspection of grids for stations with the highest residuals in each month.
Days of Rain $\geq 0.2\text{mm}$	Visual inspection of grids for stations with the highest residuals in each month.
Days of Rain $\geq 1\text{mm}$	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*	NCM stations: Mean Absolute Residual > 0.72 days. 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 .

† 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.

* For days of thunder, NCM stations operate a genuine 24-hour observing regime and should be unbiased. DLY3208 stations do not always observe throughout the whole day and as such may have a tendency to under-report the frequency of thunder. Both types of stations 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.

Table 5: Numbers of stations excluded from the production of the LTA grids.

Climate Variable	Number of Excluded Stations	Percentage of Excluded Stations
Maximum temperature	12	0.9
Minimum temperature	65	5.2
Mean temperature	72	5.7
Grass minimum temperature	103	10.1
30cm 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 6: Summary of the average performance of the infilling method.

Climate Variable	Mean Absolute Error	Root Mean Square Error
Maximum Temperature ($^{\circ}$ C)	0.19	0.25
Sunshine (%)	8	9
Rainfall (%)	15	21
Days of Rain \geq 1 mm (days)	1.19	1.61

Table 7: RMS Gridding Errors from excluded Verification Stations (10% sample)

Climate Variable	January	April	July	October	Average
Sunshine (hrs per 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
30cm 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 0.2mm	1.22	1.10	1.07	1.17	1.14
Days of Rain \geq 1mm	0.93	0.81	0.80	0.88	0.85
Days of Thunder	0.27	0.19	0.54	0.20	0.30
	January	March	November		Average
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

Table 8: Summary of the characteristics of the gridding methodology for different climate variables.

CLIMATE VARIABLE	INTERPOLATION SCHEME			<i>easting & northing</i>
	Version	Power	Radius	polynomial order
Sunshine	Custom	2	100	3
Maximum Temperature	Standard	3	100	3
Mean Temperature	Standard	3	100	3
Minimum Temperature	Standard	3	100	3
Grass Minimum Temperature	Custom	2	100	3
30cm Soil Temperature	Custom	2	100	2
Rainfall	Standard	3	50	3
Days of Rain \geq 0.2mm	Custom	2	75	3
Days of Rain \geq 1.0mm	Custom	2	75	3
Days of Air Frost	Custom	2	100	3
Days of Ground Frost	Custom	2	100	3
Days of Snow Lying	Standard	3	75	2
Days of Thunder	Custom	2	100	2

Table 9: 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
30cm Soil Temperature	392	24	0.08
Sunshine (hrs per day)	604	24	0.16
Rainfall (mm)	9669	24	2.7
Days of Rain \geq 0.2mm	9029	24	1.07
Days of Rain \geq 1.0mm	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 10: Average correlation coefficients and regression parameters for each climate variable.

Climate Variable	r^2	GEOGRAPHIC VARIABLES – REGRESSION COEFFICIENTS						
		<i>sea</i>	<i>altitude</i>	<i>urban</i>	<i>north</i>	<i>south</i>	<i>east</i>	<i>west</i>
Sunshine	0.71	0.12	-0.01	-0.04	0.00	-0.01	-0.01	-0.11
Maximum Temperature	0.93	-0.31	-0.85	0.05	—	—	—	—
Mean Temperature	0.94	0.11	-0.61	0.14	0.08	-0.03	-0.01	-0.02
Minimum Temperature	0.87	0.49	-0.41	0.20	0.04	-0.05	-0.10	-0.01
Grass Min Temp	0.75	0.67	-0.14	0.13	0.12	-0.06	-0.14	0.01
30cm Soil Temperature	0.87	0.05	-0.54	0.06	—	—	—	—
Rainfall	0.64	—	10.5	—	-5.1	7.2	0.6	7.6
Days of Rain \geq 0.2mm	0.72	—	0.57	—	-0.07	0.23	-0.02	0.28
Days of Rain \geq 1.0mm	0.77	—	0.55	—	-0.16	0.33	-0.13	0.46
Days of Air Frost	0.46	-0.88	0.38/0.58†	-0.28	-0.23 (mean altitude in all directions)			
Days of Ground Frost	0.44	-1.50	0.34	-0.30	—	—	—	—
Days of Snow Lying	0.48	—	0.67	—	0.15	0.25	0.03	0.17
Days of Thunder	0.29	—	—	—	—	—	—	—

† For days of air frost the best model involved a quadratic relationship with altitude. The coefficients are those for the linear and quadratic terms respectively (i.e. a and b in $\dots + aX + bX^2 + \dots$).

Figure 1: 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 (hours per day).

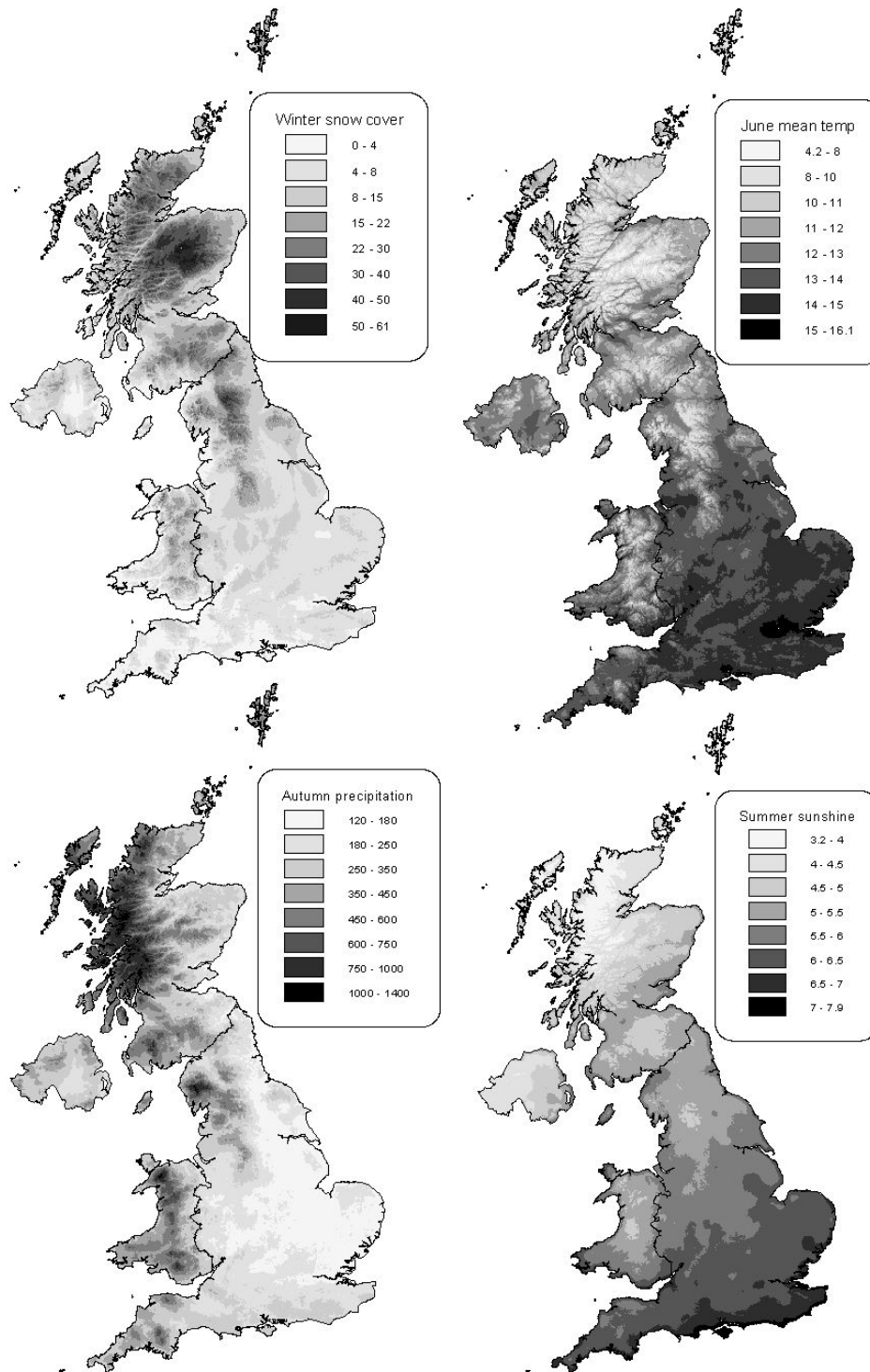


Figure 2: Change between 1961-90 and 1971-2000 periods of summer and winter precipitation and mean temperature.

