Extremes atlas

A climatology of extremes for the UK.

A baseline for UKCP09

Simon Brown, Penny Boorman, Erasmo Buonomo, Eleanor Burke, John Caesar, Robin Clark, Ruth McDonald, Matthew Perry

The Met Office Hadley Centre

June 2008

1 Executive summary

A climatology of extreme weather events has been produced for the UK to support UKCIP 2008. This provides both a baseline from which to apply future projections of extreme events for the UK and also a demonstration of the extreme methodology that will be used for determining future risk of extreme events from in the UKCIP model projections. For variables for which there are no suitable observations (drought, gusts, storms), simulated values are used from the regional climate model forced with historically observed reanalysis data. The climatology of extreme events is presented for the following:

Temperature: Daily extremes by season, hottest day, coldest day, hottest night, coldest night. Heatwaves and coldwaves of 7 and 14 days duration. The 40 year return level for hottest and coldest monthly mean of the daily maximum and minimum temperature by season

Precipitation:1-day, 5-day accumulations by season and 30-day accumulations and monthly totals for whole year.

Drought: Information not yet available.

Gusts: 90th percentile and 5, 10, 20 and 50 year return levels of daily maximum gust

Storms: Maps of the genesis, track, decay and strength of storms and anticyclones, together with blocking frequencies are presented by season.

Daily data cover the period 1960 to 2006, monthly data 1924 to 2005 and the model generated data 1957 to 2002.

2 Contents

3 Introduction

As the reality of human induced climate change gains general acceptance the concern for what climate change might be bringing to extreme weather events grows. This is not surprising as the economic and social impact of extreme weather can be high, such as the recent flooding in the UK during the summer of 2007 or the heatwave in Europe of 2003. It is also likely that the public will first notice the effects of climate change through extreme events. It is estimated that the likelihood of a summer of the severity experienced in 2003 in Europe has at least doubled and probably quadrupled due to the emission of greenhouse gasses (Stott 2004). It is therefore of national importance that any future changes in risk of extreme weather is well quantified. To this end, this report provides a baseline from which to ascertain changes in future risk by providing an assessment of the present day likelihood of a selection of extreme weather events for the UK.

The following extreme events are catalogued.

i) Temperature: Daily extremes by season, hottest day, coldest day, hottest night, coldest night. Heatwaves and coldwaves of 7 and 14 days duration. The 40 year return level for hottest and coldest monthly mean of the daily maximum and minimum temperature by season

ii) Precipitation: 1-day, 5-day accumulations by season and 30-day accumulations and monthly totals for whole year.

iii) Drought: Information not yet available.

iv) Gusts: 90th percentile and 5, 10, 20 and 50 year return levels of daily maximum gust.

v) Weather systems: Maps of the genesis, track, decay and strength of storms and anticyclones, together with blocking frequencies by season.

Detailed descriptions of the data and the methodologies are given within each section. Two genres of data are used for the analysis. Firstly that which is derived directly from observations and are aggregated to the 25km grid of the UK Climate Impacts Project (UKCP09) regional climate model output. This is data for which the density and longevity of the observing stations are sufficient to perform gridding. The temperature and precipitation results are derived from such data, with some individual station data being used for verification of gust extremes. The second type of data is that for which the observations are not sufficient to produce gridded data, eg for gusts and weather systems. In this case, data are generated by forcing the regional UKCP09 climate model with a time series of boundary conditions from the ERA-40 reanalysis project. Such a model setup is often termed a "perfect boundary condition" experiment where the large scale flow is constrained to be near reality. This helps to reduce regional simulation errors over the UK and Europe, however biases are still liable to arise from deficiencies in the representations of processes within the regional model domain (e.g. Noguer et al., 1998). Products derived from such regional model simulations are therefore constrained by the observed synoptic scale atmospheric circulation, but should not be interpreted as observed climatologies: This should be borne in mind when interpreting the results.

The maturity of the subject is different for each meteorological variable studied, so the approach and style of each section has been chosen to best match the state of understanding. The areas of temperature and precipitation extremes are well established and so the sections are more direct, cataloguing the extreme values for the UK, whereas the other extremes take an approach appropriate for reporting research results as new data and new methods are being explored.

Due to the number of graphics for some sections, figures and tables can be located in annexes where appropriate.

3.1 References

Stott, P. A., Stone D. A. and Allen M. R. (2004), Human contribution to the European heatwave of 2003, Nature, 432, 610 – 614

Noguer, M., Jones R. G. and Murphy J.M. (1998), Sources of systematic errors in the climatology of a regional climate model over Europe. Clim Dyn, 14, 691-712.

4 Datasets

4.1 Observations

Extremes are generally considered in terms of daily timescales and daily data inform a significant part of this study. Analysis of monthly extremes is also included here for temperature as these are often used to assess the unusualness of an extended period of weather, for example the general wet nature of the 2000 autumn that preceded the flooding in the winter. Details of the daily and monthly datasets are as follows.

4.1.1 Daily data

Data Source

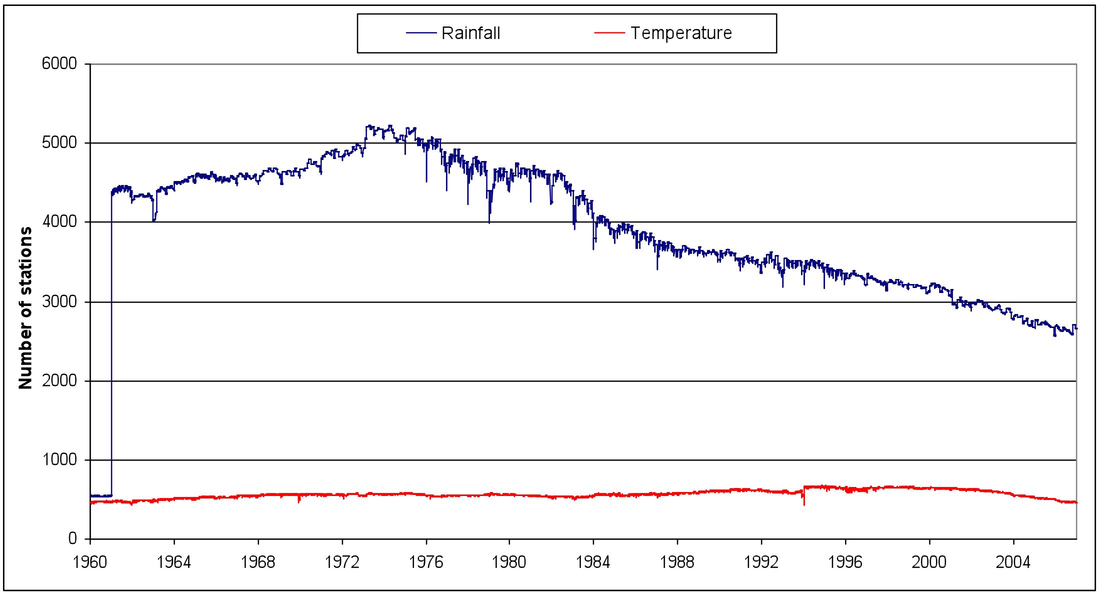
The station data was extracted from the Data Components database, which is a part of the Met Office climate data archive and contains a simplified version of the raw observations generated according to well-defined rules. The components include daily rainfall amount, and daily maximum and minimum temperature, and the tables contain single values for each day at each station, together with a flag column to indicate whether the data is suspect or estimated, and a column giving the percentage of data which was available for the calculation.

Although the data has undergone some quality checking, the extent and effectiveness of this has changed through time since the 1960's. Therefore, further checks on the extracted station data were carried out in order to identify and exclude data which was clearly in error. The mean and standard deviation across all stations was calculated for each day, and station data which exceeded a set threshold number of standard deviations away from the mean was inspected to see if there was any obvious error, either in relation to other surrounding stations on that day or in relation to the record for that station over surrounding months. The thresholds set were 5 or 6 standard deviations for temperature, and 20 standard deviations above the mean and 1.5 below the mean for rainfall. Sometimes, this led to the discovery of a station record which was wrong for a period of time, for example due to rainfall accumulations, or data being misaligned by a day. No attempt was made to correct data and any data found to be in error was excluded.

All available stations are used regardless of record length, in order to make the maximum use of the data. The number of stations used as input to the gridding

All available stations are used regardless of record length, in order to make the maximum use of the data. The number of stations used as input to the gridding varies throughout the period, as can be seen in Figure 1. These variations are due primarily to changes in the size of the observing network.

Figure 1: The number of stations used in the gridding analysis for rainfall and temperature from 1960 to 2006.



Gridding Method

The methods used are similar to those used to generate monthly datasets (Perry and Hollis, 2005b). The EWB climate data analysis software is again used to create the grids – this is a customised version of the ArcView Geographical Information System (GIS) software package. The method takes a set of irregularly distributed (in time and space) observations made at meteorological stations as its main input. The station data is normalised with respect to the monthly 1961-1990 climate normal, using the 1km x 1km gridded datasets whose generation is described in Perry and Hollis (2005a).

Inverse-distance weighting (IDW) is then used to interpolate the irregular station data to a regular grid, initially at a 5km x 5km resolution. In the case of temperature, the data which is interpolated are the residuals from a regression model relating the normalised temperature values to latitude, longitude, altitude, coastal influence and density of urban land use. In the case of rainfall, the interpolated values are the normalised rainfall values themselves as no regression model was used.

The set-up of the analysis, such as whether to use a regression model, which variables to include in the regression, and the power and radius of the IDW interpolation, were determined by testing on a sample year which was run with a set of 10% of the stations left out of the analysis for verification. The root-mean-squared error (RMSE) for each day was calculated from the differences between the observed values at the verification stations and the grid values at the station locations. The RSME averaged across all days in the test period was compared for different versions of the analysis.

Table 1 shows the RMSE at verification stations, averaged over each day of the test year, for each of the variables.   This gives an indication of the quality and accuracy of the grid values. Table 2 shows the analysis set-up which performed the best, and was thus used for the final versions of the gridding. Latitude and longitude are represented by a cubic cross-polynomial. The coastal influence variables used are the percentage of land within a radius of either 10 km or 30 km. The urban variable used was the proportion of urban land use within a 5 km radius.

Table 1: RMSE at verification stations for each climate variable.

Variable	Verification RMSE
Rainfall	1.23 mm
Mean Temperature	0.94 deg C
Maximum Temperature	1.06 deg C
Minimum Temperature	1.27 deg C

Table 2: Analysis regression model and interpolation settings used

Variable	Regression variables				IDW settings	
	Lat / Long	Altitude	Coastal	Urban	Power	Radius
Rainfall	n/a	n/a	n/a	n/a	3	50
Mean Temperature	X*Y^3	Yes	%land 30km	No	2	100
Maximum Temperature	X*Y^3	Yes	%land 30km	No	2	100
Minimum Temperature	X*Y^3	Yes	%land 10km	5km	2	100

The results for mean temperature are for gridded datasets generated by interpolating station values of daily mean temperature (obtained by averaging the observations of maximum and minimum temperature, where both are available). Testing is ongoing to determine whether more accurate mean temperature grids could be generated by averaging the final gridded datasets for maximum and minimum temperature.

Gridding Output

The gridding was run in 6-month blocks according to the methods described above. The average variance explained by the regression procedure was 0.55 for maximum temperature and 0.51 for minimum temperature.

Further labour-intensive quality checking was carried on the output from the gridding process in order to improve the quality of the final product. For temperature, the checking focussed on an analysis of the regression residuals at each station for each day. High regression residuals are often caused by data that is in error. However, they may also be caused by the regression surface being a poor representation of the actual temperature, especially around the edges of the UK, in areas of complex terrain, or on days with unusual temperature patterns, so care needs to be taken before excluding values. Two tests were used to identify suspect values:

- A large absolute value of the regression residual (>= 5 °C)

- A regression residual that is an outlier (based on the size of the gaps between the stations with the 5 highest and lowest regression residual values on each day)

Station values fulfilling both criteria were inspected in the array of station data before deciding whether to exclude them from the analysis.

For rainfall, no regression residuals were available because regression was not used in the analysis. Instead, software was written in ArcView to enable the relatively rapid and efficient eyeball checking of gridded output for each day. When bulls-eyes or clear inconsistencies in the grid were spotted, station values could be plotted in a zoomed view, to enable a decision to be made on whether to exclude a station from the analysis. The array of station data could also be inspected to help with the decision.

The gridding was then re-run without the excluded data, and the final version of the output was archived ready for further analysis and the generation of derived products. The derived product used for this report is a re-gridded 25km x 25km product where the grid matches that of the regional climate model (RCM) used in the UKCIP scenarios. Regridding is achieved through simple averaging of the 5km grid boxes falling within each grid box of the 25km RCM grid.

4.1.2 Monthly data

Climate data on 5km grids at the monthly timescale were originally created as part of the UKCIP02 national climate scenarios prepared for UKCIP by the Tyndall Centre and Met Office Hadley Centre. These are described in detail in Perry and Hollis (2005a,b), though the methodology is similar to that used in the creation of the daily grids (Section 4.1.1).

This atlas makes use of the mean monthly maximum and minimum temperature datasets, and also the total monthly precipitation dataset based on the archive of UK weather observations held at the Met Office. Originally covering the years 1961-2000, the grids have subsequently been updated by the Met Office Hadley Centre to cover the period 1914-2005, and therefore cover a longer period than the daily 5km grids.

Approximately 500 temperature stations and 3500 rainfall stations were utilised to produce the grids. The regression and interpolation process used to obtain the 5km grids alleviates the impacts of station openings and closures on homogeneity but the impacts of a changing station network cannot be removed entirely, especially in topographically variable areas.

4.2 Reanalysis and reanalysis forced RCM

The ECMWF ERA40 reanalysis dataset (Uppala et al. 2005, ERA40) covers the period 1957-2002. These are taken to be representative of the historical weather as they provide a comprehensive analysis of past observations in a dynamically consistent way.

The HadRM3 RCM is used in the UKCIP 2008 climate projections to provide downscaling information from a large ensemble of perturbed physics global circulation models (see the main UKCP09 report for details). The RCM uses a rotated pole grid with a resolution of 25km and 19 levels in the vertical and is scientifically equivalent to HadAM3 (Pope et al. 2000) except for the resolution dependent parameters that have been tuned to 25km and the addition of a representation of the sulphur cycle (Jones et al. 2001). The synoptic variability of the RCM is largely determined by the boundary forcing, therefore, we drive the RCM with boundary conditions derived from ERA40 data, in this context termed “perfect boundary conditions”, and assume that the resultant output for the UK is an acceptable measure of reality. In truth, there will be biases introduced by the limitations of the RCM. For example its 25km resolution will not capture the full details of storms and fronts, and errors in its physical parameterisations can introduce biases in the simulation in surface temperature, precipitation and related variables, particularly in summer when the influence of the driving synoptic circulation is weaker (e.g. Vidale et al., 2003). Therefore, some allowance for the potential influence of model bias may be needed when attempting to determine from this data the current risk of a particular extreme event.

4.3 References

Jones A, Roberts DL, Woodage MJ, Johnson CE. 2001. Indirect sulphate aerosol forcing in a climate model with an interactive sulphur cycle. Journal of Geophysical Reasearch – Atmospheres, 106, D17: 20293-20310

Perry MC, Hollis DM. 2005a. The development of a new set of long-term average climate averages for the UK. International Journal of Climatology 25: 1023-1039.

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

Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parametrizations in the Hadley Centre climate model – HadAM3. Clim Dynam 16: 123–146, DOI 10.1007/s003820050009

Uppala, S.M., Kållberg, P.W., Simmons, A.J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J.K., Haseler, J., Hernandez, A., Kelly, G.A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R.P., Andersson, E., Arpe, K., Balmaseda, M.A., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B.J., Isaksen, L., Janssen, P.A.E.M., Jenne, R., McNally, A.P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N.A., Saunders, R.W., Simon, P., Sterl, A., Trenberth, K.E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J. (2005) The ERA-40 re-analysis. Quart. J. R. Meteorol. Soc., 131: 2961-3012. DOI10.1256/qj.04.176

Vidale, L., Lüthi, D., Frei, C., Seneviratne, S.I and Schär, C. (2003) Predictability and uncertainty in a regional climate model. J Gepohys Res 108, 4586.