

Statistical downscaling of precipitation for a selection of sites in Ireland employing a generalised linear modelling approach

Rowan Fealy* and John Sweeney

ICARUS, Department of Geography, NUI Maynooth, Co Kildare, Ireland

Abstract:

Producing plausible outputs from statistically downscaled models of daily precipitation remains a significant challenge when producing climate scenarios for impact assessments. This challenge arises primarily as a consequence of a number of difficulties encountered when working with higher resolution daily precipitation data. Precipitation receipts at a site tend to be largely heterogeneous over space and time as local factors, such as relief, play an important role in determining whether it rains or not. More crucially from the point of view of statistical analysis, daily precipitation data is rarely, if ever, normally distributed, resulting from a high frequency occurrence of low-fall events and a low frequency of high-fall events. Additionally, modelling precipitation requires a two-step procedure. First, precipitation occurrence must be modelled: then a model is fitted to precipitation quantities which describes the rainfall distribution for days on which precipitation occurs. This paper presents a technique that overcomes some of the difficulties encountered and produces plausible precipitation amounts for a selection of 14 sites in Ireland. Difficulties still exist with predicting extreme precipitation events, which tend to be underestimated by the methodology employed. Copyright © 2007 Royal Meteorological Society

KEY WORDS statistical downscaling; Ireland; synoptic stations; generalised linear model; gamma distribution; logistic regression

Received 21 November 2005; Revised 15 January 2007; Accepted 21 January 2007

INTRODUCTION

Time series of precipitation are generally required as input to impacts models when assessing potential impacts arising as a consequence of global climate change. In order to produce plausible scenarios of change, it is of considerable importance that realistic series of precipitation, which reflect the large-scale changes in the atmosphere, are employed. As a consequence of these requirements, a number of statistical methods have been developed to produce downscaled time series of precipitation, which adequately take into account the thermal forcing of the atmosphere under various scenarios of increasing global concentrations of atmospheric CO₂.

One such technique, empirical statistical downscaling, has been widely employed to downscale climate information from the global scale. The goal of downscaling is to adequately describe the relationship between atmospheric circulation and the surface environment, with attention being focused more on model parsimony and accuracy, rather than understanding the relationship between them (Yarnal *et al.*, 2001). As a consequence of the relative ease of implementation, the use of statistical downscaling methodologies to produce climate scenarios from global

climate models (GCMs) is a preferred technique for many researchers.

Empirical statistical downscaling is based on the development of mathematical transfer functions or relationships between observed large-scale atmospheric variables and the surface environmental variable of interest. The use of statistical downscaling requires that a number of assumptions are made, the most fundamental of which, assumes that the derived relationships between the observed predictor and predictand will remain constant under conditions of climate change and that the relationships are time-invariant (Yarnal *et al.*, 2001). It also assumes that the employed large-scale predictor variables are adequately modelled by the GCM for the resultant scenarios to be valid. Busuioc *et al.* (1998), in their verification of the validity of empirical downscaling techniques, found that in the case considered, GCMs were reliable at the regional scale, with respect to precipitation in their study area, and that the assumptions of validity of predictor–predictand relationship held up under changed climate conditions.

If statistical downscaling is to be of any use, the relationship between predictor and predictand should explain a large part of the observed variability and that the expected changes in the mean climate should lie within the range of its natural variability (Von Storch *et al.*, 1993). However, due to the influence of ‘local’ factors on precipitation occurrence and volumes,

* Correspondence to: Rowan Fealy, ICARUS, Department of Geography, Rhetoric House, NUI Maynooth, Co. Kildare, Ireland.
E-mail: rowan.fealy@nuim.ie

the relationship between the large-scale predictors used when calibrating the statistical model and site-specific variability is often obscured and hence, only reflect a small part of the actual observed variability. This situation is further complicated in areas such as Ireland due to relief effects on precipitation. The selection of an optimum predictor set of atmospheric variables has been the focus of much research. However, no one technique or predictor set has come to the fore and there has been little research in evaluating the skill of various atmospheric predictor sets between studies and regions. Cross comparisons between predictors and evaluation of skill has been complicated by the fact that different studies have utilised different techniques and atmospheric predictor combinations for different regions. A number of studies have shown that choice of technique (Wilby *et al.*, 1998; Huth, 2003) and predictors can have an impact on the resulting downscaled scenarios (Winkler *et al.*, 1997; Huth, 2003).

Cavazos and Hewitson (2002) try to address what they see as a lack of systematic study in the evaluation of the relative performance of predictors used in downscaling. In their analysis, they examined the skill and errors of a large number of individual atmospheric predictors of daily precipitation as applied to a range of different locations. They also tried to determine the best combination of predictors for examined locations for winter and summer. Their results indicated that a humidity variable and mid-tropospheric geopotential heights were two of the most relevant controls on daily precipitation for all locations and seasons analysed (Cavazos and Hewitson, 2002). The incorporation of an atmospheric moisture variable in downscaling precipitation has proved crucial, as changes are likely to occur in the moisture capacity of warmed air which may not be reflected in circulation changes alone (Murphy, 2000). This was first highlighted by Karl *et al.* (1990) and Wigley *et al.* (1990) but it was common, until recently, to find studies that did not include some measure of atmospheric moisture content or humidity.

Selection of domain size is also important from the point of view of GCM output as the predictive capability or skill of the model is expected to increase with increasing domain size (Goodess and Palutikof, 1998). However, too large a domain size can add unnecessary noise and result in spurious results being produced. In order to overcome some of the issues associated with the skill level of various domain sizes, the use of mean sea-level pressure or variables derived from mean sea-level pressure have formed the centrepiece of many downscaling studies due to its relatively conservative variability and hence, predictability (Wilby, 1997, 1998; Goodess and Palutikof, 1998; Kilsby *et al.*, 1998; Chen, 2000; Trigo and DaCamara, 2000). Much of the circulation-based downscaling work has focused attention on the use of lamb weather types (LWTs) or the derived objective classification technique of Jenkinson and Collison (1977) and Jones *et al.* (1993) in an extension of the methods used in synoptic climatology.

Modelled mesoscale predictor variables, such as, mean sea-level pressure and geopotential heights, are also considered to have a much improved skill level in comparison to grid precipitation which depends on sub grid scale processes such as clouds being adequately modelled (Wilby and Wigley, 2000). Processes operating at the sub grid scale provide an important justification for requiring further downscaling, particularly for a variable like precipitation, which is so dependent on the processes which occur at this smaller scale. Additionally, most GCMs display a 'dribble effect' with regards to grid scale precipitation, which results in the number of dry days not being adequately modelled. As only one GCM grid covers the entire land area of Ireland, downscaling to individual stations in order to assess the projected changes in the spatial variation of precipitation receipts is likely to be important in providing local scale changes crucial for future water management on the island.

In an analysis of the synoptic origins of precipitation in Ireland, Sweeney (1985) found that the largest mean daily precipitation receipts for 34 stations were associated with southerly, cyclonic and westerly Lamb Circulation types. As the main storm tracks pass to the north-west of the island, greater and more frequent falls of precipitation in the north western counties are experienced than in the southeast (Keane and Sheridan, 2004). Frontal systems, associated with these Atlantic depressions, and airflow interaction with local topography, result in precipitation being one of the most variable meteorological parameters measured in Ireland, with large variations in receipts both spatially and temporally (Figure 1; Table I) (Rohan, 1975; Keane and Sheridan, 2004). Additionally, daily precipitation tends to be characterised by long-duration, low-intensity events. As a consequence of the large monthly and seasonal variations in precipitation, no clearly defined 'wet' and 'dry' seasons exist. However, the period February to July tends to be drier than August to January (Sweeney, 1985). Year to year variability also tends to be quite large.

The aim of this paper is to present a technique that addresses some of the key issues that arise when downscaling abnormally distributed climatological series, such as precipitation, in a challenging location, such as Ireland. To this end, a generalised linear model (GLM) was employed to model precipitation occurrence and amounts conditional on a range of atmospheric variables. The technique, developed by McCullagh and Nelder (1989), has been previously applied to model climatological series in a number of studies, mainly by Chandler and Wheeler (1998), (2002); Yan *et al.* (2002); Chandler (2003). GLMs are particularly useful for modelling climatological series, as they do not require the dependent to be normally distributed and can be applied to any variable that falls into the exponential family of distributions. Therefore, when modelling climatological series like precipitation or wind speeds, no prior normalisation of the dependent variable is required, minimising any data loss.

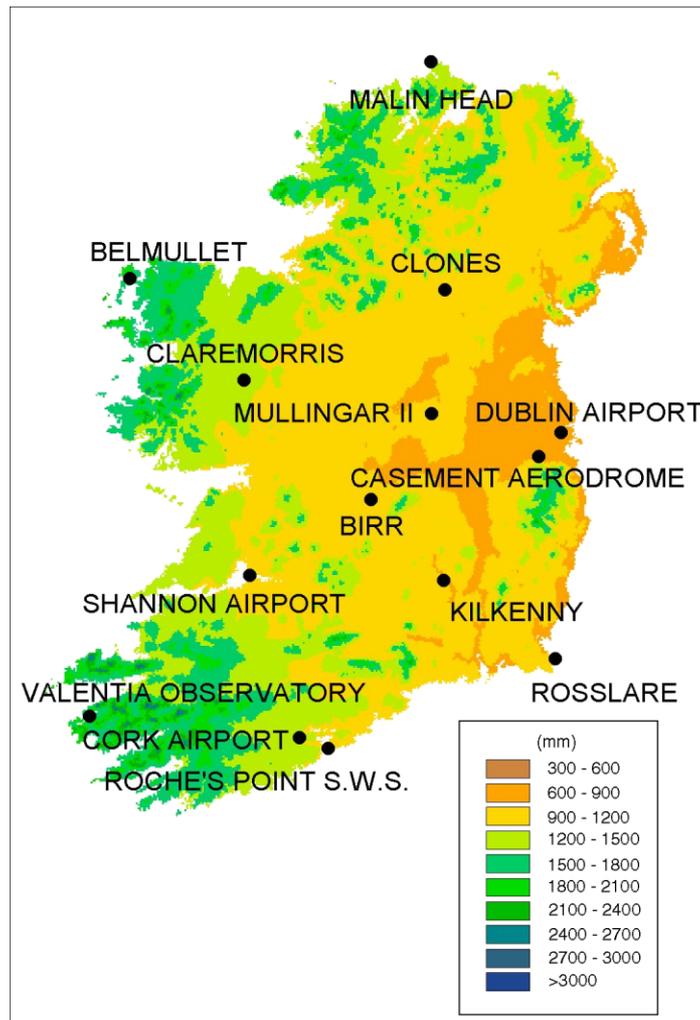


Figure 1. Station locations and annual precipitation. This figure is available in colour online at www.interscience.wiley.com/ijoc

Table I. 1961–1990 monthly averages of precipitation for a selection of synoptic stations (Data source: Met Éireann).

Station	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Kilkenny	86.3	66.1	63.9	51.4	61.9	50.5	52.5	69.4	73.5	84.9	73.8	88.6	822.8
Valentia	166.6	123.0	122.9	76.2	89.6	79.2	74.0	110.8	123.8	156.4	148.3	159.2	1430.1
Malin Hd.	114.4	76.3	85.9	58.4	59.2	64.4	72.4	91.3	102.1	118.0	114.9	103.2	1060.6
Casement	68.7	50.7	53.8	49.9	56.6	53.0	48.9	63.7	58.7	67.2	67.2	73.1	711.4

GLMs have the added advantage in that they fit probability distributions to the variable being modelled, which should offer an improvement over normal linear regression techniques, which only model the mean of the distribution. This is likely to be of significant benefit in climate change research, that a change in the distribution is likely to occur as a consequence of climate change (Yan *et al.*, 2002). Fitting probability distributions, in this manner, should also improve how extreme values in the tails of the distributions are handled within the modelling framework. A constant coefficient of variation, where the standard deviation is proportional to the mean, also ensures that variance of wet-day amounts increase with the expected value, a far more realistic assumption than

that of constant variance for precipitation (Beckmann and Buishand, 2002).

METHODOLOGY

Data sources

Observed daily precipitation data from 14 synoptic stations were obtained from the Irish Meteorological Service, Met Éireann, for the period 1961–2000. The synoptic stations, which are geographically dispersed around the island, represent low-lying conditions for a mixture of coastal and interior locations (Figure 1). No prior homogeneity analysis of the daily data has been performed.

However, the data obtained is from the synoptic network, manned by experienced meteorological officers, and is considered to be of good quality. The data is provided with quality control flags, indicating whether the measurement is the value as read, accumulated, trace or otherwise, thereby enabling the researcher to decide on a suitable threshold for accepting the data as valid. In the present research, all values not directly measured by the observer were removed from the analysis. According to this criterion, seven synoptic stations reported no missing values, four stations had less than 2.5% missing and two stations had 20% missing values while one station, Mullingar, had 42% missing values. Despite the high percentage of missing values from this station, it was decided to retain it within the analysis due to its location in the midlands where relatively few stations exist.

Large-scale surface and atmospheric data, representing both observed (1961–2000) and modelled data (1961–2100), were obtained from the SDSM (Statistical DownScaling Model) data archive (Wilby and Dawson, 2004). The observed data were derived from the National Centres for Environmental Predictions/National Centre for Atmospheric Research (NCEP/NCAR) Reanalysis project, while the modelled data, for both the A2 and B2 emissions scenarios, were derived from three GCMs, namely the Hadley Centre (HadCM3), Canadian Centre for Climate Modelling and Analysis (CCCma), coupled global climate model 2 (CGCM2) and the Commonwealth Scientific and Industrial Research Organisation (CSIRO) (Mark 2). Even though the reanalysis data are essentially modelled data, they are constrained by observations from the global monitoring network and are a modelled, gridded replicate of the observed data. Relationships between grid box values for a selection of variables from the reanalysis data were found to be significantly correlated with that of the actual observed upper air variable measured at the two upper air stations in Ireland, Valentia in the south west and Aldergrove in the north.

All the gridded datasets exist on a common resolution, that of $2.5^\circ \times 3.75^\circ$, and predictor variables were obtained for the grid box representing Ireland in the GCM domain. A common spatial resolution is crucial to overcome any mismatch in scales that may exist between observed and modelled predictor datasets. As advocated by Karl *et al.* (1990), all candidate predictor variables were standardised to account for any biases that may occur within the modelled data. In addition to the primary predictor variables, secondary airflow indices, calculated from a larger domain of nine grid boxes (3×3) centred over Ireland (after Jones *et al.*, 1993), were obtained from the SDSM data archive (Table II). The 1-day lead and lag of each standardised predictor was also calculated to allow for the temporal offset between the NCEP/NCAR Reanalysis data, which is averaged daily over the 0:00–24:00 h period, and reporting of daily precipitation, calculated over the 9:00–9:00 h period. This also allows for a temporal lag or lead between the predictor and predictand.

Table II. List of candidate predictor variables for use in the analysis from the SDSM data archive. *Italics* indicate secondary airflow indices calculated from pressure fields (surface, 500 and 850 hPa).

Variable
Mean temperature
Mean sea-level pressure
500 hPa geopotential height
850 hPa geopotential height
Near surface relative humidity
Relative humidity at 500 hPa height
Relative humidity at 850 hPa
Near surface specific humidity
<i>Geostrophic airflow velocity</i>
<i>Vorticity</i>
<i>Zonal velocity component</i>
<i>Meridional velocity component</i>
<i>Divergence</i>

Predictor selection

The precipitation series from each station were first split into occurrence (a binary sequence of 0s and 1s indicating the absence or presence of rain on a particular day) and wet-day sequences (days on which rainfall occurred). Predictor–precipitation relations were then assessed based on the Spearman correlation coefficients for both the occurrence and wet-day series and for each station and season. The Spearman correlation is a nonparametric test, which makes no assumptions about the underlying distribution of the values and is therefore suited to analysing precipitation series, which are non-gaussian.

A total of 53 candidate predictor variables were assessed for their influence on precipitation at each of the 14 synoptic stations and for each season. Predictors were then selected based on a number of criteria, such as, significance and strength of correlations and consistency of predictor across the stations. Incorporation of an atmospheric moisture predictor variable was also considered critical. In order to minimise issues associated with multicollinearity, a parsimonious set of predictors was selected which were then assessed using the product–moment correlation coefficients.

The selected NCEP/NCAR reanalysis predictor variables were then used to calibrate the transfer functions, linking the large-scale surface and atmospheric variables to the daily precipitation series for each of the 14 synoptic stations.

Precipitation occurrence model

For the purposes of the present study, logistic regression, which is a particular type of GLM, was employed to model wet and dry day sequences of precipitation. The logistic regression model, or logit, can be written as

$$\ln \left(\frac{P}{1-P} \right) = B_0 + B_1 x_1 + \dots + B_{n+1} x_{n+1}$$

which can be rewritten in terms of odds rather than log odds

$$\frac{P}{1 - P} = e^{B_0 + B_1x_1 + \dots + B_{n+1}x_{n+1}}$$

P = probability of an event

e = base of the natural logarithms

x = independent variable

B_0, B_1 = coefficients estimated from the data

The maximum likelihood method is used to select parameters within the logistic regression. Use of the logistic regression approach offers a significant improvement over multiple linear regression as the distribution of errors are normally distributed, and additionally, the predicted values can be interpreted as probabilities which ensures that P lies the interval between 0 and 1.

Precipitation amounts model

Wet-day sequences were modelled using a GLM which relates the response variable (Y), whose distribution has a vector mean $\mu = (\mu_1, \dots, \mu_n)$ to one or more covariates (x) via the relationship

$$\mu = E(y)$$

$$g(\mu) = \nu$$

$$\nu = a_0 + a_1x_1 + \dots + a_nx_n$$

A log link function, $g(\mu)$, and gamma distribution were employed for the purposes of modelling precipitation amounts. While the mixed exponential distribution has been found to provide a better fit to precipitation amounts (Wilks and Wilby, 1999) the relationship between the mean and variance for this distribution makes it difficult to incorporate into a GLM. Nonetheless, the gamma distribution GLM has been found to be a good fit to precipitation amounts in a number of regions (pers. comm. Richard Chandler).

RESULTS

To adequately assess the ability of the techniques employed to capture the underlying relationships between

the large-scale atmospheric predictors and precipitation occurrence and amounts, the data was split into two periods, one for calibration and one for independent verification purposes. The calibration periods for both the occurrence and amounts models were selected as 1961–1978 and 1994–2000, with the period 1979–1993 being withheld for verification purposes. These time periods were selected subjectively as they coincided with the calibration periods being employed by Statistical and Regional Dynamical Downscaling of Extremes (STARDEX) for European regions.

When calibrating the logistic regression, outliers were found to have a large impact on the resulting models and were excluded from subsequent analysis. Table III indicates the seasonal importance of the more important predictors used to model precipitation occurrence. Humidity, either relative or specific, is shown to be important in all seasons and at all 14 stations. Geopotential heights also appear to be important for precipitation occurrence for all seasons. Zonal airflow velocity and vorticity, at the surface or 850 hPa level, would appear to be important predictors of precipitation occurrence during the autumn and winter months. While there is a strong seasonal consistency between stations for a number of variables such as, geopotential heights and humidity, seasonal specific variables also play an important role, such as, surface divergence during the summer months.

In order to assess the relative skill of the fitted models, the Heidke skill score was employed, a method commonly used to summarise square contingency tables (Wilks, 1995). The Heidke skill score (HSS) is calculated as follows

$$HSS = \frac{2(ad - bc)}{(a + c)(c + d) + (a + b)(b + d)}$$

where

	Yes	No
Yes	a	b
No	c	d

(Wilks, 1995)

A score of one indicates a perfect forecast, while forecasts equivalent to the reference forecast produce a score of zero. Scores less than zero indicate that forecasts were worse than the reference forecast (Wilks, 1995).

Table III. Number of stations in which a particular predictor was selected by season to calibrate the occurrence models. (l) indicates the lag of the variable.

Season	Relative humidity (l)	Specific humidity	Geopotential (l)	Geostrophic airflow 500 hPa	Zonal velocity	Surface vorticity (l)	Divergence
DJF	–	14	13 (500)	–	14 (850)	13 (850)	1
MAM	14	–	14 (850)	11	–	–	–
JJA	14	–	13 (850)	10	4	1	9
SON	14	–	14 (850)	14	13	13	4

Results for the seasonal occurrence models are shown in Table V. Results indicate that all models are substantially better than the reference forecast. Values of the HSS for the verification period are comparable at all stations and for all seasons to the scores for the calibration periods, indicating that the predictors at least adequately capture some of the more important mechanisms responsible for precipitation occurrence at the stations employed in the analysis.

Table IV displays the main predictors in the seasonal amounts models for all stations. Again, key variables such as sea-level pressure, vorticity and an atmospheric moisture variable are shown to be important predictors for all seasons and stations in determining precipitation amounts. The addition of each new predictor to a model was tested using a stepwise procedure and assessing the change in the Akaike Information Criterion (AIC), a measure of the relative goodness-of-fit based on the maximised log likelihood and number of parameters employed in the model. Table VI shows the explained variance for all stations for both calibration and verification periods. The months of March, April and May appear to produce the best results for the calibration period, with a R^2 value of 0.44 being attained for the model for Roche's Point, in the south of Ireland. However, all results are comparable between seasons and

stations and between both calibration and verification periods.

Despite the apparently low explained variance for the precipitation amounts models, a measure which is sensitive to the tails of the distribution, Figures 2 and 3 show the comparison between monthly precipitation amounts for both the observed and modelled series for the verification period which demonstrate a good degree of correspondence. Figure 4 shows the interannual variability for two stations, one west coast and one east coast, between the observed and modelled series again for the 1979–1993 period. While the correlation between day to day variability may be low, monthly and yearly accumulations would appear to have been adequately captured by the model, an important requirement when assessing climate impacts on such systems as the hydrological system.

SIMULATED FUTURE CHANGES IN PRECIPITATION FOR IRELAND

To assess possible future changes in precipitation, the derived models for both precipitation occurrence and amounts were used in conjunction with data from three GCMs, namely, CCMA, CSIRO and HadCM3. Data for both the A2 and B2 scenarios from each of the GCMs were used as input to the calibrated occurrence

Table IV. Number of stations in which a particular predictor was selected by season to calibrate the rainfall amount models. Ln and exp indicate the log and exponential, respectively, of the variable.

Season	Mslp (exp)	Surface Vorticity (ln)	Meridional velocity	Zonal velocity	Specific humidity	Relative humidity	Temperature
DJF	14	14 (850)	13	8 (850)	14	–	–
MAM	13	12	13	10	–	14	1
JJA	12	9	12	6	–	14	4
SON	14	12	14	7	3	7	

Table V. Heidke skill scores for the seasonal precipitation occurrence models for both calibration (1961–1978; 1994–2000) and verification (1979–1993) periods.

Stations	DJF		MAM		JJA		SON	
	Calib.	Verif.	Calib.	Verif.	Calib.	Verif.	Calib.	Verif.
Valentia observatory	0.68	0.66	0.64	0.63	0.61	0.61	0.60	0.69
Shannon airport	0.63	0.67	0.64	0.65	0.60	0.63	0.64	0.65
Dublin airport	0.51	0.52	0.60	0.57	0.56	0.56	0.54	0.49
Malin Head	0.58	0.62	0.58	0.58	0.60	0.52	0.57	0.63
Roche's Point	0.64	0.58	0.60	0.60	0.59	0.58	0.57	0.59
Belmullet	0.58	0.58	0.59	0.62	0.56	0.51	0.57	0.59
Clones	0.62	0.64	0.65	0.64	0.66	0.62	0.60	0.62
Rosslare	0.56	0.54	0.64	0.57	0.56	0.55	0.58	0.57
Claremorris	0.65	0.66	0.62	0.61	0.65	0.56	0.62	0.63
Mullingar	0.58	0.58	0.68	0.65	0.64	0.60	0.54	0.62
Kilkenny	0.59	0.59	0.65	0.65	0.58	0.58	0.57	0.59
Casement aerodrome	0.53	0.58	0.64	0.59	0.55	0.57	0.54	0.54
Cork airport	0.62	0.58	0.61	0.59	0.58	0.60	0.58	0.62
Birr	0.59	0.63	0.65	0.65	0.58	0.55	0.60	0.62

Table VI. R-squared for the seasonal precipitation amounts models for both calibration (cal.) (1961–1978; 1994–2000) and verification (ver.) (1979–1993) periods.

Station	DJF		MAM		JJA		SON	
	Cal.	Ver.	Cal.	Ver.	Cal.	Ver.	Cal.	Ver.
Valentia observatory	0.26	0.33	0.31	0.28	0.35	0.36	0.30	0.26
Shannon airport	0.27	0.26	0.22	0.21	0.22	0.24	0.24	0.14
Dublin airport	0.20	0.20	0.20	0.17	0.21	0.18	0.18	0.11
Malin Head	0.21	0.27	0.22	0.16	0.19	0.18	0.22	0.16
Roche's Point	0.23	0.24	0.44	0.24	0.34	0.31	0.20	0.12
Belmullet	0.29	0.27	0.25	0.27	0.22	0.25	0.27	0.28
Clones	0.24	0.30	0.24	0.23	0.23	0.23	0.20	0.17
Rosslare	0.24	0.23	0.25	0.19	0.14	0.19	0.21	0.15
Claremorris	0.25	0.29	0.29	0.28	0.24	0.27	0.25	0.21
Mullingar	0.25	0.27	0.29	0.17	0.18	0.26	0.22	0.19
Kilkenny	0.32	0.24	0.28	0.25	0.22	0.16	0.23	0.24
Casement aerodrome	0.21	0.17	0.15	0.13	0.16	0.21	0.23	0.11
Cork airport	0.30	0.30	0.36	0.31	0.29	0.37	0.26	0.24
Birr	0.23	0.23	0.21	0.17	0.22	0.16	0.21	0.19

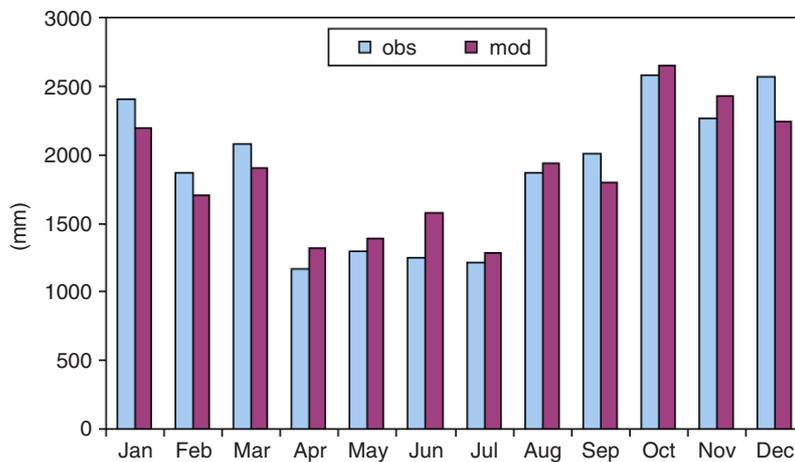


Figure 2. Comparison between monthly observed and modelled precipitation data from Valentia, a west coast station with high annual receipts. Modelled data is based on the independent verification period of 1979–1993. This figure is available in colour online at www.interscience.wiley.com/ijoc

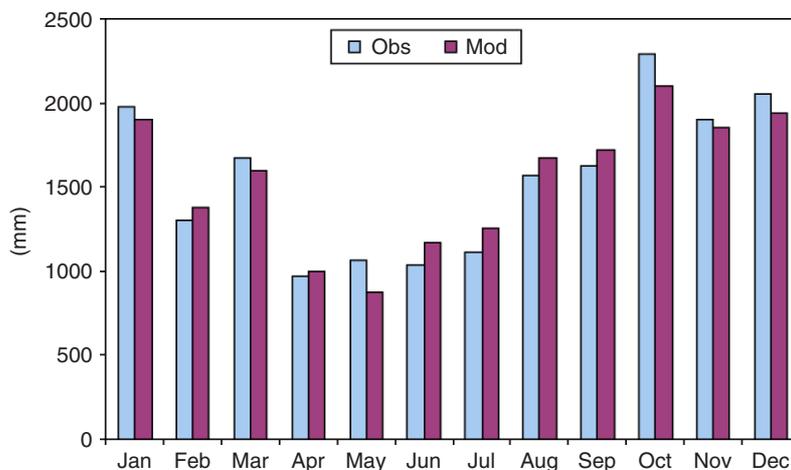


Figure 3. Comparison of observed and modelled monthly precipitation from Belmullet, a west coast station, for the independent verification period 1979–1993. This figure is available in colour online at www.interscience.wiley.com/ijoc

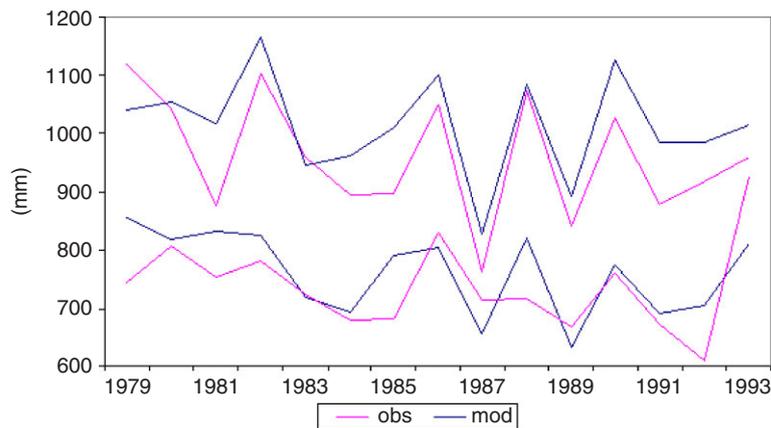


Figure 4. Interannual variability for observed and modelled annual precipitation from Shannon airport, west coast, (top) and Case-moment aerodrome, east coast, (bottom) for the independent verification period 1979–1993. This figure is available in colour online at www.interscience.wiley.com/ijoc

and amounts models discussed in previous sections. Data from a number of GCMs and emissions scenarios were employed as a number of significant uncertainties exist with regards to climate sensitivity, model uncertainty and emissions uncertainty.

Downscaling of daily occurrences and amounts of precipitation were performed for each of the fourteen synoptic stations and for each season. Seasonal and regional differences in precipitation changes are evident for all three time periods. The percentage change for each station, represented as the difference between the future time period of interest and the model control period for the different GCMs are illustrated in Figures 5–7 for three, 30-year time periods centred on the 2020s, 2050s and 2080s and for each season for the A2 emissions scenario. The range between these stations varies for each of the GCMs, with larger percentage and positive increases being demonstrated by the CCMA GCM, while the downscaled data from the CSIRO GCM

suggests that some stations will increase while others will experience a decrease in winter precipitation by the 2020s. The HadCM3-based data indicate that all stations will experience a slight decrease in winter precipitation for this period. The summer months are the only period in which all models agree that there will be a decrease in receipts, but again the changes vary between models. A clearer seasonal picture emerges for the winter and summer periods by the 2050s, with all models again suggesting an increase in winter and a decrease in summer, but again the ranges between the ‘driest’ stations and ‘wettest’ stations and models are large. Similar results are found for the 2080s. Again, these results illustrate the large seasonal and spatial ranges that can occur even over an area the size of Ireland.

Figures 5–7 also illustrate the raw, non-downscaled precipitation output for the grid cell representing Ireland from each of the three GCMs employed in the analysis for the A2 emissions scenario. Inter-model differences of the

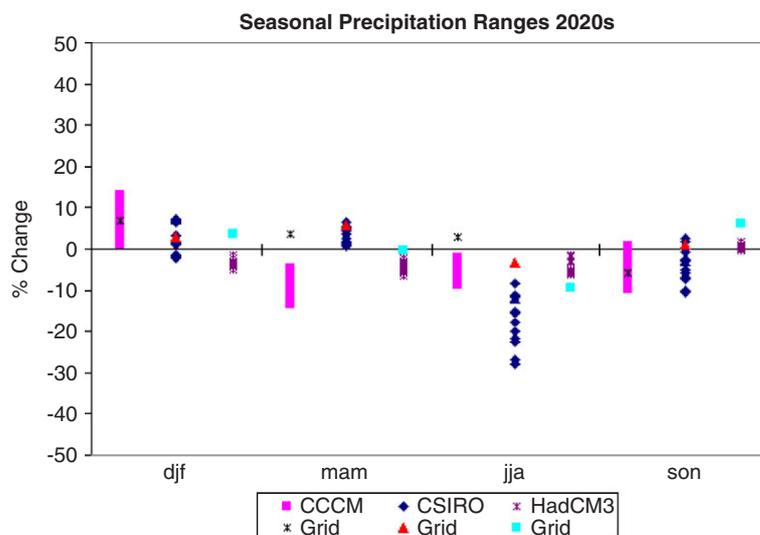


Figure 5. Downscaled seasonal precipitation ranges for 30-year period centred on the 2020s for the three GCMs and the A2 emissions scenarios employed in the present analysis. Raw Precipitation from the GCM grid representing Ireland for the A2 scenario is also included (Grid) for comparison purposes. This figure is available in colour online at www.interscience.wiley.com/ijoc

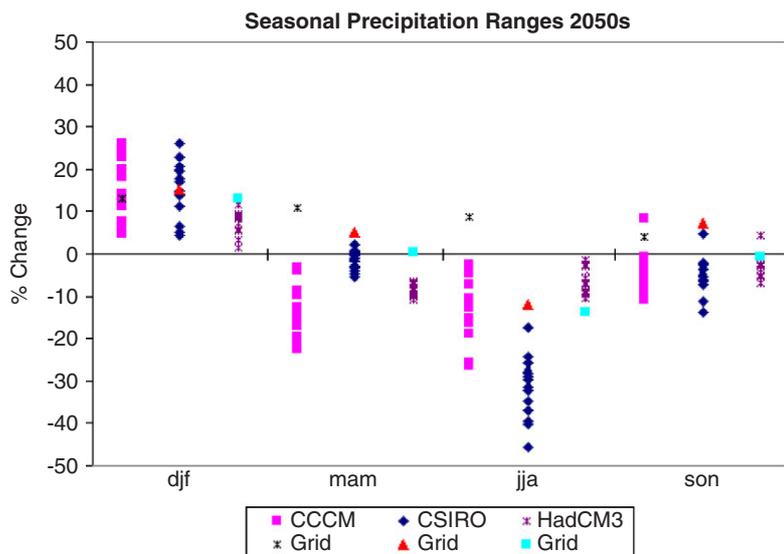


Figure 6. Downscaled seasonal precipitation ranges for 30-year period centred on the 2050s for the three GCMs and the A2 emissions scenarios employed in the present analysis. Raw Precipitation from the GCM grid representing Ireland for the A2 scenario is also included (Grid) for comparison purposes. This figure is available in colour online at www.interscience.wiley.com/ijoc

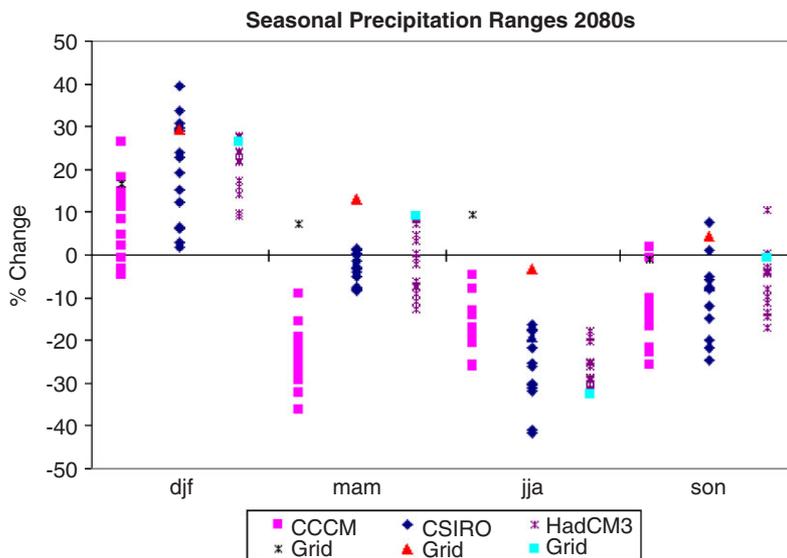


Figure 7. Downscaled seasonal precipitation ranges for 30-year period centred on the 2080s for the three GCMs and the A2 emissions scenarios employed in the present analysis. Raw Precipitation from the GCM grid representing Ireland for the A2 scenario is also included (Grid) for comparison purposes. This figure is available in colour online at www.interscience.wiley.com/ijoc

raw output tend to be greatest during the summer months and increase substantially from the 2020s to the 2080s. Raw precipitation projections for the summer months during the 2020s range from an increase of 2.9%, from the CCCM GCM, to a decrease of 9.5%, projected by the HadCM3 GCM. By the 2080s, projected model changes in precipitation for the summer months range from an increase of 9.4%, again associated with the CCCM GCM, to a substantial decrease of 32.7% projected by the HadCM3 GCM. During the winter season for all-time periods, raw GCM output tends to be more consistent with all models projecting a similar direction of change but varying in magnitude. While substantial differences in the inter-GCM model ranges arise for a

number of reasons; internal chaotic variations in the GCM being employed, different model parameterisations and initial forcing conditions and due to model uncertainties, precipitation values employed directly from a GCM are considered to have a very poor skill value in the context of this analysis and justified the effort in downscaling from the more conservative GCM outputs which have a higher skill level than those which are dependant on a process that occur at sub grid scale and hence, not adequately resolved by current GCM simulations.

In order to try and account for different model and emissions uncertainties, ensembles of the downscaled results were produced. Results from the weighted ensemble mean, based on the Climate Prediction Index (CPI)

(Murphy *et al.*, 2004) and modified by Wilby and Harris (2006) for application to a narrower suite of GCM outputs, will be discussed in the remainder of this section. The modified CPI index or impacts relevant-climate prediction index (IR-CPI) is weighted based on the individual GCMs ability to reproduce the properties of the observed climate, derived from the NCEP data, and is derived from the root-mean-square difference between modelled and observed seasonal climatological means, assessed over the baseline period (Wilby and Harris, 2006). Weights are derived based on the individual contribution of each GCM to overall error (Wilby and Harris, 2006). Weights were calculated for each season and model with the HadCM3 consistently performing well throughout three of the four seasons. The IR-CPI weights are then used to weight the relevant downscaled output from the different GCMs in order to produce an ensemble mean for both the A2 and B2 emissions separately and all emissions, together.

For space purposes, the weighted ensemble mean of the three GCMs and both the A2 and B2 emissions scenarios are illustrated for three future time periods, 2020s, 2050s and 2080s, while the upper and lower ranges, shown as deviations, represent the overall maximum and minimum prediction from the individual GCMs and emissions scenarios (Figures 8–10). Data represent seasonal averages across the 14 synoptic stations.

Based on the ensemble A2 and B2 scenarios, winter precipitation is likely to increase marginally by between 1.8–4.2% by the 2020s. The largest seasonal changes are suggested for summer, with a reduction of between 0.2 and 6.7% with an ensemble mean reduction of 3.2%; however, reductions of between 10 and 16% are suggested for stations along the southern and eastern coasts. With the exception of autumn, increases in precipitation are projected to occur in all other seasons by the 2020s based on the more conservative B2 scenario. While those based on the A2 emissions scenario suggest decreases in all seasons except winter.

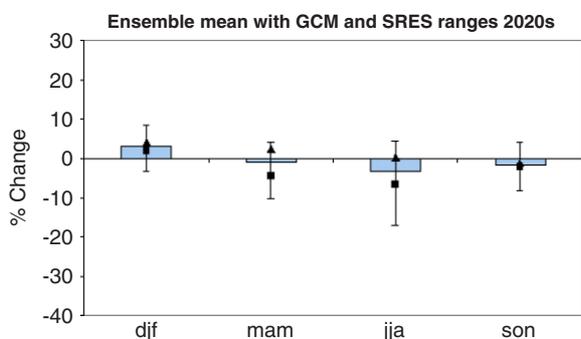


Figure 8. Ensemble mean precipitation for the 2020s produced from the weighted ensemble of all GCMs and emissions scenarios (bars). Upper and lower ranges (lines) are the results from the individual GCMs and emissions scenarios. Ensemble A2 scenario (■) and B2 scenario (▲). This figure is available in colour online at www.interscience.wiley.com/ijoc

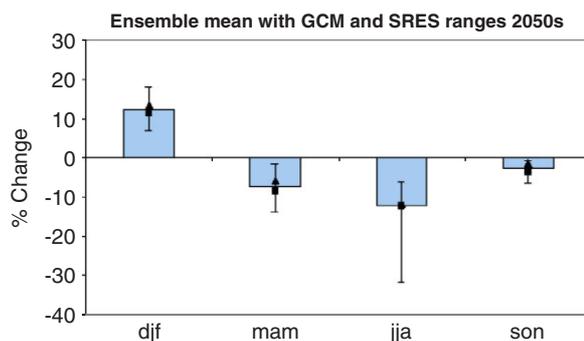


Figure 9. Ensemble mean precipitation for the 2050s produced from the weighted ensemble of all GCMs and emissions scenarios (bars). Upper and lower ranges (lines) are the results from the individual GCMs and emissions scenarios. Ensemble A2 scenario (■) and B2 scenario (▲). This figure is available in colour online at www.interscience.wiley.com/ijoc

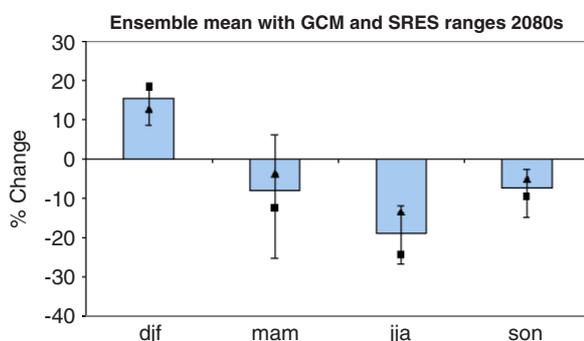


Figure 10. Ensemble mean precipitation for the 2080s produced from the weighted ensemble of all GCMs and emissions scenarios (bars). Upper and lower ranges (lines) are the results from the individual GCMs and emissions scenarios. Ensemble A2 scenario (■) and B2 scenario (▲). This figure is available in colour online at www.interscience.wiley.com/ijoc

A greater degree of consistency is evident in the changes for the 2050s, with all GCMs and ensembles suggesting a similar direction of change, but with differences in the magnitude of this change. Again, the winter and summer periods experience the largest percentage changes in receipt, ranging from a 12% increase in winter to reductions of over 12% in summer in the ensemble mean scenario. While increases are projected along the east coast and midlands during winter, reductions of between 20 and 28% are projected to occur along the southern and eastern coast during the summer season. If realised, these changes are likely to have a large impact on agriculture and hydrology in Ireland (Sweeney, 2003). These seasonal changes in precipitation are further enhanced by the 2080s, with winter increases of 15% and summer reductions of 20%. The largest percentage increases in winter precipitation are projected to occur in the midlands, of upto 20%, while the largest reductions during the summer months are again projected to occur along the southern and eastern coasts, which are likely to experience decreases of between 30 and 40% during these months. Between models, differences are greatest for the spring months during the 2080s, with two GCMs

suggesting a slight increase in spring precipitation. These increases, however, are associated with the lower emissions B2 scenario and show 0.5–6% increases in spring receipts.

CONCLUSIONS

This paper presented a technique to link large-scale circulation and atmospheric variables to precipitation in Ireland. Precipitation is inherently difficult to model in this way due to the importance of local forcing not captured at the scale of the predictors used, which is ultimately constrained by the scale of GCMs when the aim is to produce scenarios of future changes in precipitation. Additionally, daily precipitation receipts generally do not conform to the normal distribution, due to the high frequency of low receipts and low frequency of high receipts, which act to skew the distribution. This process is further complicated due to the terrain effects in Ireland resulting in large orographic enhancement on the west coast of Ireland, not strictly related to the large-scale forcing mechanisms. Both the logistic regression, due to its probabilistic component, and the generalised linear model, probability based and not constrained by requirements of normality, offer a methodological approach that is ideally suited to modelling precipitation. Results from the calibration and verification would suggest that this is the case and highlighted in Figure 4, where the derived models would appear to have captured an important component of precipitation, that of interannual variability.

Having calibrated seasonal logistic and GLM models for daily precipitation for 14 synoptic stations, the derived models were then used to produce scenarios of changes in precipitation based on the output of three GCMs. While uncertainties arising from the derived models were not accounted for, GCM and emissions uncertainties could be tentatively approached by employing a number of GCMs. The results also highlight the importance of using multiple GCMs when conducting climate change research, as the magnitude of change can be vastly different between GCMs and in some cases even different in direction. If the suggested changes in precipitation are realised, then large sectoral impacts of these changes are likely to be felt in Ireland.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support provided by the Irish Environmental Protection Agency, as part of the Environmental RTDI Programme 2000–2006. The authors would also like to thank Met Éireann for supplying the observational data. A special thanks to Dr. Richard Chandler, for advice on GLMs while this research was in its initial stages, and Jason Doran, for providing data manipulation routines. We would also like to thank the two anonymous reviewers for their helpful and insightful comments.

REFERENCES

- Beckmann BR, Buishand TA. 2002. Downscaling relationships for precipitation for the Netherlands and North Germany. *International Journal of Climatology* **22**: 15–32.
- Busuioac A, Von Storch H, Schnur R. 1998. Verification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. *Journal of Climate* **12**: 258–272.
- Cavazos T, Hewitson B. 2002. Relative performance of empirical predictors of daily precipitation. In *Integrated Assessment and Decision Support, Proceedings of the First Biennial Meeting of the International Environmental Modelling and Software Society*, Vol. 1, Rizzoli AE, Jakeman AJ (eds). Manno, Switzerland: iEMS; 334–339.
- Chandler RE. 2005. On the use of generalized linear models for interpreting climate variability. *Environmetrics*, **16**: 699–715.
- Chandler RE, Wheeler HS. 1998. Climate change detection using Generalized Linear Models for rainfall – a case study from the West of Ireland I. Preliminary analysis and modelling of rainfall occurrence. Research Report No. 194, Department of Statistical Sciences, University College London.
- Chandler RE, Wheeler HS. 2002. Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. *Water Resources Research* **38**(10): 1192.
- Chen D. 2000. A monthly circulation climatology for Sweden and its application to a winter temperature case study. *International Journal of Climatology* **20**: 1067–1076.
- Goodess CM, Palutikof JP. 1998. Development of daily rainfall scenarios for Southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology* **10**: 1051–1083.
- Huth R. 2003. Sensitivity of local daily temperature change estimates to the selection of downscaling models and predictors. *Journal of Climate* **17**: 640–652.
- Jenkinson AF, Collison P. 1977. *An Initial Climatology of Gales Over the North Sea, Synoptic Climatology Branch Memorandum No. 62*. Meteorological Office: London.
- Jones P, Hulme M, Briffa KR. 1993. A comparison of lamb circulation types with an objective classification scheme. *International Journal of Climatology* **13**: 655–663.
- Karl TR, Wang W-C, Schlesinger ME, Knight RW, Portman D. 1990. A method of relating general circulation model simulated climate to the observed local climate. Part I: seasonal statistics. *Journal of Climate* **3**: 1053–1079.
- Keane T, Sheridan T. 2004. Climate of Ireland. In *Climate, Weather and Irish Agriculture*, Keane T, Collins JF (eds). AGMET, Dublin.
- Kilsby CG, Cowpertwait PSP, O'Connell PE, Jones PD. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology* **18**: 523–539.
- McCullagh P, Nelder JA. 1989. *Generalized Linear Models*. Chapman and Hall: London.
- Murphy J. 2000. Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *International Journal of Climatology* **20**: 489–501.
- Murphy JM, Sexton DMH, Barnett DN, Jones GS, Webb MJ, Collins M, Stainforth DA. 2004. Quantification of modelling uncertainties in a large ensemble of climate change simulations. *Nature* **430**: 68–772.
- Rohan PK. 1975. *The Climate of Ireland*. The Stationery Office: Dublin.
- Sweeney JC. 1985. The changing synoptic origins of Irish precipitation. *Transactions of the Institute of British Geographers* **10**: 467–480.
- Sweeney J (ed.). 2003. *Climate Change: Scenarios and Impacts for Ireland*. Environmental Protection Agency: Johnstown Castle, Wexford; 229.
- Trigo RM, DaCamara CC. 2000. Circulation weather types and their influence on the precipitation regime in Portugal. *International Journal of Climatology* **20**: 1559–1581.
- Von Storch H, Zorita E, Cusbach U. 1993. Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *Journal of Climate* **6**: 1161–1171.
- Wigley TML, Jones PD, Briffa KR, Smith G. 1990. Obtaining sub-grid scale information from coarse-resolution general circulation model output. *Journal of Geophysical Research* **95**(D2): 1943–1953.
- Wilby RL. 1997. Non-stationarity in daily precipitation series: implications for GCM downscaling using atmospheric circulation indices. *International Journal of Climatology* **17**: 439–454.
- Wilby RL. 1998. Modeling low-frequency rainfall events using airflow indices, weather patterns and frontal frequencies. *Journal of Hydrology* **212–213**: 380–392.

- Wilby RL, Wigley TML. 2000. Precipitation predictors for downscaling: Observed and general circulation model relationships. *International Journal of Climatology* **20**: 641–661.
- Wilby RL, Dawson CW. 2004. *Using SDSM Version 3.1 – A decision support tool for the assessment of regional climate change impacts*. User Manual.
- Wilby RL, Harris I. 2006. A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. *Water Resources Research* **42**: W02419, DOI: 10.1029/2005WR004065.
- Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS. 1998. Statistical downscaling of general circulation model output: a comparison of methods. *Water Resources Research* **34**: 2995–3008.
- Wilks D. 1995. *Statistical Methods in the Atmospheric Sciences*. Academic Press Limited: London.
- Wilks DS, Wilby RL. 1999. The weather generation game: a review of stochastic weather models. *Progress in Physical Geography* **23**(3): 329–357.
- Winkler JA, Palutikof JP, Andresen JA, Goodess CM. 1997. The simulation of daily temperature series from GCM output. Part II: sensitivity analysis of an empirical transfer function methodology. *Journal of Climate* **10**: 2514–2532.
- Yarnal B, Comrie AC, Frakes B, Brown DP. 2001. Developments and prospects in synoptic climatology. *International Journal of Climatology* **21**: 1923–1950.
- Yan ZW, Bate S, Chandler RE, Isham V, Wheeler H. 2002. An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. *Journal of Climate* **15**: 2073–2088.