Climatic Change (2008) 88:157–185
DOI 10.1007/s10584-007-9382-9

Evaluating regional climate model estimates against site-specific observed data in the UK

M. Rivington · D. Miller · K. B. Matthews · G. Russell · G. Bellocchi · K. Buchan

Received: 27 June 2006 / Accepted: 6 November 2007 / Published online: 19 February 2008

Abstract This paper compares precipitation, maximum and minimum air temperature and solar radiation estimates from the Hadley Centre's HadRM3 regional climate model (RCM), (50×50 km grid cells), with observed data from 15 meteorological station in the UK, for the period 1960–90. The aim was to investigate how well the HadRM3 is able to represent weather characteristics for a historical period (hindcast) for which validation data exist. The rationale was to determine if the HadRM3 data contain systematic errors and to investigate how suitable the data are for use in climate change impact studies at particular locations. Comparing modelled and observed data helps assess and quantify the uncertainty introduced to climate impact studies. The results show that the model performs very well for some locations and weather variable combinations, but poorly for others. Maximum temperature estimations are generally good, but minimum temperature is overestimated and extreme cold events are not represented well. For precipitation, the model produces too many small events leading to a serious under estimation of the number of dry days (zero precipitation), whilst also over- or underestimating the mean annual total. Estimates represent well the temporal distribution of precipitation events. The model systematically over-estimates solar radiation, but does produce good quality estimates at some locations. It is concluded that the HadRM3 data are unsuitable for detailed (i.e. daily time step simulation model based) site-specific impacts studies in their current form. However, the close similarity between modelled and observed data for the historical case raises the potential for using simple adjustment methods and applying these to future projection data.

M. Rivington (✉) · D. Miller · K. B. Matthews · K. Buchan
Macaulay Institute, Craigiebuckler, Aberdeen AB15 8QH, UK
e-mail: m.rivington@macaulay.ac.uk

G. Russell
School of Geosciences, The Crew Building, University of Edinburgh, West Mains Road, Edinburgh EH9 3JN, UK

G. Bellocchi
Agrichiana Farming, Montepulciano, Siena, Italy

1 Introduction

A major limitation in conducting site-specific climate change (CC) impacts studies is the difficulty of determining daily data that are representative of the future climate for the site. Data currently produced by climate models driven by global and regional scale land, ocean and atmospheric processes aim to be representative at scales greater than those where impacts studies may be required, such as farms, catchments or ecozones. Though it is possible to generalise about regional level impacts, using mean monthly data provided by global circulation models (GCM) and regional climate models (RCM), it is desirable to assess the impacts at individual locations using data with daily time steps. In order to properly assess potential CC impacts (CCI) and the responses of the subject represented in the impacts study, it is necessary to downscale from global to local scales (Droogers and Aerts 2005). Spatial and temporal differences between the coarse scale GCM and RCM data and the fine scale requirements of site-specific CCI, particularly for natural systems, is seen as a major limitation on the utility of such studies (Zhang 2005). Estimates of alternative future climates derived from GCMs are associated with both significant *scenario uncertainty* (Jenkins and Lowe 2003) and *modelling uncertainty* (Murphy et al. 2004). Examples of scenario uncertainty include greenhouse gas emissions, economic and policy environment and population growth. Modelling uncertainty includes factors such as uncertainty about model parameter values and errors resulting from model structure, both of which are reflected in the quality of individual weather variables. Moberg and Jones (2004) stress the importance of knowing how well RCM's estimate the present climate in order to interpret projected data for future scenarios. It can therefore be argued that the principal limitation in assessing CC impacts and adaptation strategies, is that the uncertainties in the climate model estimates, arising from either scenario and/or modelling uncertainty are either unquantified or so large that meaningful conclusions should not be drawn from them. However, quantifying the modelling uncertainty for the past climate becomes feasible by comparison between the models' hindcast estimates with observed data. This provides indications as to how, where and when biases in future projections may appear, and importantly either what adjustments could be made to correct them, or how CCI outputs should be interpreted when the degree of uncertainty in data inputs are known.

Whilst RCMs are run for long time periods (i.e. hundreds of years) at fine time-scales (i.e. 30 min steps), and representing processes at a range of spatial scales, disseminated output data is generally in an aggregated form and presented as daily estimates for past (i.e. 1960–90) and future projections at set time slices (i.e. 2070–2100). Whilst the RCM aim to represent the climate at the regional scale, the estimates for each cell within the model (typically 50×50 km grid cells) aim to be representative of the mean weather conditions for the mean topographical and geographical characteristics within it. As such, it is beyond the RCM design remit to produce hindcast data identical to that for specific locations within each cell. That said, it would be reasonable to expect that the RCM hindcast data at the grid cell scale would be 'characteristic' of observed data from 'typical' individual sites within the cell (i.e. having variables with similar temporal distribution patterns and value ranges). Where cells contain large topographical diversity however, it may be expected that data from meteorological stations at the extremes of that diversity are unlikely to show close similarity with modelled data. Any discrepancies will be partially due to the site and cell differing in terms of topography, altitude, aspect or distance to the sea. Differences may also be attributable to the RCMs inability to adequately represent a particular cell due to necessary assumptions or simplifications within the RCM modelling process.

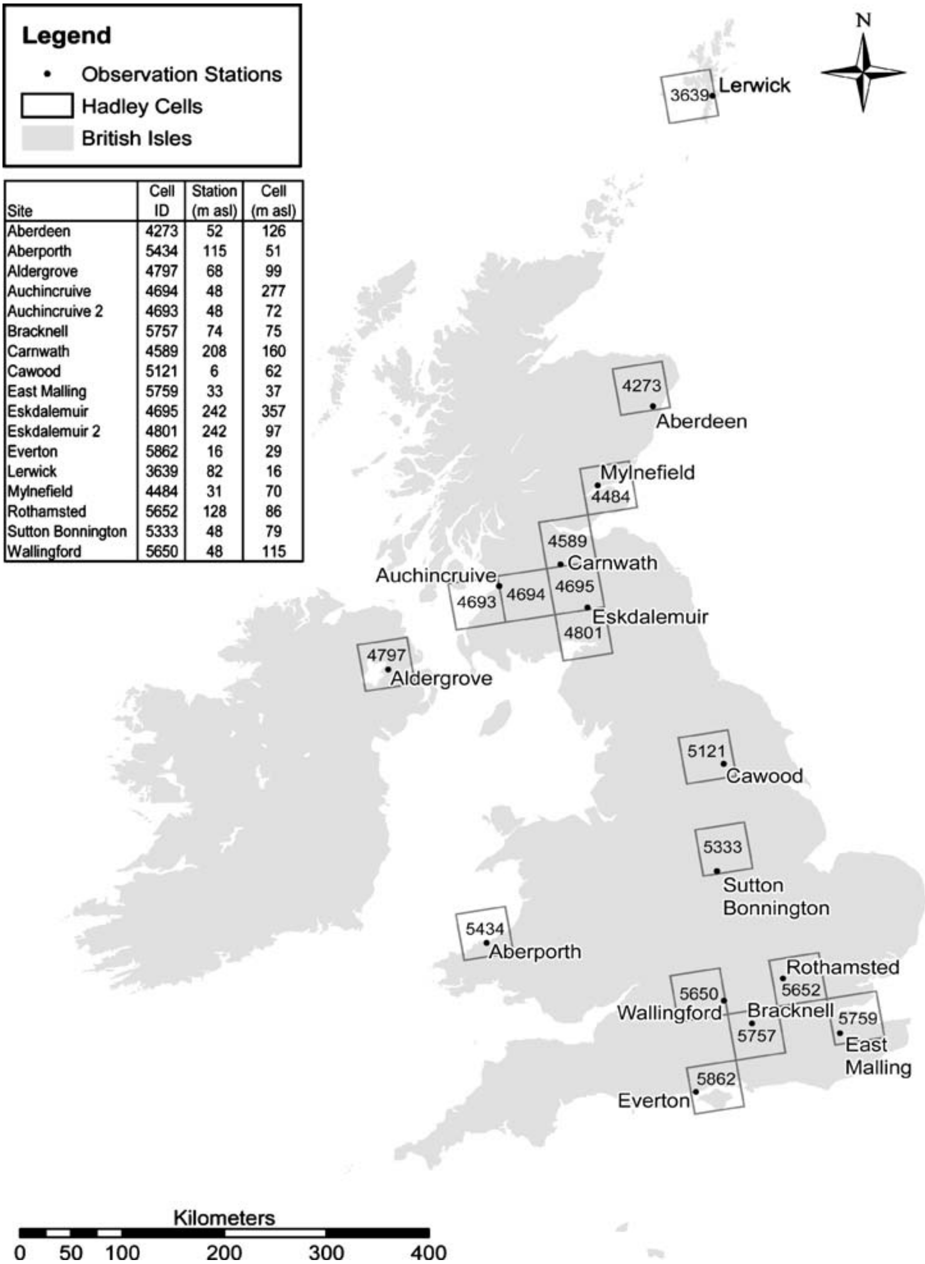
Improving the quality of daily weather variable estimates for future scenarios is vital in order to better understand how biogeochemical processes will function under new climate conditions. If appropriate CC mitigation strategies are to be developed, for example in managing land uses to reduce greenhouse gas (GHG) emissions and increasing carbon sequestration, it is essential to be able to predict reliably the dynamic responses (spatial and temporal) of the key biogeochemical processes. Many CCI studies will take the form of predictive modelling experiments, where simulations use RCM data as inputs. Unless errors and other uncertainties in modelled weather data can be identified and quantified, the reliability and utility of the projections and use within CCI will be less certain. Government strategic plans to cope with CC would hence be based on potentially incorrect evidence from impacts studies.

In this paper we compare the Hadley Centre's RCM HadRM3 hindcast estimates of precipitation, maximum ($T_{\max}$) and minimum ($T_{\min}$) air temperature and total downward surface shortwave flux (direct and diffuse solar radiation, MJ m^2 day^{-1}), here referred to as solar radiation (S_o), with observed data from 15 meteorological stations in the UK for the period 1960–90 (Fig. 1.). For the reasons given earlier, we expect there to be differences between the two data sets, as this is not a 'like with like' comparison, rather a 'grid with point' one. However, if the differences are conservative, they can be used to identify potential adjustments to future grid cell projections, and provide information as to how errors may appear when used in CCI studies. It is argued that it is necessary to assess the quality of the modelled climate estimates in advance, in order to determine the uncertainty that will be introduced to any CCI study. If the assumption is made that the same modelling errors existing in estimates of the past climate will also be present in future climate projections, it is possible to appraise the usefulness of the future estimates, and potentially adjust biases using appropriate methods. Once these steps have been taken more reliable CCI studies should be possible.

2 Related research

Studies have sought to assess the performance of GCMs and RCMs at a range of spatial and temporal scales (i.e. Peng et al. 2002; Antic et al. 2006). However, to date little work has been done to compare RCM hindcast estimates with site-specific multiple variable observed data (Moberg and Jones 2004). Exceptions include Bell et al. (2004), who performed a model versus observed validation exercise as part of a larger study of growing season length, extreme temperatures and precipitation in California. Long term data from 16 stations were compared with 15 years of modelled data from a modified version of RegCM2, for sites where the actual and modelled elevation differed by no more than 100 m. These authors concluded that the RCM was able to make good estimates of seasonal temperature and precipitation. One limitation of this study was that temperature based assessments were distorted by the need to use proxy values for maximum and minimum temperature (the model output values at midnight and midday, rather than providing the absolute daily values). This indicates the need for careful consideration of what and how data from RCMs are output and archived.

Moberg and Jones (2004) tested the HadRM3P model (closely related to the HadRM3 assessed here) estimates of daily maximum and minimum near-surface temperatures for the period 1961–90 for 185 meteorological stations across Europe. The analysis was primarily based on the model-minus-observed values for mean annual and seasonal temperature differences, though results for daily differences (forming the annual temperature cycle)



Site	Cell ID	Station (m asl)	Cell (m asl)
Aberdeen	4273	52	126
Aberporth	5434	115	51
Aldergrove	4797	68	99
Auchincruive	4694	48	277
Auchincruive 2	4693	48	72
Bracknell	5757	74	75
Carnwath	4589	208	160
Cawood	5121	6	62
East Malling	5759	33	37
Eskdalemuir	4695	242	357
Eskdalemuir 2	4801	242	97
Everton	5862	16	29
Lerwick	3639	82	16
Mylnefield	4484	31	70
Rothamsted	5652	128	86
Sutton Bonnington	5333	48	79
Wallingford	5650	48	115

Fig. 1 Meteorological stations providing observed data and the position of their associated HadRm3 50× 50 km grid cell, with the station and mean cell elevations (*m a.s.l.*)

were given for six locations. These authors found large spatial variations in the ability of the model to reproduce the historical weather well. It performed well in the UK and some other locations between 50 and 55°N, with differences generally being ±0.5°C, but other areas showed differences of up to ±15°C. This study provided valuable information about the degree of spatial variability in the quality of mean annual and seasonal temperature differences at the regional scale, but did not cover site-specific multiple variable assessment.

Studies have compared estimates with observed data for individual weather variables at larger spatial and temporal scales. For example, Mearns et al. (1995) assessed the quality of estimates of precipitation by RegCM for a 42 month period. They stressed the importance of models being able to reproduce the frequency and intensity of precipitation events, not just the daily means. They also highlighted the limitations of statistical analysis of data sets for periods of only a few years. Evans et al. (2005) tested four RCMs over a period of two years at a site in Kansas, USA, and found no clear distinction in performance between the models, which all had positive and negative attributes. Fowler et al. (2005) tested the HadRM3 RCM for extreme rainfall events at 204 sites in the UK. Although the model provided good estimates of return periods for up to 50 years, it exaggerated the west to east rainfall gradient, leading to overestimations in some higher elevation western areas, and underestimation in eastern rain shadow areas.

RCM estimates have been compared with regional scale aggregations of observed data (i.e. Frei et al. 2002; Huntingford et al. 2003), or for time scales greater than individual days (i.e. Vidale et al. 2003). In testing the Rossby Centre Atmospheric RCM, RCA2, Jones et al. (2004) found that the model tended to overestimate the number of small precipitation events, which impacted on surface temperatures and cloud-radiation interactions. Differences are not only found for temperature and precipitation. Kim and Lee (2003) found that surface insolation was generally overestimated in an eight year hindcast simulation for the Western USA with the differences being smaller over land than over the sea.

2.1 Importance of daily site-specific data

As many biological and chemical processes can only be studied effectively at a scale of a few hectares or less, there is a need to measure or otherwise provide weather data for the exact location (Hoogenboom 2000). However, policy makers are typically concerned with the outcomes of key elements such as production, i.e. crop yields, and processes like GHG emissions, soil water balances and carbon sequestration, at regional (Holman et al. 2005), national (Sperow et al. 2003) or even supranational (Nijkamp et al. 2005) scales. The reliability of estimates of such process outcomes, however, depends on robustly parameterised relationships between the driving climatic variables and the outcomes of interest whilst incorporating anthropogenic factors such as adaptations of management regimes (Rivington et al. 2007). Without site- or plot-specific data, unquantified uncertainties are introduced into CCI studies and projections, making decisions based on evidence from such studies as either unreliable, or if the uncertainties are unrecognised, introduce biases that lead to erroneous decisions being made.

The uncertainty introduced into systems-model estimates due to the weather data source can be significant (Rivington et al. 2006). Nonhebel (1994a) found that inaccuracies in solar radiation measurement of 10% and of daily temperature of 1°C in data used within a crop simulation model resulted in yield estimation errors of up to 1 t ha^{-1}. Maintaining meteorologically appropriate, synchronised relationships between individual weather variables is essential for models that represent entities with non-linear responses to driving variables such as biological systems (Nonhebel 1994b) and hydro-chemical processes (Soulsby 1995; Creed et al. 1996). Thermal time accumulation, which depends not only on the mean daily temperature but the difference between daily maximum and minimum temperatures, is the key driver of plant and insect phenological development (Arnold and Monteith 1974; Jarvis et al. 2003, respectively). Systematic errors in the estimation or synchronisation of either T_{max} and T_{min} will result in predictions of either faster (earlier) or slower (later) development, with corresponding impacts on associated management (i.e. crop) or behavioural (i.e. plant–insect–predator) responses.

While the examples above are drawn from the agro-climatic rather than climate change literature, it seems reasonable to draw conclusions that, when using estimates of future climate derived from RCMs, researchers should be as concerned with issues of driving variable data quality. Hindcast RCM data provide a unique opportunity to assess the nature of the uncertainty introduced to CCI studies, particularly systems models' predictions by the use of RCM rather than site specific information.

3 Materials and methods

Observed precipitation (mm), maximum (T_{max}) and minimum (T_{min}) air temperature (°C) and total downward surface shortwave flux (direct and diffuse solar radiation, S_o, MJ m^2 day^{-1}) data for the period 1960–90 were provided by the British Atmospheric Data Centre (BADC 2005) for 15 meteorological stations in the UK (Fig. 1). The criteria for selection of sites was that their data record contained the maximum number of complete years for all weather variables, and were sufficiently geographically dispersed to give a reasonable spatial representation of the UK, but also did not exist at the extremes of topography within each cell. The number of sites available for assessment was limited by the availability of S_o data. Carnwath, despite not having S_o data, was included as it is a site of on-going CCI modelling (paper in preparation). Observed data for precipitation, T_{max} and T_{min}, and S_o were compiled within an Oracle database, with errors, duplicates and anomalies in the original data being identified and corrected during the database loading process. Missing observed values were filled using a search and optimisation method (LADSS 2007).

Modelled data used in this assessment is based on the hindcast simulations of the Hadley Centre's HadRM3 RCM, as used in the UKCIP02 climate change scenarios report for the UK (Hulme et al. 2002). As an initial condition ensemble, five hindcast simulations (starting from 1860) were produced by the HadRM3 in order to establish the 1960–90 climate normal period 'baseline' against which future projections were compared in the UKCIP02 report. Each hindcast simulation varied slightly in their starting conditions, but atmospheric CO_2 and other GHG concentrations were varied to match the historical concentrations. Future projections of GHGs, as per the Special Report on Emissions Scenarios (SRES; IPCC 2000) were not applied until after 1990. This paper assesses two of the hindcast simulation data sets that were used to compare with the SRES A2c (medium-high) and B2 (medium low) future GHG emissions scenarios used in the UKCIP02 report. These two hindcast data sets are herein referred to as the A2c$_{IRH}$ and B2$_{IRH}$, where IRH is the initial realisation hindcast (observed historical GHG concentrations). As such, this paper assesses only two examples of the hindcast configuration of the HadRM3. The A2c$_{IRH}$ and B2$_{IRH}$ (1960–90) data were also provided by the BADC.

Daily climate data for each variable were derived from the HadRM3 archive for 50× 50 km grid cells (the extent of each RCM cell used is shown in Fig. 1). Each meteorological station was matched with its corresponding cell except in two cases where the stations were within 2 km of the cell boundary (Auchincruive and Eskdalemuir), in which case the opportunity was taken to use the closest neighbouring RCM cells for comparison as well.

The hindcast data produced by the RCM do not attempt to recreate synoptic conditions for specific locations or years in the period 1960–90. Instead, the RCM outputs are similar to those from weather-generators such as LARS (Semenov 2002) and CLIMGEN (Stöckle et al. 1999), in that they consist of time-series of data with the correct statistical properties including correlations between variables. The RCM outputs represent the 50×50 km grid cell as a whole rather than a specific site within it. RCM do not aim to reproduce the actual weather for a specific day or year in the past, rather they aim to produce values of a variable that are representative (by magnitude, variability and synoptic synchronisation) of a day at any specific time of year. As such direct day or year specific model versus observed data comparisons are impractical i.e. observed data for April 1st 1970 cannot be compared with modelled data for April 1st 1970. Instead, mean daily, annual totals or maximum and minimum values were used for comparisons between observed and RCM data. As the HadRM3 model treats a year as having 360 days (i.e. twelve months of 30 days), the last five days of the observed data were omitted from the analyses.

In this work, no *a priori* adjustments were made to the modelled data to take account of differences in elevation or other topographically significant differences between the meteorological station and the mean for the grid cell. Moberg and Jones (2004) found that adjustments to modelled data based on temperature lapse rates resulted in changes of just a few tenths of a degree K for the majority of sites they tested. The mean elevation for each grid cell was estimated (see Fig. 1) and used as one of the explanatory factors for the differences observed.

3.1 Precipitation

Histograms were created to show the frequency distribution of the magnitude of precipitation events for all precipitation events (Fig. 2). The probability of excedence (P_e), as a percentage, was calculated following Weibull (1961) for each precipitation event:

$$P_e(\%) = m/(n+1) \times 100. \tag{1}$$

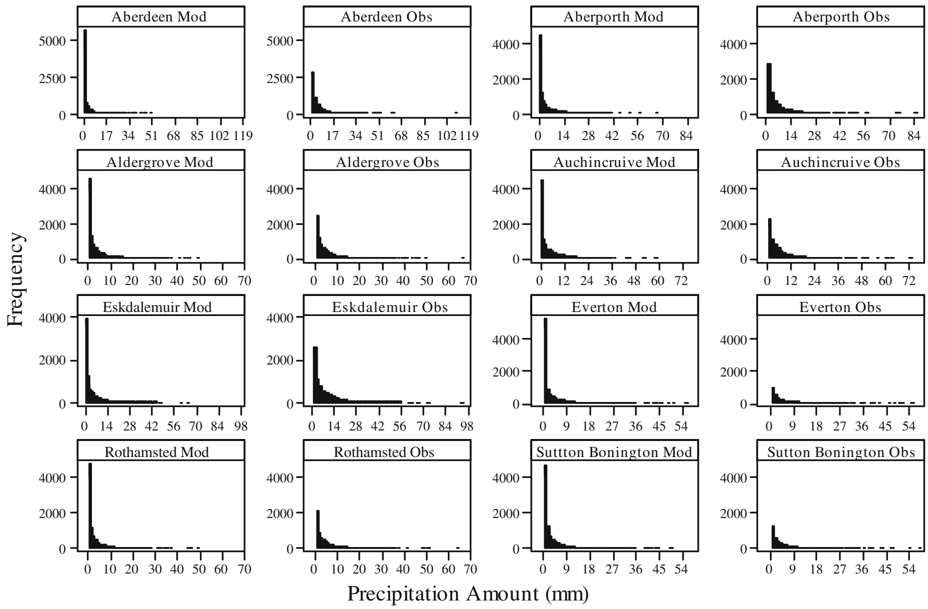


Fig. 2 Histograms of precipitation magnitude frequency for all events (not including dry days) for modelled (*Mod*) and observed (*Obs*) data at eight selected locations

Where m is the rank order of each precipitation event, with $m=1$ as the largest event and $m=n$ for the lowest, with n being the number of observations (in this case $n=360$ days× 31 years=11,160). This comparison enables the probability of occurrence to be determined for each precipitation amount (Fig. 3a) while avoiding the problem of asynchronicity between the observed and hindcast data. Subsequent to this, the actual difference (mm) and proportional difference against observed events was estimated by ranking the events in decreasing order of magnitude and taking the difference (modelled − observed; Fig. 3c) then dividing by the observed value (Fig. 3b). The annual total, magnitude of largest event and the number of days with no precipitation (dry days) were calculated for each year at all locations tested. To assess the temporal distribution of events, plots of the 7-day (weekly) means were made (Fig. 4).

3.2 Temperature

For T_{max} and T_{min} the mean daily values (for the 31-year period) were calculated and plotted for the observed and estimated data (eight examples shown in Fig. 5). This enabled the magnitude of differences to be visually identified, and their temporal distribution to be observed. The differences between mean daily T_{max} and T_{min} were calculated and plotted (Fig. 6), in order to assess the models' ability to represent the daily temperature range. The highest and lowest values for daily T_{max} and T_{min} were found and plotted (Fig. 7), to evaluate the models' ability to represent temperature daily variability and extreme ranges. Accumulated thermal time (°Cday) was calculated as $(T_{max} + T_{min})/2$ added to the previous days' accumulated thermal time (with a base temperature of 0°C; Fig. 8).

The annual total of T_{max}, T_{min}, highest and lowest temperatures, mean number of days with $T_{max}>15$°C, $T_{min}<0$°C and $T_{min}-5$°C were calculated (Tables 2 and 3). The annual mean, standard deviation and paired Student's t test of probability of equal means ($P(t)$) (where $P(t)=1$ shows equal means and $P(t)=0$ shows no similarity), were estimated for daily values to determine their statistical similarity (Table 4).

3.3 Solar radiation

Observed solar radiation data records are often incomplete for the 1960–90 period, hence analysis was limited to graphical representations using the difference D between mean daily observed versus estimated solar radiation, which was calculated from all available years at each site:

$$D \text{ has elements } d_i = \overline{e_i} - \overline{o_i} \tag{2}$$

where $\overline{e_i}$ is the mean estimated solar radiation for day i over n years, and $\overline{o_i}$ is the mean observed solar radiation for day i over n years with

$$\overline{e_i} = \frac{1}{n} \sum_{j=1,n} e_{ji} \tag{3}$$

and

$$\overline{o_i} = \frac{1}{n} \sum_{j=1,n} o_{ji} \tag{4}$$

where e_{ji} is the estimated solar radiation on day i of year j, and o_{ji} is the observed solar radiation on day i of year j. This difference in daily means helps to illustrate the temporal

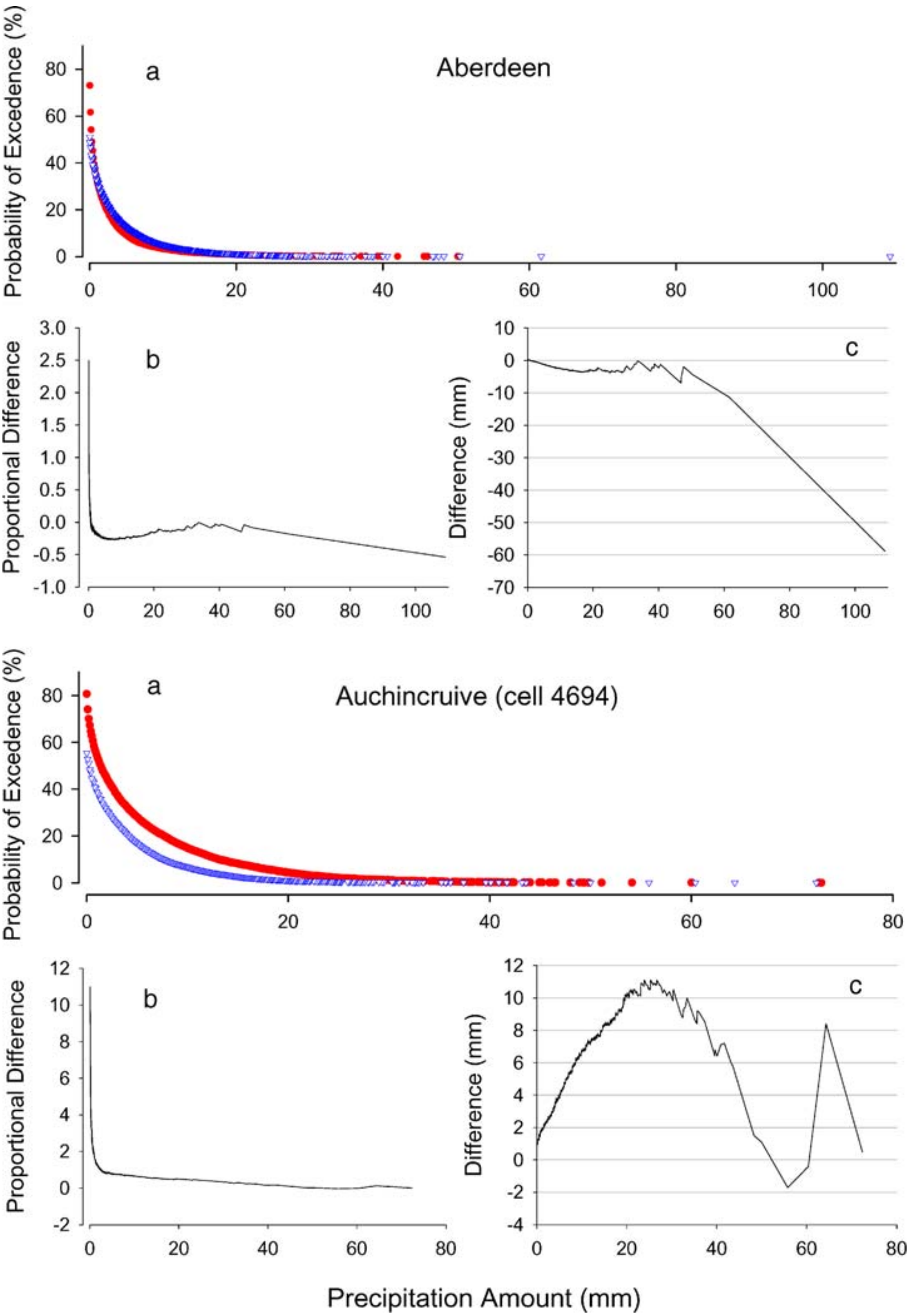


Fig. 3 Probability of excedence (%) of modelled (*red dots*) versus observed (*blue triangles*) for individual precipitation events (**a**); proportional difference ((modelled − observed)/observed) (**b**); and difference (mm; modelled − observed) (**c**) plots against observed precipitation amounts (mm) for four example locations

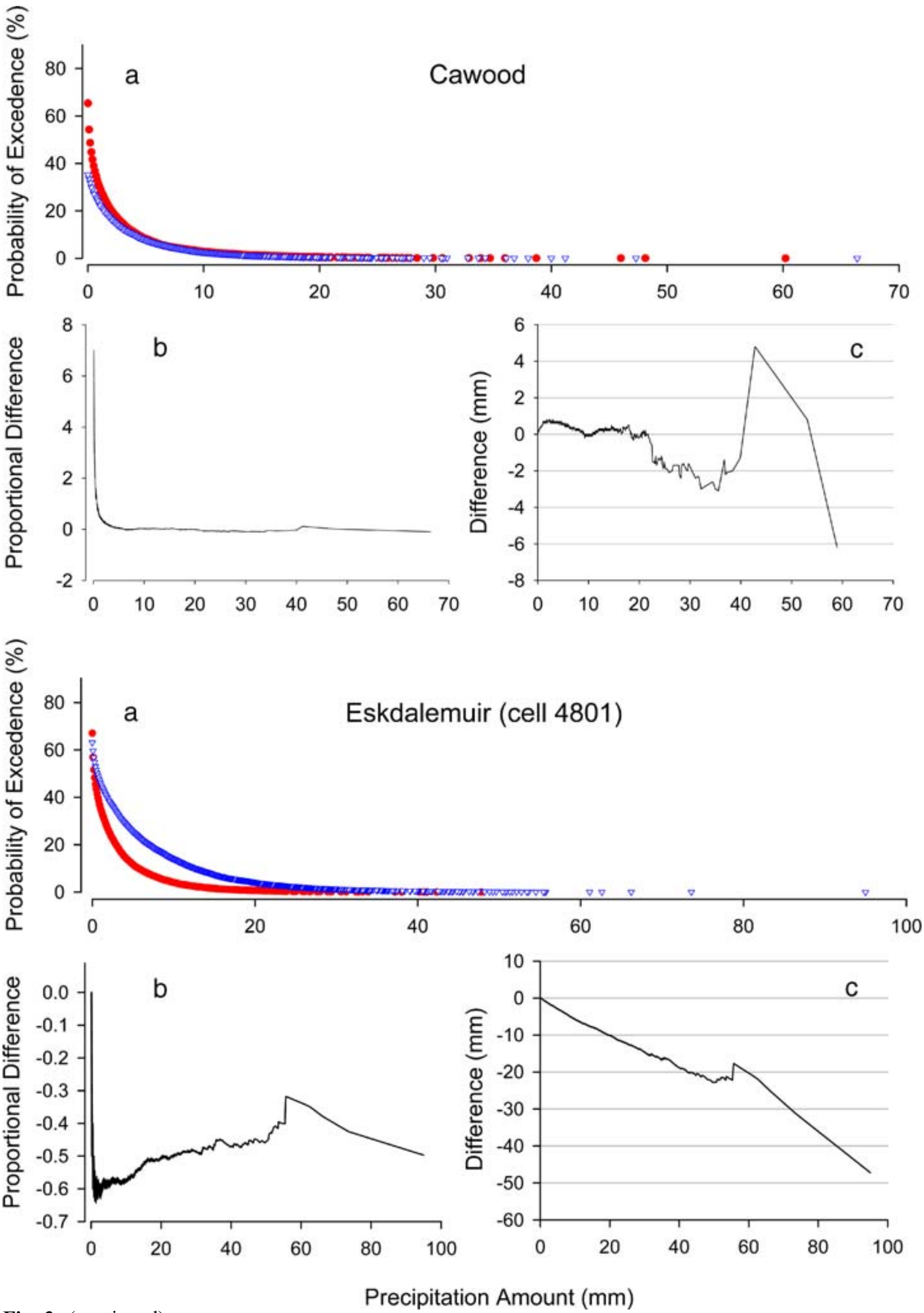


Fig. 3 (continued)

distribution of mean daily errors (over- and underestimations) over the period of a year, indicating systematic model behaviour (Fig. 9). This approach was taken to allow direct comparison of results with a previous study of solar radiation model performance by Rivington et al. (2005).

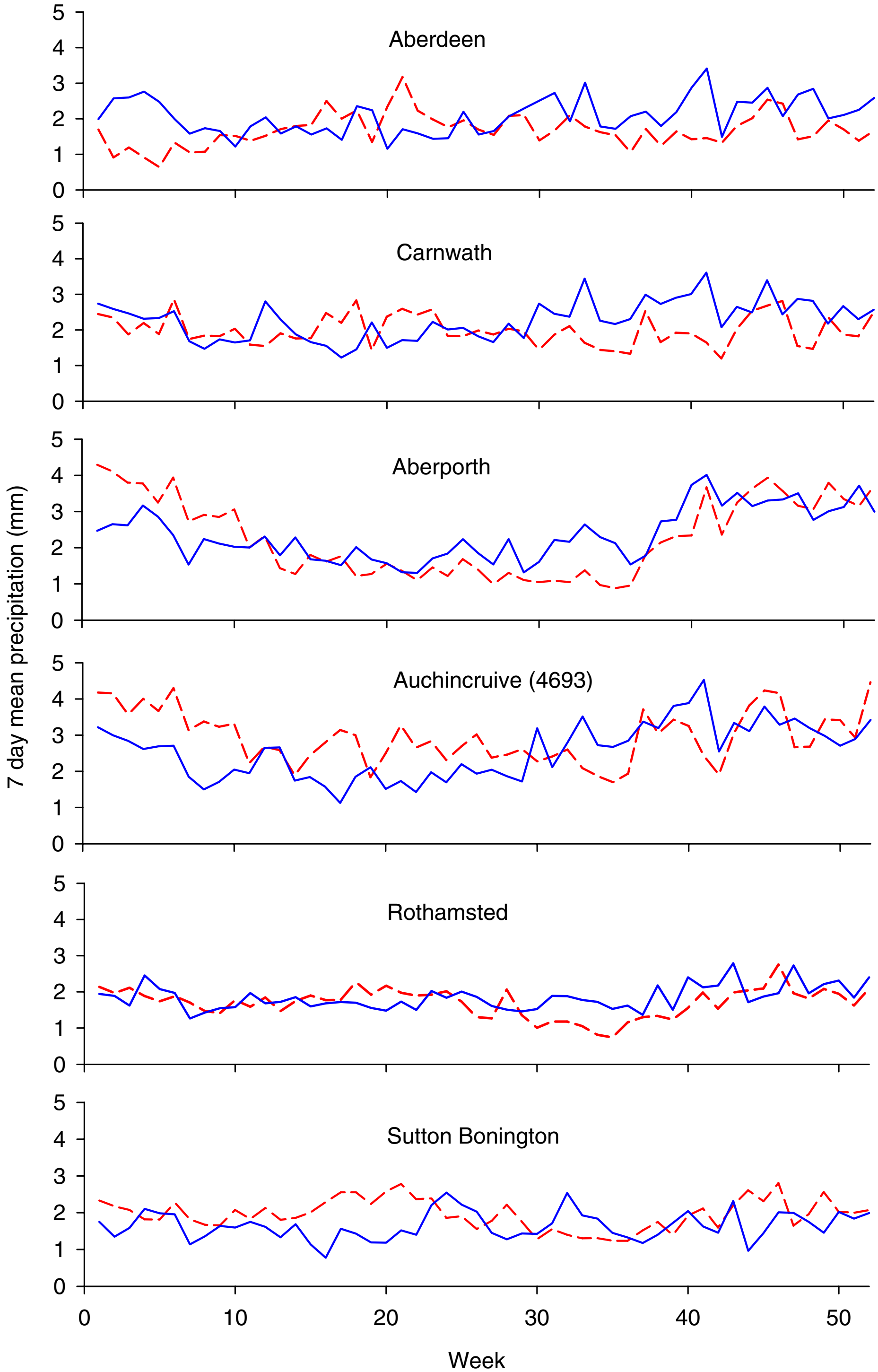


Fig. 4 Seven day (weekly) mean temporal distribution of modelled (*red dashed line*) versus observed (*blue solid line*) precipitation (n=30 years) over a 1-year period from six selected locations

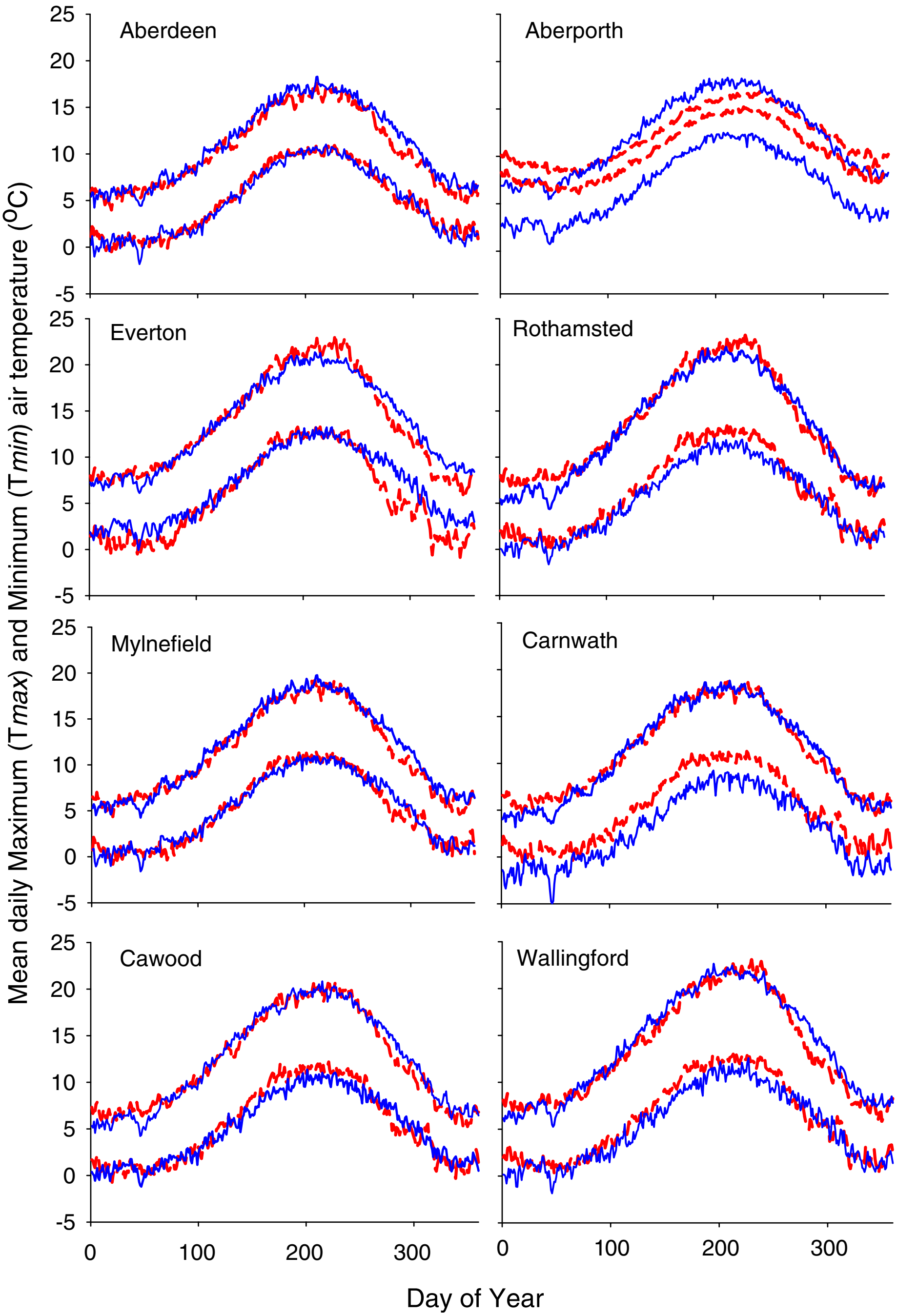


Fig. 5 Modelled (*red dashed line*) versus observed (*blue solid line*) mean daily maximum ($T_{\max}$ *upper lines*) and minimum ($T_{\min}$ *lower lines*) air temperature

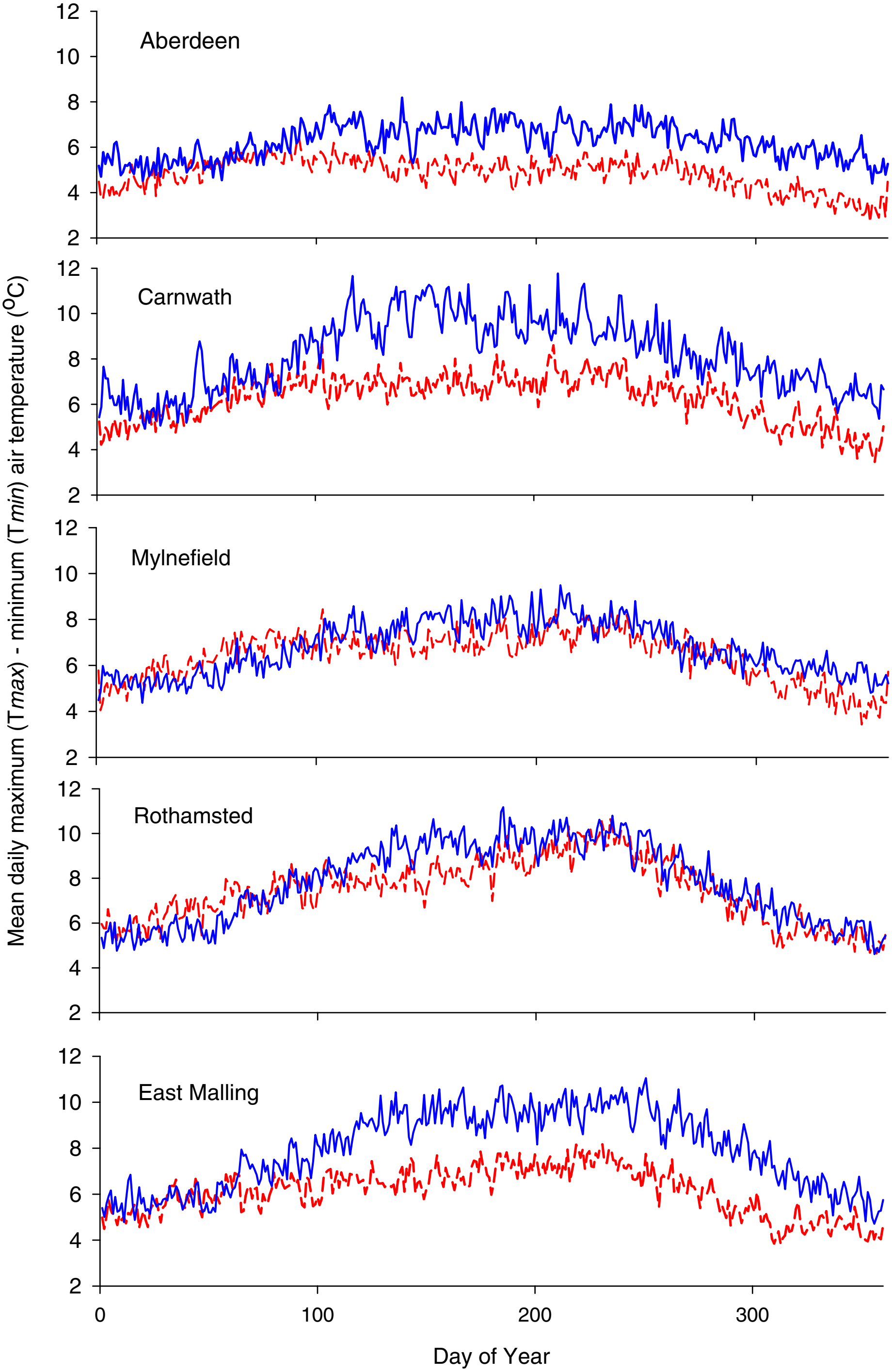


Fig. 6 Difference between mean daily maximum and minimum temperature ($T_{max} - T_{min}$) for modelled (*red dashed line*) versus observed (*blue solid line*) data

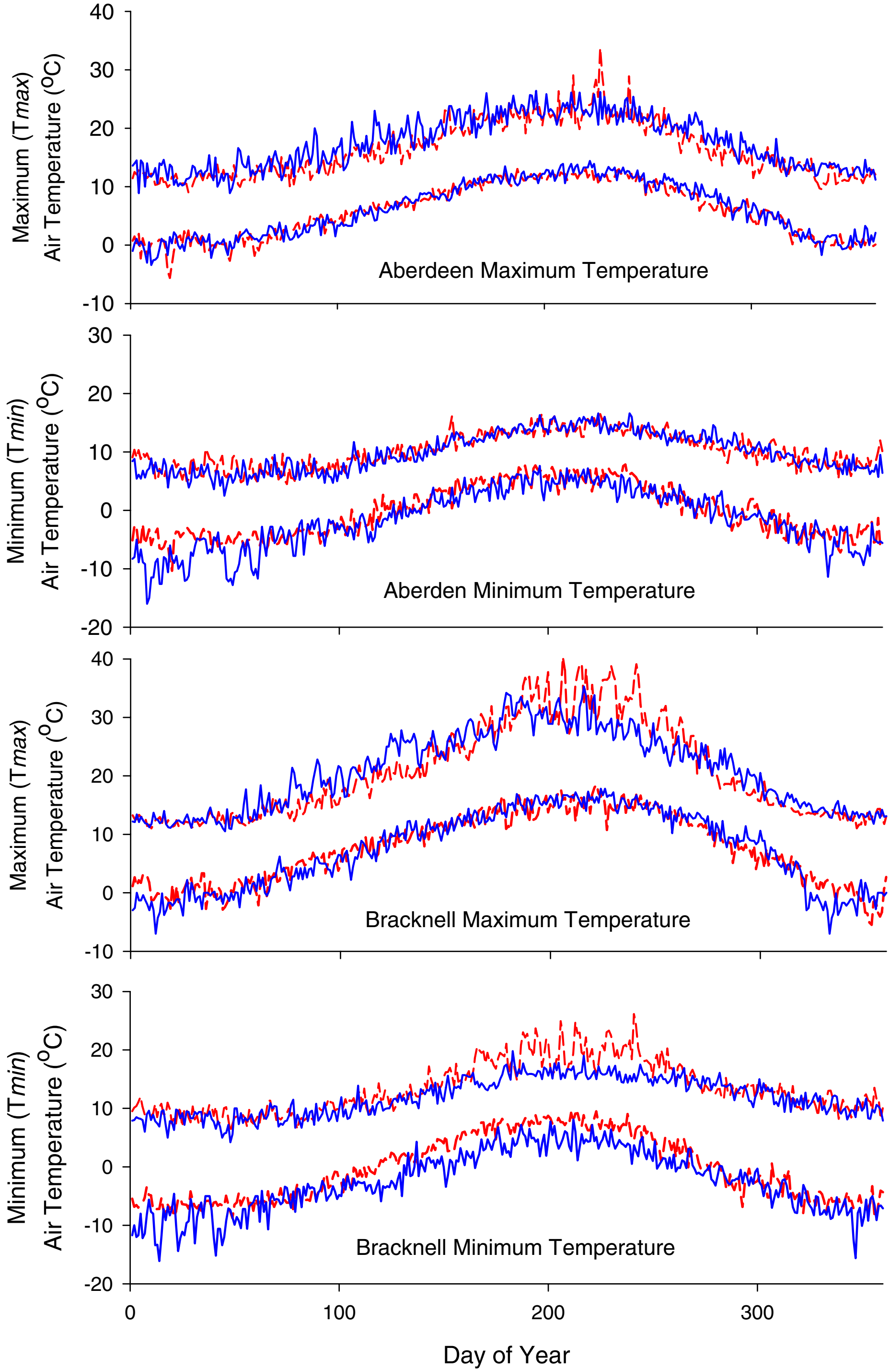


Fig. 7 Highest (*upper lines*) and lowest (*lower lines*) of maximum ($T_{\max}$) air temperature and highest (*upper lines*) and lowest (*lower lines*) of minimum ($T_{\min}$) air temperature values for modelled (*red dashed line*) versus observed (*blue solid line*) data

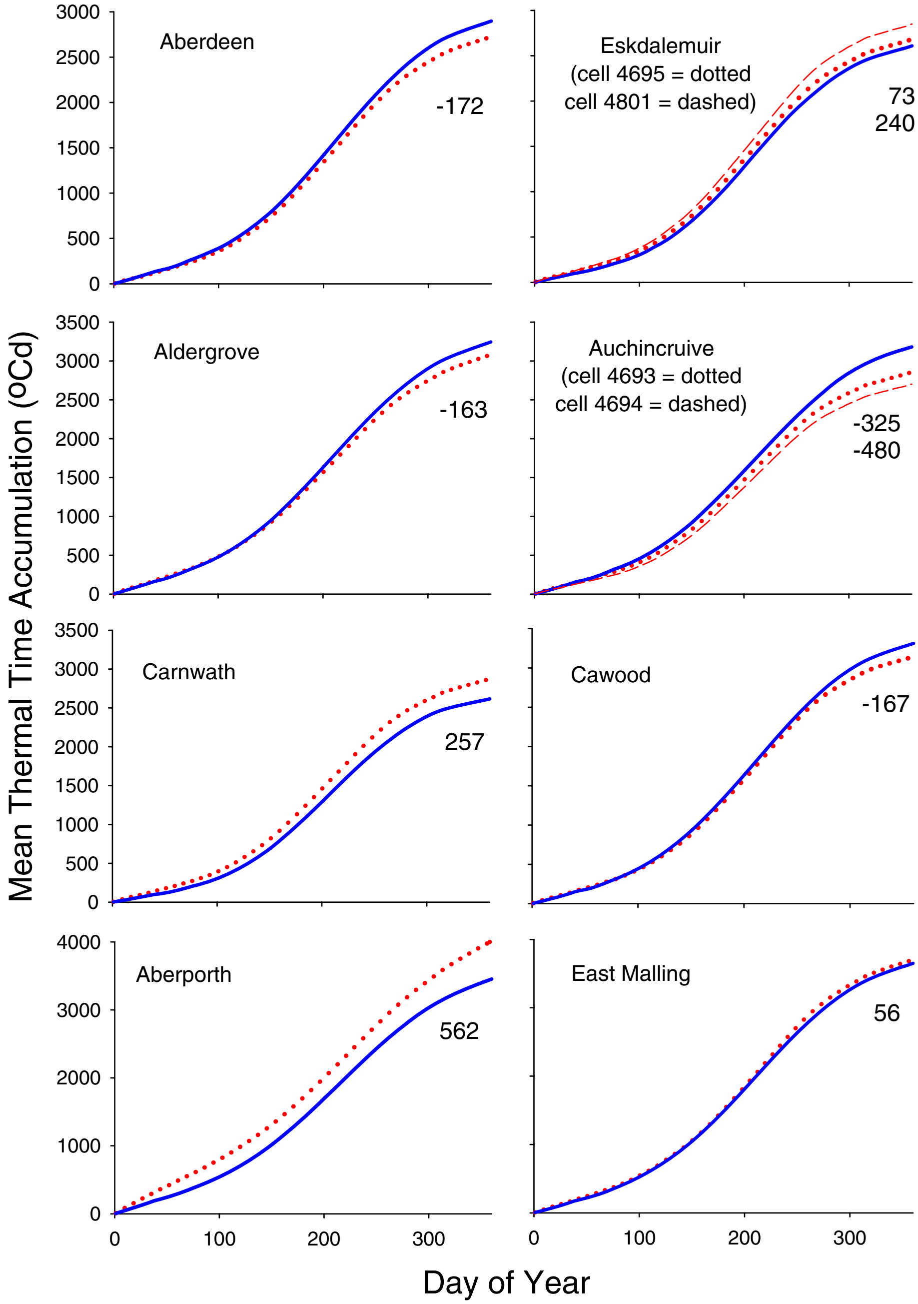


Fig. 8 Thermal time accumulation (°Cd) for modelled (*red dashed line*) and observed (*blue solid line*) data. Values shown are difference in total thermal time accumulation on the last day of the year

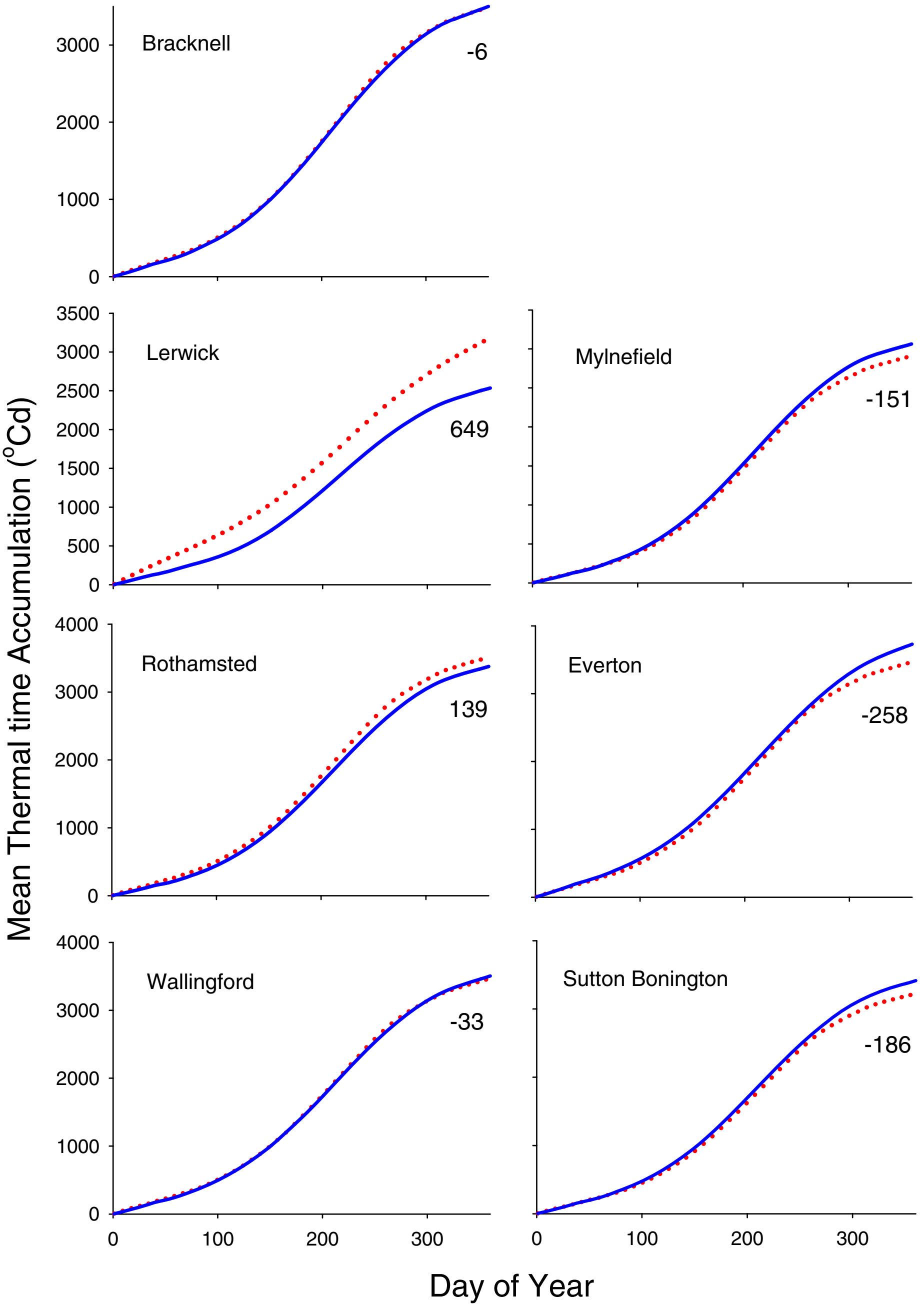


Fig. 8 (continued)

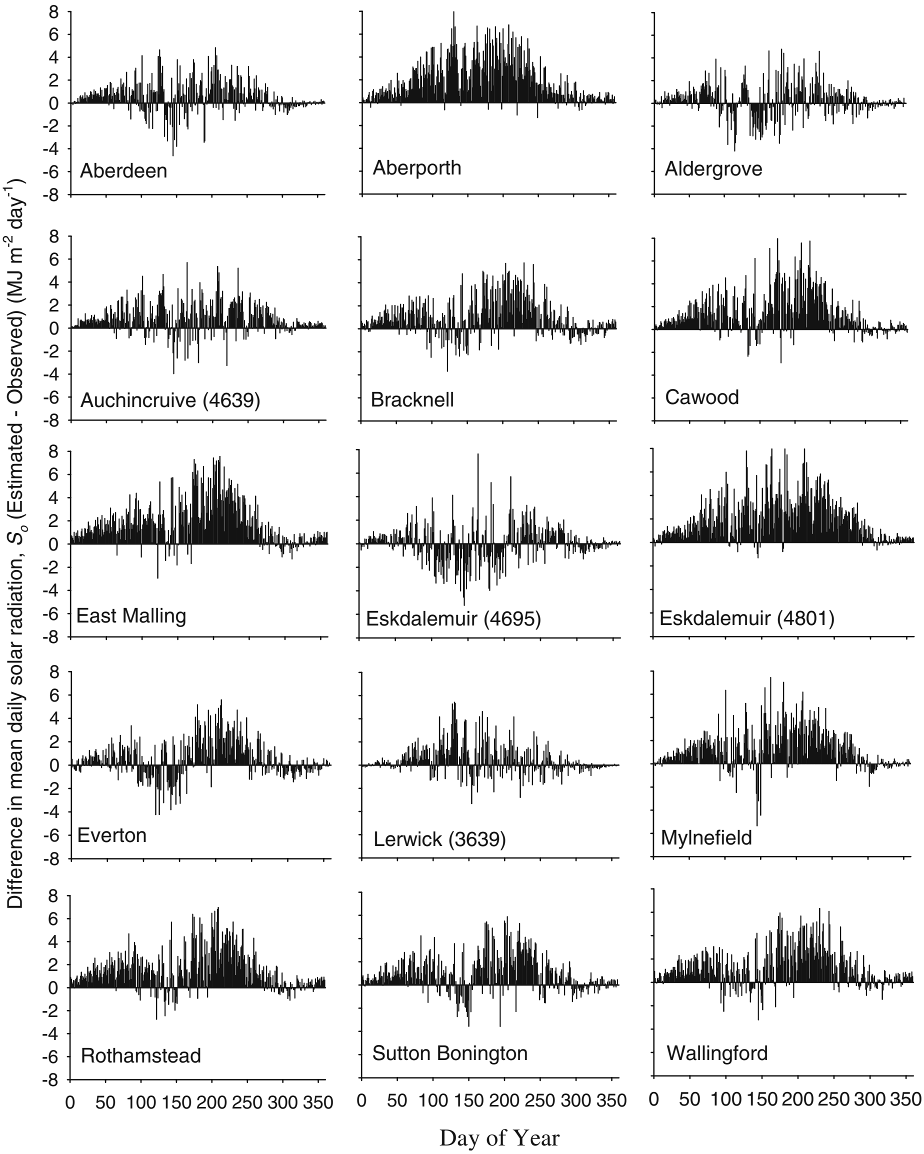


Fig. 9 Difference (estimated − observed) in mean daily solar radiation (S_o; MJ m^{-2} day^{-1}) over the period of a year, where bars above 0 indicate model overestimation

4 Results

In comparing the A2c$_{IRH}$ and B2$_{IRH}$ configuration data, temperature estimates were similar and only one location (Auchencruive: cells 4693 and 4694) showed substantially different precipitation totals. Solar radiation estimates were similar between both scenarios. On the basis of this similarity, only the A2c$_{IRH}$ graphical analyses results are presented in this paper. Graphs illustrating the results of the B2$_{IRH}$ analyses can be found at http://www.macaulay.ac.uk/LADSS/climate. The A2c$_{IRH}$ and B2$_{IRH}$ configuration data represents a two-member ensemble experiment for the 1960–90 period, which would provide useful insights into the numerical uncertainty in modelling the climate system. However, it has been beyond the scope of this paper to perform such a detailed evaluation of the numerical differences between the two, but will form the basis for further research. Modelled T_{max} data for 1984 and 1985 were missing from the hindcast data set, and observed data for each variable were missing for several years at some locations; hence sample sizes vary for some of the comparisons. Few meteorological stations in the UK have recorded precipitation, temperature and solar radiation together, particularly in the early part of the study period.

4.1 Precipitation

At all sites the model produces an excess of small precipitation events (<1 mm; Fig. 2), the number of dry days being underestimated by an average of 60% (Table 1) for both $A2c_{IRH}$ and $B2_{IRH}$. This reflects the 'drizzle effect' whereby the model attempts to capture the spatial distribution of very light rain across the entire cell. Whilst these are important in terms of impact on factors such as soil wetting, they distort the daily distribution of rainfall events. The model was able to make very good estimates, i.e. at Cawood there was a difference of only 1 mm between observed and modelled $A2c_{IRH}$ mean annual total. Where the model over-estimates the mean annual total, i.e. Auchincruive cell 4694 (Table 1), there is a corresponding higher probability of modelled events occurring (Fig. 3a: Auchincruive),

Table 1 Difference between modelled and observed precipitation (mm; for $A2c_{IRH}$ and $B2_{IRH}$ configurations) for mean annual total, maximum (largest) single event and number of days without precipitation (dry days)

	Mean annual total (mm)					Maximum single event (mm)			Dry days (0 mm)		
		$A2c_{IRH}$		$B2_{IRH}$			$A2c_{IRH}$	$B2_{IRH}$		$A2c_{IRH}$	$B2_{IRH}$
Meteorological station (cell)	Obs	Model	Diff	Model	Diff	Obs	Model	Model	Obs	Model	Model
Aberdeen (4273)	761	604	−157	595	−166	109	50	52	173	57	58
Aberporth (5434)	870	838	−31	845	−24	85	66	45	163	76	78
Aldergrove (4797)	845	814	−31	866	21	66	49	47	130	64	61
Auchincruive (4693)	936	1074	138	1691	756	72	59	73	156	48	45
Auchincruive (4694)	936	1597	662	1748	812	72	73	73	152	47	45
Bracknell (5757)	663	761	98	765	102	71	56	56	193	78	79
Carnwath (4589)	832	723	−109	745	−88	59	64	54	135	63	62
Cawood (5121)	536	535	−1	564	28	66	60	57	183	83	82
East Malling (5759)	650	547	−103	546	−104	82	63	42	193	95	98
Eskdalemuir (4695)	1534	1215	−319	1250	−284	95	66	70	127	48	46
Eskdalemuir (4801)	1534	681	−854	699	−836	95	48	57	127	77	74
Everton (5862)	738	777	40	780	42	56	55	97	203	63	65
Lerwick (3639)	1201	1057	−144	1079	−122	59	42	75	96	23	23
Mylnefield (4484)	692	500	−192	504	−188	49	73	37	175	79	77
Rothamsted (5652)	674	619	−55	637	−37	64	50	59	178	79	81
Sutton Bonington (5333)	601	711	110	759	158	59	50	52	191	70	70
Wallingford (5650)	577	693	116	708	131	65	61	55	204	79	80
Mean	858	809	−49	869	12	72	58	59	163	67	66

Table 2 Difference between modelled and observed maximum temperature (for $A2c_{IRH}$ and $B2_{IRH}$ configurations) for mean annual total, highest and lowest single event, and the number of days >15°C

	Maximum temperature (°C)													
	Mean annual total (°C)					Highest single event (°C)			Lowest single event (°C)			Mean days >15°C		
Meteorological station (cell)	Obs	$A2_{IRH}$	Diff	$B2_{IRH}$	Diff	Obs	$A2_{IRH}$	$B2_{IRH}$	Obs	$A2_{IRH}$	$B2_{IRH}$	Obs	$A2_{IRH}$	$B2_{IRH}$
Aberdeen (4273)	4031	3765	−266	3838	−193	26	34	29	−3	−6	−2	86	70	72
Aberporth (5434)	4412	4362	−50	4394	−18	32	23	22	−5	1	1	106	82	81
Aldergrove (4797)	4484	4212	−272	4254	−230	25	24	32	0	1	−1	121	93	93
Auchincruive (4693)	4270	4039	−231	3835	−435	29	32	34	−3	−8	−5	107	88	71
Auchincruive (4694)	4270	3787	−483	3760	−509	29	31	34	−3	−6	−5	107	72	71
Bracknell (5757)	5022	4877	−145	4927	−95	35	40	42	−7	−5	−4	156	138	142
Carnwath (4589)	4023	4029	7	4084	62	30	35	36	−12	−3	−4	96	94	94
Cawood (5121)	4707	4472	−235	4513	−194	34	39	39	−5	−4	−5	141	121	126
East Malling (5759)	5097	5038	−59	5097	0	35	41	41	−6	−6	−2	160	146	150
Eskdalemuir (4695)	3911	3764	−147	3820	−91	30	34	34	−10	−6	−6	85	76	75
Eskdalemuir (4801)	3911	4151	240	4202	291	30	35	39	−10	−6	−5	85	104	105
Everton (5862)	4998	4940	−57	4997	−1	34	39	42	−5	−7	−4	151	141	145
Lerwick (3639)	3347	3416	69	3443	97	22	16	16	−3	−1	1	17	1	0
Mylnefield (4484)	4290	4109	−181	4174	−116	29	36	33	−9	−3	−5	112	101	103
Rothamsted (5652)	4776	4917	141	4964	188	34	41	42	−7	−6	−5	145	141	145
Sutton Bonington (5333)	4770	4553	−217	4586	−185	35	40	42	−7	−5	−6	144	125	128
Wallingford (5650)	5047	4869	−177	4910	−137	35	41	42	−9	−5	−3	158	137	141
Mean	4433	4312	−121	4341	−92	31	34	35	−6	−4	−3	116	102	102

Table 3 Difference between modelled ($A2c_{IRH}$ and $B2_{IRH}$ configurations) and observed minimum temperature (T_{min} °C) for mean annual total, highest and lowest single event, and the number of days <0°C and −5°C

	Minimum temperature (°C)																
	Mean annual total (°C)					Highest single event (°C)			Lowest single event (°C)			Mean days <0 (°C)			Mean days <−5 (°C)		
Meteorological station (cell)	Obs	$A2_{IRH}$	Diff	$B2_{IRH}$	Diff	Obs	$A2_{IRH}$	$B2_{IRH}$	Obs	$A2_{IRH}$	$B2_{IRH}$	Obs	$A2_{IRH}$	$B2_{IRH}$	Obs	$A2_{IRH}$	$B2_{IRH}$
Aberdeen (4273)	1750	1782	32	1828	78	17	17	17	−16	−10	−8	55	53	45	7	2	2
Aberporth (5434)	2518	3711	1193	3748	1230	20	19	19	−10	−1	−1	21	0	0	1	0	0
Aldergrove (4797)	2014	2069	56	2127	114	18	20	22	−11	−8	−7	44	40	35	4	2	2
Auchencruive (4693)	1964	1677	−287	1719	−245	18	20	22	−13	−14	−11	47	72	63	6	21	13
Auchencruive (4694)	1964	1644	−320	1739	−225	18	19	22	−13	−12	−11	40	70	63	5	14	13
Bracknell (5757)	1965	2227	262	2268	303	20	26	26	−16	−9	−9	62	53	48	9	7	6
Carnwath (4589)	1044	1784	740	1845	801	18	19	21	−25	−12	−11	103	65	58	28	9	8
Cawood (5121)	1888	1921	33	1976	88	18	23	24	−15	−10	−10	53	62	56	6	7	6
East Malling (5759)	2182	2508	326	2551	369	19	25	26	−18	−8	−8	48	41	34	5	2	2
Eskdalemuir (4695)	1219	1626	407	1687	468	16	19	21	−19	−13	−13	89	72	64	17	13	10
Eskdalemuir (4801)	1219	1584	364	1663	444	16	20	25	−19	−13	−12	89	81	73	17	19	17
Everton (5862)	2473	2087	−386	2150	−323	19	26	28	−11	−15	−12	37	68	60	3	20	15
Lerwick (3639)	1715	2980	1264	3008	1292	14	15	15	−8	−3	−2	45	0	0	2	0	0
Mylnefield (4484)	1813	1804	−9	1856	43	18	20	19	−17	−9	−9	54	64	57	7	4	4
Rothamsted (5652)	1925	2237	312	2277	353	18	26	26	−17	−9	−10	57	53	47	7	6	5
Sutton Bonington (5333)	1967	2011	44	2058	92	18	24	25	−16	−10	−10	53	58	53	8	7	6
Wallingford (5650)	1904	2198	294	2239	334	19	26	26	−21	−9	−10	62	53	48	11	7	6
Mean	1854	2109	254	2161	307	18	21	23	−16	−10	−9	56	53	47	8	8	7

Table 4 Comparisons of modelled ($A2_{IRH}$ configuration) and observed maximum (T_{max}) and minimum (T_{min}) air temperature for means, standard deviation and probability of equal means ($P(t)$) using the paired Student's t test

	Maximum air temperature (°C)					Minimum air temperature (°C)				
	Mean		St Dev		P(t)	Mean		St Dev		P(t)
Meteorological station (cell)	$A2_{IRH}$	Obs	$A2_{IRH}$	Obs		$A2_{IRH}$	Obs	$A2_{IRH}$	Obs	
Aberdeen (4273)	10.38	11.08	4.11	4.22	0.010	4.95	4.83	3.63	4.83	0.659
Aberporth (5434)	12.26	12.12	3.04	3.99	0.558	10.31	6.29	2.92	3.53	0.000
Aldergrove (4797)	11.72	12.36	3.95	4.45	0.022	5.75	5.57	3.48	3.65	0.491
Auchincruive (4693)	11.18	9.71	4.42	3.54	0.000	4.56	4.50	3.88	2.98	0.808
Auchincruive (4694)	10.52	9.71	4.36	3.54	0.001	4.57	4.50	3.81	2.98	0.782
Bracknell (5757)	13.58	13.62	5.40	5.45	0.893	6.19	5.37	4.27	3.88	0.007
Carnwath (4589)	10.90	11.08	4.51	4.96	0.548	4.74	2.87	3.86	3.77	0.000
Cawood (5121)	12.42	12.51	4.99	5.14	0.811	5.34	5.03	4.11	3.66	0.291
East Malling (5759)	13.96	14.02	5.41	5.44	0.863	6.97	6.05	4.39	3.95	0.003
Eskdalemuir (4695)	9.09	9.13	4.17	4.46	0.887	4.52	2.86	3.85	3.10	0.000
Eskdalemuir (4801)	11.54	9.13	4.62	4.46	0.000	4.40	2.86	4.12	3.10	0.000
Everton (5862)	13.71	13.75	5.30	4.81	0.886	5.80	6.81	4.52	3.85	0.001
Lerwick (3639)	9.58	9.19	2.18	3.27	0.044	8.28	4.71	2.28	3.12	0.000
Lerwick (3640)	9.45	9.19	2.81	3.27	0.215	8.25	4.71	2.35	3.12	0.000
Mylnefield (4484)	11.39	11.46	4.62	4.50	0.788	5.01	4.94	3.94	3.74	0.764
Rothamsted (5652)	13.68	13.15	5.49	5.55	0.136	6.21	5.32	4.33	4.00	0.004
Sutton Bonington (5333)	12.69	13.12	5.09	5.19	0.201	5.59	5.43	4.13	3.80	0.607
Wallingford (5650)	11.79	11.08	5.01	4.22	0.005	6.11	4.83	4.21	3.69	0.000

with a linear increase in error up to the 25 mm size events (Fig. 3c: Auchincruive). Conversely for sites where the model under-estimates the mean annual total, i.e. Eskdalemuir cell 4801, there is a reduction in the probability of events occurring (Fig. 3a: Eskdalemuir), particularly in the range of 3–20 mm, with a linear increase in the model-observed difference (Fig. 3c: Eskdalemuir).

For the mean annual totals, the accuracy ranges from very good (i.e. Cawood, with −1 mm and 28 mm difference for the $A2c_{IRH}$ and $B2_{IRH}$ scenarios respectively), to very poor (i.e. Eskdalemuir cell 4801 with −854 and −836 and cell 4695 with −319 mm and −284 mm difference for the $A2c_{IRH}$ and $B2_{IRH}$ scenarios respectively). There is a general underestimation by the model at 10 of the 17 cells assessed (Table 1). Therefore, for sites where the model under-estimates the mean annual total, whilst also over-estimating the number of days on which precipitation events occur, the magnitude of each event (particularly in the range of about 2–30 mm) is too small. The distribution of erroneous modelled events is biased towards those of small magnitude. Considering the proportionality of differences to the annual total amount, the errors occurring at the larger events are smaller in relation to the importance of the errors for the much more frequent small to mid-range events.

The largest single observed event was at Aberdeen (109.2 mm) where the model estimated only 50 and 52 mm ($A2c_{IRH}$ and $B2_{IRH}$ respectively). The largest single event was generally underestimated across all sites, with an observed mean of 72 mm compared with 58 and 59 mm for the $A2c_{IRH}$ and $B2_{IRH}$ configurations respectively. The largest modelled single event was 97 mm at Everton ($B2_{IRH}$), where the observed was 56 mm. The model's ability to estimate the magnitude of the largest single event varied between sites, with a general under-estimation by about 20% of the observed, with 11 of the 17 sites having lower maximum single events than the observed. Only at Mylnefield ($A2c_{IRH}$), and Everton and Lerwick ($B2_{IRH}$) did the model overestimate the largest single precipitation event by more than 10 mm (Table 1).

The model was generally able to replicate well the patterns of mean weekly precipitation (Fig. 4), i.e. Aberporth, Rothamsted and to a lesser extent Auchincruive (cell 4693). No sites had substantially different patterns of mean weekly precipitation.

4.2 Air temperature

The annual cycle of mean daily air temperatures shows that the difference between the RCM estimates and the observations ranges from very low, e.g. Aberdeen and Mylnefield, to high, e.g. Aberporth (Fig. 5). Overall the model tends to estimate T_{max} well but to overestimate T_{min}, although this is not true of all sites. The diurnal range is too small (Fig. 6), particularly in the spring and summer. The main discrepancies in T_{max} are under-estimates in the autumn and over-estimates at the beginning of the year. At Aldergrove, however, the modelled T_{min} matches the observed values well, but T_{max} is under estimated, except in January and February.

Both $A2c_{IRH}$ and $B2_{IRH}$ tended to overestimate the highest T_{max} event at most locations, by an average of 5.5°C (excluding Aberporth and Lerwick), though at some, i.e. Aldergrove, the estimates were very close. For the lowest estimates of T_{max}, the model under-estimates by an average of 1.5°C, but does not manage to replicate the lower observed T_{max} values, i.e. at Carnwath (Table 2). The model underestimated the number of days when T_{max} was >15°C by an average of 13–14 days, but by as much as 36 days (Auchincruive: cell 4694).

The highest T_{min} events tend to be overestimated by an average of 4°C, but as high as 9°C, i.e. Eskdalemuir and Everton. For Aberdeen both the $A2c_{IRH}$ and $B2_{IRH}$ were exactly right (Table 3). However, for the lowest T_{min} events, neither configuration managed to represent the most extreme observed low values, being on average 6–7°C to high. From Figs. 5 and 7, the T_{min} often do not match those of the observed mean daily temperatures in the winter period. Conversely Everton (Fig. 5), over-estimates the $A2c_{IRH}$ the lowest T_{min} by 2°C and generally produces T_{min} data that is too low (cold) in the winter, which is surprising considering the cell contains approx. 30% sea coverage. The lowest observed T_{min} event of −25°C was at Carnwath, where the model only managed a −12°C estimate. The model underestimated the total number of days below 0°C in some locations and overestimated in others. Deviations ranged from −45 (Carnwath) to +31 days (Everton). A similar pattern is seen in the estimates of days below −5°C, with under- and over-estimates of −21 days (Carnwath) and +17 days (Everton).

Figure 7 uses two locations' results to illustrate that there are locations where the model is able to represent the temporal distribution and magnitude of highest and lowest values of T_{max} and T_{min} well. For Aberdeen there is a very close match between the observed and modelled lowest values of T_{max}. The model generally slightly under-estimates the highest values for the majority of the year, but with several larger over-estimates in early August. Also at Aberdeen, the model performs well for the highest values of T_{min} throughout the year and for the lowest values except during the winter period. At Bracknell the model over-estimates the highest values of T_{max} during the summer but under-estimates them in the early spring, whilst there is a very good match for the lowest T_{max} values (other than during early winter). For the highest T_{min} values, the model again over-estimates in the summer but shows a good match throughout the rest of the year. The modelled lowest T_{min}

values do not represent well the extreme observed lows at Bracknell and the spring and summer values are generally overestimated. These two locations represent better examples from the 17 assessed.

The match between observed and modelled mean thermal time accumulation (TTA; Fig. 8) varies considerably, i.e. Bracknell and Wallingford having only a −6 and −33°Cday difference in mean TTA on the last day of the year, respectively. Carnwath was overestimated by 257°Cday, and Auchencruive (cell 4694 – coastal) underestimated by −480°Cday. However, data from 13 of the 17 cells showed a close match with the observed TTA rate during the key growing season period. The largest error was at Lerwick (island location) which overestimated by 649°Cday.

The mean annual totals of T_{max} and $T_{\min}$, whilst not meaningful in terms of detail of model estimates, do provide a quantifiable indication of any substantial differences between the modelled and observed data. Annual total $T_{\max}$ is generally underestimated by a small amount (121°C for the $A2c_{IRH}$ and 92°C for the B2 $_{IRH}$ configurations), but the total $T_{\min}$ is overestimated by an average of 254°C for $A2c_{IRH}$ and 307°C for $B2_{IRH}$ (Tables 2 and 3).

The mean annual $T_{\max}$ and $T_{\min}$ (Table 4), as assessed by the paired Student's *t* test ($P(t)$) confirms that the model is better able to represent $T_{\max}$ than $T_{\min}$. Five locations have $P(t)$ values exceeding 0.80 for $T_{\max}$ compared with only one for $T_{\min}$. The higher $P(t)$ values for $T_{\max}$ are generally found at coastal locations (i.e. Everton) or those in lowland central England (i.e. Bracknell), but some nearby locations also had low $P(t)$ values (e.g. Wallingford). Locations that had higher $P(t)$ values for $T_{\max}$ tended to have very low ones for $T_{\min}$, with the opposite occurring when $P(t)$ values for $T_{\min}$ were high. Only Mylnefield had high $P(t)$ values (>0.78) for both $T_{\max}$ and $T_{\min}$.

Generally the temporal distribution of mean daily $T_{\max}$ and $T_{\min}$ is modelled adequately, based on the synchronisation of temporal distributions seen in Fig. 5, but with some exceptions, particularly for Aberporth and Lerwick. In both these cases the cells can be classified as 'sea cells' as they contain large areas of sea. The modelled data show the characteristics of a sea area rather than that of land, with a small range between $T_{\max}$ and $T_{\min}$. As such, it is not appropriate to make direct comparisons between land based observation station data and modelled grid cell data where the cell contains a certain percentage of sea. Further work is required to determine what that critical percentage of sea cover is.

Locations on the boundary between two cells could show contrasting results. For example Eskdalemuir (cells 4695 and 4801) had similar temperature results (not shown), but a marked difference in precipitation (Table 1). Hence care has to be taken in deciding which cells' data are most representative of sites on cell boundaries, i.e. through the use of pre-defined criteria, or multiple cell analysis.

4.3 Solar radiation

The HadRM3 model systematically over-estimates S_o (Fig. 9). However, the model does perform very well at some locations, such as Aberdeen and Aldergrove. At these locations the distribution of estimate errors is similar to that from data derived from specialist radiation models, i.e. the Donatelli–Bellocchi model (Donatelli and Bellocchi 2001; Rivington et al. 2005). The HadRM3 model estimates at these locations are in the order of only ±1 MJ m^{-2} day^{-1} larger than those from models like the Donatelli–Bellocchi model, but are much larger at other locations, i.e. Eskdalemuir (cell 4695) where the largest single error was 7.65 MJ m^{-2} day^{-1}. However, specialised models such as the Donatelli–Bellocchi make both over- and under-estimates, leading to compensating errors (i.e. when used in a crop model), whereas the HadRM3 estimates are consistently overestimated. The model over-estimates S_o particularly in the late summer to autumn period, when actual values are likely to be high, but there appears to be a characteristic shift towards either accurate or under-estimates in the spring to early summer period (i.e. Everton, Sutton Bonington, Wallingford).

4.4 Relationships between variables

The comparisons were not designed to estimate the quality of the correlation between daily variables, rather variables were assessed individually. However, the results produced do provide some evidence of the models' overall performance and relationships between variables. There does not appear to be a consistent pattern whereby if the model makes good estimates of one variable it makes equally good estimates of another. At no location does the model produce high quality estimates for all variables assessed. For example, the model estimates T_{max} and T_{min} very well at Mylnefield, but under-estimates precipitation and over-estimates S_o. Similarly at Cawood, the estimates of total annual precipitation are nearly exact with a close approximation of the largest single event, but under-estimates the number of dry days by 100, whilst T_{max} and T_{min} are estimated well (except T_{min} in the summer) but S_o is overestimated. At Aberdeen the model performs very well for T_{max}, T_{min} and S_o but under-estimates the total amount of precipitation and was particularly bad for producing too many days when precipitation occurs (Table 1).

5 Discussion

5.1 Implications for interpreting climate change projections

These results have implications for the interpretation of future projections of climate change and impacts studies. One of the aims of this work has been to identify differences between RCM estimates at the grid cell scale and site-specific observed data, and as such indicate that potential exists to correct biases. We recognise that the comparison is not a 'like with like' one, but it reflects the importance of being able to provide appropriate data for site-specific climate change impacts studies. Bridging the spatial gap between grid cell scale and specific locations within the cell in terms of data quality will present many challenges. Assessing RCM estimates of the past climate is an essential first step in order to evaluate the utility of future projections in CCI studies. This work has highlighted that the use of RCM data in CCI studies may only be appropriate when some form of bias correction has been conducted and when the specific site is similar to the mean topographical characteristics of the cell.

A fundamental issue with evaluating the quality of future projections by comparing estimates made at the grid cell scale of the past climate with site-specific observed data, is that of knowing whether errors existing in past climate estimates are maintained (or even propagated) into the future projections. Given the aim of identifying bias correction potential between the cell and specific sites within it, then if the assumption holds true that errors present in the hindcast estimates will also exist in the same approximate form in future projections, then using a standard approach of taking the differences between hindcast and future modelled data and applying those differences to observed data will result in the transfer of the errors to the adjusted observed data. Hence we argue that it is better to identify differences between modelled hindcast and site-specific observed data, then adjust the future projections based on those differences. For example, if there is a mean over-estimation of minimum temperature of 1.5°C in the hindcast data for a particular location and at a certain time of year, a refined form of estimate for future projections will be the modelled estimate minus 1.5°C. This method is independent of whether the source of the error is either *structural* (within the RCM), *representational* (the difference between the 50 km cell and the site attributes), or RCM input data or parameterisation. However, it does not take account of the dynamics affecting model response to greenhouse gas forcing. It simply provides a means by which empirically derived correction factors can be derived. However, if the assumption is unfounded, then adjustment of future projection data based on hindcast versus observed differences may not be appropriate.

On the basis that the assumption is correct (but not considering GHG forcing responses), when the bias based constraints of the modelled data have been identified, it becomes possible to determine where and when it is appropriate to use the future projection data in site-specific impacts studies. Based on our analysis, future projections for the RCM and scenarios tested, as currently published, of precipitation, extreme summer T_{max}, mean T_{min}, lowest T_{min} and S_{o} are potentially unreliable at some locations. Conversely the indication is that mean T_{max}, the lowest T_{max} and highest T_{min} estimates are reliable. These issues indicate that there is a need for more comparisons between RCM estimates and observed data, to identify the characteristics of combinations of weather variables and locations where RCMs perform poorly and to suggest corrections that can be applied.

5.2 Precipitation

The occurrence of too many modelled small precipitation events (<0.3 mm) may not be significant in terms of the overall soil water balance, as the amount of water added to the surface layer is very small but sufficient to give a surface 'wetting' effect. However, it is likely that they will adversely affect estimates of evapotranspiration due to cooling the soil surface and vegetation canopy temperatures. This in turn will affect soil water balances as it will be the wetted surface water that is evaporated rather than water drawn up from lower levels. Estimates of pest and pathogen responses will also be distorted by the inaccuracy of dry day estimates. The occurrence of such large numbers of small events indicates an issue with the model's handling of such events. One possible explanation is that errors were made acceptable during the original model validation process, when observed data was spatially aggregated within a cell, giving a 'drizzle effect'.

The fact that the model did not estimate the largest single events does not indicate a failure of the model, but that the thirty year coverage of the hindcast may not be sufficient to capture the more rare extreme events with longer return periods. In conjunction with this, the aim of the model is to represent the mean conditions for a grid cell, rather than specific extreme events recorded at individual stations. However, the consistency with which the model underestimated the largest single event across all sites does indicate a limitation.

In order to increase the utility of the data for impact studies, we recommend a reduction in the number of days on which precipitation occurs (i.e. removing estimates less than about 0.3 mm) whilst also increasing the magnitude of events >1 mm at locations where the model is shown to underestimate mean annual totals. The ability of the model to estimate the largest single precipitation event raises questions as to how useful the future projection data, in its original form, would be for use in flood risk assessment. However, the time slice for the hindcast period is 30 years, with the possibility that this is too short in order to capture the largest precipitation events. Therefore, the assessment should cover a longer hindcast versus observed period in order to be able to properly assess the ability of the model to estimate the largest event.

5.3 Temperature

The net result of the models' tendency to overestimate T_{min}, whilst performing well for T_{max}, is that projected data will be unsuitable (location dependent) for some CCI studies, as errors will be introduced to estimates of an entity's temperature response, i.e. due to thermal time accumulation, diurnal ranges, biophysical processes etc. In considering the daily variability of temperature, then mean values are not the best indicators of representation for accuracy. However, the results presented here for mean daily T_{max} and T_{min}, their highest and lowest values, indicate that the model does perform well in producing data that represents the natural temperature variability on a daily and seasonal basis. Thermal time accumulation (TTA) at some sites is very good, but it is possible to achieve the same rates of accumulation but with data of very different magnitudes, i.e. different values of T_{max} and T_{min} data can produce the same average value added to the previous day's accumulation. Hence modelled TTA rates derived from the HadRM3 hindcast data can be similar to observed TTA, but potentially for the wrong reasons. In some case, such as Carnwath, the differences between T_{max} and T_{min} (Fig. 6) produces substantially different rates of TTA (Fig. 8) from the observed, due to the overestimation of T_{min}. This effects interpretations of future plant and insect phenological responses due to TTA and correlations with the actual temperature.

5.4 Solar radiation

The overestimation of S_{o} at many locations suggest that the data are unsuitable for use in impacts studies where S_{o} is a key input. However, our personal experience has shown that data containing compensating errors of the type found in the S_{o} estimates from the HadRM3 model (i.e. at Aldergrove, Fig. 9) can still result in reasonable derived estimations, i.e. when used in a crop model. When the errors fluctuate between over- and under-estimates on a daily basis (i.e. Eskdalemuir cell 4659, Aldergrove, Fig. 9), the errors can cancel themselves out in terms of their impact on crop model estimates of yield. However, the temporal distribution of errors is critical, as over-estimation in the spring and summer will result in too high a rate of biomass accumulation (more intercepted radiation). The systematic over-estimation at many sites (i.e. Rothamsted) will produce substantial errors when used in CCI studies. Though not assessed in this study, the over-estimation of S_{o} indicates a flaw in the way the model represents cloud cover.

6 Conclusions

The types and magnitude of errors within the HadRM3 hindcast data presented here could introduce substantial errors when used within site-specific climate change impacts studies, i.e. using simulation models. Identification of errors in RCM estimates of the past climate makes it possible to assess the utility of future projection weather variable data for impacts studies. The hindcast data are, however, sufficiently similar to the observed data to raise the possibility of making simple adjustments to the modelled data. On the assumption that the same errors present in the hindcast data will also exist in the future projection data, such adjustments could then be applied to future projections to reduce systematic errors. This will improve the reliability of the climate change data and reduce the uncertainty introduced to impact studies.

The assessment of the RCM data demonstrates the importance of evaluating the quality of data prior to use within impact studies and for practitioners to be aware of how the data can introduce uncertainty. The suitability of uncorrected RCM data for climate change impact studies depends ultimately on how and for what purpose the data are used. The quality of the T_{max} and T_{min} data is sufficiently good for some locations to allow studies using monthly or weekly data to be made with confidence, but studies using daily data require caution, particularly where T_{min} and extreme cold temperatures are important factors. Precipitation data are less reliable, particularly in respect the lowest (<0.3 mm) and highest magnitude events and number of dry days, and are potentially unsuitable for CC impact studies in an uncorrected form. The reliability with which additional meteorological measures can be derived depends on which data are required. There is a risk of introducing significant errors where derived values, e.g. soil water deficit, are estimated from several weather variables as the potential exists for biases to occur with individual variables at the same time. The temperature data for island locations and some coastal sites are unsuitable for use in terrestrial impact studies (depending on the amount of sea cover within the cell). Where a location exists on the boundary between two cells, care needs to be taken in determining which cells' data best represents it. The choice of cell may depend on whether precipitation, temperature or solar radiation accuracy is more important.

This assessment of the quality of estimates made by the HadRM3 RCM for the historical period of 1960–90, has primarily taken the form of graphical comparisons. Whilst more detailed statistical assessments were possible, this exploratory analysis has yielded sufficient detail to show that the data will have an affect on the results of climate change impact studies. The important message is that the type of biases identified here need to be considered when climate model data is used in impact studies, particularly when simulation models are used. The characteristics of the data and how the biases manifest themselves may not be obvious, therefore there is a need to appraise the suitability of the data prior to use. Further analysis is also required to characterise the correlation between weather variables on a daily basis, to ensure that meteorological relationships are adequately represented, i.e. relationships between diurnal temperature ranges, solar radiation and cloud cover. It would also be informative to assess the behaviour of the model in terms of short-term variability, i.e. the continuation of patterns of weather from one day to the next.

Fundamentally, the results have shown that the HadRM3 RCM produces data that have both small and large spatially and temporally variable biases in its estimates of the past climate. Practitioners using RCM estimates for climate change impacts studies need to evaluate and quantify the biases in the data in order to determine what uncertainties they will introduce.

Acknowledgements The authors would like to thank the Scottish Executive Environmental and Rural Affairs Department for their funding support of this research. Thanks to Kevin Marsh at the BADC for processing the HadRM3 model data and to the Meteorological Office and Hadley Centre for permission to use their data.

References

Antic S, Laprise R, Denis B, de Elia R (2006) Testing the downscaling ability of a one-way nested regional climate model in regions of complex topography. Clim Dyn 26:305–325

Arnold SM, Monteith JL (1974) Plant development and mean temperature in a Teesdale habitat. J Ecol 62:711–720

BADC (2005) British Atmospheric Data Centre: http://badc.nerc.ac.uk/home/index.html

Bell JL, Sloan LC, Snyder MA (2004) Regional changes in extreme climatic events: a future climate scenario. J Clim 17:81–87

Creed IF, Band LE, Foster NW, Morrison IK, Nicolson JA, Semkin RS, Jeffries DS (1996) Regulation of nitrate-N release from temperate forests: a test of the N flushing hypothesis. Water Resour Res 32:3337–3354

Donatelli M, Bellocchi G (2001) Estimates of daily global solar radiation: new developments in the software RadEst3.00. Proceedings of the 2nd International Symposium on Modelling Cropping Systems, Florence, Italy. 16–18 July 2001. Inst. for Biometeorology, CNR, Florence, Italy, pp 213–214

Droogers P, Aerts J (2005) Adaptation strategies to climate change and climate variability: a comparative study between seven contrasting river basins. Phys Chem Earth 30:339–346

Evans JP, Oglesby RJ, Lapenta WM (2005) Time series analysis of regional climate model performance. J Geophys Res Atmos 110(D4):D04104

Fowler HJ, Ekstrom M, Kilsby CG, Jones PD (2005) New estimates of future changes in extreme rainfall across the UK using regional climate model integrations. 1. Assessment of control climate. J Hydrol 300:212–233

Frei C, Christensen JH, Déqué M, Jacob D, Jones RG, Vidale PL (2002) Daily precipitation statistics in regional climate models: evaluation and intercomparison for the European Alps. J Geophys Res 108:4124–4142

Holman IP, Rounsevell MDA, Shackley S, Harrison PA, Nicholls RJ, Berry PM, Audsley E (2005) A regional, multi-sectoral and integrated assessment of the impacts of climate and socio-economic change in the UK. Clim Change 71:9–41

Hoogenboom G (2000) Contribution of agro-meteorology to the simulation of crop production and its applications. Agric For Meteorol 103:137–157

Hulme M, Jenkins GJ, Lu X, Turnpenny JR, Mitchell TD, Jones RG, Lowe J, Murphy JM, Hassell D, Boorman P, McDonald R, Hill S (2002) Climate change scenarios for the United Kingdom: The UKCIP02 Scientific Report. Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK

Huntingford C, Jones RG, Prudhomme C, Lamb R, Gash JHC, Jones DA (2003) Regional climate-model predictions of extreme rainfall for a changing climate. Q J R Meteorol Soc 129:1607–1621

IPCC (2000) Inter-governmental panel on climate change, special report on emissions scenarios. Cambridge University Press, UK

Jarvis CH, Baker RHA, Morgan D (2003) The impact of interpolated daily temperature data on landscape-wide predictions of invertebrate pest phenology. Agric Ecosyst Environ 94:169–181

Jenkins G, Lowe J (2003) Handling uncertainties in the UKCIP02 scenarios of climate change. Hadley Centre technical note 44, 20th November 2003. Meteorological Office, Exeter, UK

Jones CG, Willen U, Ullerstig A, Hansson U (2004) The Rossby Centre regional atmospheric climate model part 1: model climatology and performance for the present climate over Europe. Ambio 33:199–210

Kim J, Lee J-U (2003) A multiyear regional climate hindcast for the Western United States using the mesoscale atmospheric simulation model. J Hydrometeorol 4:878–890

LADSS (2007) Land Allocation Decision Support System website. http://www.macaulay.ac.uk/LADSS/reference.shtml

Mearns LO, Giorgi F, McDaniel L, Shields C (1995) Analysis of daily precipitation in a nested regional climate model: comparisons with observations and doubled CO_2 results. Glob Planet Change 10:55–78

Moberg A, Jones PD (2004) Regional climate model simulations of daily maximum and minimum near-surface temperatures across Europe compared with observed station data 1961–1990. Clim Dyn 23:695–715

Murphy JM, Sexton DMH, Barnett DM, Jones GS, Webb MJ, Collins M, Stainforth DA (2004) Quantification of modelling uncertainties in a large ensemble of climate change simulations. Nature 430:768–772

Nijkamp P, Wang SL, Kremers H (2005) Modeling the impacts of international climate change policies in a CGE context: the use of the GTAP-E model. Econ Model 22:955–974

Nonhebel S (1994a) The effects of use of average instead of daily weather data in crop growth simulation models. Agric Syst 44:377–396

Nonhebel S (1994b) Inaccuracies in weather data and their effects on crop growth simulation results. I. Potential production. Clim Res 4:47–60

Peng PT, Kumar A, van den Dool H, Barnston AG (2002) An analysis of multimodel ensemble predictions for seasonal climate anomalies. J Geophys Res Atmos 107(D23):4710

Rivington M, Bellocchi G, Matthews KB, Buchan K (2005) Evaluation of three model estimates of solar radiation at 24 UK stations. Agric For Meteorol 132:228–243

Rivington M, Matthews KB, Bellocchi G, Buchan K (2006) Evaluating uncertainty introduced to process-based simulation model estimates by alternative sources of meteorological data. Agric Syst 88:451–471

Rivington M, Matthews KB, Bellocchi G, Buchan K, Stöckle CO, Donatelli M (2007) An Integrated Assessment approach to conduct analyses of climate change impacts on whole-farm systems. Environ Mod Software 22:202–210

Semenov MA (2002) LARS-WG: A Stochastic Weather Generator for Use in Climate Impact Studies. http://www.rothamsted.bbsrc.ac.uk/mas-models/download/LARS-WG-Manual.pdf

Soulsby C (1995) Contrasts in storm event hydrochemistry in an acidic afforested catchment in upland Wales. J Hydrol 170:159–179

Sperow M, Eve M, Paustian K (2003) Potential soil C sequestration on US agricultural soils. Clim Change 57:319–339

Stöckle CO, Campbell GS, Nelson R (1999) ClimGen manual. Biological Systems Engineering Department, Washington State University, Pullman, WA. USA. http://c100.bsyse.wsu.edu/climgen/

Vidale PL, Luthi D, Frei C, Seneviratne SI, Schar C (2003) Predictability and uncertainty in a regional climate model. J Geophys Res 108:ACL12.1–ACL12.23

Weibull W (1961) Fatigue testing and analysis of results. Pergamon, Oxford, UK

Zhang XC (2005) Spatial downscaling of global climate model output for site-specific assessment of crop production and soil erosion. Agric For Meteorol 135:215–229