Chapter 2

Literature Review

Contents

2.1. Introduction

Some of the difficulties that are experienced by climate modelers and climate impacts groups in the generation, evaluation and application of daily rainfall have been described in Chapter 1. In this literature review these issues are explored in greater detail with reference to the existing literature, and the gaps and shortfalls in the current literature which might be addressed in order to attempt to overcome these problems.

Initially (**Section 2.2**), the physical theory and observational and model evidence for changes in spatial and temporal precipitation variability under a warmer climate are examined, with reference to the changes in statistical properties of daily precipitation series that are expected.

Section 2.3 addresses the problem of evaluating the climate model simulations of rainfall with respect to daily variability. Previous model evaluation efforts are reviewed, and literature on relationships between point and areal rainfall from other hydrological applications is drawn on and considered for its potential use in this context.

Section 2.4 considers scenarios of future climate and the approaches to the disaggregation and downscaling of areal precipitation projections for climate impacts applications. Particular attention is paid to the ability of these approaches to represent changes in both temporal and spatial precipitation variability.

Finally, **Section 2.5** summarises the findings of this review, and identifies the research objectives that have arisen to be addressed in this thesis.

2.1.1. A note on convective and synoptic precipitation events

Throughout this thesis, precipitation events are discussed in terms of ‘convective’ and ‘synoptic’ events. This distinction is made in order to distinguish between precipitating systems that are caused by local heating and convection, which result in heavy but localised precipitation (the ‘convective’ events); and those that are caused by larger-scale air mass uplift, such as frontal activity or orographic influence, and result in precipitation events that affect a much larger region (the ‘synoptic’ events).

This type of approach is, of course, a considerable simplification of the spatial characteristics of precipitation. Whilst the two processes are distinct in model simulations, in reality, convective and synoptic activity are not separate processes. For example, convection often occurs along strong cold fronts. However, this distinction is convenient in studies where the spatial scale of events is of interest, and also because model simulated precipitation is often available in its separate convective and synoptic components.

2.2. Daily Precipitation Variability in a Warmer Climate

2.2.1. Theoretical Changes to the Hydrological Cycle and their Influence on Precipitation Variability

One implication of a warmer, more dynamic, climate system is likely to be the intensification of the hydrological cycle (Fowler and Hennessy, 1995, Trenberth, 1998, 1999). The water-holding capacity of air approximately doubles with each 10°C increase in temperature, in accordance with the Clausius-Clapeyron relationship (Fowler and Hennessy, 1995). Warmer lower-atmosphere and oceans are expected to bring about enhanced evaporation rates, and therefore a global increase in atmospheric moisture content (Fowler and Hennessy, 1995; Trenberth 1998, 1999) (see Figure 2-1).

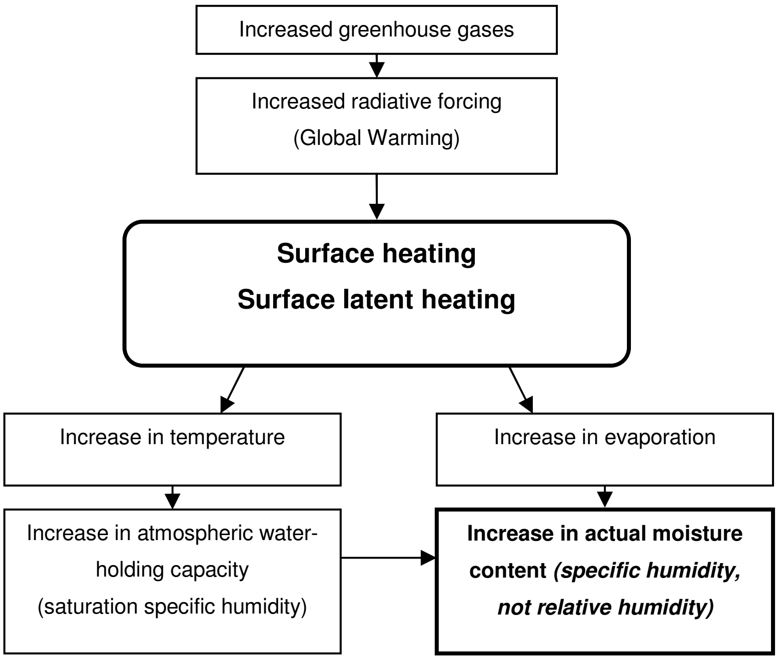

Figure 2-1: Schematic outline of the sequence of processes involved in climate change and moisture content of the atmosphere (after Trenberth, 1998).

In order to balance the enhanced evaporation, the global precipitation amount is expected also to increase (Trenberth, 1999). However, the increase in atmospheric moisture content (specific humidity, rather than relative humidity) means that that this expected increase in total precipitation is expected to occur predominantly as increases in intensity, rather than frequency, of precipitation events (Trenberth, 1999). Furthermore, in precipitating systems, latent heat release further increases the intensity of a precipitating system and the convergence of moisture into that system, resulting in a positive feedback (Trenberth, 1999) (Figure 2-2). This means that it is the heaviest events which are most affected by increases in intensity, compared to moderate and light events (Trenberth, 1998, 1999).

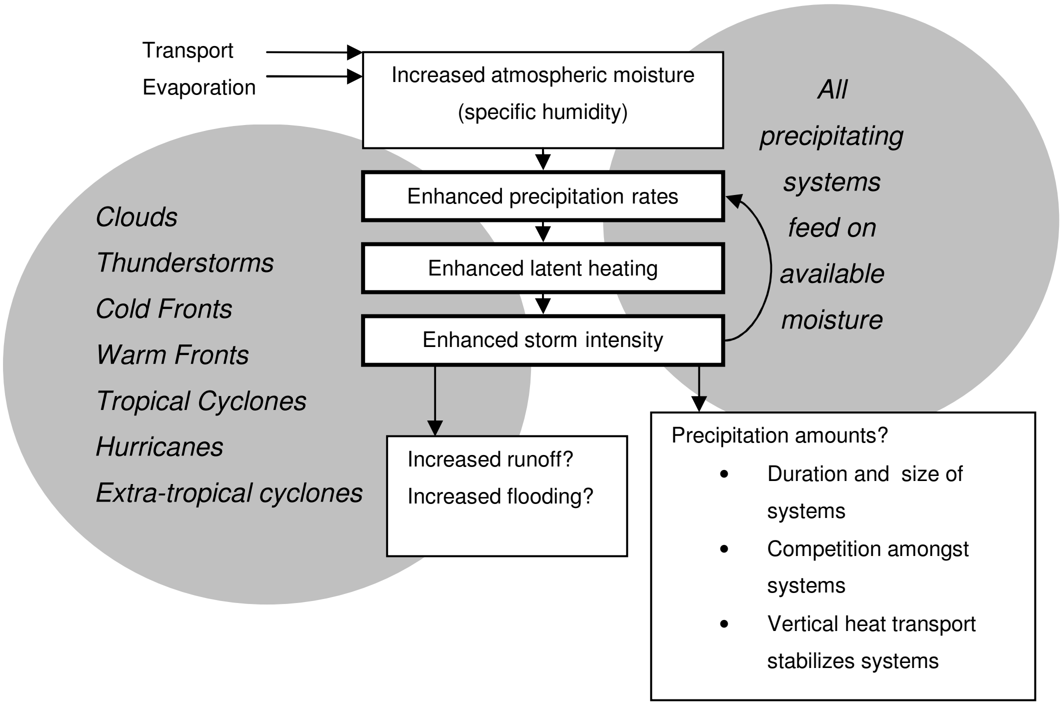

Figure 2-2: Schematic outline of sequence of processes involved in climate change and how the increasing moisture content of the atmosphere alters rainfall rates (after Trenberth, 1998, 1999).

This expected increase in total rainfall and rainfall intensity is expected to bring about overall global changes, but the effects felt at regional and local scales may not always reflect this global trend. Whilst the thermo-dynamic influences on rainfall described above are expected to affect rainfall fairly uniformly over the world, changes in atmospheric motion that are likely to arise as a secondary influence on rainfall amount and intensity of a warmer climate are likely to cause localized changes to precipitation regimes (Emori and Brown, 2005).

In addition to these overall changes in precipitation rate, the changes to the precipitation processes also indicate a 'shift' in relative amounts of precipitation which can be attributed to different mechanisms – more specifically, a shift towards more local-scale convective rainfall and less non-convective (synoptic), large scale, rainfall. Gordon *et al*. (1992) suggest that a warmer, moister atmosphere causes greater vertical instability which is conducive to stronger and more frequent convection. This explanation is expanded by Hennessy *et al*. (1997), who argue that cooling in the upper troposphere - combined with surface warming, increases in the vertical temperature gradient and increased cloud height (explored by Mitchell and Ingram, 1992) - will cause greater instability and thus more frequent, and deeper, convective activity. Subsidence associated with convection generally dries the boundary layer and troposphere, reducing the frequency of super-saturation and thus reducing the occurrence/amount of synoptic precipitation (Mitchell and Ingram, 1992; Hennessy *et al*., 1997).

These changes in rainfall characteristics suggest that important changes to both the temporal and spatial variability will occur under a warmer climate. Whilst the potential changes in temporal variability of rainfall that arise from these disproportionate increases in the heaviest precipitation events have been the subject of a vast body of observational and model-based research in recent years (explored below), changes in the degree of local spatial variability, which might be expected to accompany a shift towards more local convective rainfall, have received considerably less attention.

The importance of the potential changes in the spatial characteristics of rainfall, however, is a significant factor to consider when relating temporal variability from grid-box average rainfall to the temporal variability at points within that box. If the spatial extent of rainfall events within a grid box decreases, then the mean intensity experienced at points within that grid box will increase, because the same volume of rainfall is distributed over a smaller fraction of the grid box. This relationship between spatial and temporal variability is demonstrated by Osborn (1997), where proportional changes in convective/non-convective precipitation from model simulations were used to investigate the intensity of rainfall that would be experienced at points within a grid box, based on arbitrarily assigned values of fractional grid-box coverage for convective and synoptic events (0.4 and 1 respectively). The mean intensities at point scale, when different fractional coverage was applied, were demonstrated to differ significantly, even showing changes of opposite sign.

The inter-dependence of spatial and temporal variability means that the expected increases in intensity and temporal variability in precipitation may be enhanced further when looking at point or local scales than are indicated by the grid-box average, if the spatial characteristics of rainfall change. The following examination of model and observational evidence for changes in precipitation variability will therefore address both temporal and spatial variability.

2.2.2. Changes to Precipitation Variability in Model Simulations

2.2.2.1. Changes in Precipitation Amount, Intensity and Variability

GCM simulations of climate under enhanced atmospheric greenhouse gas concentrations are broadly consistent in indicating global trends of increased precipitation, which occur due to increases in intensity, rather than the frequency, of rainfall events (Noda and Tokioka, 1989; Gordon *et al.*, 1992; Fowler and Hennessy, 1995; Hennessy *et al.*, 1997; Gregory *et al.*, 1997; McGuffie *et al.*, 1999; Kharin and Zwiers, 2000; Semenov and Bengtsson, 2002; Watterson and Dix , 2003, Tebaldi *et al.*, 2006; Barnett *et al.*, 2006).

It is evident from model projections that the increases in intensity of heavy events are considerably greater in magnitude and spatial coherency than the changes in total precipitation (Kharin and Zwiers, 2000; Kharin and Zwiers, 2005, Emori and Brown, 2005). Experiments with CGCM1 (Kharin and Zwiers, 2000), for example, show increases in extreme precipitation everywhere in the world, by about 8% by 2040-60 and 14% by 2080-2100 compared to increases of 1% and 4% in mean annual precipitation. Emori and Brown (2005) find global mean changes of mean and extreme rainfall alter by 6.0% and 13.0% respectively between the periods of 1981-2000 and 2081-2100, based on an ensemble of six models. Even for some regions where mean precipitation decreases, increases in the heaviest events are still indicated (Gordon *et al.*, 1992; Christensen and Christensen, 2004; Kharin and Zwiers, 2005; Tebaldi *et al.*, 2006). Changes in extreme precipitation (events with a 20-year return period) compared to changes in mean annual rainfall are shown for one simulation using CGCM2 in Figure 2-3 (Kharin and Zwiers, 2005).

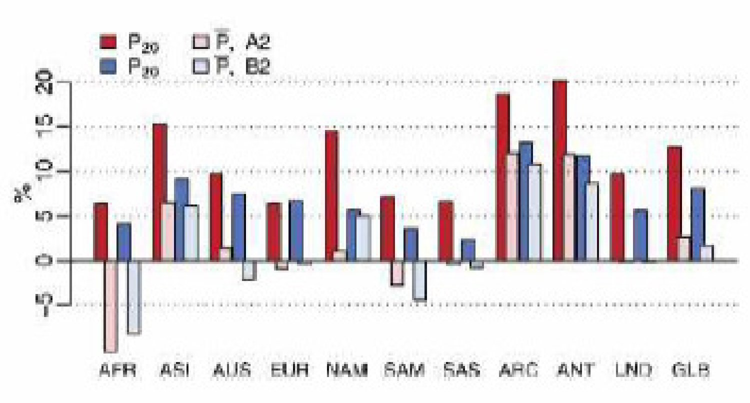

Figure 2-3: Regionally averaged changes in 20-yr return period values of extreme precipitation, and corresponding changes in annual means expressed by percentage, simulated by CGCM2 in 2090 compared to 2000. Changes in extremes are indicated by darker colours, changes in means in lighter colours (Kharin and Zwiers, 2005).

Whilst the increases in extreme rainfall are predominantly caused by thermodynamic mechanisms (i.e. the change in atmospheric moisture content) which affect the globe relatively uniformly, the global increase in mean precipitation is not distributed evenly over the globe because the mean precipitation amount and distribution is influenced more heavily by the secondary influences of changes in atmospheric circulation (Emori and Brown, 2005). This means that some regions of the world experience substantial increases in mean precipitation in model simulations (the mid-to-high latitudes and equatorial regions), while others experience drying (the subtropics) (Figure 2-4).

Regional Climate Model (RCM) experiments generally show similar changes to the temporal variability of daily precipitation as the global GCM experiments, with overall increases in intensity and/or extreme daily rainfall found for the American mid-west (Kothavala, 1997), Europe (Frei *et al.*, 1998; Arpe and Roeckner, 1999; Raisanen and Joelsson, 2001; Raisanen *et al.*, 2004, Christensen and Christensen, 2004; Giorgi *et al.*, 2004; Semmler and Jacob, 2004, Frei *et al.*, 2006, Gao *et al.*, 2006); UK (Jones and Reid, 2001; Ekström *et al.*, 2005); India (Kumar *et al.*, 2006); Korea (Boo *et al.*, 2006) and the USA (Wilby and Wigley, 2002).

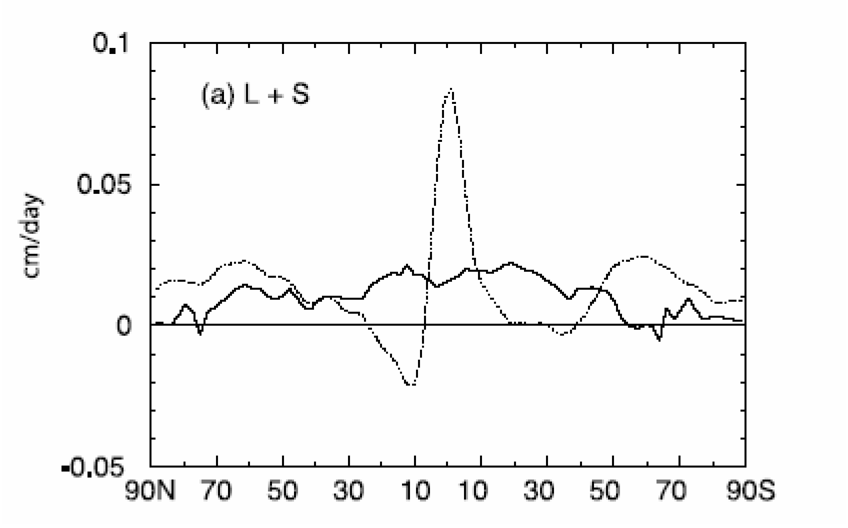

Figure 2-4: Latitudinal profiles of the simulated zonal mean changes in the annual mean rates of precipitation (dotted line) and evaporation (solid line) by 2035-2065 under scenario IS92a using a coupled GCM (Wetherald and Manabe, 2002).

Changes in rainfall intensity, variability and extremes are expressed using a number of different statistical parameters in different studies. Meehl *et al.* (2005) use the simple daily intensity index (the mean rainfall amount on days when more than 1mm falls) to identify changes in mean intensity over the world in future climate, whilst Barnett *et al.* (2006) use fixed value thresholds, determined by the 99th percentile value for current climate, and calculate changes in the frequency in which the threshold is exceeded in future climate. Changes in rainfall characteristics have been expressed in several of the experiments by fitting a gamma distribution to the wet-day rainfall totals (Semenov and Bengtsson, 2002; Wilby and Wigley, 2002; Watterson and Dix, 2003). The changes in the distribution of daily precipitation are characterised by an increase in the scale parameter, essentially a 'stretching' of the distribution, which embodies the increases in variance and increases in magnitude of the heaviest events.

Changes in the rarer, or 'more extreme' events are often expressed by the probability by which a daily rainfall amount will be exceeded (e.g. Kharin and Zwiers, 2005; Semmler and Jacob, 2004). Kharin and Zwiers (2005) use the Generalised Extreme Value (GEV) distribution to estimate daily rainfall amounts with return values of 10, 20 and 50 years in precipitation simulated under the A2 scenario. The study finds that the GEV distribution for projected future climate of 2090 experiences a 'stretch' towards higher daily rainfall values, such that events that only occur every 20 years in current climate occur twice as often by the end of the 21st century.

A range of 'indicators' of extreme values introduced by Frich *et al.* (2002) have been used in more recent studies of climatic extremes and variability in an attempt to unify studies and allow simpler comparisons between studies. The precipitation-based indices of extremes and variability include the simple daily intensity (total annual precipitation amount divided by the total number of wet days, >1mm, in a year), number of days per year with greater than 10mm rainfall, maximum number of consecutive dry days and the fraction of total rainfall which occurs in events exceeding the 95th percentile of the distribution for wet day amounts. Studies which have used some or all of these indicators include Tebaldi *et al.* (2006), Meehl *et al.* (2005) and Kiktev *et al.* (2003).

2.2.2.2. Changes to Precipitation Type and Spatial Variability

Several of the studies have also considered changes in the proportions of convective and synoptic components of total precipitation (Noda and Tokioka, 1989; Gordon *et al.*, 1992, Hennessy *et al.*, 1997; Chen *et al.*, 2005). Whilst the general changes in mean intensity and extremes appear to affect all areas of the globe, a shift towards a higher proportion of convective precipitation is more seasonally and regionally dependent. Hennessy *et al.* (1997) suggest that the mid-to-low latitudes are affected by such changes whilst in the high latitudes, precipitation of both types simply gets more intense. Gordon *et al.* (1992), however, found widespread increases in penetrating convective rainfall and decreases in large-scale non-convective rainfall in all but the high latitudes. Chen *et al.* (2005) find increases in the proportion of convective rainfall year-round in RCM simulations over North America under a climate with doubled CO_2 (Figure 2-5).

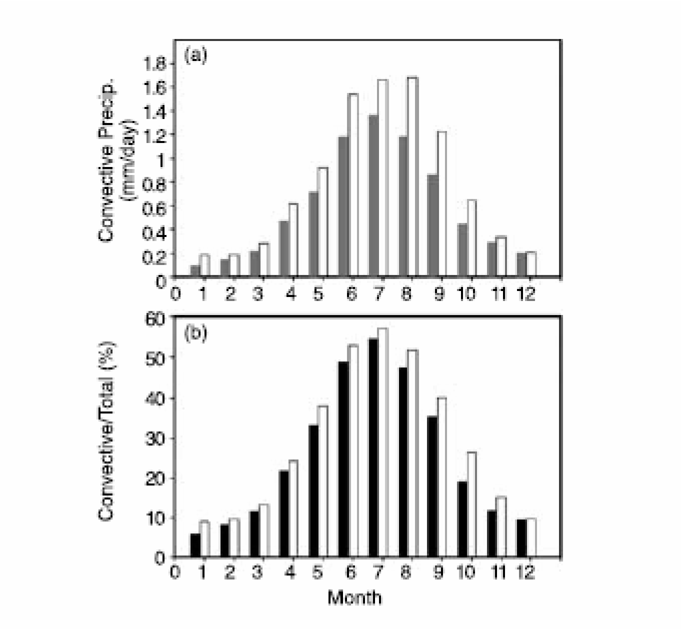

Figure 2-5: Annual cycle of (a) monthly mean convective precipitation and (b) the ratio of convective to total precipitation (%) over North America simulated using RCM MM5. Solid bars are for present-day climate (1990-1999) and non-filled bars are for future climate (2090-2099) under 2xCO_2 (Chen *et al.*, 2005)

Given this shift towards a higher proportion of convective precipitation, it might be expected that future rainfall will be more spatially variable. Determining fine-scale spatial variability from model simulations is difficult because even at the higher spatial resolutions that RCMs afford, the finite resolution limits the representation of small scale processes.

A study by Booij (2002b) attempts to quantify a change in spatial correlation for a number of GCMs and suggests an overall *increase* of 30-40 percent in correlation decay length (i.e. that overall correlation is higher between two stations at any given separation distance) in the GCMs investigated; the opposite of what might be expected. However, the spatial correlation between GCM grid boxes represents only very large separation distances between points, and is not necessarily indicative of spatial correlation at smaller spatial scales.

2.2.3. Observed Changes to Daily Precipitation Characteristics

Given the theoretical basis for an intensified hydrological cycle, it might be expected that observed records of precipitation will show evidence of changes related to the degree of warming which has already been experienced. The IPCC AR4 report includes a thorough review of the changes in climate seen in observations of recent climate (Trenberth *et al.*, *in press*). The report found strong evidence for overall global increases in tropospheric water vapour e.g Dai, 2006a). Dai's (2006a) global study of humidity from stations, ships and buoys found statistically significant increasing trends in global and Northern Hemisphere average specific humidity since 1970, while trends in relative humidity were small. These trends in humidity were found to correspond to the observed changes in temperature in a quantitatively similar way to that which would be expected due to the Clausius-Clapeyron relationship.

The physical theory and model evidence explored in Section 2.2.1 suggests that increase in the water-holding capacity of the atmosphere, determined via the Clausius-Clapeyron relationship, will affect precipitation amount through increases in rainfall intensity, rather than increases in frequency, and that this increase will be amplified in the intensity of the heaviest events (Gordon *et al.*, 1992; Hennessy *et al.*, 1997; Trenberth, 1999). It is therefore reasonable to expect that the climate signal in observed data will be stronger, and more spatially coherent, in the changes in variability and extremes of rainfall than in the changes in rainfall amount.

The detection of trends in variability and extremes in observed climate data, however, is limited by the requirement for long and homogenous observed datasets which are needed to identify shifts in frequencies of rare extreme events (Frei and Schär, 2001; Klein Tank and Konnen, 2003). Such data are often limited to developed and well populated regions such that a bias in global coverage causes uneven global representation, making it difficult to detect global trends. Empirical studies of changes in variability and extremes for regions where data are available are, however, numerous but diverse with respect to the region and size of region studied, data quality and statistical approach taken. Variations in the approach to the problem of detecting changes in extremes and variability include the definition of an 'extreme' or 'heavy' event; which might be, for example, the 10%, 5%, 1% most intense events; exceedence frequency of a set threshold; threshold value based on a return period; or the distribution parameters of annual maxima. Changes might be assessed as a trend over a continuous period or a difference between two time slices; applied to hourly, daily or multi-day accumulations; and over records of varying length, spatial coverage and quality. This can make quantitative comparisons between studies, and the formulation of general conclusions about trends in variability and extremes difficult, as different analysis techniques can yield different conclusions (Zhang *et al.*, 2004).

2.2.3.1. Observed Changes in Precipitation Amount and Intensity

The IPCC Third and Fourth Assessment Reports (Folland *et al.*, 2001; Trenberth *et al.*, *in press*) have both concluded that whilst studies of changes in total rainfall show substantial variation in sign and magnitude for different regions of the world and periods in history, the evidence for a global increasing trend in precipitation variability and extremes is more compelling. Studies of the extremes and variability in observed precipitation have generally been more statistically significant and spatially coherent than those of total rainfall, with increasing trends in extremes and variability occurring even in regions where total rainfall decreases.

Attempts to globally consolidate observational trends in variability and extremes include those by Groisman *et al.* (1999, 2005), Frich *et al.* (2002) and Alexander *et al.* (2006), which have applied statistical analyses to global datasets. Groisman *et al.* (1999) looked at the gamma distribution of daily precipitation over 8 countries (covering 80% of the extra-tropical land area), finding changes in the scale parameter which indicate that the increase in total precipitation has resulted in a disproportionate increase in the frequency of heavy (upper fifth and tenth percentiles) daily events. Frich *et al.* (2002) have also conducted a global study, looking at both regional variation and global trends. Rather than considering the distribution of values, a broad range of ‘indicators’ of extremes, including mean intensity, fraction of events exceeding baseline 95th percentile, and fixed 10mm threshold exceedence, were applied. Trends in simple daily intensity were mixed, such that the global increase was not found to be statistically significant due to local variability, changes in frequency of events over 10mm, maximum 5-day total and the fraction of total precipitation exceeding the 95th percentile all showed more significant and coherent increasing trends. Alexander *et al.* (2006) used gridded indices of precipitation extremes for 1953-2003 to identify widespread increases in those indices, although not all changes were statistically significant.

Regional studies also showing statistically significant trends in all or some indices of precipitation variability and extremes, for at least some seasons and regions, include United States (Karl and Knight, 1998; Trenberth, 1998; Kunkel *et al.*, 1999); parts of Canada (Stone *et al.*, 2000); Switzerland (Frei and Schär, 2001); winter in the UK (Osborn *et al.*, 2000; Fowler and Kilsby, 2003); Europe (Moberg *et al.*, 2006); India (Sen Roy and Balling, 2004; Goswami *et al.*, 2006); China (Endo *et al.*, 2005); Italy (Brunetti *et al.*, 2004), South Africa (Fauchereau *et al.*, 2003; New *et al*, 2006), Central and South America (Liebmann *et al.*, 2004; Aguilar *et al.*, 2005; Haylock *et al.*, 2006).

Whilst the weight of evidence supports the physical theory that precipitation variability and extremes will increase as the atmosphere warms, there remain a number of areas for which results are mixed or trends are too weak to be conclusive. Canada is one such example. Whilst Stone *et al.* (2000) find that in some regions and seasons, heavy events are significantly more frequent, Zhang *et al.* (2001) and Kunkel (2003) find no discernable trend. Kunkel *et al.* (2003) find that the seemingly unambiguous trends found for extremes in the USA in earlier studies are less significant when a longer dataset is examined, and precipitation extremes are found to be similarly high at the beginning as at the end of the 20th century. For Japan, Iwashima and Yamamoto (1993) and Yamamoto and Sakurai (1999) find increases in extremes which are large relative to the increase in mean, but Easterling *et al.* (2000) suggest that whilst Southern Japan has experienced decreases in mean, and Northern Japan has experienced increases in total precipitation, both regions have experienced decreases in events above the 90th percentile. A number of other regional studies which have found mixed, weak or spatially incoherent trends include Central and Southern Asia (Klein Tank *et al.*, 2006) and Africa (Kruger, 2006).

Some studies have also demonstrated trends opposite to those expected. Easterling *et al.* (2000) found that in areas such as Ethiopia, western Kenya and Thailand, overall precipitation decreases were accompanied by amplified decreases in heavy events. Studies of the Iberian Peninsula (Gallego *et al.*, 2006; Garcia, 2007) find a general shift towards lighter precipitation, with some regions showing no trend and some with a negative trend in extreme precipitation events.

2.2.3.2. Observed Changes to Type and Spatial Variability of Precipitation

Whilst the numerous studies detailed above provide information regarding changes in the temporal variability at gauging stations, the detection of changes in spatial variability, or rainfall type, has not received the same attention.

Detecting change in the type (i.e. convective or synoptic) rainfall is not a straightforward task because the two types cannot easily be separated by their mechanism in the same way that modeled precipitation can be. Some studies have separated satellite observed precipitation into convective and synoptic (or 'stratiform') components (e.g Tremblay, 2005 and Dai, 2006b), but satellite records are currently of insufficient length to identify long-term trends in the contribution of each rainfall type. Studies of the spatial correlation of rainfall in stations data would require very dense, long and high quality daily precipitation data to identify any significant trends that might exist.

2.2.4. Summary

Climate model simulations indicate that an intensification of the hydrological cycle is expected to occur under warmer climate conditions and cause increases in precipitation intensity, particularly in heavy events. This is supported by observational evidence for many regions of the world, which have shown greater intensity and daily variability in rain-gauge records.

Trends in the variability, intensity or extreme values of precipitation have, on the whole, been more significant and coherent than the changes in precipitation amount in recent studies of observed precipitation records. Whilst studies of total rainfall have shown substantial geographical and temporal variation, the changes in extremes have generally tended to be of larger magnitude, more statistically significant and more spatially coherent. There remain, however, exceptions to this.

Models have also suggested that the mid-to-low latitudes may experience a shift towards a higher proportion of local convective precipitation, although it is more difficult to find sources of observed data against which to verify this. The possibility of such a change towards more localised rainfall suggests that the assumption that spatial variability of rainfall will be unchanged in the future climate may not be valid. For regional/local impact assessment, this means that the increases in temporal variability simulated at coarse grid-scale may be enhanced at sub-grid-scale.

2.3. Evaluating Characteristics of Daily Precipitation in Climate Simulations

Assessing the ability of climate models to accurately reproduce characteristics of present climate is an important stage in their continued development and improvement, and in assessing the uncertainty which surrounds the projections of future climate that are applied in climate impact assessment.

Evaluation of precipitation simulated by GCM has largely, though not exclusively, focused on the mean values of precipitation, and the geographical and seasonal patterns in these amounts (e.g. Srinivasan *et al.* 1995; Gates *et al.*, 1999; Dai *et al.*, 2001; Delworth *et al.*, 2002; Covey *et al.*, 2003, Rasch *et al.*, 2006) (Dai, 2006b). Relatively few studies have attempted the evaluation of the spatial and temporal variability of simulated daily rainfall, such as the frequency, intensity, extremes and areal extent of events. This is largely due to the mismatch between the levels of temporal variability found in point-scale station observations and the grid-scale model output, which means that variability and extremes in climate simulations of daily precipitation are difficult to evaluate quantitatively. The discrepancy is eased slightly for high-resolution Regional Climate Models (RCMs) as the smaller grid boxes are more similar to the point values, but the problem still remains to a lesser extent.

For regions where a dense station network is available, gridded station data can provide a source of areal observations which can be assumed to be representative of the ‘true’ areal mean, and therefore provide a reliable benchmark against which to evaluate model-simulated precipitation. However, gridded datasets suffer from temporal and spatial variations in station coverage (Hulme and New, 1997) not only causing biases in the mean values, but also inconsistencies in the levels of variability, which can result in misleading results if they are used for model evaluation.

Remotely-sensed precipitation data from satellites and radar are becoming increasingly more available, and projects such as CMAP (Xie and Arkin, 1997) and GPCP (Adler *et al.*, 2003) provide a source of areal-average precipitation observations which merge remotely-sensed and station data to maximize the benefits of ground-based station observations and the indirect precipitation measurements from satellites or radar (New *et al.*, 2001). However, satellite and radar observations of precipitations are indirect and can suffer from errors, particularly in regions where the station networks on which they are calibrated are sparse (e.g. Matsuyama *et al.*, 2002, Gebremichael *et al.*, 2003). Precipitation data from satellites can also vary with region depending on the particular source of satellite data used; the GPCP daily dataset is more reliable at lower latitudes where it is based on geostationary satellite and microwave data, while data at higher latitudes, TIROS Operational Vertical Sounder data is used (Emori *et al*, 2005). Perhaps most importantly, the availability of satellite data at daily resolution is limited to recent years only; GPCP daily data begins only in 1997, limiting its use in model evaluations.

The problems associated with determining properties of areal rainfall are not exclusive to the application of model evaluation. The need for information about areal rainfall extremes for hydrological modeling and engineering is an ongoing problem which has led to many investigations into the construction of areal precipitation time series from point data (see, for example, Omolayo (1993) for a comparison of techniques) and modeling of the spatial and temporal variability of rainfall extremes (e.g. Veneziano *et al.*, 2006).

The following review will examine some of the attempts to date in the literature that have been made to evaluate precipitation variability and extremes in climate simulations, and look at approaches taken to resolving differences between point and areal rainfall in other applications.

2.3.1. Climate Models and Precipitation

Whilst GCMs have demonstrated skill in reproducing large-scale patterns and characteristics of rainfall, the smaller-scale properties (temporally and spatially) are widely acknowledged to be unreliable (Prudhomme *et al.*, 2002). Precipitation is a particularly complex atmospheric process to model, relying on the accurate representation of difficult physical processes including cloud micro-physics, cumulous convection, planetary boundary layer processes, and large-scale circulations (Dai, 2006b). Errors in the simulated precipitation are therefore often representative of errors in the representation of processes, (Dai, 2006b).

Climate models simulate precipitation via two main mechanisms. The first are synoptic (or stratiform) precipitating systems which occur due to vertical uplift of air by developing pressure systems, orography or monsoonal circulations (Stocker *et al.*, 2001). The second are convective precipitating systems, where more vigorous uplift of air, driven by latent heat, causes more localised but intense rainfall (Stocker et al., 2001). The approaches taken to modelling of the larger-scale (synoptic/stratiform) precipitation systems usually include a mixture of dynamical process modelling and parameterisation, and can be considered 'fairly successful' in simulating continental and sub-continental scale precipitation distributions, including orographic influences, when the large-scale circulation is reliably represented (Stocker *et al.*, 2001). The convective systems however, have to be fully parameterised due to their smaller spatial and temporal scale, and those parameterisations are often tuned to match observations (Kharin *et al.,* 2005). These schemes have been considerably less successful in reproducing characteristics of observed climate than those which deal with larger-scale precipitation (Stocker *et al.*, 2001). Many of these moist convection schemes tend to initiate convection prematurely such that instability does not develop sufficiently (Yang and Slingo, 1998; Dai *et al.*, 1999). Furthermore, inter-model comparison studies between models using different parameterisation schemes have found large differences between simulations from different models (e.g. Takle *et al.*, 1999).

2.3.2. Previous Evaluations of Variability and Extremes in Climate Model Experiments

2.3.2.1. Direct Evaluation Against Station Observations

A number of studies have used point-scale station observations to make validations of extremes in precipitation simulations from GCMs (Kharin and Zwiers, 2000) and RCMs (Mearns *et al.*, 1995; Raisanen and Joelsson, 2001; Jones and Reid, 2001; Yiping Guo and Senior, 2006).

Kharin and Zwiers' (2000) study of simulations from the Canadian Coupled Global Climate Model (CGCM1) compares the coarse scale model simulations with station observations. While the mean precipitation rate is found to be well reproduced by the GCM, the 20-year return values are found to be underestimated compared to the station observations. This underestimation of the extreme values is stated to be 'expected', but quantitative assessment of the skill in simulating extremes is, therefore, not possible without an observed areal-mean extreme values with which the model can be compared.

For high-resolution RCM evaluations, it is argued that the spatial resolution of the model is fine enough that variability and extremes are comparable to those found in station observations (e.g. Yiping Guo and Senior, 2006). Mearns *et al.* (1995) perform an evaluation of Regional Climate Model RegCM and the driving GCM (NCAR MM4) using station values to evaluate properties of daily rainfall including mean daily amount, mean intensity and rainday frequency. This study, however, included a pilot assessment of the suitability of individual stations for RCM evaluation. An average series of four stations, from one RCM grid box was compared to the single-station series, and found that the aggregation increased the wet-day frequency, decreased mean intensity but did not affect the daily mean. The disparity between the four-station and single-station series corresponded to the error found between RCM and corresponding station values in direction, but did not account for the full magnitude of the error. This highlights the importance of quantifying the disparity between point and areal rainfall variability even in high resolution studies.

Raisanen and Joelsson (2001) evaluate simulations of mean and annual maxima of daily precipitation from two RCMs from one grid box in Sweden. It is found that whilst the daily mean precipitation is over-estimated by 10-16%, the annual maxima are closer to those observed.

Raisanen and Joelsson (2001) suggest that the disagreement is largely a result of gauge error due to under-catch, which affects the mean more than the extremes, and conclude that the level of agreement is therefore ‘reasonable’. However, it could be argued that, if the spatial scale of the simulations is taken into account, the models might be found to overestimate both the mean and annual maxima compared to ‘true areal mean’ observations. This study demonstrates the ambiguity of model evaluation results, in both magnitude and direction, when the effects of spatial scale on variability are not quantified.

In all these examples, the degree of model skill is obscured by uncertainties in the ‘observed’ areal precipitation. Without quantification of the effects of spatial scale, the ‘target’ values for areal extremes are unknown, making it very difficult to quantify and compare fairly between models the level of skill in simulating precipitation variability and extremes.

2.3.2.2. Evaluation Against Gridded Station Observations

Many evaluations rely on gridded station datasets to provide ‘observed’ areal rainfall (Durman *et al.*, 2001; Frei *et al.*, 2003; Semmler and Jacob, 2004; Tolika *et al.*, 2006; Sun *et al.*, 2006).

In some cases, the station density in the gridded dataset used to represent observed areal rainfall may be sufficient to assume the level of temporal variability in the observed data is comparable to that of the model. Frei *et al.* (2003), for example, use a gridded dataset based on over 6000 stations in the European Alps, translating to 10 to 50 stations per RCM grid box. However, as Osborn and Hulme (1997) point out, the variability in an areal average is heavily dependent on the number of stations which make up that average. Station coverage in datasets used to construct gridded data varies with region, and through time. Whilst various gridding techniques have been developed to give the most reliable values on each day (or at whichever temporal interval the measurements are taken), the level of temporal variability in the resulting areal-average time-series for each grid box is determined by the number of available stations averaged to give that series. This means that for some regions and/or time periods in a gridded dataset, the level of temporal variability may be higher than would be found in the ‘true’ areal mean.

The number of stations required to give a reliable estimate of temporal variability and extremes in an areal average depends on the grid box size and shape, station distribution over the area, station distribution over time, and the spatial variability in the region. This makes it difficult to suggest a 'ball-park' estimate of how dense a station network should be to give a reliable areal-average temporal variability. High-elevation regions with varied topography such as the Alps, for example, require a higher station density in order to give a reliable representation of their climate characteristics than lower lying, more homogeneous regions (Frei and Schär, 1998). Durman *et al.* (2001) use work by Osborn and Hulme (1997) on the variance and wet-day probability in *n*-station means from UK precipitation data to suggest that 10-15 well spaced stations is adequate to represent areal rainfall for GCM evaluation for Europe. Whilst this has been demonstrated to be appropriate for the region and grid-scale, it may not be transferable to other regions where precipitation is more spatially variable, or to finer grid-scales for RCM validation.

Climate model evaluations based 'observed' areal averages from gridded datasets should therefore be interpreted with caution. Regions and time periods with insufficient station coverage are likely to have too high a level of spatial variability compared to the 'true' areal mean, which may affect the magnitude and even the sign of the apparent model error, and lead to bias in assessments of model performance. It is important to have some knowledge of the methods and station density used to construct the dataset in order to make an assessment of how well it might represent a true areal mean.

2.3.2.3. Evaluation Against Satellite Observations

Improvements in availability of satellite data have allowed some more recent evaluations of model performance to use these areal rainfall observations to evaluate some aspects of simulated precipitation (e.g. May, 2004; Iorio *et al*., 2004; Emori *et al*., 2005; Sun *et al*., 2006; Dai, 2006b; Wilcox and Donner, 2007).

Emori *et al.* (2005) have used the GPCP 1DD (daily resolution) merged dataset to compare how well precipitation extremes are simulated when two different cumulous parameterisation schemes are included in an atmospheric GCM, demonstrating that precipitation extremes are highly dependent on the parametisations used. May (2004) has also used the GPCP 1DD dataset, evaluating RCM ECHAM4's ability to reproduce daily variability and extremes in the Indian summer monsoon. The daily GPCP data has yet to be used for a more extensive inter-model-comparison of variability and extremes in simulated rainfall.

Satellite data from the tropical rainfall measurement mission (TRMM) have been used by Dai (2006b) and Wilcox and Donner (2007). The high temporal and spatial resolution (e.g half-hourly and 0.25°) of TRMM data has allowed detailed evaluations of modeled precipitation which are useful in comparing the dynamics of different model formulations (e.g. Wilcox and Donner, 2007). Dai (2006b) used this data at daily resolution to evaluate precipitation characteristics of intensity and frequency for four GCMs for all regions between 45N and 45S, finding that all four models tended to overestimate the proportion of total rainfall that fall in light events (<5mm) and underestimate the amount falling in heavier events (>20mm).

Some studies have used satellite data to evaluate the proportion of convective and synoptic rainfall (e.g. Iorio *et al.*, 2004; Dai, 2006b). An algorithm is used to separate the convective and synoptic components of rainfall in satellite observations of rainfall, which although not fully comparable to the separate convective and synoptic fractions from model simulations, can be usefully applied to learn more about model dynamics (Dai, 2006b).

2.3.2.4. Evaluations which Explicitly Resolve the Differences Between Point and Areal Rainfall Variability for Model Evaluation

Several studies have developed and applied statistical approaches to resolving the differences in temporal variability between point and areal rainfall, allowing a more quantitative evaluation to be made using the available station network.

Fowler *et al.* (2005a) applied Areal Reduction Factors (ARF) to observed station annual precipitation maxima in order to make the values comparable to the 50km resolution HadRM3H values. ARFs are empirically-determined relationships between point and areal rainfall for a specified duration, and are used for a number of hydrological applications, such as flood-defence design. The ARF values are assumed to be applicable in any region of the UK and to an event of any return period and can be found listed in the Flood Studies Report (NERC, 1974). Fowler *et al.* (2005a) apply the ARF value 0.82 to annual maxima of daily rainfall, and find that HadRM3 provides a good representation of extreme rainfalls up to a return period of 50 years for most regions of the UK.

Areal Reduction Factors resolve the differences between point and real extreme rainfall empirically, and thus while they are a useful facility for model evaluation at this regional scale, the approach has not been applied globally to GCM simulations because the derived ARF values are specific to the region for which they are calculated.

Osborn and Hulme (1997) develop a more theoretically-based method to determine variance and wet-day frequency in an areal mean from a limited number of stations. Their method makes use of a known relationship between the variance in a spatially-averaged series, the variance in individual timeseries and the mean inter-station correlation between the individual timeseries. This relationship essentially determines the amount of variability which is lost from a number of timeseries when they are averaged together, based on the amount of variance that is shared by the individual series' (i.e. the correlation between stations). The approach can be used to estimate variance in a 'true' areal mean by estimating the variance in an average of an infinite number of stations, and using mean inter-station correlation for an infinitely sampled grid box using a correlation decay curve. This curve is fitted to the separation distance and correlation values between pairs of stations, allowing the prediction of the correlation between two stations of given distance, and hence the prediction of the mean inter-station correlation for a grid box based on the distribution of possible separation distances. This approach allows the use of stations which fall both inside and outside the grid box to determine the mean inter-station correlation, maximizing the use of available stations.

The approach is extended to wet-day frequency. Whilst no pre-existing relationship can be used for this, a similar relationship can be empirically derived based on the change in the wet-day frequency when a number of stations are averaged is determined by the spatial scale of rainfall events. Osborn and Hulme (1997) measure the spatial scale of events for a region of season using the probability of coincident dry-days between stations (the conditional probability that a second station will be 'dry' on a given day, given that the first station is 'dry' on that day). These values are used to fit 'probability decay curves' which are used similarly to the correlation decay curves. These values are then used to fit an empirical function that can be used to predict the wet/dry-day probability for an average of a number of stations. However, the conditional probability values used by Osborn and Hulme (1997), used as a measure of the spatial dependence of wet-or-dry-day occurrence between a pair of stations are biased by the wet/dry day probability that exists at those stations. The probability that dry-days will coincide at two stations is dependent on two factors: the first is the dry-day probability at each station, individually and the second is the degree of dependence between the two stations (i.e. a pair of stations in a dry region, that generally experience frequent dry-days, will be more likely to experience coincidences of dry-days, even if they are independent of one another, compared to stations in wetter regions that experience fewer dry-days). This caveat in Osborn and Hulme's (1997) approach reduces the robustness of their empirical relationship. In fact, if the conditional probability values used are adjusted to remove the influence of station dry-day probability, leaving a measure of inter-station dependence only, the empirical relationships are considerably less strong (Osborn, *Pers comm.*.)

Osborn and Hulme (1998) go on to apply their techniques to the validation of twelve models in the Atmospheric Model Inter-comparison Project (AMIP). The study finds that most of the models show too much day-to-day variability, and often too many rain-days, particularly in winter. Average daily intensity was either similar to, or slightly less than, the observed.

Booij (2002a, 2002b) investigated the properties of extreme daily precipitation events at point and areal scale for the purposes of model evaluation. This work applies geostatistical techniques, which are used to determine properties of areal precipitation (particularly extremes) for other hydrological applications, such as design rainfalls. Where Osborn and Hulme (1997) used correlation decay length (distance at which correlation falls to 1/e) as a measure of station inter-dependence, Booij (2002a, 2002b) recognises that extreme events might show greater spatial variability (because very heavy precipitation events are often localised), such that the correlation between stations for the complete precipitation series is not a good indication of the spatial structure on a day when extreme rainfall is recorded. The variogram, used in geostatistics, can be used to identify the degree of spatial variability in a single realisation of the rainfall process (i.e. one day). Booij applies this to the annual maxima, and calculates spatial correlation length of these extreme events to be about 80km, compared to 300km for the complete series in this study, and 200km in summer and 300km in winter for Europe by Osborn and Hulme (1997).

The paper goes on to apply this in determining the parameters of the Gumbel distribution to describe the distributions of point and areal extreme rainfall. Booij (2002a, 2002b) makes use of work by Sivapalan and Bloeschl (1998) which proposes a method for transforming between point and areal extreme precipitation. This involves application of a 'Variance Inflation Factor', based on the size, shape and correlation structure of a precipitation area, to transform the parameters of the Gumbel distribution from those which are observed at point scale to those which could be expected at areal scale. However, Booij (2002a, 2002b) does not test the approach, assuming that combining the approach of Sivapalan and Bloeschl (1998) with the geostatistical approach to measuring spatial correlation will give reliable indications of the properties of the extreme values distribution, and proceeds to use the technique to evaluate two re-analysis models, three GCMs and two RCMs, with respect to their ability to produce accurate 20-year return values.

2.3.3. Evaluation of Variability and Extremes in Climate Model Experiments: Model Inter-comparison and General Results

A number of model inter-comparison studies have investigated the realism of daily precipitation characteristics in GCM simulations (e.g. Osborn and Hulme 1998; Kharin *et al*., 2005; Dai, 2006b; Sun *et al.*, 2006). These studies have highlighted general strengths and weaknesses which affect models similarly. A common conclusion from such studies is that models tend to over-estimate the amount of light or 'drizzly' rainfall and underestimate more moderate and heavy rainfall (Osborn and Hulme, 1998; Iorio *et al*., 2004; Dai, 2006b; Sun *et al*., 2006). This bias towards rainfall that is too light and too frequent is generally attributed to problems with convective parameterisations (Kharin *et al*., 2005). This weakness in many models means that rainfall characteristics are often least well simulated in the tropics (Kharin *et al*., 2005). Model skill has generally been found to improve with increasing spatial resolution as this allows more of the rainfall processes to be resolved dynamically rather than by parameterisation (Iorio, 2004, Kimoto, 2005).

Despite the tendency for models to produce too much rainfall at light intensities and too little at heavy intensities, studies of the stratiform and convective components of rainfall have shown that, at least some, models tend to produce too much convective rainfall relative to stratiform (Dai *et al*, 2006b; Sun *et al*., 2006).

In a slightly different approach to evaluation than has been taken in the other studies, Kiktev *et al.* (2003) investigated the ability of the atmospheric GCM HadAM3 to reproduce the same trends in precipitation extremes that were found at stations, and found that the model showed very little skill in this respect.

2.3.4. Other Studies Addressing the Spatial Scaling Properties of Rainfall Variability and Extremes

There is a large body of research in the hydrological literature which has explored the issue of regionalisation of precipitation observations, fuelled by the need for areal precipitation totals (particularly extreme events) for hydrological modeling and engineering projects such as dam spillways, drainage and flood defence.

It has long been recognised that uneven and/or sparse station distribution can bias catchment or gridded precipitation averages, and this has lead to the development of a variety of different interpolation and regionalisation techniques, including Theissen's polygons, Triangulation and Kriging (Omolayo, 1993). Whilst these techniques can help to redress issues of accurate representation of rainfall total for a region, more complex approaches have been taken to determine the representation of areal extremes.

The following section summarises some of these other approaches, and their potential suitability for application to climate model evaluation.

2.3.4.1. Areal Reduction Factors (ARFs)

Fixed area Areal Reduction Factors (ARFs) have traditionally been used to derive the magnitude of an areal event expected to recur every *T* years from the magnitude of a point event of the same return period. There have been a number of different approaches taken to calculate these transformation functions which can be divided (e.g.by Omolayo, 1993) into those which are empirically based and those which are theoretically based.

Empirically based methods include the US Weather bureau method and the UK method (Omolayo, 1993), and Bell's Method (Bell, 1976) which essentially apply a ratio between observed point extremes and 'observed' areal extremes. These approaches all come with their own set of caveats (see Omolayo, 1993), but perhaps the greatest disadvantage is that being based on a rain gauge network for a particular region, they cannot necessarily be transferred to climatically different regions, or to a different climate regime for the same region. It is often the practice to transpose ARFs calculated for one area to another, and often without testing that they are appropriate (Omolayo, 1993). Although this may be legitimate for areas with similar precipitation characteristics, this is not a sensible approach for the purposes of overcoming the problems of model evaluation of variability and extremes, as the areas with sparse networks are often climatically very different to more economically developed regions which do have sufficiently dense observation networks to calculate the ARF.

2.3.4.2. Transforming the Statistical Distributions of Extremes of Point and Areal Precipitation

A number of studies have built on the concept of the Areal Reduction Factor in order to transform the parameters of the extreme value distribution, rather than simply to adjust the individual values by a fixed proportion.

The work of Sivapalan and Bloeschl (1998) has already been mentioned above as it was applied by Booij (2002a, 2002b). Sivapalan and Bloeschl proposed a theoretical approach to transforming between the extreme value distributions of point and areal rainfall distributions based on the inter-station correlation in a region, but this was considered by Booij (2002a, 2002b) to be flawed because the transformation of extreme values was based on inter-station correlation for the full precipitation timeseries. Booij (2002a, 2002b) demonstrated, using geostatistical techniques, that the spatial correlation of extremes was generally considerably lower for extreme precipitation events compared to that of all events, and therefore the size of the necessary transformation was comparably larger.

The use of geostatistics to characterise the structure of spatial variability of the extreme fraction of the precipitation process was used in a much earlier paper by Lebel and Laborde (1988), who use geostatistical techniques to infer parameters of the distribution of areal extremes. This approach, however, assumes that the distribution of point and areal rainfall stem from the same class of distribution and that the distribution is characterized only by two parameters: the mean and the variance. This approach quantifies the change in distribution, but cannot account for changes in the distribution shape. Skaugen *et al*. (1996) point out that basic changes in the shape of the distribution also occur between spatial scales. In accordance with the central limit theorem, the distribution of rainfall is likely to tend towards normality when stations are aggregated to the areal scale (Rodriguez-Iturbe and Mejia, 1974; Skaugen *et al.*, 1996). Because stations are unlikely to be completely independent, it is reasonable to assume that the areal rainfall distribution will lie somewhere between the positively skewed point distribution and the Gaussian (Skaugen *et al.*, 1996). Skaugen *et al.* (1996) tested this hypothesis and found it to be the case. Coles and Tawn (1996) improve on these studies, developing a theoretical approach to transforming the parameters of the Generalised Extreme Value (GEV) distribution for areal rainfall, thus making use of contemporary work on extreme value theory and allowing for changes in the shape of the extreme value distribution, but also explicitly accounting for the spatial characteristics of extreme rainfall.

The limitation of these approaches, in respect to climate model evaluation, is that hydrological approaches such as these exclusively deal with the extremes of the distribution. Whilst this might have very interesting applications to the evaluation of extreme values in model-simulated precipitation, climate model evaluation concerns also the values throughout the parent distribution. The 'extremes' that are often considered tend to be 'less extreme' than the 20-and-50-year return period values, such as 95th or 99th percentile values which occur 3-6 times per year. A gap remains in the current literature whereby values throughout the distribution of daily rainfall need to be estimated for areal rainfall.

2.3.5. Summary

An increasing body of research has assessed the ability of GCMs and RCMs to reproduce the daily variability of rainfall, not just the mean values. Gridded datasets based on high density station networks or available satellite data can provide good estimates of observed areal rainfall for some regions and time periods for model evaluation. However, for areas and time periods where such data are not available, other approaches to resolving differences between point and areal scale variability are required to make quantitative evaluations of model performance in respect to daily variability.

There are currently few examples of evaluations of daily precipitation from GCM simulations which explicitly address the issue of reduced variability between point observations and areal model simulations. Those which have include evaluation of standard deviation and wet-day frequency (Osborn and Hulme, 1997) and parameters of the Gumbel distribution of extreme values for point and areal precipitation (Booij, 2002a, 2002b).

Although there have been a number of other studies which have addressed the statistical characteristics of areal rainfall to meet the needs of other hydrological fields, these have dealt mainly with extreme events. For climate model evaluation, more reliable estimates of the parent distributions of areal rainfall, covering statistics such as distribution shape and wet-day frequency, and rainfall intensity as well as the extremes values, are required.

2.4. Representing Sub-grid-scale Precipitation Variability in Climate Change Scenarios

---

2.4.1. Areal Rainfall Projections: Problems for Regional Hydrological Impact Assessment

While climate models can be used directly to generate possible scenarios of changes in day-to-day precipitation variability at coarse grid-box scale, it is the variability at a more local scale which is usually of more interest for climate impact assessment. The sub-grid-scale variability of precipitation, both temporal and spatial, is of utmost importance in several impacts sectors. One area where this is particularly important is the estimation of future flood frequency. There are concerns that the changes in rainfall intensity and increases in very heavy rainfall events projected under future climate change scenarios might result in increased incidences and/or severity of river flooding.

The application of GCM (and even RCM) simulated precipitation to hydrological models is limited by spatial and temporal resolution in climate models, and by the lack of skill in simulating precipitation reliably. The distribution of rainfall in space, as well as time, has a significant influence on peak river flows, and hence on the possibility that a given event will result in a flood (Bell and Moore, 2000). Furthermore, whilst GCMs have demonstrated skill in reproducing large-scale patterns and characteristics of rainfall, the smaller-scale properties (temporally and spatially) are widely acknowledged to be unreliable (Prudhomme *et al.*, 2002).

Various approaches to the 'downscaling' of precipitation have been developed in an effort to overcome either or both of these limitations, and thus provide estimates of rainfall for the future which are reliable in their magnitude and characteristics, and represent variability at an appropriate spatial scale for hydrological modeling.

2.4.2. Downscaling Techniques

2.4.2.1. Dynamical Downscaling

Regional climate models (RCMs) potentially provide the most realistic rainfall estimates at regional and local scale because they resolve processes dynamically. However, whilst higher resolution models have been demonstrated to produce characteristics of rainfall more reliably than coarser scale GCMs (Iorio, 2004; Kimoto, 2005), significant biases in quantity and/or variability are still often evident (Prudhomme *et al.*, 2002).

The ongoing improvements in dynamical modeling have, however, meant that some recent studies have been able to successfully apply RCM data directly to hydrological models after some bias correction to the mean values (e.g. Wilby *et al.*, 2000; Wood *et al*, 2004) and have demonstrated skill in reproducing historical flows, although Hay *et al.* (2002) suggest that variability is also underestimated and may present a limitation to this approach. Fowler and Kilsby (2007) recently applied HadRM3 simulated daily precipitation, after bias correction, to catchments in Northwest England, and found an increase in magnitude by 25% of high flows (the 5% highest flows) by 2100 under the SRES A2 scenario (see Nakic´enovic´ *et al.*, 2000) , thus presenting an increased flood risk during winter months.

Kay *et al*. (2006a, 2006b) also applied HadRM3 simulated precipitation to modeling of 15 catchments around the UK for climate change under SRES scenario A2. This study found that 8 of the 15 catchments showed an increased flood frequency by the 2080s, despite the fact that many of these showed an overall decrease in mean precipitation over that period, demonstrating that changes to the distribution of rainfall, in space and/or time, can affect flood frequency.

2.4.2.2. Statistical Downscaling

Statistical downscaling provides a less computationally demanding approach to determining local scale climate, and includes deterministic and stochastic models, and models which are a hybrid of the two. Deterministic methods include regression-based and weather-typing-based approaches. These techniques rely on the greater reliability of the climate model at simulating large scale atmospheric conditions compared to local surface variables. They use large-scale atmospheric conditions to predict surface weather, such as daily precipitation, using statistical relationships with selected 'predictor' variables or weather type categories. An assumption of these approaches is that these relationships are physically-based, and remain stationary through time, although this represents a weakness of the technique as this cannot be guaranteed under changed climatic conditions. Stochastic downscaling approaches are essentially random number generators which produce output resembling weather recordings at a location (Wilks and Wilby, 1999). At single sites, these 'weather-generator' types of model can reproduce wet-day probability and strings of wet and dry days effectively, using a two-stage statistical model. The first stage models precipitation occurrence, based on the autocorrelation in the series (for example, by first or second order Markov chains) and the second stage models precipitation amount on wet days (for example, sampling from the Gamma distribution).

In climate change applications, statistical downscaling models are more often some combination of the two approaches to maximize skill. Wilby and Dawson's (2002) Statistical DownScaling Model, SDSM, for example is regression based but includes a stochastic component to add in the variance which cannot be explained by the predictor variables. Fowler *et al.* (2000, 2005b) use weather-typing together with a stochastic model (Cowperwait *et al*, 2002). The stochastic part of the model is based on Neyman Scott Rectangular Pulses (NSRP) which model daily rainfall at a site based on rectangles representing the occurrence, duration and intensity of individual rainfall cells. An NSRP model is therefore associated with each weather type for each season and used to generate daily rainfall based on the classification of weather types in simulation of future climate.

Evaluation of statistical downscaling models have generally found that their level of skill is comparable to that of RCMs, which are considerably more complex and demanding in computer power (Kidson and Thompson, 1998; Murphy, 1999; Wilby and Wigley, 2000; Hellström *et al.*, 2001), although different models shown divergence between their projections of future climate at the regional scale (Cubasch *et al.*, 1996; Wilby and Wigley, 1997; Wilby *et al.*, 1998; Mearns *et al.*, 1999; Murphy, 2000).

2.4.2.3. Multi-site downscaling models

The extension of downscaling models to simultaneously generate weather at multiple sites is of interest for hydrological applications which require multiple spatially-related precipitation for realistic modeling studies. The approaches of Cowpertwait *et al.* (2002), Wilby *et al* (2003), Hughes and Guttorp (1999) and Wilks (1999) are all examples of multi-site downscaling models.

Most of the multi-site approaches to statistical downscaling use patterns of spatial correlation in existing climate to train the spatial components of the model (e.g. Wilks, 1999; Wilby *et al*, 2003). This means that the models can reproduce much more realistic patterns of temporal and spatial variability in rainfall for the observed period over which the model is trained (Wilby *et al*, 2003). When applying these multi-site models to future climate, it is implicitly assumed that the spatial dependence between sites is stationary and does not alter in a warmer climate. Cowperwait *et al.* (2002) is one example of a multi-site approach which may be able to capture some change in spatial correlation. This extension to the weather-type conditioned stochastic model described above (applied by Fowler *et al.*, 2005b), assigns a spatial correlation parameter to each weather type, such that a change in the relative occurrence frequencies of different weather types in future climate might cause a change in overall spatial correlation.

2.4.3. The Relevance of Scaling Relationships Under Future Climate Conditions

A number of the studies reviewed in Sections 2.3.1.4 and 2.3.3 have addressed the scaling properties of precipitation variability and extremes in order to determine the proprieties of the areal mean from point observations (e.g. Omolayo, 1993; Coles and Tawn, 1996; Osborn and Hulme, 1997). These approaches share common ground in using some measure of spatial correlation (using either traditional correlation, or geostatistical approaches to measuring spatial dependence), and an implicit assumption shared by all these approaches is that the spatial dependence is stationary. A similar problem occurs here, as does for the multi-site downscaling approaches, in that this assumption may not hold in future if the spatial variability of precipitation changes as rainfall becomes more convective and localised, as climate model simulations have suggested (Noda and Tokioka, 1989; Gordon *et al.*, 1992, Hennessy *et al.*, 1997; Chen *et al.*, 2005).

The application of scaling relationships in reverse to estimate the variability of point precipitation from the areal precipitation generated by climate models under future scenarios is therefore subject to this uncertainty. The quantification of this uncertainty via some estimate of the magnitude of any change in spatial correlation at sub-grid-scale is, to date, an area of research which has yet to be filled.

2.4.4. Summary

A variety of downscaling approaches have been applied to GCM (and RCM) simulated precipitation in order to try to determine the level of temporal variability at the sub-grid scales that are required for hydrological models used in flood estimation. Many of these approaches have been fairly successful in determining temporal characteristics for points or small-grid scales, but reliable flood estimation requires precipitation estimates which represent the appropriate levels of both temporal and spatial variability, the latter of which is much more difficult to estimate under future climate conditions. Whilst dynamical models (RCMs) potentially have the ability to resolve both the spatial and temporal variability of rainfall, and represent changes in spatial variability as well as temporal, their application is currently limited by the simulation errors in mean and/or variability found at local scales. Multi-site downscaling models are currently mostly limited to reproducing rainfall with existing levels of spatial dependence between sites.

The prediction, identification or quantification of any likely change in the spatial correlation of rainfall under a warmer climate is currently a gap in the existing literature. This type of study is made difficult by the limited resolution of climate models, and their inability to resolve small-scale processes important to precipitation explicitly.

2.5. Summary of Literature Review and Research Objectives

The literature review has led to the identification of a number of gaps in literature, and some areas which would benefit from further research and development of new techniques. Here follows a list of the main summary points.

- Physical theory suggests that a warmer atmosphere has a greater capacity for holding water, such that increased atmospheric temperatures will cause the hydrological cycle to become more intense resulting in increased intensity of rainfall, affecting particularly the heaviest events. This is supported by evidence from modelling experiments, where precipitation simulated under global warming scenarios are broadly consistent in indicating global trend of increasing precipitation, which occur due to increases in intensity of rainfall events. Analysis of the variability and extremes in observations of rainfall have suggested that there is some evidence that, in some regions, such changes are already occurring due to the warming that has been experienced in recent decades.

- There is also some model evidence that suggest that the type of rainfall experienced in a warmer climate might shift towards a greater proportion of convective rainfall in the mid-to-low-latitudes compared to the present day. This might alter the overall spatial characteristics of rainfall in these regions. There is little in the way of observed evidence to support this.

- Evaluations of model skill in simulating rainfall have suggested that, on the whole, models tend to simulate realistic mean rainfall amounts, but when the daily variability of that rainfall is often unrealistic. Model evaluations which have looked at daily variability have generally found that models often simulate light, or ‘drizzly’, rainfall on too many days.

- A variety of methods have been used to resolve the problem of comparing grid-box average daily rainfall variability with station observations. Many studies rely on gridded datasets for areal rainfall estimates (e.g. Sun *et al.*, 2006), but in regions where the station network is sparse, these are unlikely to give a good indication of spatial variability in the 'true' areal mean. Others have used scaling approaches to estimating the extremes of the distribution for areal rainfall (e.g. Booij, 2002a, 2002b). Only Osborn and Hulme (1997, 1998) propose and/or apply an approach to estimating characteristics of the distribution of all rainfall days, rather than only the extremes, which can be used for region that is not very densely gauged. This uses a theoretical approach to estimate variance of an areal mean precipitation series, and an empirical approach for estimating the wet-day frequency. The development of approaches for estimating other characteristics of the distribution (for example, the shape/skewness of the distribution, or parameters of a suitable distribution model) of areal-rainfall amounts for regions where a dense station network is not available would allow for more explicit and quantitative evaluation of the level of daily variability in simulated rainfall.

- The scaling relationships used in these studies of existing climate might also be applied to future climate (e.g. in order to estimate the characteristics of point rainfall from the grid-box averages rainfall simulated by GCMS) *if* the spatial characteristics of rainfall can be assumed to remain stationary under a changed climate regime. This may not be the case if a shift towards a greater proportion of convective, rather than synoptic, rainfall occurs in a warmer climate. Changes to the spatial correlation of rainfall under future climate have received little attention in the literature to-date.

- A variety of 'downscaling' techniques have been developed in order to estimate characteristics of point, or local-scale, rainfall for future climate for climate impact assessment. The use of these models to produce multi-site rainfall series often relies on the assumption that the spatial correlation of rainfall will remain stationary under future climate.

Given the conclusions drawn above, the following two areas are identified for further research and development, and are addressed in this thesis.

1.) **The development of robust approaches for estimating characteristics of the variability of areal rainfall where only a few stations are available, and are suitable for application to climate model evaluation.** Whilst a number of existing approaches address the extremes of the areal rainfall process, the characteristics of the parent distribution is less commonly addressed. The approach of Osborn and Hulme (1997) will be used as a basis for developing new approaches to estimating the dry-day probability and gamma parameters of wet-day amounts for areal rainfall.

2.) **To investigate further the suggestion that future rainfall might be more convective in nature, whether this might cause a significant change in the spatial correlation of daily rainfall and hence whether this will affect the validity of point/areal scaling relationships for rainfall variability and extremes in the future**. This issue has significant implications for the application of techniques developed above, and existing downscaling approaches. If such a change is found to be likely, the quantification of a change in spatial correlation would allow the application of the above techniques simulations of future climate, as well as to existing multi-site downscaling techniques, and provide valuable indications of the temporal variability which might be expected to occur at points or local areas under climate change scenarios.