Tyndall°Centre
for Climate Change Research

The identification and evaluation of suitable scenario development methods for the estimation of future probabilities of extreme weather events

Clare Goodess, Mike Hulme and Tim Osborne

July 2001

Tyndall Centre for Climate Change Research Working Paper 6

The identification and evaluation of suitable scenario development methods for the estimation of future probabilities of extreme weather events

Clare Goodess, Mike Hulme and Tim Osborn

School of Environmental Sciences
University of East Anglia
Norwich NR4 7TJ

Tyndall Centre Working Paper No. 6
20 July 2001

CONTENTS

1. INTRODUCTION

Many impacts of climate change will be realised as the result of a change in the frequency of occurrence of extreme weather events such as windstorms, heavy precipitation or extreme temperatures over a few hours to a few days (Mearns *et al.*, 1997). Thus a specific research objective of the Tyndall Centre Research Theme 3 (Adapting to Climate Change) is to estimate past and future probabilities of selected extreme weather events and rapid climate changes. To date, however, few scenarios consider changes in the magnitude and occurrence of extremes, preferring to focus on changes in mean climate (see Section 2). The reason for this is twofold: first, the lack of suitable tested methods for developing scenarios that include information about climate extremes and variability changes; and second, the limited availability of climate model simulations with reliable output at the necessary spatio-temporal resolution.

An overriding problem – and hence the need for more sophisticated scenario development methods – is that output from climate model simulations cannot, in general, be used to directly quantify future variability and extremes because of bias in simulated means and variability of present-day climate and weather (see Section 4). This bias may originate from systematic model errors, from spatial scale incompatibilities (area-mean grid-box output has different statistical properties to station data; see Section 3.6) and due to the exclusion of sub-grid-scale processes.

These problems are addressed by the three key research objectives of Tyndall Project IT1.16:

(i) To **identify** the range of scenario development methods that are most suitable for extremes of weather and climate, to expertly **assess their limitations/problems** when applied to the range of UK-oriented events and climate variables which are most important for impact assessment studies, and to **explore solutions** for overcoming these limitations/problems.

(ii) To **quantitatively test and intercompare** the most promising scenario development methods using a number of specific **UK case studies** (such as drought) which are relevant to the Tyndall Centre research objectives.

(iii) To **develop guidelines** for the subsequent construction of reliable and consistent scenarios of extremes for the UK as part of future Tyndall Centre research activities, identifying suitable methods and best practise, and outlining strategies to allow scenarios of extremes to be used in probability-based impact assessments and subsequently in integrated assessment models (IAMs).

The starting point for a comprehensive assessment of methods that are most suitable for the estimation of future probabilities of extreme weather events is recent work on scenario development, including that reported in the IPCC Third Assessment Report (TAR), which is outlined in Section 2.

In order to assess the most suitable methods, it is first important to identify the range of scenario information that is required by stakeholders and others with requirements for climate change impact assessments. This issue was a major focus of the project workshop held in Norwich on 4-5 June 2001. A number of relevant questions and topics debated during this workshop, including issues which are particularly relevant to the need to develop probabilistic approaches and integrated assessment models, are presented in Section 3, together with a list of proposed indicators of temperature and rainfall weather extremes.

The second major focus of the project workshop was the identification of the available scenario development methods and their limitations. A number of potential methods are identified in Section 4 and the limitations of each evaluated (based, in part, on evidence from the studies identified in Section 2), both generally, and for the specific problems of inter-annual variability, multi-variate correlations, multi-site correlations, spatial-scale dependence and scaling by simple models to obtain a range of scenarios.

In Section 5, consideration is given to the evaluation and incorporation into assessment studies of low probability, but high impact events, such as an abrupt reorganisation of the North Atlantic thermohaline circulation. There are two aspects of such events that must be addressed: (i) the possibility of the event occurring, and (ii) the response of the climate system to the event. Both aspects are subject to considerable uncertainty which means that a more preliminary and subjective approach has to be taken compared with the more conventional extremes considered in Section 4.

All the necessary questions concerning scenarios of extreme events cannot be answered on the basis of past work and the critical reviews presented in this report alone. Thus specific applications to test cases will be undertaken as part of the project in order to rigorously and, where possible, quantitatively, compare and evaluate the various methods, and to further develop methods in order to overcome the problems identified in Section 4. The proposed case-study work is outlined in Section 6.

A draft version of this working paper formed the basis of discussion at the project workshop. A number of additional issues which arose in the workshop discussions, relating to indicators of extremes, methodological approaches and case studies, are summarised in Appendix 1.

The case-study work forms Task 3 of the project [*Quantitative testing of the most promising scenario development methods (identified as part of Task 2) for selected case studies (identified as part of Task 1)*]. Task 1 (*Identification of the weather extremes scenario needs of the UK impacts community*) and Task 2 (*Identification and evaluation of the suitable scenario development methods*) are addressed by Sections 2-4 of this report and are the foci of the project workshop. The final project Task, Task 4 (*Development of guidelines to underpin the future development of scenarios that include information about climate/weather extremes and variability*) will be addressed following completion of the case-study work.

The key outputs of this project will be (i) the evaluation of scenario development methods and (ii) resulting guidelines for generating scenarios of weather variability and extremes. The former, the subject of this preliminary internal report, will be disseminated via the peer-reviewed literature and scientific conferences, focusing on the case studies, while the later will be disseminated primarily via the final project report (May 2002).

2. RECENT WORK ON SCENARIO DEVELOPMENT FOR EXTREMES

2.1 Scenarios of extremes

Relatively few studies have focused specifically on the construction of scenarios of extremes rather than mean climate, in part, because of the problems associated with the reliability and availability of high (spatial and temporal) resolution climate model output (Kharin and Zwiers, 2000; Meehl *et al*., 2000). Initial studies, particularly of precipitation, tend to focus more on changes in variability (Buishand and Beersma, 1996; Beersma and Buishand, 1999; Boer *et al*., 2000; Giorgi and Francisco, 2000a) or changes in distributions rather than specific extreme events. Gregory and Mitchell (1995), for example, examined changes in the parameters characterising daily temperature and precipitation simulated by the Hadley Centre UKHI model, while Hennessy *et al*. (1997) investigated changes in precipitation frequency distributions and in the relative contributions of convective and non-convective precipitation mechanisms in UKHI and CSIRO9. Wetherald and Manabe (1999) focused on changes in soil moisture (particularly as an indicator of summer dryness) simulated in four GFDL experiments.

With the greater availability of daily output from GCMs, this output has begun to be used directly to construct scenarios of specific extremes. A number of these studies are summarised in Table 2.1. Some of them use time slices which are relatively short for the analysis of extremes (10 or 20 years) and many are based on equilibrium rather than the newer generation of transient, coupled AOGCMs.

Most recently, a few studies have used output from regional climate models (RCMs) to construct scenarios of extremes (Schär *et al*., 1996; Frei *et al*., 1998; Mearns *et al*., 1999; Durman *et al*., 2001; Jones and Reid, 2001). Durman *et al*. (2001), for example, focus on the occurrence of intense precipitation events over Europe and the UK defined using two different thresholds (15 mm per day and the upper 1% percentile calculated from the model control run). A comparison is made of future scenarios (for 2080-2100) constructed from HadCM2 GCM and RCM output in order to determine the added value of using high-resolution model output (Durman *et al*., 2001; see also Section 4.2.3). Jones and Reid (2001) also use the HadCM2 RCM to construct future scenarios of extreme precipitation for the UK, in this case focusing on the occurrence of the top 10% quantile events (calculated using the method of Osborn *et al*., 2000) and 5, 10, 20 and 50 year return period events. Both these studies provide support for the increased frequency of heavy precipitation events, particularly in winter (see Section 2.3).

RCMs allow the dynamical downscaling of GCM output to the higher spatial resolutions which are more appropriate for the construction of scenarios of extremes. Statistical downscaling provides a less computer-intensive method of downscaling (Hewitson and Crane, 1996; Wilby *et al*., 1998; Zorita and von Storch, 1999; see Section 4.2.4), but has not been widely used to investigate changes in extremes. Wagner (1996) used a probability model based on thresholds to show that changes in temperature extremes are more sensitive to changes in variability than mean climate (using output from the ECHAM GCM and daily temperature data for Berlin). A similar conclusion was reached using a first order autoregressive Markov Chain model

(Wagner, 1999). Although statistical downscaling has rarely been used specifically to construct scenarios of extremes (one exception is Brandsma and Buishand, 1998), a number of studies do include analyses of relevant indicators, particularly as part of the validation of the methodology (Table 2.2).

Table 2.1: Summary of recent studies which use GCM output directly to construct scenarios of extremes.

Study	***Extremes***	***Region***	***GCM***
Dai *et al*., 2001	Frequency and persistence of 'hot' days (>80th percentile) Storm activity	Global USA	NCAR CSM Coupled model 2 scenarios 20 year time slices
Delworth *et al*., 1999	Steadman heat index (based on monthly temperature and atmospheric moisture)	Global	GFDL Coupled model 3 simulations 30 year time slices
Huth *et al*., 2000	Heat waves/dry spells	Czech Republic	ECHAM3 Equilibrium model 30 year time slices
Kharin and Zwiers, 2000	Temperature, precipitation, wind: 20 year return periods, thresholds, cooling & heating degree days	Global Canada	CGCM1 Coupled model 3 ensembles 21 year time slices
Kothavala, 1997	Precipitation: return periods, percentiles and Palmer Drought Severity Index (PDSI)	Midwest USA	CCM1-OZ Equilibrium model 10 year time slices
Kothavala, 1999	PDSI (based on monthly temperature and precipitation)	Eastern Australia	CCM0 Coupled model 30 year time slices
McGuffie *et al*., 1999	Temperature and precipitation: return periods and range of descriptive regional statistics	Global 5 IPCC regions	5 equilibrium GCMs 10 year time slices
Zwiers and Kharin, 1998	Temperature, precipitation and wind: 20 year return periods and thresholds	Global Canada	CCC GCM2 Equilibrium model 20 year time slices

2.2 Observed changes in extremes

In parallel to the move towards the development of scenarios of extremes (Section 2.1), a number of recent studies reflect the move towards a more rigorous exploration of observed changes in extremes. Three such studies, for example, provide evidence of trends towards more intense precipitation events in the UK (Jones *et al*., 1999; Osborn *et al*., 2000; Alexander and Jones, 2001). Other recent studies attempt to provide a more global synthesis of observed changes in extremes (Karl *et al*., 1999; Easterling *et al*., 2000a,b; Frich *et al*., 2001). Such studies provide the opportunity to determine whether projected changes in extremes are consistent with the observed changes (Section 3.3), though this must be done with statistical rigour due to the weak signal-to-noise ratio of the expected changes to date, especially at the local scale.

Table 2.2: Statistical downscaling studies which include analysis of extreme event indicators.

Indicators studied	***Study***
Studies which include analysis of precipitation-related extreme indicators, e.g. length of (longest) wet/dry spells, return period events, ranked extremes	Bardossy and Plate, 1991; 1992 Bates *et al*., 1998 Bogardi *et al*., 1993 Brandsma and Buishand, 1998 Charles *et al*., 1999 Conway and Jones, 1998 Corte-Real *et al*., 1999 Goodess, 2000 Hay *et al*., 1991; 1992 Hughes *et al*., 1999 Semenov *et al*., 1998 Weichert and Bürger, 1998 Wilby, 1998 Wilby *et al*., 1994; 1998 Wilks, 1999 Wilson *et al*., 1991; 1992
Studies which include analysis of storm-related indicators, e.g. storm length, inter-storm arrival time	Hughes *et al*., 1993; Hughes and Guttorp, 1994 Schnur and Lettenmaier, 1998
Studies which include analysis of temperature-related extreme indicators, e.g. annual maximima/minima, heat waves and cold spells, frosts, threshold exceedence	Hayhoe, 2000 Palutikof *et al*., 2001 Schubert, 1998 Schubert and Henderson-Sellers, 1997 Trigo and Palutikof, 1999 Winkler *et al*., 1997

2.3 Extremes in the IPCC Third Assessment Report

The IPCC Third Assessment Report gives greater consideration to extreme events than previous assessments and cites many of the references identified in Section 2.1 above. Sections of the Working Group 1 and 2 reports which deal specifically with extremes are listed in Table 2.3.

Table 2.3: Sections of the IPCC Third Assessment Report dealing specifically with extremes.

WG	*Chapter*	*Section*	*Subject*
1	Summary for Policy Makers	Table 1	Estimates of confidence in observed and projected changes in extremes
1	2: Observed climate variability and change	2.1, 2.2, 2.5, 2.7	Observed changes in extremes over the 20th century are assessed
1	9: Projections of future climate change	9.3.6	Possible future changes in extreme events (temperature, precipitation, extratropical storms, tropical cyclones) are assessed using global model output
1	10: Regional climate information – evaluation and projections	10.3.1.2 10.3.2.2	AOGCM simulations of regional extremes are assessed for current and future climate
		10.4.1.2 10.4.2.2	Variable/increased horizontal GCM simulations of extremes are assessed for current and future climate
		10.5.1.3 10.5.2.2	RCM simulations of regional extremes are assessed for current and future climate
1	13: Climate scenario development	13.4.2.1	Techniques for incorporating changes in daily variability: daily to interannual time-scales (i.e. ‘conventional’ extremes)
		13.4.2.2	Techniques for incorporating other extremes (e.g. hurricanes and tornadoes) into climate scenarios
2	3: Developing and applying scenarios	3.5.4.6	Problems of incorporating extremes such as hurricanes and tornadoes in scenarios
		3.5.4.7	Need to consider ‘surprises’, i.e. low-probability high-impact events

The most concise summary of observed and projected changes in extreme events is provided by Table 1 from the Working Group 1 Summary for Policymakers. The table provides an assessment of confidence in the observed and projected changes based on observational and modelling studies, as well as the physical plausibility of future projections across all commonly-used scenarios, based on expert judgement. Higher maximum temperatures and more hot days over nearly all land areas, together with higher minimum temperatures, fewer cold days and frost days over nearly all land areas, for example, are considered ‘very likely’ (i.e. 90-99% chance) during the 21st century. More intense precipitation events are also projected to be ‘very likely’ over many areas, while increased summer continental drying and the associated risk of drought are considered ‘likely’ (i.e. 66-90% chance) over most mid-latitude continental interiors.

While reflecting the advances that have been made in the study of observed and projected changes in extremes since the Second Assessment, the Third Assessment Report stresses the continuing problems and uncertainties associated with extreme events and recommends further research effort to address these issues, which are discussed further in Section 4 of this report.

2.4 UK scenarios for impact assessment

The University of East Anglia (through the Climatic Research Unit and Tyndall Centre) has been involved in the production of climate change scenarios for Tyndall Centre stakeholders. In particular, it produced the UKCIP98 climate change scenarios for the United Kingdom (Hulme and Jenkins, 1998). More recently, regional climate change scenarios for Scotland (Hulme *et al*., 2001) have been produced for the Scottish Executive and the UKCIP2002 climate change scenarios for the United Kingdom will be published in February 2002. The information on extreme events provided by these scenarios is summarised below.

UKCIP98 scenarios (Hulme and Jenkins, 1998):

- **Based on direct HadCM2 output.**
- Changes in interannual and day-to-day variability of temperature and precipitation are described.
- Changes in some seasonal extremes are presented in tables, i.e., ‘hot’ summers, ‘warm’ years, ‘dry’ summers and ‘dry’ years, defined using the monthly Central England Temperature and England and Wales Precipitation records.
- Maps are shown for accumulated degree days (Tmin <0°C, Tmax >25°C, Tmean >5.5°C).
- Probabilities of wind speeds >10 m/s and daily precipitation totals >5 mm are also considered.
- Some ‘derived’ indices are presented, i.e., gales/severe gales/very severe gales and air-flow indices.

Regional climate change scenarios for Scotland (Hulme *et al*., 2001)

- **Based on direct HadRM2 output.**
- Maps are shown for changes in ‘intense’ daily precipitation thresholds defined using the quantile method and for changes in the frequency of these events.
- Maps are shown for changes in the 2-year return period daily precipitation intensity.
- Maps are shown for changes in ‘hot’ days defined using the percentile method.
- Maps are shown for changes in the 20-year return period daily mean temperature.
- Maps are shown for changes in ‘windy’ days defined using the percentile method.
- Maps are shown for changes in the 20-year return period daily mean wind speed

The UKCIP2002 scenarios which are currently being constructed, based on HadCM3 and HadRM3 output, will use a similar set of extreme event indicators to the recent Scottish regional scenarios.

Although the UKCIP98, UKCIP2002 and Scottish regional scenarios all provide some information about extremes, they do not consider the full range of indicators that may be of interest for impact assessment, particularly for future integrated assessment models (see Section 3.2) and the scenario construction methods used are not specifically designed for the construction of scenarios of extremes. The UKCIP2002 and Scottish regional scenarios are based on output from one dynamical downscaling model (i.e., HadRM3). They do not consider statistical downscaling (Section 4.2.4) or inter-model variability (Section 4.1.2). Nor will they provide daily station time series. Thus this Tyndall Centre project addresses a wider range of development issues than is possible for UKCIP2002 and will lead to advances in future stakeholder-led projects.

3. EXTREME WEATHER SCENARIO NEEDS OF THE TYNDALL CENTRE AND THE WIDER IMPACTS COMMUNITY

3.1 Introduction

One of the major objectives of the Tyndall Centre research programme is to develop a range of new and existing integrated assessment methodologies. These make new and very specific demands of scenario construction methods, which are outlined in Section 3.2.

It is recognised that the extreme events for which scenarios are required need to be stakeholder-defined (Beersma *et al*., 2000). Thus this issue will be addressed by the project workshop which will have strong involvement of stakeholder representatives (Section 3.3). However, in order to construct more robust and reliable scenarios, it is useful to consider a number of climatic-based criteria for selecting indicators of extremes (Section 3.4). On the basis of these criteria, a number of potential indicators of temperature and rainfall weather extremes are proposed (Section 3.5). Feedback on these proposed indicators was sought from stakeholders and other participants in the project workshop.

3.2 The extreme weather scenario needs of integrated assessment

One of the essential characteristics of integrated assessment (IA) is the simultaneous consideration of the multiple dimensions of an environmental problem, in this case climate change. A number of formal IA models for climate change have been designed over the last decade, IMAGE 1.0 and ESCAPE being perhaps the first two in the early 1990s. The essence of these types of IA models is that they contain modules that are reduced-form versions of more complex simulation models – whether, for example, of the economy, of the climate system or of ecosystems. Quite often the climate modules, or ‘engines’, involved generate one (global; e.g., PAGE) or two (zonal; IMAGE) dimensional descriptions of future climate usually at mean (e.g. 30-year average) seasonal or annual resolution. Some IA models (e.g., IMAGE 2.4 and AIM) are then capable of generating spatially explicit descriptions of future climate, usually by accessing stored patterns of climate change extracted from more complex GCM experiments.

These different approaches to generating future climate descriptions in IA models are versatile, efficient and allow multiple experiments to be easily conducted in an integrated framework. The climate drivers may then be input into an ecosystem, agriculture or health module (e.g., AIM), or used directly to calculate estimates of climate damage due to a look-up climate damage function (e.g., DICE). In either case, the lack of any information about changes in daily or extreme weather in these IA scenario generators is potentially a major constraint on their application. Agriculture, for example, may well be more sensitive to changes in daily weather sequences than to changes in mean monthly climate; climate damage functions that express the economic impact of climate change as a function of global- (or regional-) mean climate are likely to underestimate the economic damage associated with climate change. The lack of daily weather scenarios in IA models also greatly limits what may be incorporated into them about simulating the adaptive process in either social institutions or in environmental systems. Yet including an assessment of adaptive capacity remains an important objective of third generation IA models.

The above summary indicates clearly that it would be desirable to have efficient and robust algorithms that would allow daily weather scenarios to be generated inside an IA model, driven perhaps by one or more large-scale indicators of future climate generated by the climate engine of the IA model. These daily weather scenarios should be spatially explicit, capable of transferring across a range of spatial scales, and maintain realistic temporal and spatial autocorrelation. Designing and implementing such an algorithm(s) may well be one of the objectives of the Tyndall Centre integrated assessment programme.

3.3 Questions for the impacts community

One of the aims of the project workshop (4-5 June 2001) was to identify the range of scenario information required (i.e., what are the important extremes?). In particular, the workshop addressed the questions listed below.

- What are the important **temperature and precipitation extremes** for particular impact studies?
- On what **temporal scale(s)** are scenarios of extremes required for particular impact studies (e.g., daily, sub-daily)? It is generally assumed that information at the daily time scale is necessary to investigate extreme events. However, are there extreme events which can be usefully defined using monthly data? Monthly data are more widely available from climate models (e.g., from the IPCC DDC), facilitating the investigation of inter-model variability (see Section 4.1.2), and are considered to be more reliably simulated by climate models than daily data.
- At what **spatial scale(s)** are scenarios of extremes required for particular impact studies?
- What are the important **non-temperature/precipitation extremes** for particular impact studies (e.g., wind, fog, lightning, storm surges)?
- How important are **joint-probability events** (e.g., wind storms with snow/rain, heavy snow followed by rapid thaw, intense rainfall on dry/frozen or already saturated ground, storm surge with river flood)?
- How important is it to know about the **persistence and sequence** of extreme events (e.g., sequences of long dry/hot summers)?
- How important is it to know about **seasonal changes in the timing** of extremes (e.g., changes in the season of maximum frequency of occurrence)?

- For what extremes and impacts is it important to have **self-consistent multi-site and/or multi-variate scenarios**?

- How should scenarios of extremes be **presented** for particular impact studies (e.g., maps, probability distributions)? Is it sufficient to provide information about relative changes or should these be added to an observed base-line climatology? Are daily time series required for input to some impact studies/models?

- **How much data** can realistically be handled in impact assessments (daily, high spatial resolution data sets for a number of different extreme parameters/scenarios/ensembles will be very large)?

- Is it possible to identify a **standard set of extremes** of interest to the widest possible range of impact assessment sectors?

- How should the uncertainties by represented (e.g., should **probabilities** be attached to the scenarios)?

- What **low-probability high-impact events** (see Section 5) should be considered (e.g., abrupt reorganisation of the thermohaline circulation, collapse of the West Antarctic ice sheet, large and rapid releases of methane trapped below the seafloor and in permafrost)?

3.4 Climatic criteria for selecting indicators of extremes

Four climatic-based criteria have been identified for selecting indicators of extremes which should help to ensure that more useful, robust and reliable scenarios can be constructed:

- Indicators should be based on climatic variables for which appropriate high-quality and long observed time series are available for the construction of base-line climatologies, climate model validation and construction of statistical downscaling models.

- In order to reduce problems associated with statistical uncertainty and sample size, relatively moderate definitions of extremes, rather than the most ‘extreme’ extremes, are preferred.

- A range of indicators which are not highly correlated and which therefore provide information about different aspects of changes in extremes (e.g., changes in different seasons, changes in the magnitude of individual events and changes in persistence) should be used.

- In order to facilitate comparison of projected and observed changes in extremes, indicators used for scenario construction should be consistent with those used in observational studies, particularly recent and ongoing studies moving towards a global synthesis of changes in extremes.

3.5 Proposed indicators of temperature and rainfall weather extremes

Based on the climatic selection criteria listed in the previous section, a number of indicators of temperature and rainfall weather extremes are proposed (Table 3.1). It is intended to use a number of these indicators during the quantitative testing of the most promising scenario development methods (Section 6).

Table 3.1: Proposed indicators of temperature and rainfall weather extremes (based on Karl et al., 1999; Frich et al., 2001; Hulme et al., 2001).

Temperature extremes
• Daily measures of Tmax/Tmin/Tday using 10th/90th percentiles
• Diurnal temperature range
• Intra-annual extreme temperature range (difference between the highest temperature value of any given calendar year and the lowest value of the same calendar year)
• Number of frost days (Tmin < 0°C)
• Frost severity index (percentage of time with Tmin < 0°C)
• Growing season length (period between Tday > 5°C for > 5 days and Tday < 5°C for > 5 days)
• Heat wave duration index (maximum period > 5 consecutive days with Tmax > 5°C above long-term mean – or based on percentile threshold)
• Magnitude of the 2, 5, 10 and 20 year return period event
Indices calculated from daily rainfall data
• Simple daily intensity index (total rainfall/number of rain days)
• Magnitude of the 90th/95th percentile/quantile
• Frequency of exceeding 90th/95th percentile/quantile (number of days)
• Percentage of rainfall falling on days with rainfall above 90th/95th percentile/quantile
• Percentage of region with rainfall in lowest/highest 5/10% percentile/quantile
• Maximum length of wet/dry spell
• Number of days with rainfall ≥ 10 mm
• Maximum 5-day rainfall total
• Magnitude of the 2, 5, 10 and 20 year return period event

Percentiles calculated by fitting the (gamma) distribution using the maximum likelihood method.
Quantiles calculated using the approach of Osborn et al. (2000).
All indices calculated for calendar year and standard seasons using daily data.

3.6 Comparability of observed and scenario extremes

It is important to recognise the difference between grid-box (GCM or RCM) and station values, particularly for precipitation. While it has been suggested that simulated precipitation can be interpreted as point values (Skelly and Henderson-Sellers, 1996), it is more generally accepted that what are being simulated are actually true (i.e.

infinitely-sampled) area-average means for each grid box (Osborn, 1997; Osborn and Hulme, 1997; 1998). Statistical techniques have been developed to construct 'true' area-average means from observed precipitation data for use in climate model validation studies (Osborn, 1997; Osborn and Hulme, 1997). Similarly, recent studies of observed changes in extremes focus on area-averaged, rather than station, time series (Karl *et al.*, 1999; Easterling *et al.*, 2000a,b; Frich *et al.*, 2001).

The different nature of grid box and point values raises questions about the representativeness and reliability of scenarios derived solely from raw GCM output. Particularly for extreme events on daily timescales, GCMs generate variables that are quite different from those measured at individual stations (Hulme and Jenkins, 1998). In order to produce more directly comparable statistics, some form of either downscaling or disaggregation is required.

A number of questions concerning the issue of grid box *vs* point values are raised below. The first was discussed briefly during the project workshop, while the others are considered in Section 4 and will be explored further as part of the case-study work (Section 6).

- How useful are statistics or extremes of spatially-aggregated variables for impact studies?

- How much information is lost, particularly at the tails of the distributions, in spatial averaging?

- To what extent do RCMs (i.e., dynamical downscaling methods) overcome the problems of grid box *vs* point values?

- Can statistical downscaling be used to produce more reliable station/point values of extremes?

4. REVIEW OF SCENARIO DEVELOPMENT METHODS

4.1 Suitability of scenarios of extremes

4.1.1 Introduction

Chapter 13 of the IPCC TAR lists five criteria, adapted from Smith and Hulme (1998), for assessing the suitability of each type of climate change scenario for use in impact assessment:

1. Consistency at regional level with global projections. Scenario changes in regional climate may lie outside the range of global mean changes but should be consistent with theory and model-based results.

2. Physical plausibility and realism. Changes in climate should be physically plausible, such that changes in different climatic variables are mutually consistent and credible.

3. Appropriateness of information for impact assessments. Scenarios should present climate changes at an appropriate temporal and spatial scale, for a sufficient number of variables, and over an adequate time horizon to allow for impact assessments.

4. Representativeness of the potential range of future regional climate change.

5. Accessibility. The information required for developing climate scenarios should be readily available and easily accessible for use in impact assessments.

These criteria are all applicable to scenarios of extremes. The appropriateness of information, and the associated issues of scale, are of particular concern for extremes. Extremes place additional demands on climate models and scenario development methods: higher-order statistics (such as standard deviations and skewness), together with the tails of distributions, not just mean values, must be well reproduced. Different extreme events/variables exhibit different characteristics that must be reliably captured by climate-change scenarios, and hence place different demands upon climate models and scenario development methods. For example, the demands arising from the need to reproduce daily precipitation totals which are typically described by a mixed statistical distribution are different to the case of variables with a quasi-Gaussian distribution such as daily temperature. In both cases, extreme events are highly dependent on spatial scale, while multi-month drought occurrence is less so, but instead introduces a need to reproduce the correct persistence levels. Certain sectors also require that realistic inter-site relationships are maintained (e.g., hydrological modelling), while others (e.g., certain crop models; Semenov and Brooks, 1999) require realistic inter-variable relationships. Other sectors may require joint probabilities (Section 3.3).

While the need to generate scenarios that successfully reproduce present-day climate variability and extremes and that also give reliable and plausible estimates of climate change is paramount, a number of other issues must also be addressed. Ideally, scenarios should have estimates of their associated uncertainty (Section 4.1.2) and should be able to be scaled to reflect a range of possible greenhouse gas emissions pathways (Section 4.1.3), allowing probabilistic impacts and integrated assessments to be undertaken (Section 3.2). The issues of uncertainty and scaling are discussed below, before summarising the advantages (✔) and disadvantages (✘) of various methods for the construction of scenarios of extremes (Section 4.2) in the light of these issues and the criteria identified above.

4.1.2 Uncertainty

There is growing recognition of the need to take into account the full range of uncertainties in scenario construction and impacts assessment and, at the same time, to distinguish between the inherent unpredictability of climate and model deficiencies (Hulme and Brown, 1998; Hulme and Carter, 1999; Hulme *et al*., 1999; Katz, 1999a; Mitchell and Hulme, 1999; Giorgi and Francisco, 2000a,b; Jones, 2000a,b; New and Hulme, 2000; Visser *et al*., 2000; Räisänen and Palmer, 2001). The IPCC TAR and many of the references cited above refer to a cascade of uncertainty related to:

- The forcing scenarios, i.e., inter-scenario variability.
- The use of different climate models, i.e., inter-model variability.
- Different realizations of a given scenario with a given climate model, i.e., internal model variability (which is, in part, a reflection of natural climate variability).
- Sub-grid scale forcings and processes.

Techniques for handling the first three sources of uncertainty have been identified, but are not yet widely used in impact assessments. In theory, these techniques, listed below, can be applied to extremes, although additional issues may arise:

- Uncertainties due to inter-scenario variability can be handled by using more than one forcing scenario and scaling climate model output (see Section 4.1.3).
- Uncertainties due to inter-model variability can be handled by using output from more than one climate model. For extremes, this depends on the availability of daily output from a number of different modelling centres. However, daily output is less widely available than monthly output (the IPCC DDC, for example, only holds monthly data). The case-study work undertaken as part of this project will be based on Hadley Centre model data available through the Climate Impacts LINK project (Section 6.5) and will not, therefore, explore uncertainties due to inter-model variability in any quantitative way. Methods of incorporating this potentially important source of uncertainty in future Tyndall Centre projects will, however, be considered.
- Uncertainties due to internal model variability and thus, in part, natural variability, can be handled by using intra-model ensembles (i.e., simulations performed with the same climate models and forcing, but starting from different initial conditions). It is, for example, proposed to use the three member ensemble of HadCM3 SRESA2 (and, if possible, HadRM3 SRESA2) simulations for the case-study work (Section

6.5). For the investigation of extremes, output from individual ensemble members must be pooled (e.g., 30 years of output from three ensemble simulations gives an intra-model ensemble of 90 years).

Comparative studies of the first three sources of uncertainty listed above indicate that, for mean climate, inter-model variability tends to be greater than inter-scenario or internal model variability, particularly over the earlier part of the 21st century, (Dutton and Barron, 2000; Giorgi and Francisco, 2000a,b; Bergstrom *et al*., 2001). However, different sources of uncertainty (e.g., internal model variability) may dominate for extremes.

The fourth source of uncertainty identified above, sub-grid scale forcings and processes, has not yet been adequately addressed in the literature, but may be particularly important for extreme events with high temporal and spatial resolutions. RCMs provide information at the sub-GCM grid scale, so ensemble RCM output provides one way of exploring this issue (Dutton and Barron, 2000). However, the current resolution of RCMs, 50 km x 50 km for HadRM3, is still relatively coarse for some extreme event processes, such as convective precipitation. Statistical downscaling methods (Section 4.2.4) provide station or point values and thus may provide another way of exploring this issue, but introduce additional uncertainties due to the methods themselves. Similarly, statistical manipulation methods which attempt to correct for model biases (Sections 4.2.2.3 and 4.2.3.3) are also likely to introduce new uncertainties. Thus techniques for handling uncertainties in sub-grid scale forcings and processes will be considered as part of the case-study work (Section 6).

Another aspect of uncertainty which needs to be addressed concerns the relationship between the future climate-change uncertainties discussed above and multi-decadal climate variability (Hulme *et al*., 1999), i.e., the issue of signal-to-noise ratios (which is important for determining the significance of projected climate changes and for detection and attribution studies). One approach is to quantify multi-decadal natural variability by taking 30-year time slices from the GCM control (i.e., unforced) simulation (Hulme and Brown, 1998; Hulme and Jenkins, 1998; Hulme *et al*., 1999). Hulme and colleagues, for example, used 240 or 1400 years of the HadCM2 control simulation for this purpose and will use a similar approach for the UKCIP2002 scenarios. The unforced 30-year time slices can either be used directly as input to impact models and the impacts of climate change and natural variability compared (Hulme *et al*., 1999) or used to define the range of natural climate variability (e.g., the two standard deviation range; Hulme and Brown, 1998; Hulme and Jenkins, 1998).

While the HadCM2 control simulation has been shown to contain substantial multi-decadal variability of mean climate (Hulme *et al*., 1999), the extent to which GCMs capture the multi-decadal variability of extreme events needs to be investigated. However, the shortage of reliable, sufficiently long daily records, makes it difficult to quantify this variability using observed data. Another issue which needs to be addressed is the legitimacy of quantifying natural variability using GCM output and then comparing this variability with climate changes derived from dynamical or statistical downscaling. No long RCM control simulations are available, so there may be no alternative in this case. However, where stochastic statistical downscaling methods are used (Section 4.2.4), it may be possible to generate multiple 30-year time slices for this purpose.

Changes in extreme events may be non-linear and greater than changes in mean climate (Mearns *et al*., 1997; Wagner, 1999). However, the natural variability of extremes is also greater than that of mean climate, thus the signal-to-noise ratio may be lower for extremes than for mean climate, making it more difficult to identify significant changes in extremes.

4.1.3 Scaling

Scaling the pattern of climate change taken from a GCM by global temperature change taken from a simple climate model, such as MAGICC, (Santer *et al.*, 1990; Hulme and Jenkins, 1998; Mitchell *et al*., 1999; Huntingford and Cox, 2000; Mitchell, 2001) allows a much wider range of forcing scenarios and climate sensitivities to be considered than is possible from the limited number of GCM simulations which have been performed using a limited range of emissions scenarios. Thus it provides a way of addressing uncertainties due to inter-scenario and inter-model variability (Section 4.1.2) and is also an approach that may be employed in integrated assessment models (Section 3.2).

Recent investigations of this technique indicate that it is a legitimate approach, certainly so far as mean climate is concerned (Mitchell *et al*., 1999; Huntingford and Cox, 2000; Mitchell, 2001). However, the applicability of this method to RCM output, to statistical downscaling and to extremes has not been investigated. While it is possible to scale the statistics derived from time series, and changes in these statistics, the question as to whether it is possible to scale the time series themselves is open to debate.

The most detailed assessment of pattern scaling to date is that of Mitchell (2001). A number of conclusions from this study which may be relevant to scaling extremes are listed below:

- The sources of error include non-linear global-mean responses to radiative forcing and non-linear grid-box responses to radiative forcing.

- It is possible to scale interannual variability (i.e., standard deviations), although the errors are larger than for mean values.

- If it is assumed that the variable in question has a Gaussian distribution, changes in the entire probability distribution may be estimated by individually scaling the mean and standard deviation. It may also be possible to scale non-Gaussian variables by estimating secondary parameters: for example, it may be possible to estimate precipitation by scaling the mean and standard deviation and estimating from them the alpha and beta parameters of the gamma distribution.

- It is better to take the pattern from a regression based on the full model period (240 years) rather than from a 20 year period at the end of the simulation. If there are any non-linearities in the responses, a regression based on the full model period will minimise the magnitudes of any instantaneous errors, though they may change sign over the course of the simulation period. A pattern taken from 240 years may also be defined more robustly than a pattern taken from a 20-year period.

- There is a weak linear relationship between signal-to-noise ratio and spatial scale, so signal-to-noise ratios are slightly larger on larger spatial scales. However, although the errors from scaling decrease with spatial scale, so does internal variability, so the statistically significant errors actually increase with spatial scale. Therefore, there is a trade-off between the size of errors and the robustness of the pattern as the spatial scale increases/decreases.

- It is better to scale down, rather than to scale up, i.e., to calculate the pattern for the largest possible global temperature change.

- It may be possible to incorporate non-linear responses by using a quadratic rather than linear regression or by using a second scaling parameter.

- It is, in theory, possible to statistically downscale from the scaled changes or to scale RCM patterns.

The legitimacy of using scaling techniques for extremes, which may change non-linearly (Mearns *et al*., 1997; Wagner, 1999), will be investigated as part of the case-study work (Section 6). These techniques do, however, have the major advantage of facilitating a probabilistic approach to scenario construction and integrated assessment (Section 3.2).

4.2 Scenario methods

4.2.1 Introduction

An overriding problem – and hence the need for more sophisticated scenario development methods – is that output from global climate model simulations cannot, in general, be used to directly quantify future variability and extremes, particularly at the station level, because of bias in simulated means and variability of present-day climate and weather (Zwiers and Kharin, 1998; Meehl *et al*., 2000). This bias may originate from systematic model errors (Gregory and Mitchell, 1995; Hennessy *et al*., 1997), from spatial scale incompatibilities (area-mean grid-box output has different statistical properties to station data; Osborn and Hulme, 1997; Section 3.6) and due to the exclusion of sub-grid-scale processes.

In the development of scenarios of mean climate change, the bias in simulated means is the main difficulty and can be "overcome" by assuming that the climate change is independent of these mean biases and, therefore, applying climate change fields to appropriate observed baseline climatologies. A similar approach can be used for the development of scenarios that focus on climate variability and extremes (Wilks, 1992; Bates *et al*., 1994), i.e., by assuming that changes in higher-order statistical parameters (variance, skewness, persistence, etc.) are reliable, despite differences between observed and simulated present-day values of these parameters (Sections 4.2.2.1 and 4.2.2.2). Additionally, it may be possible to use appropriate statistical manipulation to reproduce the present-day climate characteristics (Section 4.2.2.3). Alternatively, extremes and their changes can be defined in a relative rather than absolute sense (i.e., percentile or quantile values rather than absolute thresholds) in order to reduce model biases (Section 4.2.2.4). The advantages and disadvantages of these various approaches to using global model output are summarised below.

For global climate models to become more reliable in their simulation of variability and extremes as well as means, requires not only an improvement in the models, but also a solution to the two major spatial resolution problems of scale incompatibilities and sub-grid-scale processes. These can only be overcome by an increase in the spatial resolution of climate models together with improvement in their reliability. Higher-resolution global (such as timeslice, Cubasch *et al*., 1996) or regional (nested within a global model, Christensen *et al*., 1997; Jones *et al*., 1997) models have begun to address this problem. Output from regional climate models (RCMs) is more widely available than output from timeslice simulations and thus provides a more viable approach to dynamical downscaling from GCM output. It should not, however, be assumed that the higher resolution of RCMs automatically provides more meaningful or reliable spatial detail (Durman *et al*., 2001; Hulme *et al*., 2001; Rummukainen *et al*., 2001). The advantages and disadvantages of using RCM output for constructing scenarios of extremes are outlined in Section 4.2.3.

Statistical downscaling provides a less computer-intensive method of downscaling and the potential of this approach to the construction of scenarios of extremes is reviewed in Section 4.2.4.

In Sections 4.2.2 to 4.2.4, the following annotations are used:

✔ Advantage of the method
✘ Disadvantage of the method
? Advantage/disadvantage of the method is uncertain

4.2.2 Direct use of global climate model output.

✘ Spatial-scale problems arise, i.e., grid box rather than point values

✘ Even area-averaged extremes (i.e., grid-box values) may not be reliably simulated

✔ Provides physically-consistent multi-variate information

4.2.2.1 Diagnosed changes in statistical parameters (mean, plus higher-order parameters, such as variance, scale and shape, etc.) applied to observed baseline time series

✘ Non-realistic scenarios, e.g., negative precipitation, may occur when the changes are applied to the baseline climatology

✘ Assumes biases will be unchanged in the future

✔ Simple method

✔ Suitable for scaling

4.2.2.2 As 4.2.2.1, but changes are applied to weather generator parameters, previously tuned to reproduce observed climate

✘ Weather generators tend to underestimate variability and persistence, e.g., length of wet/dry spells

✘ May be difficult to adjust weather generator parameters in a consistent way

✔ Long and/or multiple time series can be generated for analysis of extremes/uncertainties

✔ Suitable for scaling

4.2.2.3 Direct model time series used, after appropriate statistical manipulation to reproduce present-day climate characteristics

✔ May overcome some model biases

✘ May be more difficult to manipulate extremes than mean values

✘ Assumes model biases will be unchanged in the future

? Either ‘un-intelligent’ or ‘informed’ manipulation may be applied, the latter using validation/statistical downscaling approaches to adjust model output for specific physically-identified biases

? Less suitable for scaling

4.2.2.4 Model output used to assess specific extremes (via percentile or extreme value distribution approaches), which are defined in a relative rather than absolute sense

✔ May overcome some systematic model deficiencies and facilitates model inter-comparisons

✔ May overcome some spatial-scale incompatibilities

✘ Assumes model biases will be unchanged in the future (because percentiles or thresholds are defined from the control period)

? Less suitable for scaling

4.2.3 Direct use of regional climate model output

✘ Affected by biases in the underlying GCM

✔ Provides physically-consistent multi-variate information

✔ Higher spatial resolution should reduce biases (e.g., more intense extremes)

✘ Relatively short runs make it difficult to assess multi-decadal natural variability

✘ Relatively few simulations/ensembles available

? Added value of higher resolution needs to be demonstrated

? Scaling may be more difficult, in part, because of shorter model simulations

4.2.3.1 Diagnosed changes in statistical parameters (mean, plus higher-order parameters, such as variance, scale and shape, etc.) applied to observed baseline time series

✔/✘ See 4.2.2.1

4.2.3.2 As 4.2.2.1, but changes are applied to weather generator parameters, previously tuned to reproduce observed climate

✔/✘ See 4.2.2.2

4.2.3.3 Direct model time series used, after appropriate statistical manipulation to reproduce present-day climate characteristics

✔/✘ See 4.2.2.3

4.2.3.4 Model output used to assess specific extremes (via percentile or extreme value distribution approaches), which are defined in a relative rather than absolute sense

✔/✘ See 4.2.2.4

4.2.4 Statistical downscaling of climate model output.

4.2.4.1 Introduction

A range of statistical downscaling methods has been developed in recent years (see reviews by Hewitson and Crane, 1996; Wilby and Wigley, 1997; Wilby *et al*., 1998; Wilks and Wilby, 1999; Zorita and von Storch, 1999; Goodess, 2000), though these have not been designed specifically for downscaling extremes (Section 2.1). Nonetheless, these methods provide an alternative approach to obtaining information about climate variability and extremes. Relationships between larger-scale climate variables (such as atmospheric circulation) and local surface climate variables (such as daily temperature and precipitation), derived empirically using observed data, can be applied to the generation of climate-change scenarios, under the two assumptions that the larger-scale climate variables are more reliably simulated by climate models, and that the relationships remain valid under a changed climate. Theoretically, the latter assumption (of stationarity) should be valid if all the necessary predictor variables [such as atmospheric circulation, temperature and humidity (Buishand and Brandsma, 1999; Wilks and Wilby, 1999)] are used. In practice, however, this may be limited by the availability of sufficiently long data series to determine the important predictors on all necessary time scales (a problem that is exacerbated if the variables that generate inter-daily to inter-annual variability are different to those that cause climate change).

Nevertheless, given adequate data, statistical downscaling has sufficient advantages to warrant consideration as a scenario-generation method.

From the range of available methods, two approaches (Goodess, 2000; Palutikof *et al.*, 2001) are considered to be most appropriate for the scenario needs of the Tyndall Centre, particularly the need for probabilistic scenarios, and are assessed in Sections 4.2.4.2 and 4.2.4.3. First, however, the common advantages/disadvantages of all statistical downscaling methods are summarised:

✔ Provide station/point values of extremes

✔ Less computer intensive than dynamical downscaling

✔ Can be applied to GCM and/or RCM output

✘ Assume that predictor/predictand relationships will be unchanged in the future (the stationarity issue)

✘ Require long/reliable observed data series

✘ Affected by biases in the underlying GCM

? May be possible to correct predictors for systematic model biases

? Scenarios may indicate changes which differ substantially in magnitude, and even in direction, from those based directly on model output

? Ideally downscaling methods should reflect the underlying physical mechanisms and processes, but statistical downscaling is unlikely, for example, to treat convective rainfall events in a physically realistic way

? Suitability for scaling needs to be investigated

4.2.4.2 Resampling of observed data conditioned by large-scale climate variables

✔ Provides self-consistent multi-site, multi-variate scenarios

✔ Multiple time series can be generated

✔ Relatively simple method

✘ Magnitude of the largest extreme is limited by the observations

✘ Difficult to extend to multiple predictors if sample size is limited

4.2.4.3 Weather generator, with parameters conditional upon similar large-scale climate variables

✔ Long/multiple time series can be generated

- ✘ Variability and persistence tend to be underestimated [the overdispersion problem (Katz and Zheng, 1999)]
- ? Methods are being developed for the production of self-consistent multi-site, multi-variate scenarios (Goodess, 2000), but tend to be more complex and subject to technical/statistical problems

4.2.5 Cross-cutting issues

There are certain cross-cutting issues that must be considered for all methods, related to the implicit or explicit 'building-in' of certain statistical properties derived from the present-day climate. If these properties are prevented from changing as climate changes, by constraints in the structure of the scenario method, then the climate change scenario may be biased. This is a particularly important issue for scenarios of extremes because such constraints are more likely to be imposed out of necessity than when developing scenarios of mean climate. For example, the manipulation of model output to resemble station observations (Section 4.2.2.3) is necessary to overcome model bias and to relate an area-mean behaviour to a point behaviour. The latter adjustment depends upon the spatial coherence of, for example, precipitation, and this may vary with climate change (Osborn, 1997). Thus, the statistical manipulation might be inappropriate under climate change. This could be tested by evaluating whether there is a change in spatial coherence in RCM simulations. There are similar properties that might be built implicitly into any of the scenario methods and which should, where possible, be assessed.

4.3 Non-temperature/precipitation extremes

The methods considered in Section 4.2 provide a range of possible approaches for the construction of scenarios of temperature and precipitation-based extremes (such as those listed in Table 3.1). However, scenarios of other types of extreme event may also be required for impact assessments. The IPCC TAR, for example, identifies other extreme phenomena (such as cyclonic storms, together with very small-scale phenomena such as thunderstorms, tornadoes, hail and lightning which are not simulated in climate models), many of which may have important impacts on the environment and society, but for which there is currently insufficient information to assess recent trends and for which climate models currently lack the spatial detail to make confident projections. Some of these phenomena, such as hurricanes, are not highly relevant to the UK, while others, such as North Atlantic cyclonic storms, clearly are.

A number of recent modelling studies have investigated potential changes in the occurrence of North Atlantic cyclones and storm tracks (Carnell and Senior, 1998; Schubert *et al*., 1998; Ulbrich and Christoph, 1999; Knippertz *et al*., 2000). This issue is also the focus of Tyndall Centre project IT1.4 which will carry out an integrated assessment of the potential for change in storm activity over Europe, focusing on the implications for insurance and forestry in the UK. Storm activity will not, therefore, be considered as part of this project.

One of the questions to be addressed by the project workshop was, what are the important non-temperature and precipitation extremes for particular impact studies (e.g., wind, lightning, fog, storm surges)? The final project report will include an assessment of the extent to which it may be possible to provide information about each of the extremes which are identified as being important. Suitable approaches to scenario construction will be proposed and, where possible, assessed. Approaches to the treatment of low-probability high-impact events, such as thermohaline circulation reorganisation, are discussed in Section 5.

5. LOW-PROBABILITY HIGH-IMPACT EVENTS

5.1 Introduction

Climate change assessment studies, particularly probabilistic studies and integrated assessments, should, but do not yet (with rare exceptions, e.g., Keller *et al.*, 2000), consider scenarios of low probability but high impact events, such as an abrupt reorganisation of the thermohaline circulation or a collapse of the West Antarctic ice sheet, that could arise due to non-linearities in the climate system (Hulme and Carter, 1999).

Such events have also been referred to as climate 'surprises' and there has been a tendency to focus on 'surprises' with negative rather than positive impacts (Schneider and Root, 1996; Jones, 2000a; Streets and Glantz, 2000; Visser *et al.*, 2000). The concept of climate 'surprise' is somewhat subjective and does not distinguish between events which are truly unpredictable and those which could be anticipated (Streets and Glantz, 2000). Thus the term low-probability high-impact event is preferred (Jones, 2000a).

The IPCC TAR identifies a number of such events:

- Abrupt reorganisation of the thermohaline circulation.
- Collapse of the West Antarctic ice sheet.
- Fast changes to the carbon or methane cycle, such as large and rapid releases of methane trapped below the sea floor and in permafrost.

There are two aspects of such events that must be addressed: (i) the probability of the event occurring, and (ii) the response of the climate system to the event. These aspects are somewhat better, but still not well, understood for thermohaline reorganisation (Section 5.2) and West Antarctic ice sheet collapse (Section 5.3) than for abrupt carbon or methane cycle changes (Buffett, 2000).

5.2 Abrupt collapse of the thermohaline circulation

Many experiments have been performed with the most recent generation of models to explore the sensitivity of the thermohaline circulation system to freshening and increased meltwater discharges in the North Atlantic (Tziperman, 1997; Rahmstorf and Ganopolski, 1999; Wang *et al.*, 1999a,b; Aeberhardt *et al.*, 2000; Cubasch *et al.*, 2000; Manabe and Stouffer, 2000; Rahmstorf, 2000; Ganopolski and Rahmstorf, 2001; Hall and Stouffer, 2001).

The IPCC TAR concludes that most models show weakening of the ocean thermohaline circulation over the 21st century. The model projections reviewed in the TAR do not indicate a complete shut-down of the thermohaline circulation by 2100 but it is acknowledged that the thermohaline circulation could shut-down completely if the change in radiative (and, by implication, freshwater) forcing is large enough and applied for long enough. Thus, there is a necessity to determine how large and persistent the forcing needs to be to cause the thermohaline circulation to collapse.

Keller *et al*. (2000) have estimated critical atmospheric CO_2 concentrations (e.g., 776 ppmv for a climate sensitivity of 3.5°C) beyond which the thermohaline circulation is 'supposed' to collapse, based on the model results of Stocker and Schmittner (1997). Other model results, however, are likely to provide different estimates of the critical concentration. Thus one approach might be to construct a distribution function of critical thresholds based on published studies and/or on expert judgement. Atmospheric CO_2 concentration may not, however, be the most suitable or feasible parameter to use for this purpose.

The impacts of thermohaline circulation changes on Northwest European climate have not been widely explored. In the climate models reviewed for the IPCC TAR where the thermohaline circulation weakens but does not collapse, there is still a warming over Europe. However, in simulations in which the thermohaline circulation is forced to collapse, cooling and drying occur over Europe (Klein Tank and Können, 1997; Rahmstorf and Ganopolski, 1999; Vellinga and Wood, 2001). While these experiments provide useful guidance as to the climatic changes that might be expected due to the collapse of the thermohaline circulation, the published results do not provide any information about extreme events. Note also, that the recent HadCM3 experiment (Vellinga and Wood, 2001), does not incorporate greenhouse gas forcing. Even if daily output from all the completed experiments was available for use in scenario construction, the limited number of simulations makes it difficult to investigate uncertainty due to internal and inter-model variability (Section 4.1.2). The probability of thermohaline circulation collapse is dependent on both the magnitude and rate of warming, thus it is considered particularly important to consider multiple forcing scenarios and ensembles (Mitchell and Hulme, 1999).

5.3 West Antarctic ice sheet collapse

The stability of the West Antarctic ice sheet during a period of global warming is a matter of concern because it is grounded below sea level. It is, therefore, the subject of many glaciological studies and reviews (Bentley, 1997; Bentley, 1998; Bindschadler, 1998; Oppenheimer, 1998; Conway *et al*., 1999; Thorne *et al*., 2000; Poore and Dowsett, 2001; Vaughan and Spouge, 2001).

The IPCC TAR concludes that loss of grounded ice from the West Antarctic ice sheet is now widely agreed to be 'very unlikely' (i.e., 1-10% chance) during the 21st century, although its dynamics are still inadequately understood, especially for projections on longer time scales. However, the scientific uncertainty and unpredictability of the physical system are such that input from an interdisciplinary panel of experts recently evaluated by Vaughan and Spouge (2001) indicates a 5% probability of West Antarctic ice sheet collapse causing sea level rise of at least 10 mm/year within 200 years.

For the UK, the major impact of total collapse of the West Antarctic ice sheet would be the additional contribution of about 5 m to global sea level rise (which, depending on the location, would be amplified or reduced by local isostatic effects).

5.4 Methods for incorporating low-probability high-impact events

For UK impact assessments, collapse of the thermohaline circulation is considered to be the low-probability high-impact event of greatest concern and thus will be the focus of project work in this area. Within this Tyndall Centre project, it will not be possible to obtain improved estimates of either the probability of, or climatic response to, thermohaline collapse. Thermohaline circulation changes are the subject of current or planned major NERC (http://www.nerc.ac.uk/funding/thematics/prescient/index.shtml) and DETR initiatives which may eventually lead to such improved estimates. In the meantime, the project will consider how such information might be incorporated in scenarios and risk assessment studies. Thus this area of work will provide an interface between the basic science and the impact assessment research communities. It will require subjective expert judgement, based on literature review (Section 5.2), and, given the uncertainties and problems identified in Section 5.2, it is likely that the resulting scenarios will be 'conceptual' or semi-quantitative rather than quantitative. They will, however, take into account the scientific plausibility of the underlying assumptions. The extent to which the climatic response to thermohaline weakening can be treated as a linear or non-linear response, and hence the appropriateness of using pattern scaling techniques (Section 4.1.3), will also be explored.

The IPCC TAR uses a number of categories to indicate judgmental estimates of confidence, for example, 'very likely' is defined as 90-99% chance and 'very unlikely' as 1-10% chance. The applicability of such definitions, and methods by which they might be applied, in studies of low-probability high-impact events will be considered.

6. THE PROJECT CASE STUDIES

6.1 Introduction

Task 3 of the project is the ‘Quantitative testing of the most promising scenario development methods for selected case studies’ and will be carried out between July 2001 and March 2002. The most promising methods still have potential disadvantages (Section 4.2), thus the case-study work will investigate ways of overcoming or minimising these problems. Preliminary suggestions as to the indicators on which the case studies could be based (rather than the issues – which are listed in Section 4.2) were made at the project workshop. Guidance on the selection of case studies was sought from the participants and a number of additional proposals were made. While it will not be possible to carry out full case studies for all the proposed extremes (which included wind, floods, subsidence and snowmelt), the final project report will include recommendations on approaches for constructing scenarios of these events. The two case studies which it is proposed to carry out are outlined in Section 6.2.

The general methodology which will be used is outlined in Section 6.3, while the observed and model data sets that will be used are described in Sections 6.4 and 6.5 respectively.

6.2 The selected case studies

Two illustrative examples will be investigated using observed UK data (see Section 6.3), NCEP reanalysis data (Kalnay *et al*., 1996: available in the Climatic Research Unit for a European window, see http://www.cru.uea.ac.uk/cru/data/ncep/), and Hadley Centre global and regional climate model output (see Section 6.4). Each case study will focus on a particular event or extreme indicator, or suite of multi-variates, and on a limited number of UK study areas.

The first case study will focus on drought because the other relevant Tyndall Centre RT3 projects (IT1.4 on storm activity over Europe and IT1.8 on temperature extremes) do not cover this important extreme. It is likely that climate model output and station data for two UK study regions representative of lowland and upland climate regimes (~10-15 stations within each ~200-300 km by 200-300 km region) will be used to construct scenarios of drought for the present day and the future (i.e., the period 2071-2100) using the most promising methods identified in Section 4 and during the project workshop. Monthly, rather than daily, indices are considered appropriate for this case study. A drought index based on precipitation only will be used, together with one employing additional variables such as soil moisture.

The second case study will be based on changes in the intensity of daily precipitation. On average, the intensity of daily precipitation has increased over the UK in winter, and decreased in summer, over the period 1961-1995 (Osborn *et al*., 2000). More intense precipitation events are considered ‘very likely’ over many areas during the 21st century (Section 2.3). However, confidence in regional projections is currently much lower. Thus the case-study work will test methods for the construction of more reliable, quantitative scenarios of high rainfall events, focusing on indicators which are most relevant to the occurrence of flooding (which is of particular concern following the October 2000 UK floods). Although this case study will be based on intense precipitation scenarios, the focus will be on joint probability events, i.e., the joint probability of high river flows, high tides and storm surges, which are likely to result in severe flooding. Thus this case study will require hydrological modelling and input from experts on storm surges and sea-level change.

6.3 The case-study methodology

Given the proposed spatial scales of the case studies, some form of downscaling (dynamical or statistical) will be required. However, it will also be useful to compare downscaled scenarios with those based directly on GCM output in order to determine the added value, if any, of downscaling.

The performance of each method will be validated by comparing the results with station data, focusing on the ability to reproduce the statistics of observed indicators (such as the persistence of precipitation) and the frequency of events (with varying magnitude or return period definitions). The plausibility of the future scenarios will be evaluated by reference to (i) results from the present-day validation studies, (ii) inter-comparison of the scenarios obtained via the various methods, and (iii) analysis of the ability of the climate models to reproduce the statistics and inter-relationships of the observed predictor variables. An example of the latter might be a comparison of the relationship between specific humidity and precipitation, found to be simulated too strongly in some climate models (Murphy, 1999; Wilby and Wigley, 2000), with obvious implications for future scenarios.

Whilst it is not the aim of this project to develop new statistical methods for the analysis of extremes, care will to be taken to ensure that appropriate methods are used for the case studies, for example, for choosing and fitting extreme value distributions (Katz, 1999b; Brabson and Palutikof, 2001). Recent developments in the statistics of extremes and their relevance to scenario construction will be reviewed in the final project report.

The case-study work will contribute to the assessment of the suitability of each scenario construction method and feed into the development of guidelines to underpin the future development of scenarios of extremes.

6.4 Observed data

Previous and current UK scenario studies (Section 2.4) use monthly gridded temperature and precipitation data sets as baseline climatologies. The UKCIP2002 scenarios, for example, will use a new 5 km x 5 km monthly climatology. Gridded data sets tend to average out local extremes and thus are likely to be of less value for the construction of scenarios of extremes, particularly where station or point values are required (Section 3.6).

Thus the case-study work will make extensive use of monthly and daily station data held in the Climatic Research Unit. For example, the data set of 110 daily precipitation records for 1961-1995 used by Osborn *et al.* (2000) is available for the second case study which is based on intense precipitation. A number of these precipitation records have recently been updated and extended further back in time. Most of these stations also have Tmin and Tmax records for 1961-1995. In order to investigate changes over longer time periods, use may also be made of the monthly Central England Temperature (CET) and England and Wales Precipitation (EWP) series.

6.5 Climate model data

Daily output from the Hadley Centre global coupled and regional climate models available through the Climate Impacts LINK Project will be used for the case-study work.

It is proposed to use output from the following recent simulations performed with the HadCM3 global coupled atmosphere ocean model using two of the SRES emissions scenarios (A2 and B2):

- HadCM3SRESA2a
- HadCM3SRESA2b
- HadCM3SRESA2c
- HadCM3SRESB2a

These simulations were run for the period 1861-2100. Two 30-year periods, 1961-1990 and 2071-2100 will be used for scenario construction. For the calculation of some extreme indicators, such as the magnitude of the 20-year return period event, it would be preferable to use longer time series. However, if longer time slices are extracted from transiently-forced simulations, the data may be strongly non-stationary, breaking the assumptions underlying a number of statistical methods for extremes. A 30-year time period provides a reasonable balance between these two considerations, as well as being widely used in other UK scenario studies (Section 2.4). Output from the earlier part of the simulations (1861-1960) will be used to provide information about natural variability (Section 4.1.2). It may also be desirable to use a longer time slice for scaling model output (Section 4.1.3).

The HadCM3 simulations which it is proposed to use include a three member ensemble for the SRES A2 scenario together with one simulation for a second scenario, SRES B2. Thus these four simulations provide some information about uncertainties arising from internal model variability and inter-scenario variability (Section 4.1.2). The SRESA2 simulations, for example, provide a 90-year intra-model ensemble. They do not, however, provide any information about inter-model variability.

In order to allow direct comparison of scenarios based on GCM and RCM output, it is hoped that daily output from the HadCM3 European regional climate model (HadRM3), which has a spatial resolution of 50 km x 50 km, forced by the above global model simulations can be used for the case-study work. Otherwise, output from the earlier HadCM2 RCM will be used.

7. REFERENCES

Aeberhardt, M., Blatter, M. and Stocker, T.F., 2000: 'Variability on the century time scale and regime changes in a stochastically forced zonally averaged ocean-atmosphere model', *Geophysical Research Letters*, **27**, 1303-1306.

Alexander, L.V. and Jones, P.D., 2001: 'Updated precipitation series for the UK and discussion of recent extremes', *Atmospheric Science Letters*, **1**, doi:10.1006/asle.2001.0025.

Bardossy, A. and Plate, E.J., 1991: 'Modeling daily rainfall using a semi-Markov representation of circulation pattern occurrence', *Journal of Hydrology*, **122**, 33-47.

Bardossy, A. and Plate, E.J., 1992: 'Space-time model for daily rainfall using atmospheric circulation patterns', *Water Resources Research,* **28**, 1247-1259.

Bates, B.C., Charles, S.P., Sumner, N.R. and Fleming, P.M., 1994: 'Climate change and its hydrological implications for South Australia', *Transactions of the Royal Society of South Australia*, **118**, 35-43.

Bates, B.C., Charles, S.P. and Hughes, J.P., 1998: 'Stochastic downscaling of numerical climate model simulations', *Environmental Modelling and Software*, **13**, 325-331.

Beersma, J.J. and Buishand, T.A., 1999: 'A simple test for equality of variances in monthly climate data', *Journal of Climate*, **12**, 1770-1779.

Beersma, J., Agnew, M., Viner, D. and Hulme, M., (eds), 2000: *Climate Scenarios for Water-Related and Coastal Impacts*, ECLAT-2 KNMI Workshop Report No. 3, KNMI, the Netherlands, 10-12 May 2000, Climatic Research Unit, UEA, Norwich, UK.

Bentley, C.R., 1997: 'Rapid sea-level rise soon from West Antarctic ice sheet collapse?', *Science*, **275**, 1077-1078.

Bentley, C.R., 1998: 'Rapid sea-level rise from a West Antarctic ice-sheet collapse: a short-term perspective', *Journal of Glaciology*, **44**, 157-163.

Bergström, S., Carlsson, B., Gardelin, M., Lindström, G., Pettersson, A. and Rummukainen, M., 2001: 'Climate change impacts on runoff in Sweden – assessments by global climate models, dynamical downscaling and hydrological modelling', *Climate Research*, **16**, 101-112.

Bindschadler, R., 1998: 'Future of the West Antarctic ice sheet', *Science*, **282**, 428-429.

Boer, G.J., Flato, G.M. and Ramsden, D., 2000: 'A transient climate change simulation with greenhouse gas and aerosol forcing: projected climate to the twenty-first century', *Climate Dynamics*, **16**, 427-450.

Bogardi, I., Matyasovszky, I., Bardossy, A. and Duckstein, L., 1993: 'Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns', *Journal of Geophysical Research*, **98**, 16653-16667.

Brabson, B. and Palutikof, J.P., 2001: 'The evolution of extreme temperatures in the Central England temperature record', in preparation.

Brandsma, T. and Buishand, T.A., 1998: 'Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling', *Hydrology and Earth System Sciences*, **2**, 195-209.

Buffet, B.A., 2000: 'Clathrate hydrates', *Annual Reviews of Earth and Planetary Science*, **28**, 477-507.

Buishand, T.A. and Beersma, J.J., 1996: 'Statistical tests for comparison of daily variability in observed and simulated climates', *Journal of Climate*, **9**, 2538-2550.

Buishand, T.A. and Brandsma, T., 1999: 'Dependence of precipitation on temperature at Florence and Livorno (Italy)', *Climate Research*, **12**, 53-63.

Carnell, R.E. and Senior, C.A., 1998: 'Changes in mid-latitude variability due to increasing greenhouse gases and sulphate aerosols', *Climate Dynamics*, **14**, 369-383.

Charles, S.P., Bates, B.C. and Hughes, J.P., 1999: 'A spatiotemporal model for downscaling precipitation occurrence and amounts', *Journal of Geophysical Research*, **104**, 31657-31669.

Christensen, J.H., Machenhauer, B., Jones, R.G., Schär, C., Ruti, P.M., Castro, M. and Visconti, G., 1997: 'Validation of present-day regional climate simulations over Europe: LAM simulations with observed boundary conditions', *Climate Dynamics*, **13**, 489-506.

Conway, D. and Jones, P.D., 1998: 'The use of weather types and air flow indices for GCM downscaling', *Journal of Hydrology*, **212-213**, 348-361.

Conway, H., Hall, B.L., Denton, G.H., Gades, A.M. and Waddinton, E.D., 1999: 'Past and future grounding-line retreat of the West Antarctic ice sheet', *Science*, **286**, 280-283.

Corte-Real, J., Xu., H. and Qian, B., 1999: 'A weather generator for obtaining daily precipitation scenarios based on circulation patterns', *Climate Research*, **13**, 61-75.

Cubasch, U., von Storch, H., Waszkewitz, J. and Zorita, E., 1996: 'Estimates of climate change in

Southern Europe derived from dynamical climate model output', *Climate Research*, **7**, 129-149.

Cubasch, U., Voss, R. and Mikolajewicz, U., 2000: 'Precipitation: a parameter changing climate and modified by climate change', *Climatic Change*, **46**, 257-276.

Dai, A., Wigley, T.M.L., Boville, B.A., Kiehl, J.T. and Buja, L.E., 2001: 'Climates of the twentieth and twenty-first centuries simulated by the NCAR climate system model', *Journal of Climate*, **14**, 485-519.

Delworth, T.L., Mahlman, J.D. and Knutson, T.R., 1999: 'Changes in heat index associated with CO_2-induced global warming', *Climatic Change*, **43**, 369-386.

Durman, C.F., Gregory, J.M., Hassell, D.C., Jones, R.G. and Murphy, J.M., 2001: 'A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates', *Quarterly Journal of the Royal Meteorological Society*, **127**, 1005-1015.

Dutton, J.F. and Barron, E.J., 2000: 'Intra-annual and interannual ensemble forcing of a regional climate model', *Journal of Geophysical Research*, **105**, 29523-29538.

Easterling, D.R., Meehl, G.A., Parmesan, C., Changnon, S.A., Karl, T.R. and Mearns, L.O., 2000a: 'Climate extremes: observations, modeling and impacts', *Science*, **289**, 2068-2074.

Easterling, D.R., Evans, J.L., Groisman, P.Y., Karl, T.R., Kunkel, K.E. and Ambenje, P., 2000b: 'Observed variability and trends in extreme climate events: a brief review', *Bulletin of the American Meteorological Society*, **81**, 417-425.

Frei, C., Schär, C., Lüthi, D. and Davies, H.C., 1998: 'Heavy precipitation processes in a warmer climate', *Geophysical Research Letters*, **25**, 1431-1434.

Frich, P., Alexander, L.V., Della-Marta, P., Gleason, B., Haylock, M., Klein Tank, A. and Peterson, T., 2001: 'Global changes in climatic extremes during 2nd half of the 20th century', *Climate Research*, accepted for publication.

Ganopolski, A. and Rahmstorf, S., 2001: 'Rapid changes of glacial climate simulated in a coupled climate model', *Nature*, **409**, 153-158.

Giorgi, F. and Francisco, R., 2000a: 'Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM', *Climate Dynamics*, **16**, 169-182.

Giorgi, F. And Francisco, R., 2000b: 'Evaluating uncertainties in the prediction of regional climate change', *Geophysical Research Letters*, **27**, 1295-1298.

Goodess, C.M., 2000: *The Construction of Daily Rainfall Scenarios for Mediterranean Sites Using a Circulation-type Approach to Downscaling*, PhD Thesis, University of East Anglia, Norwich.

Gregory, J.M. and Mitchell, J.F.B., 1995: 'Simulation of daily variability of surface temperature and precipitation over Europe in the current and $2xCO_2$ climates using the UKMO climate model', *Quarterly Journal of the Royal Meteorological Society*, **121**, 1451-1476.

Hall, A. and Stouffer, R.J., 2001: 'An abrupt climate event in a coupled ocean-atmosphere simulation without external forcing', *Nature*, **409**, 171-174.

Hay, L.E., McCabe, G.J., Wolock, D.M. and Ayers, M.A., 1991: 'Simulation of precipitation by weather type analysis', *Water Resources Research*, **27**, 493-501.

Hay, L.E., McCabe, G.J., Wolock, D.M. and Ayers, M.A., 1992: 'Use of weather types to disaggregate general circulation model predictions', *Journal of Geophysical Research,* **97**, 2781-2790.

Hayhoe, H.N., 2000: 'Improvements of stochastic weather data generators for diverse climates', *Climate Research*, **14**, 75-87.

Hennessy, K.J., Gregory, J.M. and Mitchell, J.F.B., 1997: 'Changes in daily precipitation under enhanced greenhouse conditions', *Climate Dynamics*, **13**, 667-680.

Hewitson, B.C. and Crane, R.G., 1996: 'Climate downscaling: techniques and application', *Climate Research*, **7**, 85-95.

Hughes, J.P. and Guttorp, P., 1994: 'A class of stochastic models for relating synoptic atmospheric patterns to regional hydrologic phenomena', *Water Resources Research,* **30**, 1535-1546.

Hughes, J.P., Lettenmaier, D.P. and Guttorp, P., 1993: 'A stochastic approach for assessing the effect of changes in synoptic circulation patterns on gauge precipitation', *Water Resources Research*, **29**, 3303-3315.

Hughes, J.P., Guttorp, P. and Charles, S.P., 1999: 'A non-homogeneous hidden Markov model for precipitation occurrence', *Applied Statistics,* **48**, 15-30.

Hulme, M.. and Brown, O., 1998: 'Portraying climate scenario uncertainties in relation to tolerable regional climate change', *Climate Research*, **10**, 1-14.

Hulme, M. and Carter, T.R., 1999: 'Representing uncertainty in climate change scenarios and impact studies', in Carter, T.R., Hulme, M. and Viner, D. (eds.), *Representing Uncertainty in Climate*

Change Scenarios and Impact Studies, ECLAT-2 Workshop Report No. 1, Helsinki, Finland, 14-16 April 1999, Climatic Research Unit, UEA, Norwich, UK, pp 11-37.

Hulme, M. and Jenkins, G.J., 1998: *Climate Change Scenarios for the UK: Scientific Report*, UKCIP Technical Report No. 1, Climatic Research Unit, Norwich, 80pp,

Hulme, M., Barrow, E.M., Arnell, N.W., Harrison, P.A., Johns, T.C. and Downing, T.E., 1999: 'Relative impacts of human-induced climate change and natural climate variability', *Nature*, **397**, 688-691.

Hulme, M., Crossley, J. and Lu, X., 2001: *An Exploration of Regional Climate Change Scenarios for Scotland*. The Scottish Executive Central Research Unit, ISBN 1-84268-493-0.

Huntingford, C. and Cox, P.M., 2000: 'An analogue method to derive additional climate change scenarios from existing GCM simulations', *Climate Dynamics*, **16**, 575-586.

Huth, R., Kysely, J. and Pokorna, L., 2000: 'A GCM simulation of heat waves, dry spells, and their relationships to circulation', *Climatic Change*, **46**, 29-60.

Jones, P.D. and Reid, P.A., 2001: 'Assessing future changes in extreme precipitation over Britain using regional climate model integrations', *International Journal of Climatology,* in press.

Jones, P.D., Horton, E.B., Folland, C.K., Hulme, M., Parker, D.E. and Basnett, T.A., 1999: 'The use of indices to identify changes in climatic extremes', *Climatic Change*, **42**, 131-149.

Jones, R.G., Murphy, J.M., Noguer, M. and Keen, A.B., 1997: 'Simulation of climate change over Europe using a nested regional-climate model. II: Comparison of driving and regional model responses to a doubling of carbon dioxide', *Quarterly Journal of the Royal Meteorological Society*, **123**, 265-292.

Jones, R.N., 2000a: 'Managing uncertainty in climate change projections – issues for impact assessment', *Climatic Change*, **45**, 403-419.

Jones, R.N., 2000b: 'Analysing the risk of climate change using an irrigation demand model', *Climate Research*, **14**, 89-100.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, D., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R. and Joseph, D., 1996: 'The NCEP/NCAR 40-year Reanalysis project', *Bulletin of the American Meteorological Society*, **77**, 437-471.

Karl, T.R., Nicholls, N. and Ghazi, A. (eds.), 1999: 'Weather and climate extremes: changes, variations and a perspective from the insurance industry', *Climatic Change*, **42**, 1-349.

Katz, R.W., 1999a: 'Techniques for estimating uncertainty in climate change scenarios and impact studies', in Carter, T.R., Hulme, M. and Viner, D. (eds.), *Representing Uncertainty in Climate Change Scenarios and Impact Studies*, ECLAT-2 Workshop Report No. 1, Helsinki, Finland, 14-16 April 1999, Climatic Research Unit, UEA, Norwich, UK, pp 38-53.

Katz, R.W., 1999b: 'Extreme value theory for precipitation: sensitivity analysis for climate change', *Advances in Water Resources*, **23**, 133-139.

Katz, R.W. and Zheng, X., 1999: 'Mixture model for overdispersion of precipitation', *Journal of Climate*, **12**, 2258-2537.

Keller, K., Tan, K., Morel, F.M.M. and Bradford, D.F., 2000: 'Preserving the ocean circulation: implications for climate policy', *Climatic Change*, **47**, 17-43.

Kharin, V.V. and Zwiers, F.W., 2000: 'Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM', *Journal of Climate*, **13**, 3760-3788.

Klein Tank, A.M.G. and Können, G.P., 1997: 'Simple temperature scenario for a Gulf Stream induced climate change', *Climatic Change*, **37**, 505-512.

Knippertz, P., Ulbrich, U. and Speth, P., 2000: 'Changing cyclones and surface wind speeds over the North Atlantic and Europe in a transient GHG experiment', *Climate Research*, **15**, 109-122.

Kothavala, Z., 1997: 'Extreme precipitation events and the applicability of global climate models to the study of floods and droughts', *Mathematics and Computers in Simulation*, **43**, 261-268.

Kothavala, Z., 1999: 'The duration and severity of drought over eastern Australia simulated by a coupled ocean-atmosphere GCM with a transient increase in CO_2', *Environmental Modelling and Software*, **14**, 243-252.

Manabe, S. and Stouffer, R.J., 2000: 'Study of abrupt climate change by a coupled ocean-atmosphere model', *Quaternary Science Reviews,* **19**, 285-299.

McGuffie, K., Henderson-Sellers, A., Holbrook, N., Kothavala, Z., Balachova, O. and Hoekstra, J., 1999: 'Assessing simulations of daily temperature and precipitation variability with global climate models for present and enhanced greenhouse climates', *International Journal of Climatology*, **19**, 1-26.

Mearns, L.O., Rosenzweig, C. and Goldberg, R., 1997: 'Mean and variance change in climate scenarios: methods, agricultural applications, and measures of uncertainty', *Climatic Change*, **35**, 367-396.

Mearns, L.O., Bogardi, I., Giorgi, F., Matyasovszky, I. and Palecki, M., 1999: 'Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling', *Journal of Geophysical Research*, **104**, 6603-6621.

Meehl, G.A., Zwiers, F., Evans, J., Knutson, T., Mearns, L. and Whetton, P., 2000: 'Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change', *Bulletin of the American Meteorological Society*, **81**, 427-436.

Mitchell, J.F.B., Johns, T.C., Eagles, M., Ingram, W.J. and Davis, R.A., 1999: 'Towards the construction of climate change scenarios', *Climatic Change*, **41**, 547-581.

Mitchell, T.D., 2001: *An Investigation of the Pattern Scaling Technique for Describing Future Climates*, PhD Thesis, University of East Anglia, Norwich.

Mitchell, T.D. and Hulme, M., 1999: 'Predicting regional climate change: living with uncertainty', *Progress in Physical Geography*, **23**, 57-78,

Murphy, J., 1999: 'An evaluation of statistical and dynamical techniques for downscaling local climate', *Journal of Climate,* **12**, 2256-2284.

New, M. and Hulme, M., 2000: 'Representing uncertainty in climate change scenarios: a Monte-Carlo approach', *Integrated Assessment*, **1**, 203-213.

Oppenheimer, M., 1998: 'Global warming and the stability of the West Antarctic ice sheet', *Nature*, **393**, 325-332.

Osborn, T.J., 1997: 'Areal and point precipitation intensity changes: Implications for the application of climate models', *Geophysical Research Letters*, **24**, 2829-2832.

Osborn, T.J. and Hulme, M., 1997: 'Development of a relationship between stations and grid-box rainday frequencies for climate model evaluation', *Journal of Climate*, **10**, 1885-1908.

Osborn, T.J. and Hulme, M., 1998: 'Evaluation of the European daily precipitation characteristics from the Atmospheric Model Intercomparison Project', *International Journal of Climatology*, **18**, 505-522.

Osborn, T.J., Hulme, M., Jones, P.D. and Basnett, T.A., 2000: 'Observed trends in the daily intensity of United Kingdom precipitation', *International Journal of Climatology,* **20**, 347-364.

Palutikof, J.P., Goodess, C.M. and Watkins, S.J., 2001: 'A methodology to generate rainfall and temperature scenarios at multiple sites', *Journal of Climate*, submitted.

Poore, R.Z. and Dowsett, H.J., 2001: 'Pleistocene reduction of polar ice caps: Evidence from Cariaco Basin marine sediments', *Geology*, **29**, 71-74.

Rahmstorf, S., 2000: 'The thermohaline circulation: a system with dangerous thresholds?', *Climatic Change*, **46**, 247-256.

Rahmstorf, S. and Ganopolski, A., 1999: 'Long-term global warming scenarios computed with an efficient coupled climate model', *Climatic Change*, **43**, 353-367.

Räisänen, J. and Palmer, T.N., 2001: 'A probability and decision-model analysis of a multi-model ensemble of climate change simulations', *Journal of Climate*, submitted.

Rummukainen, M., Räisänen, J., Bringfelt, B., Ullerstig, A., Omstedt, A., Willén, U., Hansson, U. and Jones, C., 2001: 'A regional climate model for northern Europe: model description and results from the downscaling of two GCM control simulations', *Climate Dynamics*, **17**, 339-359.

Santer, B.D., Wigley, T.M.L., Schlesinger, M.E. and Mitchell, J.F.B., 1990: *Developing Climate Scenarios from Equilibrium GCM Results*, MPI Report 47, Hamburg.

Schär, C., Frei, C., Lüthi, D. and Davies, H.C., 1996: 'Surrogate climate-change scenarios for regional climate models', *Geophysical Research Letters*, **23**, 669-672.

Schneider, S.H. and Root, T.L., 1996: 'Ecological implications of climate change will include surprises', *Biodiversity and Conservation*, **5**, 1109-1119.

Schnur, R. and Lettenmaier, D.P., 1998: 'A case study of statistical downscaling in Australia using weather classification by recursive partitioning', *Journal of Hydrology,* **212-213**, 362-379.

Schubert, M., Perlwitz, J., Blender, R., Fraedrich, K. and Lunkeit, F., 1998: 'North Atlantic cyclones in CO_2-induced warm climate simulations: frequency, intensity, and tracks', *Climate Dynamics*, **14**, 827-837.

Schubert, S., 1998: 'Downscaling local extreme temperature changes in south-eastern Australia from the CSIRO Mark2 GCM', *International Journal of Climatology*, **18**, 1419-1438.

Schubert, S. and Henderson-Sellers, A., 1997: 'A statistical model to downscale local daily temperature extremes from synoptic-scale atmospheric circulation patterns in the Australian region', *Climate Dynamics*, **13**, 223-234.

Semenov, M.A. and Brooks, R.J., 1999: 'Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain', *Climate Research*, **11**, 137-148.

Semenov, M.A., Brooks, R.J., Barrow, E.M. and Richardson, C.W., 1998: 'Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates', *Climate Research*, **10**, 95-107.

Skelly, W.C. and Henderson-Sellers, A., 1996: 'Grid box or grid point: what types of data do GCMs deliver to climate impacts researchers?', *International Journal of Climatology*, **16**, 1079-1086.

Smith, J.B. and Hulme, M., 1998: 'Climate change scenarios', in J. Feenstra, I. Burton, J.B. Smith and R.S.J. Tol (eds.) *Handbook on Methods of Climate Change Impacts Assessment and Adaptation Strategies*, UNEP/IES, Version 2.0, October, Amsterdam, Chapter 3.

Stocker, T.F. and Schmittner, A., 1997: 'Influence of CO_2 emissions rates on the stability of the thermohaline circulation', *Nature,* **388**, 862-865.

Streets, D.G. and Glantz, M.H., 2000: 'Exploring the concept of climate surprise', *Global Environmental Change*, **10**, 97-107.

Thorne, M.C., Watkins, S.J., Goodess, C.M. and Palutikof, J.P., 2000: *The Role of the Antarctic Ice Sheets in Determining Global Sea Level*, AEAT/R/ENV/0170, Issue 2, August 2000.

Trigo, R.M. and Palutikof, J.P., 1999: 'Simulation of daily temperatures for climate change scenarios over Portugal: a neural network model approach', *Climate Research,* **13**, 45-59.

Tziperman, E., 1997: 'Inherently unstable climate behaviour due to weak thermohaline ocean circulation', *Nature*, **386**, 592-595.

Ulbrich, U. and Christoph, M., 1999: 'A shift of the NAO and increased storm track activity over Europe due to anthropogenic greenhouse gas forcing', *Climate Dynamics*, **15**, 551-559.

Vaughan, D.G. and Spouge, J.R., 2001: 'Risk estimation of collapse of the West Antarctic ice sheet', *Climatic Change*, submitted.

Vellinga, M. and Wood, R.A., 2001: 'Global climatic impacts of the Atlantic thermohaline circulation', *Climatic Change*, submitted.

Visser, H., Folkert, R.J.M., Hoekstra, J. and de Wolff, J.J., 2000: 'Identifying key sources of uncertainty in climate change projections', *Climatic Change*, **45**, 421-457.

Wagner, D., 1996: 'Scenarios of extreme temperature events', *Climatic Change*, **33**, 385-407.

Wagner, D., 1999: 'Assessment of the probability of extreme weather events and their potential effects in large conurbations', *Atmospheric Environment*, **33**, 4151-4155.

Wang, X., Stone, P.H., and Marotzke, J., 1999: 'Global thermohaline circulation. Part I: Sensitivity to atmospheric moisture transport', *Journal of Climate*, **12**, 71-82.

Wang, X., Stone, P.H. and Marotzke, J., 1999: 'Global thermohaline circulation: Part II: Sensitivity with interactive atmospheric transports', *Journal of Climate*, **12**, 83-91.

Weichert, A. and Bürger, G., 1998: 'Linear versus nonlinear techniques in downscaling', *Climate Research'*, **10**, 83-93.

Wetherald, R.T. and Manabe, S., 1999: 'Detectability of summer dryness caused by greenhouse warming', *Climatic Change*, **43**, 495-511.

Wilby, R.L., 1998: 'Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices', *Climate Research*, **10**, 163-178.

Wilby, R.L. and Wigley, T.M.L., 1997: 'Downscaling general circulation model output: a review of methods and limitations', *Progress in Physical Geography*, **21**, 530-548.

Wilby, R.L. and Wigley, T.M.L., 2000: 'Precipitation predictors for downscaling observed and general circulation model relationships', *International Journal of Climatology*, **20**, 641-661.

Wilby, R.L., Greenfield, B. and Glenny, C., 1994: 'A coupled synoptic-hydrological model for climate change impact assessment', *Journal of Hydrology*, **153**, 265-290.

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and Wilks, D.S., 1998: 'Statistical downscaling of general circulation model output: A comparison of methods', *Water Resources Research*, **34**, 2995-3008.

Wilks, D.S., 1992: 'Adapting stochastic weather generation algorithms for climate change studies', *Climatic Change*, **22**, 67-84.

Wilks, D.S., 1999: 'Interannual variability and extreme-value characteristics of several stochastic daily precipitation models', *Agricultural and Forest Meteorology*, **93**, 153-169.

Wilks, D.S. and Wilby, R.L., 1999: 'The weather generation game: a review of stochastic weather models', *Progress in Physical Geography*, **23**, 329-357.

Wilson, L.L., Lettenmaier, D.P. and Wood, E.F., 1991: 'Simulation of daily precipitation in the Pacific Northwest using a weather classification scheme', *Surveys in Geophysics*, **12**, 127-142.

Wilson, L.L., Lettenmaier, D.P. and Skyllingstad, E., 1992: 'A hierarchical stochastic model of large-scale atmospheric circulation patterns and multiple station daily precipitation', *Journal of Geophysical Research*, **97**, 2791-2809.

Winkler, J.A., Palutikof, J.P., Andresen, J.A. and Goodess, C.M., 1997: 'The simulation of daily temperature time series from GCM output. Part II: Sensitivity analysis of an empirical transfer function methodology', *Journal of Climate*, **10**, 2514-2532.

Zorita, E. and von Storch, H., 1999: 'The analog method as a simple statistical downscaling technique: comparison with more complicated methods', *Journal of Climate*, **12**, 2474-2489.

Zwiers, F.W. and Kharin, V.V., 1998: 'Changes in the extremes of the climate simulated by CCC GCM2 under CO_2 doubling', *Journal of Climate*, **11**, 2200-2222.

APPENDIX 1: ISSUES ARISING IN DISCUSSION DURING THE PROJECT WORKSHOP ON 'NEW DIMENSIONS FOR CLIMATE SCENARIOS: A WORKSHOP TO IDENTIFY THE EXTREME WEATHER SCENARIO NEEDS OF THE TYNDALL CENTRE AND THE WIDER IMPACTS COMMUNITY', 4-5 JUNE 2001, UEA, NORWICH

Indicators of extremes

'Benchmarking' against experienced events (e.g., the 1995 hot summer, the October 2000 floods) is very important for engaging stakeholder attention (as is some element of financial costing for the insurance industry).

A distinction needs to be made between meteorological information on extremes (the focus of this project) and their impacts. What, for example, is the relationship between the 99th percentile precipitation event and the 1 in 100 year flood event? The latter may be affected by land-use changes, etc.

The quality (homogeneity) of observed data is particularly important for the investigation of extremes.

For some impact sectors (e.g., air quality, building comfort), 'non-events' such as very low wind speed may be important. These events are extreme in terms of their impact although the weather event itself is not extreme.

The 2070-2100 time slice simulated by the most recent Hadley Centre models (HadAM3 and HadRM3) is very late for most impact sectors which are more interested in the next 25 years, or even the next 10-20 years, but this time slice has the advantage of maximising the signal-to-noise ratio. (The Hadley Centre may, however, run a simulation for an earlier time slice to test linearity/scaling issues.)

It is also important to know how current risks are affected by climate change. How robust are current estimates of, for example, the 1 in 100/200 year event? Are there better ways of estimating these? Is it acceptable to assume a stationary climate? How appropriate is MAFF's interim guidance on peak river flows (i.e., to add 20% to current estimates)? Will return period events become smaller/larger? These issues are all important for current planning decisions.

Methodological issues

It is important to consider the uncertainties and confidence intervals associated with scenarios of extremes. Are there real changes in the tails of distributions or are the projected changes within the noise levels? To what extent is the uncertainty of changes in extremes related to that of the mean changes?

What is the optimal size of inter-/intra-model ensemble? (One guide is the square root of the sample size.) Would it be worthwhile to use combined weather generator ensembles, e.g. 10 models run 100 times rather than one model run 1000 times?

Is non-stationarity across ensembles (e.g. output from three x 30-year SRES A2 simulations) likely to be a problem when calculating extreme value statistics? (It was concluded that these errors are probably small compared with the other sources of error and uncertainty.)

If temperature and precipitation are scaled separately, is the spatial coherence of these variables lost? Is it possible to capture the non-linearities by using a variable other than global temperature as the scalar? Is it possible to scale daily time series, i.e. streams of weather? Are previous findings (i.e., Mitchell, 2001) model dependent? There is some evidence of linearity of precipitation response in HadRM2 (seen in the ranked annual maxima, for example). In terms of thermohaline collapse, are the climatic impacts likely to be different for a slow down compared with total collapse, i.e. is scaling possible?

It is important to determine whether the changes in extremes simulated by climate models are due to changes in the shape and/or scale parameters.

Better methods are needed for evaluating inter-variable correlations because conventional measures such as covariance are not considered adequate.

Regression-based statistical downscaling methods (e.g., Conway *et al.*, 1996; Wilby, 1998; Wilby *et al.*, 1998) should be evaluated as part of the case-study work. But consideration needs to be given as to how to use these methods in a fully stochastic/probabilistic way.

Statistical downscaling methods should include some type of humidity forcing as well as more conventional circulation-related parameters as predictor variables.

At least 50 data points should be used for return period estimates (though this is not always possible). Sample size can be increased by working with point process (exceedence) models, though objective ways are needed of identifying appropriate thresholds. Confidence intervals should always be given with return period estimates.

For a direct comparison of statistical and dynamical downscaling, statistical downscaling models should really be driven by output from the HadAM3 and not the HadCM3 simulations.

Case studies

There is no standard UK definition of drought and appropriate definitions may vary by impact sector (e.g., different definitions are appropriate for agriculture and water resources). It is important to consider sequences and persistence (e.g., the occurrence of 'back-to-back' drought). A 'rolling month' index or running soil moisture deficit index calculated over a 60/90 day window could be used (e.g., Phillips and McGregor, 1998).

For some impact sectors, it is important to consider 'double' joint probabilities, e.g. the probability of a dry winter/wet summer together with water resource supply and demand.

Climate scenarios produced for the drought and flood case studies could be fed through Nigel Arnell's hydrological model (for three or four catchments, using hydrological and/or water resource measures of drought).

A case study based on subsidence would be a good test of joint probabilities (and would require a high, i.e. post-code, spatial resolution).

Snowmelt would also be a good test of joint probabilities. Any information on the changing risk of catastrophic snowmelt episodes would be very useful for flood defence and reservoir design (see the Flood Estimation Handbook). Critical catchments may be the Tay or Yorkshire Ouse.

For flooding, the timescales which need to be considered are likely to depend on catchment size. Cumulative totals and durations are likely to be more important than one-off events, e.g. in Scotland, a 72 hour timescale is important.

The case-study work on statistical downscaling should most usefully focus on aspects that are poorly simulated by the RCM. For example, if persistence is reasonably well simulated in the RCM, other diagnostic variables should be explored in statistical downscaling.

Particular consideration needs to be given to the treatment of uncertainty in scenarios of extremes, including how to handle ensembles and different emissions scenarios. A number of ways of handling stochastic weather generator output (i.e. simulation sets each consisting of 1000 runs) are proposed by Goodess (2000). Consideration also needs to be given as to how to apply probabilistic approaches to direct GCM/RCM output.

References

Conway, D., Wilby, R.L. and Jones, P.D., 1996: ‘Precipitation and air flow indices over the British Isles’, *Climate Research*, **7**, 169-183.

Goodess, C.M., 2000: *The Construction of Daily Rainfall Scenarios for Mediterranean Sites Using a Circulation-type Approach to Downscaling*, PhD Thesis, University of East Anglia, Norwich.

Mitchell, T.D., 2001: *An Investigation of the Pattern Scaling Technique for Describing Future Climates*, PhD Thesis, University of East Anglia, Norwich.

Phillips, I.D. and McGregor, G.R., 1998: ‘The utility of a drought index for assessing the drought hazard in Devon and Cornwall, South West England’, *Meteorological Applications*, **5**, 359-372.

Wilby, R.L., 1998: ‘Statistical downscaling of daily precipitation using daily airflow and seasonal teleconnection indices’, *Climate Research*, **10**, 163-178.

Wilby, R.L., Wigley, T.M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J. and Wilks, D.S., 1998: ‘Statistical downscaling of general circulation model output: A comparison of methods’, *Water Resources Research*, **34**, 2995-3008.

Tyndall°Centre
for Climate Change Research

The inter-disciplinary Tyndall Centre for Climate Change Research undertakes integrated research into the long-term consequences of climate change for society and into the development of sustainable responses that governments, business-leaders and decision-makers can evaluate and implement. Achieving these objectives brings together UK climate scientists, social scientists, engineers and economists in a unique collaborative research effort.

Research at the Tyndall Centre is organised into four research themes that collectively contribute to all aspects of the climate change issue: Integrating Frameworks; Decarbonising Modern Societies; Adapting to Climate Change; and Sustaining the Coastal Zone. All thematic fields address a clear problem posed to society by climate change, and will generate results to guide the strategic development of climate change mitigation and adaptation policies at local, national and global scales.

The Tyndall Centre is named after the 19th century UK scientist John Tyndall, who was the first to prove the Earth's natural greenhouse effect and suggested that slight changes in atmospheric composition could bring about climate variations. In addition, he was committed to improving the quality of science education and knowledge.

The Tyndall Centre is a partnership of the following institutions:
University of East Anglia
UMIST
Southampton Oceanography Centre
University of Southampton
University of Cambridge
Centre for Ecology and Hydrology
SPRU – Science and Technology Policy Research (University of Sussex)
Institute for Transport Studies (University of Leeds)
Complex Systems Management Centre (Cranfield University)
Energy Research Unit (CLRC Rutherford Appleton Laboratory)

The Centre is core funded by the following organisations:
Natural Environmental Research Council (NERC)
Economic and Social Research Council (ESRC)
Engineering and Physical Sciences Research Council (EPSRC)
UK Government Department of Trade and Industry (DTI)

For more information, visit the Tyndall Centre Web site (www.tyndall.ac.uk) or contact:
External Communications Manager
Tyndall Centre for Climate Change Research
University of East Anglia, Norwich NR4 7TJ, UK
Phone: +44 (0) 1603 59 3906; Fax: +44 (0) 1603 59 3901
Email: tyndall@uea.ac.uk

Other titles in the Tyndall Working Paper series include:

1. A country-by-country analysis of past and future warming rates, *November 2000*
2. Integrated Assessment Models, *March 2001*
3. Socio-economic futures in climate change impact assessment: using scenarios as 'learning machines', *July 2001*
4. How high are the costs of Kyoto for the US economy?, *July 2001*
5. The issue of 'Adverse Effects and the Impacts of Response Measures' in the UNFCCC, *July 2001*
6. The identification and evaluation of suitable scenario development methods for the estimation of future probabilities of extreme weather events, *July 2001*
7. Security and Climate Change, *October 2001*
8. Social Capital and Climate Change, *October 2001*
9. Climate Dangers and Atoll Countries, *October 2001*

The Tyndall Working Papers are available online at:
http://www.tyndall.ac.uk/publications/working_papers/working_papers.shtml