Climatic Change (2007) 81:283–312
DOI 10.1007/s10584-006-9159-6

Integrated assessment of changes in flooding probabilities due to climate change

Thomas Kleinen · Gerhard Petschel-Held

Received: 27 December 2004 / Accepted: 19 May 2006 / Published online: 19 January 2007

Abstract An approach to considering changes in flooding probability in the integrated assessment of climate change is introduced. A reduced-form hydrological model for flood prediction and a downscaling approach suitable for integrated assessment modeling are presented. Based on these components, the fraction of world population living in river basins affected by changes in flooding probability in the course of climate change is determined. This is then used as a climate impact response function in order to derive emission corridors limiting the population affected. This approach illustrates the consideration of probabilistic impacts within the framework of the tolerable windows approach. Based on the change in global mean temperature, as calculated by the simple climate models used in integrated assessment, spatially resolved changes in climatic variables are determined using pattern scaling, while natural variability in these variables is considered using twentieth century deviations from the climatology. Driven by the spatially resolved climate change, the hydrological model then aggregates these changes to river basin scale. The hydrological model is subjected to a sensitivity analysis with regard to the water balance, and the uncertainty arising through the different projections of changes in mean climate by differing climate models is considered by presenting results based on different models. The results suggest that up to 20% of world population live in river basins that might inevitably be affected by increased flood events in the course of global warming, depending on the climate model used to estimate the regional distribution of changes in climate.

This article is dedicated to the memory of the late Gerhard Petschel-Held. He was an inspiring colleague, as well as a good friend. His sudden departure leaves me deeply shocked, and I am sure he will sorely be missed by all who had the pleasure of meeting him. Thomas Kleinen

T. Kleinen · G. Petschel-Held
Potsdam-Institute for Climate Impact Research,
P. O. Box 601203, 14412 Potsdam, Germany

Present address:
T. Kleinen (✉)
Climatic Research Unit, School of Environmental Sciences,
University of East Anglia, Norwich NR4 7TJ, UK
e-mail: t.kleinen@uea.ac.uk

1 Introduction

The integrated assessment of climate change needs to take into account both the costs and the benefits of climate protection measures. Whereas the first mainly relates to issues of energy production, the latter is associated with avoided damages from climate change. Whereas many integrated assessment models consider the costs of mean climate change, the effects of extreme events are often neglected. This is despite the fact that there is an increasing trend of economic losses due to 'atmospheric' natural disasters (Berz 1999). The Mississippi flood of 1993, for example, has caused economic losses of between US$ 12 and 16 billion (Hipple et al. 2005), whereas the losses of the 2002 summer floods in Europe are estimated to be about EUR 18.5 billion in central Europe (Munich Re 2003). Both numbers are of the same order of magnitude as the estimated damage costs in the water sector for both regions for an increase in global mean temperature of about 1–2.5°C (Tol 2002). This indicates that extreme floods, which appear in the 'midfield' in the statistics of economic losses (Berz 1999), should be an essential component of integrated assessment.

For the recent global warming of the twentieth century no general and coherent trends could be observed with regard to increases in annual maximum flows (Kundzewicz et al. 2004). For great events, i.e.,100-year floods, however, an increasing risk was detected in 29 basins larger than 20,000 km^2 by Milly et al. (2002). In spite of major uncertainties, there are some studies, including Working Group II of the IPCC TAR, which claim an increase of major flooding probability for future warming (Kundzewicz and Schellnhuber 2004; Milly et al. 2002). Other studies show similar results with a rather heterogeneous geographical distribution of changes in flooding probabilities (Arnell 1999a; Arora and Boer 2001). Yet, in some highly vulnerable regions a significant increase of flooding probabilities has been found under global warming, e.g., for Bangladesh (Mirza 2002), central Asia and eastern China (Arnell 1999a). All of these studies are restricted to climate change induced shifts in flooding probabilities and do not take into account other major factors relevant for changes in flooding intensities and frequencies. These factors include land-use changes, modification of streamflows by various water-management schemes like dams or dykes, or, when it comes to the actual damages, the relocation of infrastructure or settlements. On the one hand, this makes assessments easier, but on the other hand it might give unreliable or biased results.

For a flood component of an integrated assessment model (IAM), it is generally not sufficient to model the shifts in flooding probabilities only. In addition, one needs to map those probabilities to actual damages, where the specific measure of damage depends on the overall framework of the IAM. In case of a cost-benefit approach, for example, the flood model needs to give a monetary output. Within other frameworks, e.g., the tolerable windows approach (TWA), damages need to be calculated in a decision relevant measure, which doesn't need to be directly related to monetary costs.

Another difficulty in developing an integrated assessment module of flood changes is due to computational requirements of those models, in particular if the overall model includes the decision making with respect to climate mitigation endogenously in the model. These computational costs ask for so-called reduced-form models, which mimic the outcomes of more detailed models, yet are much faster to compute.

Within the first part of this paper (Section 2-5), we develop such a reduced-form model, based on simplified descriptions of regional patterns of climate change and on a highly reduced scheme for runoff computation. As 'output' variable, the model computes the number of people affected by a predefined shift of flooding probability, e.g., a once in 50 years event shifts to a once in 25 years event. These shifts are computed for large river basins with an area of more than $2.5 \times 10^4\,\mathrm{km}^2$, and we also neglect the 'other major factors' affecting flooding probabilities. In the final section, we present a first application of the model within the tolerable windows approach. In the TWA the integrated assessment process starts by assessing which impacts of climate change are undesirable. These impacts are then excluded by setting normative constraints, 'guardrails' in the language of the TWA. In a subsequent step the TWA then determines sets of emission reduction strategies that are compatible with the predefined guardrails.

Seen somewhat more formally, the basic problem in IA is a control problem with a basic differential equation $\dot{x} = f(x, t; u)$ where the time evolution of the climate state x is dependent on the state x itself, time t and a control strategy u. In so-called policy evaluation modeling, e.g., the IMAGE family of models (Alcamo et al. 1998; Rotmans et al. 1989), the control strategy u is predefined and the consequences of this strategy are evaluated exogenously, i.e., by the model user. Contrary to this, the aim in cost-benefit modeling is to determine an optimal policy $\widetilde{u}$. In the TWA there are additional constraints $h(x, t; u) \leq 0$, the 'guardrails', and the aim is to solve the differential inclusion $\dot{x} \in \mathbb{F}(x, t)$ with $\mathbb{F} := \{f(x, t; u) \mid u \in \mathbb{U}\}$ under the condition $h(x, t; u) \leq 0$ in order to determine the set of emission reduction strategies that are compatible with the predefined guardrails.

Within the TWA impacts of climate change can be represented as a Climate Impact Response Function (CIRF). CIRFs indicate the relationship between climate change and the impacts of climate change. They can formally be represented as $I = I(C, S)$ with the impact I, the relevant climatic variables C and the significant socio-economic variables S. In previous assessments (Füssel 2003; Füssel et al. 2003), CIRFs were defined within a deterministic framework. The present paper will extend the concept of CIRFs to the probabilistic domain.

2 Model description

2.1 Aims and scope

We are aiming to develop a reduced-form model that is able to incorporate the probabilities of large-scale flooding in an integrated assessment modeling framework. We will use this model to determine CIRFs that can be used to estimate the effects of climate change on flooding probabilities and their consequences. While floods may have a multitude of causes, ranging from blocking of river passages by ice or debris, via land-use changes and river regulation, to large precipitation events, most of these causes are not directly related to climate change. Due to climate change the hydrological characteristics of the atmosphere may change. Higher temperatures cause an increase in evaporation, and the moisture capacity of the atmosphere increases as well. This may lead to increases in precipitation and in particular increases in intense precipitation according to the Clausius-Clapeyron law. As the non-climatic causes for flooding mentioned above cannot easily be incorporated in the model we are developing, our analysis will focus on the climate change related causes. In addition we have to restrict the type of floods we are attempting to model. Local, sudden floods ('flash floods') occur in small catchments and are mainly caused by localized intense precipitation events. While changes in the characteristics of these events are to be expected in a changed climate, we regard an integrated assessment of changes in probability of flash floods as too ambitious on a global scale for the time being. Extensive, long-lasting floods ('plain floods'), on the other hand, occur in larger catchments (Bronstert et al. 2002). These floods may be caused by extreme short-term precipitation events, especially in mountainous areas, but they may also be caused by large-scale rainfall lasting several days or weeks. The latter is the type of flood we are attempting to model.

The assessment we are conducting is global in scope. Therefore a compromise has to be made with regard to the temporal and spatial scales that can be resolved. While high spatial resolutions allow assessments on the scale of small river basins, or even sub-basins, they also lead to high requirements with respect to computing time, input data and validation data. Similarly, high temporal resolution could allow the simulation of flash floods and similarly fast events, and might generally improve the fidelity of model results, but again the requirements with respect to data and computational resources are very demanding.

For the assessment of changes in flooding probability on the scale of large river basins, a spatial resolution of 0.5° seems to be a reasonable compromise, as well as a temporal resolution of one month. Vörösmarty et al. (2000) estimate that river basins with drainage areas $\geq 2.5 \times 10^{4}\,\mathrm{km}^2$ can be modeled reasonably at a spatial resolution of 0.5°, and climate data are readily available at this resolution, e.g., the 'CRU' data by New et al. (2000), the data by Willmott and Matsuura (2001) or data by Leemans and Cramer (1991). These data have a temporal resolution of one month, which allows the resolution of the annual cycle, while fast events like flash floods cannot be investigated at this time scale. As gauge records from a large number of streamflow gauges with a global coverage also use the monthly time scale, the model uses a timestep of one month for calculation.

In addition to the choice of resolution a few other simplifications are made. Our model will neglect the temporal dynamics of river routing, as this seems hardly worthwhile at a temporal resolution of one month. At this temporal scale water traveling at $0.5\,\mathrm{m/s}$ moves approximately 1300 km during one timestep (Vörösmarty et al. 2000). The mean travel times therefore exceed one timestep for the very largest rivers only. The consideration of river routing would therefore only influence results for these river systems. In addition, river routing will not change significantly due to climate change, neglecting possible changes in the timing of flows. We are also neglecting the soil storage of moisture and evaporation from water bodies. While these factors may degrade model results, especially with regard to the simulation of the annual cycle of runoff, the sensitivity analysis (Section 4.3) suggests that the simulation of floods would not be improved by the reductions in runoff implied by these factors.

The aim of our model is therefore not the modeling of the dynamical processes of flood events. We believe that these cannot be modeled adequately at the spatial and temporal scales considered. Our assessment rather focuses on the potential for large-scale flood events. Therefore events of a very dynamical nature, such as snowmelt floods, floods due to ice jam or flash floods remain outside the scope of our assessment.

2.2 Downscaling of climate change

The climate components of many IA models, e.g., the models DICE (Nordhaus 1994), MERGE (Manne et al. 1995), MiniCAM (Edmonds et al. 1996) and SIAM (Hasselmann et al. 1997), are intended for the evaluation of large numbers of climate change scenarios. In some cases they are also coupled to economic models, which obtain solutions by optimizing some value-function. Therefore the climate models employed in such a framework must be run a large number of times. This limits the computational resources such a model may consume. Therefore a typical climate model for integrated assessment applications only calculates the change in global mean temperature ΔT_{GM}, while the spatial distribution of temperature change and changes in other climatic variables have to be inferred from this.

The impact of climate change we want to assess here not only requires a more explicit spatial resolution, but it also needs to take into account climate variability, and not just the changes in mean climate. We therefore divide the modeling approach into a 'mean' and a 'variability' part.

Geographically explicit changes in mean climate can be calculated by using the pattern scaling approach (Füssel 2003; Mitchell et al. 1999; Mitchell 2003). In this approach geographically explicit patterns of climate change obtained from GCM experiments are scaled by ΔT_{GM} calculated by the simple climate model included in the integrated assessment model. Despite the apparent simplicity of the approach, results obtained in this way are surprisingly accurate (Mitchell 2003).

We are using climate change patterns obtained by an EOF analysis of output from a number of GCM experiments (Füssel 2003). In order to reflect the pertaining uncertainty about the spatial aspects of climate change, we are using patterns of temperature and precipitation change from three different GCMs, i.e., HadCM 2 (Johns et al. 1997), ECHAM 3 (Voss et al. 1998) and ECHAM 4 (Roeckner et al. 1996). These patterns of monthly climate change are scaled by the change in global mean temperature ΔT_{GM} and applied to the climatology.

While pattern scaling gives the geographically explicit changes in the mean climate, a representation of the variability of precipitation and evaporation is also necessary for the evaluation of changes in probabilities of flooding. An estimate of variability can be obtained in a number of ways. Besides the vast uncertainties to be expected in each method, most of the approaches, e.g., high resolution GCMs (Hennessy et al. 1997; Voss et al. 2002), statistical downscaling (e.g., Xu 1999; Wilby and Wigley 1997; Wilby et al. 1998) or stochastic weather generators (e.g., Cameron et al. 2000; Hutchinson 1995; Wilks and Wilby 1999) are computationally expensive.

Therefore we chose a resampling approach, similar to the one used by Alcamo et al. (2001) for the GLASS model. This approach is based on data of observed climatic variables on a 0.5° grid with monthly resolution. Both a climatology and the deviations from the climatology are determined from the data, and the deviations from the climatology are used as 'templates' of spatio-temporal variability patterns.

As source of climate data, we are using the CRU-PIK dataset by Österle et al. (2003) (see Section 3.2). From this dataset we determined the monthly climatology for the years 1961–1990, and then determined the deviations from the climatology with $T'(m,t) = T(m,t) - T_C(m)$ and $P'(m,t) = P(m,t)/P_C(m)$ the temperature and precipitation deviation patterns for year t and month m.

In more detail, the 'complete' climate is calculated as follows. A climate model is used to calculate the change in global mean temperature $\Delta T_{GM}(t)$ in year t. We are currently using the 'ICLIPS' climate model (Petschel-Held et al. 1999; Kriegler and Bruckner 2004) for this purpose, but in principle any other climate model giving $\Delta T_{GM}(t)$ could be used as well. $\Delta T_{GM}(t)$ is then used to scale the patterns for temperature and precipitation, which are applied to the climatology in order to obtain the spatial distribution of the mean climate for $\Delta T_{GM}(t)$. This mean climate is then perturbed by a randomly drawn variability pattern in order to represent natural variability.

The global temperature and precipitation fields in a particular month m within year t are thus computed via

$$T(r,m,t) = T_C(r,m) + k\Delta T_{GM}(t) \times T_P(r,m) + T'(r,m,t') \tag{1}$$

$$P(r,m,t) = (P_C(r,m) \times (1 + k\Delta T_{GM}(t) \times P_P(r,m))) \times P'(r,m,t') \tag{2}$$

with $T_C(r,m)$ the climatological temperature in month m in location r, $P_C(r,m)$ the climatological precipitation, $T_P(r,m)$ and $P_P(r,m)$ temperature and precipitation climate change patterns obtained from GCM runs, $\Delta T_{GM}(t)$ the change in global mean temperature in year t and k the scaling factor relating the scaling of the patterns to $\Delta T_{GM}(t)$. $T'(r,m,t')$ and $P'(r,m,t')$ are the deviations from the climatology described above, where the time t' refers to a year randomly drawn from the twentieth century deviations from climatology.

Advantages of this scheme are that spatial and temporal correlations of past variability are well represented by using this approach, even though the temporal correlations are only maintained during the course of any particular year and interannual correlations are destroyed, which mainly affects the temporal correlation between December and January. Since we will mainly be using the complete original sequence of deviation patterns, this effect can be neglected in the current application.

The main drawback is that variability is assumed to stay the same in a changed climate – exactly the same for temperature due to the additivity of the deviation pattern and somewhat increased in the case of precipitation due to the multiplicity of the precipitation deviation patterns. While this drawback makes the application of the method to a future changed climate somewhat questionable, we are assuming that this approach can still give major insights into the effects of global warming on flooding probabilities. In addition, water vapor is not conserved in the modeling approach, since the precipitation calculated using the pattern scaling approach is not dependent on the evaporation determined by the model, as detailed in Section 2.3.

2.3 Runoff calculation

Runoff is calculated using the water balance equation as the difference between precipitation and evaporation

$$R(r,m,t) = P(r,m,t) - E_a(r,m,t) - \Delta S(r,m,t), \tag{3}$$

with runoff R, precipitation P, actual evaporation E_a and the change in soil storage ΔS, all in location r, month m and year t. We are assuming $\Delta S = 0$ as we are neglecting the storage of moisture in the soil. This is based on the assumption that soil will be saturated during the large precipitation events that lead to large-scale flooding.

At temperatures below 0°C, we are assuming that precipitation falls as snow, which is removed from the precipitation field and stored until temperatures rise above freezing again. At temperatures above freezing, the accumulated snow melts and is added to the precipitation field again.

Due to data constraints, the calculation of potential evaporation E_p (the evaporation that would occur, if enough water was available) has to be done by a scheme that does not depend on very detailed climatological data. We have therefore used the Hamon scheme (Hamon 1963) that is only dependent on temperature data. In intercomparisons of different evaporation schemes (Federer et al. 1996; Vörösmarty et al. 1998) the Hamon scheme was found to have comparatively little bias and to be well suited to a large range of surface types. On the other hand, the Hamon scheme is a purely empirical formulation that has been derived for present climatic conditions, which makes it questionable whether it is still applicable in a drastically changed climate (Vörösmarty et al. 1998). Nonetheless, we will use the Hamon scheme for our model since most other evaporation schemes evaluated by Federer et al. (1996) had a larger bias and requirements with regard to input data that cannot be fulfilled by present climate models suitable for integrated assessment.

In the Hamon scheme, potential evaporation E_p (in mm) is calculated as

$$E_p(T, \Lambda) = \frac{715.5 \times \Lambda \times e(T)}{T + 273.2} \tag{4}$$

with T the mean air temperature (in °C), Λ the day length as fraction of day and $e(T)$ the saturated vapor pressure (in kPa) at temperature T. As the model uses monthly timesteps and available input data have monthly resolution, we are also calculating the monthly evaporation. This choice of temporal resolution suits the assessment by Federer et al. (1996) that the scheme is not very sensitive to the use of data with low time resolution.

In principle evapotranspiration changes in a climate with elevated levels of CO_2. However, estimates of this effect vary and strongly depend on vegetation type (Lockwood 1999). We therefore disregard this effect.

Finally, we calculate the actual evaporation E_a from the potential evaporation E_p using

$$E_a = \begin{cases} E_p & \forall E_p \leq P \\ P & \forall E_p > P. \end{cases} \tag{5}$$

Once again, this formulation assumes that soil and plants have no storage capacity for moisture.

The procedure described above gives the amount of runoff per grid cell. Subsequently this is multiplied by grid cell area and summed up over all grid cells belonging to a river basin in order to obtain the total monthly runoff for each river basin considered.

3 Data and methods

3.1 River basin description

The evaluation of changes in the probability of large-scale flooding events only makes sense on the scale of river basins. The river basin description in our model is based on the STN-30p dataset, a dataset of major river basins (Fekete et al. 1999; Vörösmarty et al. 2000). It is derived from a GIS-based analysis of global topographic fields, has a resolution of 0.5°, and lists the grid cells belonging to the drainage areas of 6,152 individual river basins.

As Vörösmarty et al. (2000) estimate that the accuracy of the data is better for river basins with drainage areas $\geq 2.5 \times 10^4$ km^2, we exclude river basins below that size from our analysis.

Using a dataset of population density (CIESIN 2000), interpolated to the projected population in 2100 using the median population projection by IIASA (Lutz et al. 2004), we obtain the total population living in a river basin. This guides us in the choice of river basins for the assessment of future climates: Of those river basins large enough, we chose the river basins with the largest populations, with the exception of a few basins, like the Nile and Chang Jiang, where the assessment would not be meaningful due to large dams that limit the danger of flooding. The assessment takes place in 83 river basins, where about 50% of world population in 2100 live. These basins are listed in the Appendix.

3.2 Input and validation data

As source for climate data, we are using a dataset by Österle et al. (2003). This dataset is derived from the CRU timeseries dataset (New et al. 2000), a dataset of observed climatic variables (precipitation, daily mean temperature, diurnal temperature range, vapor pressure and cloud cover) interpolated to a 0.5° grid and covering the time range from 1901 to 1998 with monthly resolution. Österle et al. removed temporal inhomogeneities from the temperature and precipitation fields and extended the dataset to 2003. Henceforth, this dataset will be referred to as CRU-PIK.

For model validation, we make use of two datasets of streamflow gauge records. The first dataset lists monthly discharge data for world rivers excluding the former Soviet Union (Bodo 2001a), based in large parts on the UNESCO (1974) dataset. The other dataset contains information on monthly discharge data for rivers in the former Soviet Union (Bodo 2001b). These two datasets give us monthly discharge data from 6,883 streamflow gauge sites. Of these gauges, 1,226 had drainage areas $\geq 2.5 \times 10^4$ km^2, and of those gauges, 640 had records longer than 25 years, with only complete years considered.

The 640 gauge sites are located in 148 river basins. If there is more than one gauge site in a river basin we choose the site gauging the largest drainage area, unless there is another site with insignificantly smaller drainage area, but longer record length. About a third of the gauges (52) are at latitudes between 40 and 60°N, all other 20° latitude bands north of 40°S still contain between 10 and 28 gauge sites, and 26 stations are located in the southern hemisphere. The latitudinal coverage of validation records therefore appears to be adequate.

3.3 Validation of annual and monthly runoff

The validation of simulated annual and monthly runoff may seem straightforward at first glance. One would assume that it is sufficient to take precipitation and temperature measurement data, determine the model output for the river basin area upstream of a gauge site, and compare the result with gauge records.

Such a model validation would certainly be possible, if perfect measurements of streamflow, precipitation and temperature were available. If this were the case, any discrepancies between model output and streamflow measurements would have to be regarded as model error. In reality, there may be quite large errors in the measured values, especially in the precipitation measurements (Adam and Lettenmaier 2003; Fekete and Vörösmarty 2004). In addition, those areas where higher quality measurements can be expected, are just those areas where it is very likely that streamflow characteristics have been changed by human intervention, since the highest measurement quality, the longest timeseries, and the highest density of measurement networks can be expected in the industrialized countries, where extensive fluvial management has taken place.

Fekete et al. (2002) investigated this problem in some detail. They compared runoff estimates from the 'WBM' water balance model (Vörösmarty et al. 1996, 1998), driven by precipitation data from the Willmott and Matsuura (2001) climate dataset, with streamflow measurements from selected streamflow gauging stations. They report large differences between simulated and measured streamflow, including some cases where measured streamflow actually exceeded the total measured precipitation.

Therefore we test the quality of our model by comparing its results with the output of other models given similar input data. For this we determine the *bias* of the mean annual streamflow, defined as

$$bias = \frac{\overline{S} - \overline{O}}{\overline{O}} \times 100\%, \tag{6}$$

with $\overline{S}$ the mean modeled annual streamflow and $\overline{O}$ the corresponding observed annual streamflow. Though this *bias* is neglecting interannual variability of streamflows and thus is of limited use for our purpose here, it allows a far reaching comparison to other hydrological models.

In order to get better measures for model simulation quality, we also determined Willmott's index of agreement (Willmott 1982) for the annual total runoff in the validation basins. The index of agreement d is defined as

$$d = 1 - \left[\frac{\sum_{i=1}^{N} (S_i - O_i)^2}{\sum_{i=1}^{N} \left(\left| S_i - \overline{O} \right| + \left| O_i - \overline{O} \right| \right)^2} \right] \tag{7}$$

with S_i the modeled value at time t_i, O_i the observed value at time t_i and $\overline{O}$ the mean observed value. It describes model quality with respect to variations, with $d = 0$ indicating complete disagreement, while $d = 1$ indicates complete agreement. It was proposed by Willmott because the correlation coefficient often used for such investigations is not consistently related to the quality of prediction (Willmott 1982).

3.4 Validation of runoff extremes

The intended purpose of our model is not the accurate reproduction of the mean streamflows, but rather the assessment of probabilities of major flooding due to extreme precipitation. Therefore model validation will focus on the validation of model simulated runoff extremes, even though annual and monthly runoff will also be evaluated.

The magnitude of the so-called 'T-year flood' at a site, which is the amount of streamflow that has a probability $1/T$ of being exceeded in any one year, is commonly estimated by using the annual maximum series (AMS) approach (Li et al. 1999). In this method, a suitable probability distribution is fitted to the annual maxima of the timeseries in order to estimate the return period T of certain flood levels.

In principle, we regard the other possible approach for the estimation of the magnitude of the T-year flood, the peak over threshold (POT) approach (Madsen et al. 1997), as superior, but this approach requires well-defined flood peaks. As our model works on a monthly time scale, it produces just a single flood-peak per year in most river basins. Therefore, the advantage of the POT approach, the ability to use more data than just the single annual maximum, does not come into play, and we thus make use of the AMS approach.

According to a recent review of probability distributions for the AMS approach (Li et al. 1999), various distribution functions are possible. Yet it is difficult to conclude which one is the most appropriate, as the choice of distribution function in mainly dependent on type of data and other factors. Of the distributions that were evaluated favorably by Li et al. (1999), the probability distribution that gives the best fit to the streamflow records we have available is the gamma distribution.

In order to obtain a measure of model performance, we normalize streamflow data and model results and fit a gamma distribution to the annual maxima of streamflow (validation data) or runoff (model results). We use all available data for fitting the distribution, the timeframe considered therefore is variable for the validation data, while it is 100 years (1901–2000) for the model results.

From the gamma distribution, we determine the magnitude of the 50-year maximum streamflow/runoff event. The deviation

$$\Delta_{50yr} = \frac{\left(S_{50yr} - O_{50yr}\right)}{O_{50yr}} \times 100\% \tag{8}$$

of the 50-year maximum event, expressed as a percentage of O_{50yr}, shows how well the model reproduces the streamflow extremes. In this equation S_{50yr} is the magnitude of the model-generated 50-year maximum runoff event, and O_{50yr} is the magnitude of the 50-year maximum streamflow event, as estimated from the gauge records.

As we will later be calculating the change in probability of the twentieth century 50-year maximum streamflow event, this measure gives the most direct indication of simulation quality for the intended purpose of the model.

3.5 Sensitivity analysis

Methodologically, a number of causes of uncertainty in model results be identified. These are:

1. Uncertainty in the climate model
2. Uncertainty in the downscaling scheme
3. Uncertainty in the hydrological model.

These causes of uncertainty could of course be broken down further into the uncertainties in these particular parts of the assessment scheme. We will be addressing the first two causes of uncertainty by considering patterns of change in mean climate derived from different GCMs, and we will address the third point in this section.

This section will therefore focus on uncertainty in the hydrological model, which is mainly contained in the assumed runoff balance (Eq. 3). In order to assess the model sensitivity to the chosen parameterizations, we perform a sensitivity analysis. Within the runoff balance, five uncertain factors appear:

1. Some portion of precipitation may be converted to runoff instantly, without being available for evaporation.
2. Some portion of precipitation may be stored as soil water or converted to groundwater, removing it from the water balance equation.
3. Evaporation may be over – or underestimated by the simple parameterization (Eq. 4) we are using.
4. Precipitation may be over – or underestimated in the dataset.
5. The neglect of changes in soil moisture.

In order to test the first four of these possibilities, we have performed a series of five sensitivity experiments by changing the components of the runoff balance (Eq. 3). These experiments are listed in Table 1. The fifth uncertain factor in Eq. 3 is the neglect of changes in soil moisture. While this factor may have a large influence on model error, especially with respect to the monthly flows, it is not possible to take this into account

Table 1 Sensitivity experiments performed

Experiment	Equation	Formula	Reason
A	Equation (3)	$R_A = 0.1 \times P + (0.9 \times P - E)$	Direct conversion P to R
B	Equation (3)	$P_B = 0.9 \times P$	Groundwater recharge
C	Equation (3)	$P_C = 1.1 \times P$	Underestimation P
D	Equation (4)	$E_{p,D} = 0.9 \times E_p$	Overestimation E_p
E	Equation (4)	$E_{p,E} = 1.1 \times E_p$	Underestimation E_p

Listed are experiment identifier, equation modified, formula for the modification and the reason for performing the experiment.

without introducing soil dynamics into the model. We therefore had to neglect this uncertain factor, but we can make a rough estimate in which cases it may be important.

Equation 3 implies that ΔS may have two effects on the runoff balance, depending on P and E_a. If $P > E_a$ the soil could 'soak up' some P decreasing R, whereas in times of $P < E_a$ soil storage could increase the water available for evaporation, increasing E_a. This latter would not affect runoff since the water would just be evaporated away, and only the former effect could actually affect R. Since this implies a reduction in available P, this effect is therefore also partially considered in sensitivity experiment B.

For each of the five sensitivity experiments, as well as the original model configuration, the measures *bias* and Δ_{50yr} are determined and are compared with each other in Section 4.3.

4 Model validation

4.1 Verification of annual and monthly runoff

In order to validate the model performance, we determine the mean annual runoff and compare it to estimates from other models of similar scale.

Models of similar scale are the macro-scale hydrological models WBM (Vörösmarty et al. 1996), WGHM (Döll et al. 2003), VIC (Nijssen et al. 2001; Liang et al. 1994), and Macro-PDM (Arnell 1999b; Meigh et al. 1999) on the one hand. On the other hand, one could also consider the land surface model of atmospheric GCMs (Russell and Miller 1990; Oki et al. 1999), and the Dynamic General Vegetation Model LPJ (Gerten et al. 2004). Unfortunately, the publication of actual numbers for the error in single river basins, as opposed to plots summarizing the error, is not very common. We therefore have to restrict the detailed comparison of model error to the numbers published by Russell and Miller (1990) and Nijssen et al. (2001).

The simulation quality of these models varies widely, but is much improved if the model parameters are tuned on a basin scale. For example, Döll et al. (2003) report a great increase in simulation quality after model tuning, similar to Nijssen et al. (2001). Since no tuning on the river basin scale takes place in our model, as there are no validation records available for some important river basins, we limit the comparison to the published errors before model tuning.

The simulation quality of the macro-scale models, where no such tuning on a basin scale takes place, generally is worse than desirable. Nijssen et al. (2001), for example, report biases ranging from −74.6% to 424.3%, with a median value of −18.1% for the untuned model, with increasing simulation quality after tuning. Similarly, Russell and Miller (1990) report biases ranging from −62.98% to 1018% with a median value of 33.93%.

Arnell (1999b) and Meigh et al. (1999) do not publish numbers for specific river basins, but judging from their plots, the biases range from about −50% to +20% for Arnell (1999b), where some tuning takes place for the whole continent of Europe, and from at least −50% to more than + 50% for Meigh et al. (1999), but in both cases the median bias seems to be quite small.

In Table 2 we are showing the simulation error for the annual runoff in those river basins, where either Russell and Miller (1990) or Nijssen et al. (2001) publish values for their models, and a direct comparison is therefore possible. While Nijssen et al. (2001) publish values for *bias*, Russell and Miller (1990) only publish values for mean annual runoff, both simulated and observed, and the *bias* has to be inferred from these. Overall, the *bias* of our model shows a similar spread of values as both Nijssen et al. and Russell and Miller, with the exception of the very extreme values our model produces in the Colorado and Murray basins.

Table 2 Error in those river basins, where either Nijssen et al. (2001) or Russell and Miller (1990) publish values

River	Δ_{50yr}[%]	d	$bias$[%]	$bias_N$[%]	$bias_R$[%]
Amazon	11.48	0.34	−30.79	−39.80	−62.98
Amur	11.11	0.86	−8.33	−45.90	−2.77
Chang Jiang	20.45	0.43	−32.98	−14.30	44.89
Colorado	n. a.	0.10	2,120.39		315.00
Columbia	4.00	0.65	−19.90	−74.30	20.72
Danube	27.14	0.80	6.21	12.30	44.66
Dvina	0.64	0.75	−4.78	31.30	10.38
Fraser	25.18	0.75	−11.19		33.93
Indigirka	24.00	0.39	−56.75	−54.70	
Indus	3.49	0.60	40.06		26.05
Kolyma	1.12	0.50	−42.71	−32.00	376.06
Lena	9.77	0.36	−38.66	−68.20	5.84
Mackenzie	33.07	0.48	−20.28	−69.00	83.66
Magdalena	21.58	0.57	−24.58		32.49
Mekong	5.67	0.63	−12.12	−19.10	51.49
Mississippi	2.58	0.65	31.95	18.00	−10.86
Murray	−18.05	0.08	1,490.34		431.82
Niger	25.19	0.12	336.75		82.81
Nile	16.55	0.05	508.47		606.02
Ob	18.90	0.73	7.11	46.50	30.91
Olenek	20.99	0.52	−40.97	−36.70	
Parana	10.22	0.28	93.26	6.20	
Pechora	12.42	0.48	−29.26	16.30	
Senegal	34.21	0.20	144.59	424.30	
Shatt el Arab	4.92	0.80	3.53		71.74
St. Lawrence	27.87	0.25	47.24		3.36
Volga	−13.33	0.53	26.90	83.60	
Yana	32.08	0.42	−52.28	−74.60	
Yenisei	12.69	0.28	−34.19	−44.40	−10.54
Yukon	−5.63	0.34	−48.87	104.80	152.31
Zambezi	−4.74	0.16	318.49		13.45

Shown are Δ_{50yr}, index of agreement d and $bias$ for our model, $bias_N$ for Nijssen et al. and $bias_R$ for Russell and Miller.

Taking all validation basins into account, the *bias* for our model ranges from −68.8% to 2,120.4%, with a median value of 9.5%, while the index of agreement *d* ranges from 0.05 to 0.93 with a median of 0.54.

In general the model overestimates runoff, 87 gauge sites (53%) show a positive *bias*. Of the 148 gauge records, 98 show an absolute *bias* below 50% and 67 below 25%. 15 gauge records have a *bias* above 250%. A histogram of the distribution of *bias* is shown in Fig. 2, along with the results of the sensitivity analysis.

The Colorado and Murray basins, where model *bias* is particularly large, as well as the Nile and some other validation basins, are located in very dry areas, and therefore a number of processes that are not considered in our model become important. First of all there may be seepage from the river channel, and the evaporation from open water may play a major role here as well, especially if the river runs through lakes or wetlands. For the Nile, Niger, Senegal and Orange similar problems are reported by Döll et al. (2003), while Oki et al. (1999) report such problems for the Colorado and Niger. In addition to these processes, basins like the Colorado are heavily managed by humans, and as these processes are not included in the model, they cannot be represented adequately either. This latter fact may well explain the very large bias our model shows for the Colorado basin, which is one of the most heavily managed river basins.

Model simulation quality with respect to the annual total runoff and the annual cycle of runoff therefore is comparable to other models of similar scope and scale, where no tuning on a river basin scale takes place, and a better performance would be desirable. We mainly attribute these performance problems to three causes. First of all, the Hamon scheme for the parameterization of potential evaporation (Eq. 4) basically rests on the assumption of uniform soil and vegetation characteristics. This leads to the potential evaporation scheme being more suitable to some river basins than to others. In addition, the neglect of soil storage of moisture and river routing may lead to additional errors, especially with regard to the timing of the annual cycle. Similarly, the simple parameterization of snow and snowmelt introduces additional errors into the model results.

4.2 Validation of runoff extremes

As we report in the methods section (Section 3.4), the return period of extreme runoff events is commonly evaluated by fitting a suitable probability distribution to the annual maxima of runoff. In the case of the streamflow records we have available, a gamma distribution turns out to be most suitable. By performing a Kolmogorov-Smirnov test, we determine whether the gauge records are compatible with this hypothesis. At 5% significance level, only two out of the 148 gauge records are rejected. These are the Colorado and Rio Grande basins, where extensive human influence on streamflow characteristics has to be assumed. These streamflow records are excluded from the subsequent analysis, leaving us with 146 gauge records for the validation of model extremes.

As the mean flows the model simulates are biased (Section 4.1), the extremes can only be compared after a suitable normalization of the data. After normalizing streamflow data and model results to a mean annual maximum streamflow/runoff of one, the probability distributions fitted to these data are in comparatively good agreement with another. In order to give the reader an impression of model simulation quality, we show plots of the estimated probability distributions at nine gauge sites. Figure 1 shows the probability distributions for the selected verification basins, as well as histograms of the number of annual maximum runoff events for the normalized event sizes as estimated from streamflow measurements. While the probability distributions are similar in every case, some

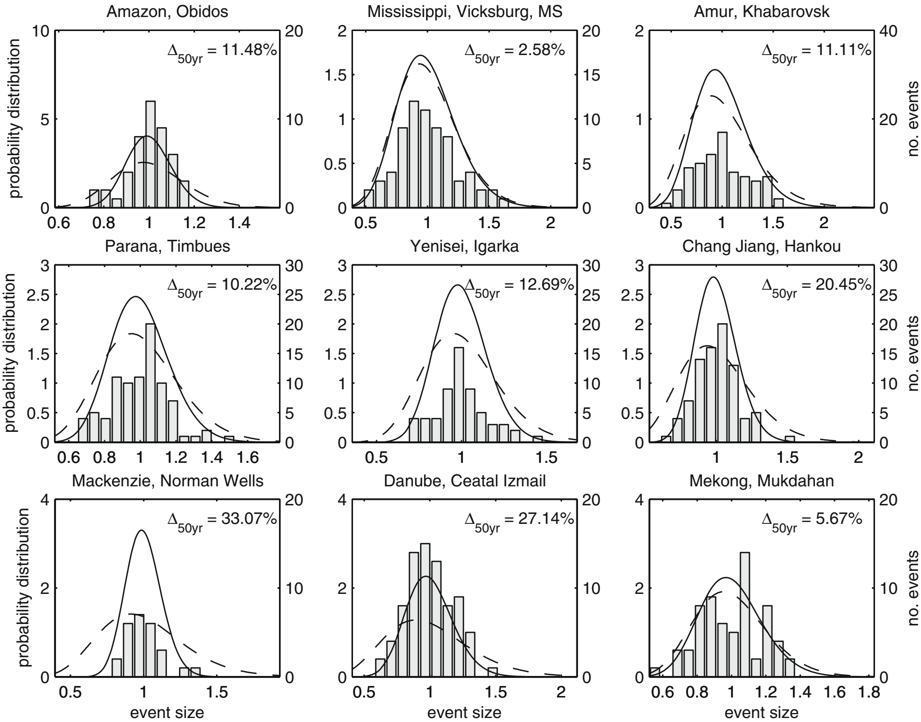

Fig. 1 Probability distributions for extremes at selected gauge sites. *Continuous line*: fit to normalized gauge record annual maxima, *dashed line*: fit to normalized model annual maxima. *Histograms* show the measured distribution of extremes. Also shown: Δ_{50yr}

differences are apparent. In all cases the probability distributions for the model-generated extremes are wider than the ones for the measured extremes. In addition, the peak of the probability distribution is higher in the case of the measured extremes. Therefore the model overestimates the probability of events that are larger or smaller than the mean event, while it underestimates the probability of the mean event sizes.

In order to quantify these errors, we determine the error Δ_{50yr} (Eq. 8) in the estimated 50-year extreme streamflow/runoff event.

Table 2 lists these values for selected river basins. The deviation of the 50-year extreme event ranges from an underestimation by −18.05% in the Murray to an overestimation by 34.21% in the Senegal. Taking all validation records considered into account, the deviation of the 50-year event between model and data is ranges from −36.11% to 47.02% with a median value of 3.53%. In 87 out of the 146 records considered, the 50-year event is overestimated. The absolute value of Δ_{50yr} stays below 10% in 66 (45%) of the 146 gauge records, and it stays below 25% in 130 cases (89%). The error was never larger than 50%. A histogram of the distribution of Δ_{50yr} is shown in Fig. 2, lower panel, along with results from the sensitivity experiments.

All in all, the agreement of the model simulated extreme events with the extreme events estimated from streamflow records is surprisingly good, considering the much larger bias in the annual and monthly flows. The error is below 10% in more than 45% of the gauge records evaluated, and no gauge displayed an error larger than 50%.

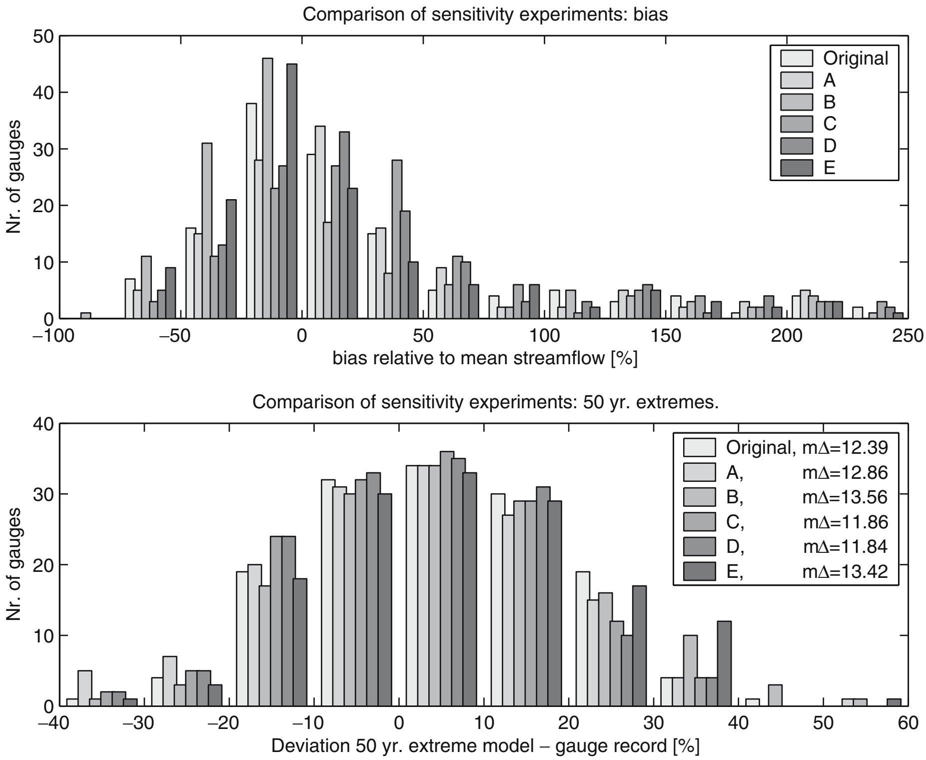

Fig. 2 Histogram of results for original model setup, as well as sensitivity experiments A–E, as defined in Table 1. *Upper panel*: *bias* relative to mean streamflow for sensitivity experiments. One hundred forty-eight gauge records considered, but between 7 and 22 (depending on experiment) not shown due to bias $>250\%$. *Lower panel*: deviation Δ_{50yr} of model simulated 50-year extremes from gauge record derived extremes, relative to gauge record derived extremes, for original configuration and sensitivity experiments. One hundred forty-six gauge records considered. Legend also shows mean absolute Δ_{50yr} as $m\Delta$

This good agreement of the probability distributions and of the 50-year maximum runoff event, after an appropriate normalization, leads us to the conclusion that the current model appears to be suitable to the evaluation of future probabilities of high runoff events, as long as the intercomparison of current and future probabilities takes place within the model results. Even though the annual and monthly flows the model simulates may be biased, the agreement of probability distributions fitted to streamflow data and model results suggests that the probability of high runoff events relative to the (biased) mean flows is estimated more or less correctly.

4.3 Sensitivity analysis

The simple model formulation allows a thorough analysis, which of the factors in the runoff balance (Eq. 3) has the largest influence on model performance. The sensitivity experiments we undertook are listed in Table 1, and the model results of the sensitivity analysis runs are subjected to the same analysis as above, namely a validation of the model extremes and of the mean flows.

Figure 2, upper half, shows a histogram of the *bias* relative to the mean streamflows at the gauge sites for all 148 gauge records considered. The mean absolute *bias* is highest (145%) in experiment A, where 10% of precipitation was instantly converted to runoff, while it is lowest (80%) in experiment B, where P was reduced by 10% to account for possible groundwater recharge. Model performance is improved in sensitivity experiments B and E (10% increase in E_p), while it is worse than the original in sensitivity experiments A, C (10% increase in P) and D (10% decrease in E_p). As the model generally overestimates runoff, this was expected. Precipitation is reduced in B and evaporation is enhanced in E, which in both cases reduces the overestimation of R.

Similarly, Fig. 2, lower half, shows a histogram of the deviations Δ_{50yr} of model simulated 50-year extremes from gauge record derived 50-year extremes, relative to the gauge record derived extremes, for the sensitivity experiments. The mean absolute Δ_{50yr} is shown as $m\Delta_{50yr}$ in the legend. Overall, the spread of the different sensitivity experiments is smaller for the extremes than for the means. The sensitivity experiments B and E performed worse than the original setup, while experiments A, C and D performed slightly better. The lowest mean absolute Δ_{50yr} (11.8%) is found in experiment D, while it is largest (13.6%) in experiment B.

Taking these results together, it seems recommendable to keep the original model setup. While sensitivity experiment D has the lowest mean absolute Δ_{50yr}, the result for the original setup is only slightly worse than that of experiment D. When looking at the mean flows, sensitivity experiments B and E perform best, while they perform worst when comparing the extremes. Setup C and D, on the other hand, would slightly improve performance with respect to the extremes, but they involve an arbitrary scaling of precipitation or evaporation and would also have a worsening effect on the mean flows.

If the neglect of soil storage was a major problem in the model, sensitivity experiment B should show improved results, as elaborated in Section 3.5. This is the case for the mean flows, but for the extremes results actually become worse. The choice of $\Delta S = 0$ in Eq. 3 therefore seems justified.

The sensitivity analysis has shown that there is no clear-cut 'best' model configuration, and it seems best not to introduce arbitrary scaling factors. Hence we will keep the original, most simple model configuration in the following assessment of changed climates.

5 Changed probabilities for extreme runoff events under climate change

5.1 A single scenario experiment

As an example of the potential changes in probability of extreme runoff events, we are showing a synthetic temperature change scenario and the corresponding timeseries of annual maximum runoff in Fig. 3. The top panel shows the change in global mean temperature, relative to the late twentieth century, in the climate change scenario. As we are using the CRU-PIK measurement data during the twentieth century, climate change is not shown during this timeframe. During the twenty-first century, global mean temperature rises rapidly and peaks in 2080 at a global mean temperature change $\Delta T = 4$K. Afterwards temperature decreases again, but in 2200 global mean temperature is still about 2K higher than during the twentieth century. For simplicity, climate variability is assumed to be the same sequence of variability patterns as measured during the twentieth century. The lower panels show annual maximum runoff in the Mississippi (*middle panel*) and Amazon (*bottom*) basins. Contrary to the runoff plots shown in Section 4.1, the runoff shown in

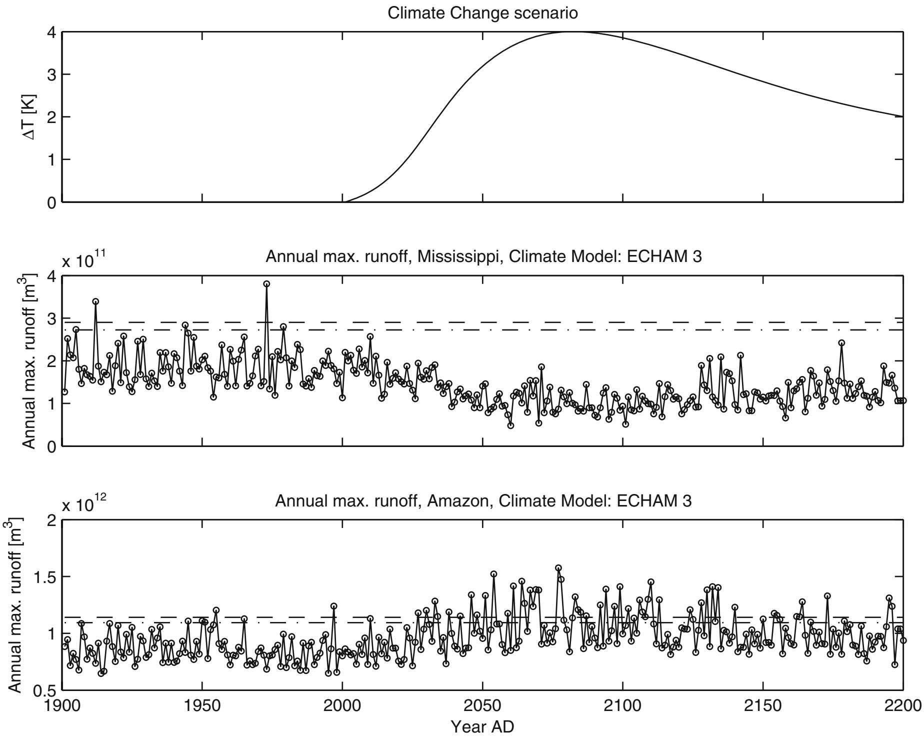

Fig. 3 Consequences of climate change in two river basins. *Top panel*: Climate change scenario, twentieth century not shown because driven by CRU-PIK data. *Lower panels*: Annual maximum runoff, model-generated, for the Mississippi (*middle*) and Amazon (*bottom*) basins. Also shown: 50-year maximum runoff event (*dashed line*) and 25-year maximum runoff event (*dash-dotted line*)

these plots is not the annual total summed up over sub-basins belonging to some streamflow gauge, but the runoff shown is the annual maximum monthly area-weighted sum of all the grid cells belonging to a drainage basin. The runoff timeseries is therefore comparable to the annual maximum streamflow timeseries given by a gauge located at the river mouth. The plots also show the level of the 50-year maximum runoff event during the twentieth century (*dashed line*) and the level of the 25-year event (*dash-dotted line*). These were derived by fitting a gamma distribution to the model-generated annual maxima of runoff. Climate change patterns for this plot were derived from ECHAM 3.

It is clearly visible in Fig. 3, that the annual maxima of runoff in the Mississippi basin decrease in magnitude. Both the 25-year and the 50-year maximum runoff events during the twentieth century are never exceeded during the next centuries. The probability of flooding therefore decreases in the Mississippi basin. In the Amazon basin, on the other hand, the picture is quite different. Here, the 25-year event is exceeded 59 times, while the 50-year event is exceeded 49 times during the twenty-first and twenty-second centuries. If the system were in a stationary state (which it clearly isn't), the 25-year event would become a 3.1-year event, while the 50-year event would become a 3.6-year event. The probability of major runoff events therefore clearly increases.

The model allows the determination of the change in flooding probability depending on the amount of global mean warming. We assess the changes in flooding probability for 83 of the largest river basins, where 50% of the projected world population in 2100 live. These basins are listed in the Appendix. In order to do this, we simulate 100 years of monthly runoff data for increased global mean temperatures, ranging from 0.1K to 5K in steps of 0.1K. The sampling sequence of the deviation patterns was as in the twentieth century. As described above, we fit a gamma distribution to the timeseries of annual maximum runoff and are thus able to assess the change in probability of a runoff event of equal magnitude to what was the 50-year maximum runoff event during the twentieth century.

Results of this assessment for nine large river basins are shown in Fig. 4. This figure also shows the uncertainty that arises through the difference in GCM projections, since we use climate change patterns generated by three different GCMs. While temperature projections by the GCMs differ only moderately, the precipitation changes by the different GCMs differ strongly. These models differ in many details, especially in their parameterizations of sub-gridscale processes which leads to quite different precipitation projections.

The changes in probability are quite heterogeneous. While the probability $P\left(Q_{50yr}\right)$ of the twentieth century 50-year event Q_{50yr} clearly increases in some river basins, there are other river basins where the magnitude Q_{50yr} of the 50-year event is never reached at all. Using the patterns generated by ECHAM 3, shown as dashed lines, the probability increases markedly with rising temperatures in the Amazon, Parana, Chang Jiang and Mekong basins. Other river basins, namely the Mississippi, Amur, Mackenzie and Danube

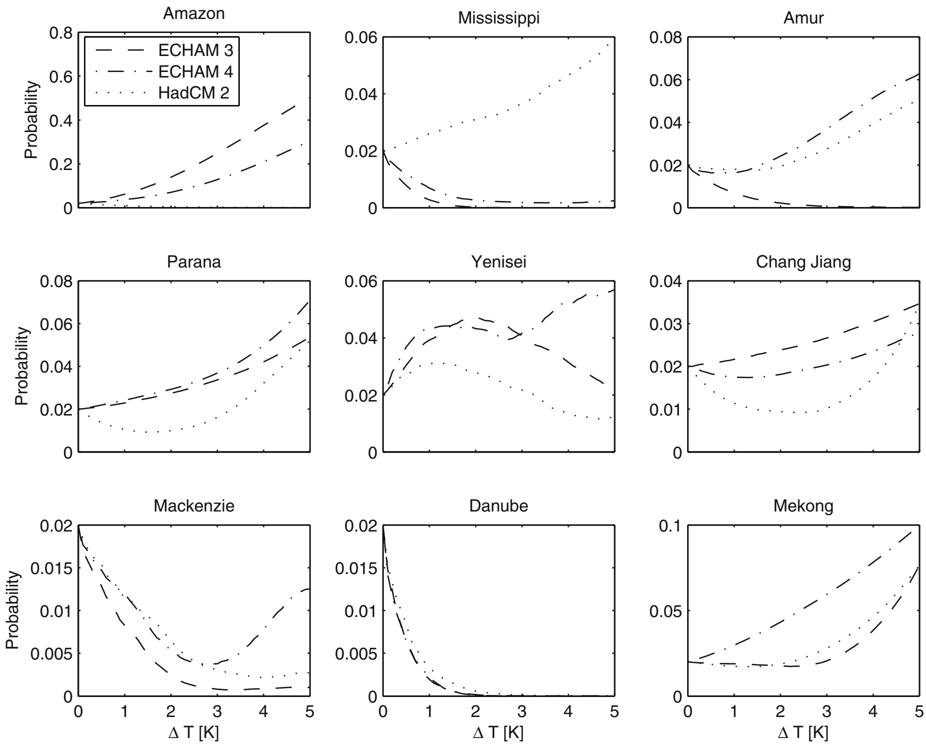

Fig. 4 Changed probabilities for the twentieth century 50-year maximum runoff event ($P = 0.02$) depending on change in global mean temperature ΔT. Determined using climate change patterns from ECHAM 3 (*dashed line*), ECHAM 4 (*dash-dotted line*) and HadCM 2 (*dotted line*)

river basins, experience a marked decrease in $P(Q_{50yr})$, while flooding probability in the Yenisei basin first increases and then decreases again. The climate change patterns produced by ECHAM 4, shown as dash-dotted lines, give a similar overall picture, with the exception of the Amur, Yenisei and Mackenzie basins. The most interesting of these cases are the Yenisei and the Mackenzie. While ECHAM 3 simulates an increase in $P(Q_{50yr})$ at temperature changes up to about 2 K for the Yenisei basin, followed by a decrease, ECHAM 4 simulates a faster initial increase followed by a short decrease, which is again followed by an increase in probability. A similar behavior is apparent in the Mackenzie basin. Here, both models project an initial decrease in $P(Q_{50yr})$, but ECHAM 4 simulates an increase in probability at climate changes larger than 2.75 K, while ECHAM 3 projects no further change in $P(Q_{50yr})$. This difference is due to changes in the annual cycle of runoff in the ECHAM 4 model. While the patterns generated by ECHAM 3 project that the annual maximum of runoff occurs in May, ECHAM 4 simulates a shift of the annual maximum of runoff to April, due to earlier snowmelt, and as evaporation is smaller in April due to both the shorter day length and lower temperatures, this generates increases in flooding probability. In the Amur basin, the different projection by the two models is simply due to different precipitation projections, with ECHAM 4 simulating increases, while ECHAM 3 produces decreases in precipitation.

Looking at the climate change generated by HadCM 2, the largest difference to the ECHAM models occurs in the Mississippi basin, where HadCM 2 projects an increase in $P(Q_{50yr})$, while the ECHAM models simulate a decrease. This is once again due to different precipitation patterns derived from the different models. HadCM 2 projects an increase in precipitation, while the ECHAM models project a decrease.

5.2 Climate impact response function

Climate impact response functions (CIRF; Füssel et al. 2003; Füssel 2003) have been developed as reduced-form models in order to enable the representation of the impacts of climate change in integrated assessment models. A CIRF is a representation of the relation between climate change and CO_2 concentration on the one hand, and the impact(s) of climate change under consideration on the other hand. While CIRFs were embedded within a deterministic framework previously, the approach presented here is the first attempt at using CIRFs in a probabilistic setting.

In order to determine a CIRF that is a suitable indicator for changes in flooding probability on a global scale, the results on the scale of single river basins have to be aggregated to the global scale in some way. Aggregating these changes in probability to a global level – after all we have performed this analysis in 83 of the largest river basins – is nontrivial, as the aggregation of the change in probability over all river basins may very well mask the severity of the problem, as decreasing probabilities in some river basins may mask the strong increases in other river basins. Therefore we determine the population affected by increasing probabilities of large runoff events. In order to do this, we use the dataset of population density by CIESIN (2000), which we extrapolate to the population in 2100 by using the regionalized IIASA median population scenario (Lutz et al. 2004), to determine the population living in the river basins analyzed.

This measure may not quite represent the number of people that are actually affected by the change in flooding probability. Not all the people living in a river basin will be affected by the changed flooding probability, but it seems safe to assume that the majority of the population living in a river basin lives close to the river and will therefore be affected by the change in flooding probability. Furthermore, the overall damage by a flood does affect an entire region, e.g., by demand for financing of the reconstruction of destroyed infrastructure. Therefore the number of people living in a river basin is a reasonable first approximation to the number of people affected by a change in $P(Q_{50yr})$.

Results for this analysis, derived using the climate change patterns from the three GCMs, are shown in Fig. 5. Using the climate change patterns obtained from ECHAM 3, shown in Fig. 5, *upper panel*, one can see that the population affected by a change in probability of the former 50-year event Q_{50yr} to a 25-year event $P(Q_{50yr}) = 1/25$ (marked by plus signs) rises steeply for a global warming $\Delta T \geq 0.3$K. The rise in fraction of world population affected then slows at a global warming $\Delta T = 0.5$K, where about 15% of world population are affected. The fraction of world population affected finally reaches about 31% at $\Delta T = 5$K. The non-smooth nature of these curves is due to the fact that once a basin crosses the threshold, its population is added to the total at once. The large initial increase in the plots for ECHAM 3 and ECHAM 4, for example, is mainly due to the Ganges basin with its projected population of 762 million in 2100 crossing the threshold.

This series of figures also highlights the uncertainty in these estimates. If one considers the fraction of population obtained using the climate change patterns from ECHAM 4, shown in Fig. 5, *middle*, the overall shape of the curves is similar to the ones obtained using ECHAM 3, while the threshold temperatures may be somewhat shifted. Using HadCM 2, shown in Fig. 5, *bottom*, the overall picture is quite different. The fractions of world population affected are significantly lower, and the increases are less steep than in the cases using the ECHAM models. This difference between the projections by the different models is largely due to the different estimates of future monsoon rainfall. While the ECHAM models project increases in monsoon precipitation, HadCM 2 projects a decrease, and due to the large population in the Ganges basin, this has a large effect on the projected population affected.

The dependence of the population affected by a change in $P(Q_{50yr})$ on climate change shown in Fig. 5 can be interpreted as a CIRF within this context. This model-derived function relates the fraction of world population affected by a change in flooding probability to the amount of climate change causing this change in flooding probability. In the final section, this CIRF is used within the TWA to calculate emission corridors, where the fraction of world population living in river basins affected by changes in flooding probability is limited.

6 Emission corridors limiting the change in flooding probability

In the tolerable windows approach (TWA) (Petschel-Held et al. 1999; Toth 2003; Bruckner et al. 2003), the aim is to determine emission corridors, i.e., the complete set of emission reduction strategies that are compatible with predefined normative constraints. These constraints are called ‘guardrails’ in the TWA.

In order to limit the population affected by a change in flooding probability, the relation between change in flooding probability and temperature change, developed in Section 5.2, can be used as a CIRF within the framework of the TWA.

In order to obtain the emission corridors, we are using the ICLIPS climate model first presented in Petschel-Held et al. (1999) and described further by Kriegler and Bruckner (2004). The model is kept as used by Kriegler and Bruckner (2004) with the exception of two changes. First of all, the reference period of the climatology we are using is 1961–1990. Therefore, this timeframe also defines the initial conditions the model uses to calculate future climate states. Secondly, as the model contains just a primitive carbon cycle and no other greenhouse gases, we are using a CO_2-equivalent formulation. In this formulation, the

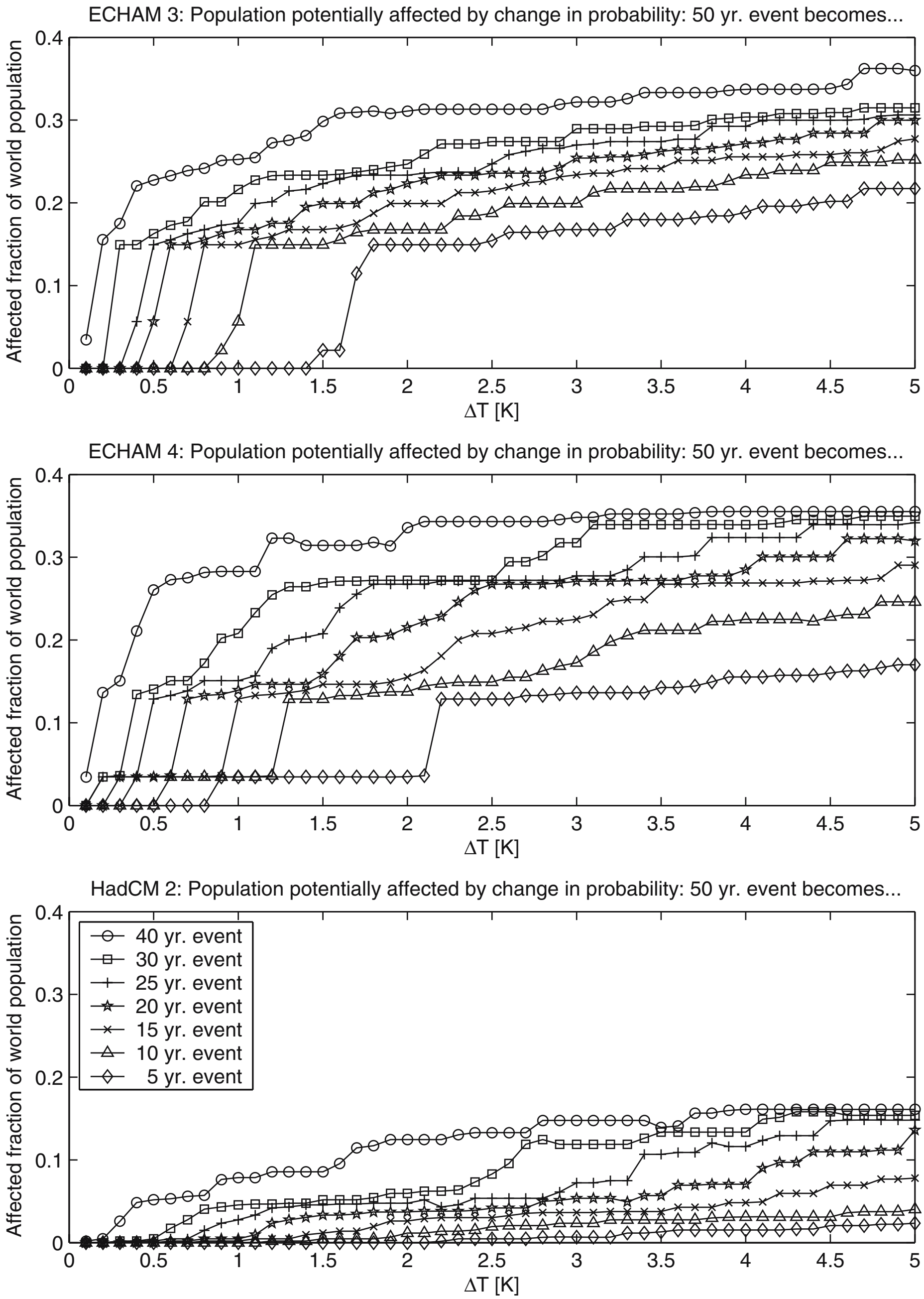

Fig. 5 Fraction of world population living in river basins affected by changed probability of 50-year maximum runoff event $P\left(Q_{50yr}\right)$, dependent on change in global mean temperature ΔT. Climate change patterns were taken from ECHAM 3 (*upper panel*), ECHAM 4 (*middle*) and HadCM 2 (*bottom*). The legend for all plots is shown in the *bottom panel*

radiative forcing by all forcing agents is converted to the CO_2 concentration that would generate the same radiative forcing. Climate sensitivity is set to 3K.

As a guardrail, a normative constraint that is not to be exceeded by climate change, various settings are possible. Here, we are concentrating on the change in probability of the 50-year maximum runoff event $P(Q_{50yr})$, as calculated by the model when forced with twentieth century observed climate, yet other events can easily be used. We are using $P(Q_{50yr})$ for two reasons. First of all, we believe that it would be misleading to estimate the size of events that have an even smaller probability from a timeseries that is just 100 years long. Second, the amount of runoff that is reached or exceeded only once in 50 years is already so large, that it seems plausible that this level will in many cases already cause major damage to infrastructure and endanger human lives, unless protection measures are undertaken. The 50-year event during the twentieth century therefore seems to be a suitable benchmark to compare future climate states with. As guardrails we are using limits to the population that live in the river basins affected by a change in $P(Q_{50yr})$.

Following Kriegler and Bruckner (2004), three further constraints are imposed on the change in emissions. The change in emissions is parameterized as $\dot{E} = gE$, and we are limiting the maximal emission reduction to 4% *p.a.*, as large emission reductions may be very costly. In addition, we are also limiting the rate of change in emission reduction, as a certain inertia in the socio-economic system has to be assumed. We are assuming a transition time scale t_{trans} of $t_{trans} = 20$ *yrs* from the initial rate of change in emissions g_0 to the maximal emission reduction $g_{max} = -0.04$. We are also assuming that the growth rate in emissions does not rise again, after emission reductions have started, for plausibility reasons. The latter two constraints can be summarized as $0 \leq \dot{g} \leq -(g_0 + g_{max})/t_{trans}$.

The corridor boundaries are then calculated by performing a constrained optimization, where the maximum (minimum) in emissions allowed by the constraints is determined for successive points in time in order to determine the upper (lower) boundary of the emission corridor (Leimbach and Bruckner 2001; Bruckner et al. 2003). The initial growth in emissions g_0 is determined by the optimization as well, but limited to be between 1% *p.a.* and 3% *p.a.*, which is close to the range of the late twentieth century growth in emissions.

Figure 6 shows such emission corridors. These corridors show the CO_2-equivalent emissions that are possible, if not more than 20% of the world population in 2100 are to be affected by a change in probability of the 50-year maximum runoff event, based on the climate change patterns generated by ECHAM 3. The plot shows the emission corridors for a change of $P(Q_{50yr})$ to the new probabilities shown in the legend. The actual emission corridor is the total shaded area between the upper boundary of the respective shaded area and the lower boundary of all the shaded areas. Please note that the upper boundaries of the 30-year, shown as a dotted line with circles, and the 25-year emission corridors, shown as a dotted line with diamonds, are actually located *below* the lower boundary. The emission corridors therefore are empty sets: only emission reduction strategies that involve emission reductions larger than 4% *p.a.* would produce a valid solution, and as we limit emission reductions to 4% *p.a.* for socio-economic reasons, this guardrail cannot be observed.

When interpreting these corridors, it is important to keep in mind that the corridors derived this way are *necessary* corridors. This means that all emission strategies that lie outside the corridor, or leave the corridor at some point in time, definitely violate the guardrail. For emission strategies that lie completely within the corridor, one has to check, whether they violate the guardrails or not. Especially emission strategies that stay close to the upper boundary of the emission corridor for most of the time are not acceptable. For further information on the interpretation of emission corridors see Kriegler and Bruckner (2004).

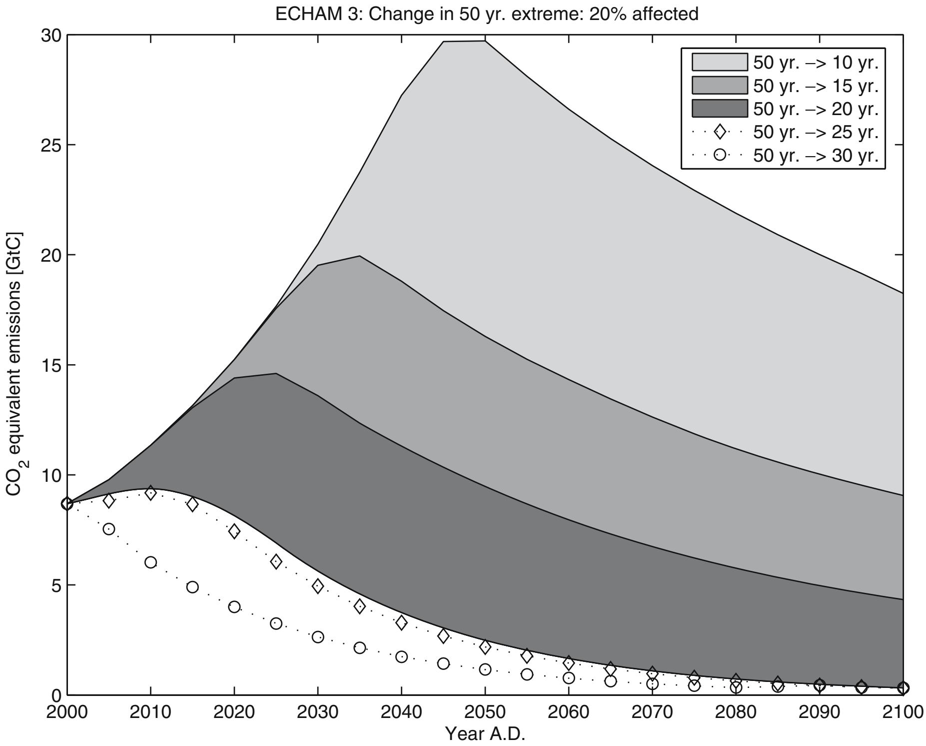

Fig. 6 Emission corridor limiting the change in $P(Q_{50yr})$. Maximal CO_2 equivalent emissions allowable, if less than 20% of world population are to be affected by a change in probability of the 50-year maximum runoff event to the new probability shown in the legend. Based on the climate model ECHAM 3

Figure 7 presents a different perspective to the emission corridors. In Fig. 7, isolines are presented that mark the maximum of the emission corridors for varying changes in probability and population affected. This figure also highlights the considerable uncertainty that is still inherent in this analysis, due to the different climate change patterns generated by the different GCMs. Shown are isoline diagrams for the GCM patterns considered, with ECHAM 3 shown on the upper left, ECHAM 4 on the upper right, and HadCM 2 on the lower left. On the lower left-hand side of the figures, no emission corridor exists that could limit the population affected by the changed flooding probability to these numbers. This is due to the fact that the maximum in emissions of the allowable minimum emissions trajectory is 9.4 GtC, due to the transition time scale and the maximum emission reductions imposed, which still implies a temperature change of about 1.3°C relative to the 1961–1990 average global mean temperature. Emissions above a maximum of 60 GtC were not evaluated, since these imply temperature changes larger than 5°C – a temperature change, where the simple climate model we are using is not applicable anymore.

If the ECHAM models should prove to be correct, it will be impossible to prevent 20% of the world population from being affected by the 50-year maximum runoff event becoming a 25-year event, and more than 10% will be affected by even larger changes in probability. This is mainly due to the large increases in precipitation that the ECHAM models project for the Ganges basin. If, on the other hand, HadCM 2 should prove to be correct, the population

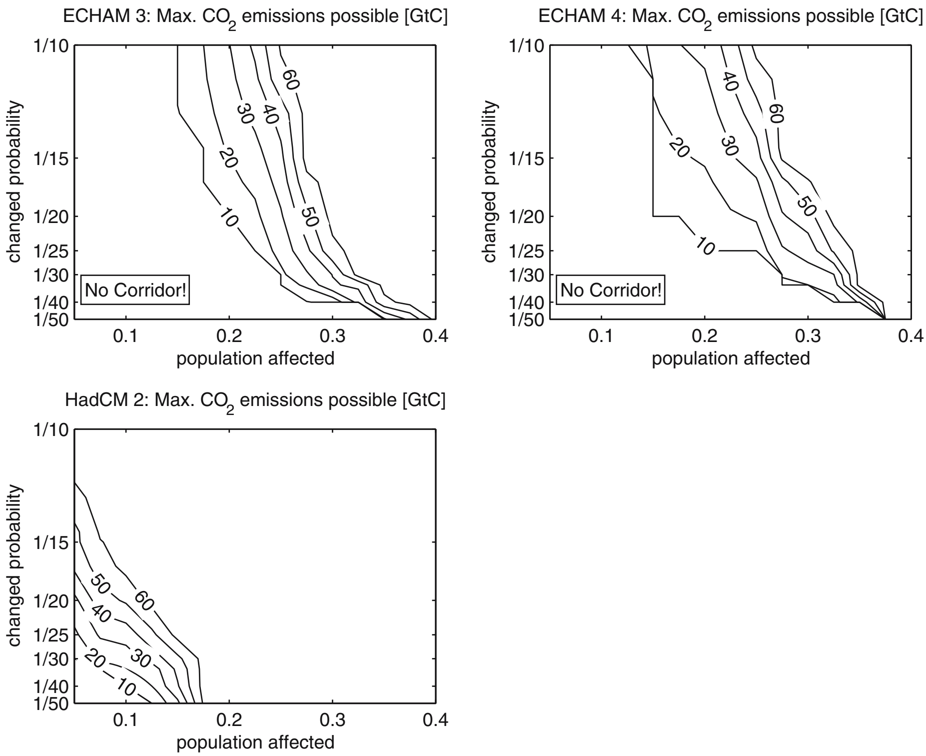

Fig. 7 Maximum CO_2 equivalent emissions [GtC] of the emission corridors for the climate change patterns generated by all three GCMs. Shown are the maximal CO_2 equivalent emissions allowed, if the fraction of world population, shown on the *abscissa*, affected by the change in $P\left(Q_{50yr}\right)$ to the new probability shown on the ordinate is to be limited. In the *lower left-hand corner* of the three plots no viable emission corridors exist

affected will be less dramatic, but it will still be impossible to prevent 10% of world population from being affected by a change of the 50-year to a 40-year event.

7 Discussion and conclusions

We have presented an approach to allow the representation of changes in probabilities of large-scale flood events in the integrated assessment of climate change. We have developed a downscaling scheme that enables us to use the changes in global mean temperature calculated by integrated assessment climate models to determine the changes in precipitation and evaporation on a river basin scale, including a representation of natural variability. These are then used to drive a hydrological model that aggregates the changes to river basin scale, and an assessment of changes in flooding probability can be performed.

Throughout the paper we have attempted to be very clear about uncertainties and shortcomings in our approach, and some need to be repeated here.

First of all, the type of flood events that can be considered is quite restricted. Due to limited temporal and spatial resolution, events on small spatial or temporal scales cannot be considered adequately. This, unfortunately, is the price one has to pay for conducting such an assessment on a global scale, which is, in our opinion, a necessity for the integrated assessment of climate change. On the other hand, the model validation shows that model performance is satisfactory for large-scale events, the so-called plain floods, and here model performance actually improves for the assessment of extreme events, as compared to the performance for the mean flows. We think, therefore, that our assessment of changes in flooding probability is meaningful.

A second major shortcoming is the uncertainty that comes from the different changes in the mean climate that are projected by different GCMs. This uncertainty has to be accepted for the time being. We try to incorporate it by using climate change patterns derived from different GCMs, but the representation of this uncertainty can still be improved.

The third problem is the assumption that spatio-temporal variability of climate will stay the same in a changed climate. It is likely that this will not be the case, and there is evidence from the diagnosis of some GCM simulations (e.g., Kharin and Zwiers 2000), that climate change may actually increase probabilities of extreme precipitation events and therefore floods. Once again there is scope for improvement of our approach, and our results may turn out to be a lower bound on the change in probability.

The shortcomings of our approach can, unfortunately, not be avoided completely when dealing with this subject matter on a global scale. On the other hand, this is the first assessment of this nature on this scale that we are aware of, and future developments will undoubtedly allow improvements to be made. At the same time we firmly believe that integrated assessment has to take into account changes in extreme events because it is through these changes that many of the most widespread consequences of climate change will be felt.

The modeling results presented in the previous sections suggest that changes in the probability of large-scale flooding due to changes in precipitation induced by future climate change might have a severe impact on a significant portion of the world's population. Not only does the simulation with a single climate change scenario as in Section 5.1 suggest an increase in probabilities for large-scale floods, but even more significant are the results obtained within the application of the tolerable windows approach.

Within this application of the TWA, the portion of the world population experiencing an increased probability of what is today a 50-year event has been implemented as a constraint for future climate change. Within this first step, a climate impact response function (CIRF) is implemented, which is based on the hydrological model presented before. This CIRF gives the proportion of world population which experiences a specified shift in flooding probabilities as a function of the global mean temperature. In a second step, the corridors of admissible emissions were calculated, which comply with this constraint and which do not exceed a reduction rate of more than 4% *p.a*. Both the climate impact response function and the resulting corridors suggest that:

– There is a significant risk that even a small increase in global mean temperature by less than 0.5°C brings about a significant increase in flooding probabilities which can affect up to 20% of the world population. Here, results differ with different spatial patterns of climate change obtained from three GCMs. More specifically, the risk depends on the fate of the Indian Monsoon, as the two ECHAM GCMs both show a strengthening. Therefore, the uncertainties associated the future behavior of the monsoon are not only of relevance for agriculture, but also for floods.
– If the changes in mean climate projected by the ECHAM models should turn out to be right, there is no reasonable emission scenario to insure that only small proportions of the world population are affected by increases in the probabilities of major floods.

If, for example, we consider a proportion of 20% of the world population, we have to reckon with shifts in probabilities, where what has been a 50-year event in the twentieth century becomes at least a 25-year event over the next 100 years.
– The danger of such unavoidable consequences of climate change implies that adaptation to increasing flooding probabilities are inevitable. Given the possibility that these shifts might happen with relatively small increases in global mean temperature, adaptation measures need to be taken soon, which calls for an increasing effort to study and understand the processes of adaptation.

Despite all the uncertainties mentioned, these conclusions are quite robust, and we consider the model as good enough to conclude that an increase in flooding probabilities is a major reason for concern about climate change. Increased modeling efforts need to be undertaken to localize the critical regions for increased flooding, in order to get improved information for adaptation priorities.

Acknowledgments We gratefully acknowledge the helpful suggestions by Zbigniew Kundzewicz and an anonymous reviewer. T. K. acknowledges support by the Volkswagen foundation and by the Deutsche Forschungsgemeinschaft.

Appendix

List of river basins considered

Table 3 River basins considered in the assessment

Name	Population 2100 [10^6]	Area [$10^5 km^2$]
Ganges	762	16.33
Indus	284	11.46
Niger	180	22.46
Zaire	157	37.09
Huang He	128	8.96
Parana	128	26.69
Huai	125	2.45
Krishna	108	2.52
Mississippi	104	32.12
Godavari	100	3.12
Hai Ho	93	2.46
Shatt el Arab	87	9.70
Zhujiang	80	4.10
Zambezi	79	19.94
St. Lawrence	71	12.70
Damodar	61	0.60
Amur	61	29.11
Mekong	60	7.76
Danube	54	7.90
Amazon	50	58.70
Balsas	48	1.23
Brahmani	46	1.42
Syr-Darya	44	10.73
Volta	44	3.99
Amu-Darya	43	6.14

Table 3 (continued)

Name	Population 2100 [10^6]	Area [$10^5 km^2$]
Limpopo	43	4.21
Magdalena	42	2.52
Rhine	41	1.66
Irrawaddy	40	4.07
Volga	35	14.67
Cauweri	35	0.79
Liao	34	2.75
Jubba	34	8.18
Narmada	32	1.14
Grande de Santiago	31	1.92
Tapti	28	0.67
Chari	27	15.76
Jordan	27	2.70
Orange	24	9.46
Orinoco	24	10.42
Fuchun Jiang	23	0.67
Hong	23	1.71
San Francisco	23	6.17
Ob	22	25.77
Chao Phraya	21	1.42
Galana	21	1.18
Elbe	20	1.49
Brahmani	19	0.58
Cross	19	0.52
Rabarmati	19	0.28
Dnepr	19	5.10
Panuco	18	0.92
Po	18	1.02
Mahi	17	0.29
Sacramento	17	1.93
Tana (Ken)	16	0.99
Kizil Irmak	15	1.10
Penner	15	0.54
Wisla	15	1.81
Seine	13	0.74
Dongjiang	13	0.34
Senegal	13	8.50
Paraiba do Sul	13	0.63
Don	12	4.24
Menjiang	12	0.66
Meuse	11	0.43
Jacui	11	0.81
Kura	11	2.20
Hudson	11	0.43
Rufiji	11	1.87
Trinity	11	0.48
Urugay	10	3.56
Farah	10	3.86
Bandama	10	1.04
Columbia	10	7.26
Cuanza	10	1.64

Table 3 (continued)

Name	Population 2100 [10^6]	Area [$10^5 km^2$]
Cheliff	9	0.58
Sebou	9	0.39
Motagua	9	0.27
Asi	9	0.28
Comoe	9	0.83
Odra	9	1.20
Sassandra	9	0.77

Listed are river basin name, population in 2100 in millions, and river basin area in $10^5 km^2$.

References

Adam JC, Lettenmaier DP (2003) Adjustment of global gridded precipitation for systematic bias. J Geophys Res 108(D9):4257

Alcamo J, Leemans R, Kreileman E (eds) (1998) Global change scenarios of the 21st century. Elsevier, Oxford, UK

Alcamo J, Endejan MB, Kaspar F, Rösch T (2001) The GLASS model: a strategy for quantifying global environmental security. Environ Sci Policy 4(1):1–12

Arnell NW (1999a) Climate change and global water resources. Glob Environ Change 9(S1):S31–S49

Arnell NW (1999b) A simple water balance model for the simulation of streamflow over a large geographic domain. J Hydrol 217(3–4):314–335

Arora VK, Boer GJ (2001) Effects of simulated climate change on the hydrology of major river basins. J Geophys Res 106(D4):3335–3348

Berz G (1999) Catastrophes and climate change: concerns and possible countermeasures of the insurance industry. Mitig Adapt Strategies Glob Chang 4(3–4):283–293

Bodo BA (2001a) Monthly discharge data for world rivers (excluding former Soviet Union). Version 1.3, available online at http://dss.ucar.edu/datasets/ds552.1

Bodo BA (2001b) Monthly discharges for 2400 rivers and streams of the former Soviet Union. Version 1.1, available online at http://dss.ucar.edu/datasets/ds553.2

Bronstert A, Niehoff D, Bürger G (2002) Effects of climate and land-use change on storm runoff generation: present knowledge and modelling capabilities. Hydrol Process 16(2):509–529

Bruckner T, Petschel-Held G, Leimbach M, Toth FL (2003) Methodological aspects of the tolerable windows approach. Clim Change 56(1–2):73–89

Cameron D, Beven K, Tawn J (2000) An evaluation of three stochastic rainfall models. J Hydrol 228 (1–2):130–149

CIESIN (2000) Gridded Population of the World (GPW), Version 2. Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); and World Resources Institute (WRI). CIESIN, Columbia University, Palisades, New York. Available at http://sedac.ciesin.columbia.edu/plue/gpw

Döll P, Kaspar F, Lehner B (2003) A global hydrological model for deriving water availability indicators: model tuning and validation. J Hydrol 270(1–2):105–134

Edmonds J, Pitcher MWH, Richels R, Wigley TML, MacCracken C (1996) An integrated assessment of climate change and the accelerated introduction of advanced energy technologies: an application of MiniCAM 1.0. Mitig Adapt Strategies Glob Chang 1:311–339

Federer CA, Vörösmarty C, Fekete B (1996) Intercomparison of methods for calculating potential evaporation in regional and global water balance models. Water Resour Res 32(7):2315–2321

Fekete BM, Vörösmarty CJ (2004) Uncertainties in precipitation and their impacts on runoff estimates. J Climate 17(2):294–304

Fekete BM, Vörösmarty CJ, Grabs W (1999) Global composite runoff fields on observed river discharge and simulated water balances. Report 22, Global Runoff Data Center

Fekete BM, Vörösmarty CJ, Grabs W (2002) High-resolution fields of global runoff combining observed river discharge and simulated water balances. Glob Biogeochem Cycles 16(3), http://dx.doi.org/10.1029/1999GB001254

Füssel H-M (2003) Impacts analysis for inverse integrated assessment of climate change. PhD thesis, Universität Potsdam

Füssel HM, Toth FL, van Minnen JG, Kaspar F (2003) Climate impact response functions as impact tools in the tolerable windows approach. Clim Change 56(1–2):91–117
Gerten D, Schaphoff S, Haberlandt U, Lucht W, Sitch S (2004) Terrestrial vegetation and water balance – hydrological evaluation of a dynamic global vegetation model. J Hydrol 286(1–4):249–270
Hamon WR (1963) Computation of direct runoff amounts from storm rainfall. Int Assoc Sci Hydrol Publ 63:52–62
Hasselmann K, Hasselmann S, Giering R, Ocana V, Storch Hv (1997) Sensitivity study of optimal CO_2 emission paths using a Simplified Structural Integrated Assessment Model (SIAM). Clim Change 37(2):345–386
Hennessy KJ, Gregory JM, Mitchell JFB (1997) Changes in daily precipitation under enhanced greenhouse conditions. Clim Dyn 13(9):667–680
Hipple JD, Drazkowski B, Thorsell PM (2005) Development in the upper Mississippi basin: 10 years after the great flood of 1993. Landsc Urban Plan 72:313–323
Hutchinson MF (1995) Stochastic space–time weather models from ground-based data. Agric For Meteorol 73(3–4):237–264
Johns TC, Carnell RE, Crossley JF, Gregory JM, Mitchell JFB, Senior CA, Tett SFB, Wood RA (1997) The second Hadley Centre coupled ocean–atmosphere GCM: model description, spinup and validation. Clim Dyn 13:103–134
Kharin VV, Zwiers FW (2000) Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. J Climate 13(21):3760–3788
Kriegler E, Bruckner T (2004) Sensitivity analysis of emissions corridors for the 21st century. Clim Change 66(3):345–387
Kundzewicz ZW, Schellnhuber H-J (2004) Floods in the IPCC TAR perspective. Nat Hazards 31:111–128
Kundzewicz ZW, Graczyk D, Maurer T, Przymusiñska I, Radziejewski M, Svensson C, Szwed M (2004) Detection of change in world-wide hydrological time series of maximum annual flow. WCASP-64, WMO Geneva
Leemans R, Cramer W (1991) The IIASA database for mean monthly values of temperature, precipitation and cloudiness of a global terrestrial grid. Rr-91-18, International Institute for Applied Systems Analysis (IIASA)
Leimbach M, Bruckner T (2001) Influence of economic constraints on the shape of emission corridors. Comput Econ 18:173–191
Li Y, Saxena KML, Cong S (1999) Estimation of the extreme flow distributions by stochastic models. Extremes 1(4):423–448
Liang X, Lettenmaier DP, Wood EF, Burges SJ (1994) A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J Geophys Res 99:14415–14428
Lockwood JG (1999) Is potential evapotranspiration and its relationship with actual evapotranspiration sensitive to elevated atmospheric CO_2 levels? Clim Change 41(2):193–212
Lutz W, Sanderson WC, Scherbov S (eds) (2004) The end of world population growth in the 21st century: new challenges for human capital formation and sustainable development. Earthscan, London
Madsen H, Rasmussen PF, Rosbjerg D (1997) Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. At-site modeling. Water Resour Res 33:747–757
Manne AS, Mendelsohn R, Richels RG (1995) MERGE – a model for evaluating regional and global effects of GHG reduction policies. Energy Policy 23:17–34
Meigh JR, McKenzie AA, Sene KJ (1999) A grid-based approach to water scarcity estimates for eastern and southern Africa. Water Resour Manag 13:85–115
Milly PCD, Wetherald RT, Dunne KA, Delworth TL (2002) Increasing risk of great floods in a changing climate. Nature 415:514–517
Mirza MMQ (2002) Global warming and changes in the probability of occurrence of floods in Bangladesh and implications. Glob Environ Change 12(2):127–138
Mitchell TD (2003) Pattern scaling: an examination of the accuracy of the technique for describing future climates. Clim Change 60(3):217–242
Mitchell JFB, Johns TC, Eagles M, Ingram WJ, Davis RA (1999) Towards the construction of climate change scenarios. Clim Change 41(3–4):547–581
Munich Re (2003) Topics-annual review: natural catastrophes 2002
New M, Hulme M, Jones P (2000) Representing twentieth-century space-time climate variability. Part II. Development of 1901–96 monthly grids of terrestrial surface climate. J Climate 13(13):2217–2238
Nijssen B, O’Donnell G, Lettenmaier DP, Lohmann D, Wood EF (2001) Predicting the discharge of global rivers. J Climate 14:3307–3323
Nordhaus WD (1994) Managing the global commons. The economics of climate change. MIT, Cambridge, Massachusetts
Oki T, Nishimura T, Dirmeyer P (1999) Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP). J Meteorol Soc Jpn 77:235–255

Österle H, Gerstengarbe F-W, Werner P-C (2003) Homogenisierung und Aktualisierung des Klimadatensatzes der Climate Research Unit of East Anglia, Norwich. In: Proc. 6. Deutsche Klimatagung, Klimavariabilität, pp 326–329

Petschel-Held G, Schellnhuber H-J, Bruckner T, Tóth FL, Hasselmann K (1999) The tolerable windows approach: theoretical and methodological foundations. Clim Change 41(3/4):303–331

Roeckner E, Arpe K, Bengtsson L, Christoph M, Claussen M, Dümenil L, Esch M, Giorgetta M, Schlese U, Schulzweida U (1996) The atmospheric general circulation model ECHAM-4: model description and simulation of present-day climate. Report 218, Max Planck Institute for Meteorology

Rotmans J, de Boois H, Swart RJ (1989) IMAGE: an integrated model to assess the greenhouse effect. Technical Report 758471009, RIVM, Bilthoven, The Netherlands

Russell GL, Miller JR (1990) Global river runoff calculated from a global atmospheric general circulation model. J Hydrol 117:241–254

Tol RSJ (2002) Estimates of the damage costs of climate change. Part I. Benchmark estimates. Environ Resour Econ 21:47–73

Toth FL (2003) Climate policy in light of climate science: the ICLIPS project. Clim Change 56(1–2):7–36

UNESCO (1974) Discharge of selected rivers of the world; a contribution to the International Hydrological Decade, Vol. 5. Unesco, Paris

Voss R, Sausen R, Cubasch U (1998) Periodically synchronously coupled integrations with the atmosphere–ocean general circulation model ECHAM3/LSG. Clim Dyn 14:249–266

Voss R, May W, Roeckner E (2002) Enhanced resolution modelling study on anthropogenic climate change: changes in extremes of the hydrological cycle. Int J Climatol 22(7):755–777

Vörösmarty CJ, Willmott CJ, Choudhury BJ, Schloss AL, Stearns TK, Robeson SM, Dorman TJ (1996) Analyzing the discharge regime of a large tropical river through remote sensing, ground-based climatic data, and modeling. Water Resour Res 32(10):3137–3150

Vörösmarty CJ, Federer CA, Schloss AL (1998) Potential evaporation functions compared on US watersheds: possible implications for global-scale water balance and terrestrial ecosystem modeling. J Hydrol 207(3–4):147–169

Vörösmarty CJ, Fekete BM, Meybeck M, Lammers RB (2000) Global system of rivers: its role in organizing continental land mass and defining land-to-ocean linkages. Glob Biogeochem Cycles 14(2):599–621

Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr 21(4):530–548

Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS (1998) Statistical downscaling of general circulation model output: a comparison of methods. Water Resour Res 34 (11):2995–3008

Wilks DS, Wilby RL (1999) The weather generation game: a review of stochastic weather models. Prog Phys Geogr 23(3):329–357

Willmott CJ (1982) Some comments on the evaluation of model performance. Bull Am Meteorol Soc 63:1309–1313

Willmott CJ, Matsuura K (2001) Terrestrial air temperature and precipitation: monthly and annual time series (1950–1999). Available online at http://climate.geog.udel.edu/~climate/html_pages/archive.html

Xu C-Y (1999) From GCMs to river flow: a review of downscaling methods and hydrologic modelling approaches. Prog Phys Geogr 23(2):229–249