Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios

M. E. Elshamy1,2, I. A. Seierstad3,4, and A. Sorteberg3,4

1Nile Forecast Center, Ministry of Water Resources and Irrigation, Egypt
2Nile Basin Research Programme, University of Bergen, Norway
3Bjerknes Center for Climate Research, University of Bergen, Norway
4Geophysical Institute, University of Bergen, Norway

Received: 7 May 2008 – Published in Hydrol. Earth Syst. Sci. Discuss.: 12 June 2008
Revised: 29 April 2009 – Accepted: 29 April 2009 – Published: 6 May 2009

Abstract. This study analyses the output of 17 general circulation models (GCMs) included in the 4th IPCC assessment report. Downscaled precipitation and potential (reference crop) evapotranspiration (PET) scenarios for the 2081–2098 period were constructed for the upper Blue Nile basin. These were used to drive a fine-scale hydrological model of the Nile Basin to assess their impacts on the flows of the upper Blue Nile at Diem, which accounts for about 60% of the mean annual discharge of the Nile at Dongola. There is no consensus among the GCMs on the direction of precipitation change. Changes in total annual precipitation range between −15% to +14% but more models report reductions (10) than those reporting increases (7). Several models (6) report small changes within 5%. The ensemble mean of all models shows almost no change in the annual total rainfall. All models predict the temperature to increase between 2°C and 5°C and consequently PET to increase by 2–14%. Changes to the water balance are assessed using the Budyko framework. The basin is shown to belong to a moisture constrained regime. However, during the wet season the basin is largely energy constrained. For no change in rainfall, increasing PET thus leads to a reduced wet season runoff coefficient. The ensemble mean runoff coefficient (about 20% for baseline simulations) is reduced by about 3.5%. Assuming no change or moderate changes in rainfall, the simulations presented here indicate that the water balance of the upper Blue Nile basin may become more moisture constrained in the future.

1 Introduction

It is widely accepted that Global Circulation Models (GCMs) are the best physically based means for devising climate scenarios. They reproduce the global and continental scale climate fairly well (e.g. Hewitson and Crane, 1996) but often fail to simulate regional climate features required by hydrological (catchment scale) and national (country scale) impact studies. The main reason for this gap, between the spatial scale of GCM output and that needed for impact studies, is the coarse spatial resolution of GCMs which restricts their usefulness at the grid-size scale and smaller (Wilby and Wigley, 1997). Other reasons include inadequate parameterization of several processes regarding cloud formation and land surface interactions with the atmosphere.

GCM experiments show very different pictures of climate change over the Nile basin. While they all agree on a temperature rise, they disagree on the direction of precipitation change. Earlier analysis of 16 transient experiments from 7 different GCMs (used for the IPCC Third Assessment Report – TAR) revealed an average increase in temperature over the basin by 2–4.3°C by 2050 (Elshamy, 2000). The study showed that temperature changes were not uniform over the basin, with larger temperature rises in the more arid regions of Northern Sudan and most of Egypt and smaller rises near the equator. Although most of the analyzed experiments showed an increase in precipitation over the basin (up to 18%), some experiments showed a reduction (up to 22%), while one experiment showed almost no change (Elshamy, 2000).

Even with increases in basin rainfall, river runoff may still decrease because of the expected increase in
evapotranspiration due to temperature rise. Based on climate change scenarios from two versions of the Hadley Centre GCM, Arnell (1999) estimated that the increase in evaporative demand would counterbalance the increase in basin rainfall so that runoff would remain virtually unchanged. Results from an earlier study (Yates and Strzepek, 1998), using 3 equilibrium experiments and one transient experiment, showed a wide range of changes. While three of the models indicated increases in natural river flow at Aswan of more than 50%, the fourth model showed a 12% reduction. The study expected that the temperature rise would increase evaporation losses from Lake Nasser as well as irrigation water demands. Considering such losses in addition to possible increases in Sudan abstractions, the study predicted changes in water availability for Egypt ranging between −11% to +61%. Later, Strzepek et al. (2001) showed that 8 out of 9 climate scenarios resulted in reductions of Nile flows during the 21st century. In one of the latest studies based on the results from 4 GCMs, Sayed (2004) expected that the change in rainfall over the Blue Nile basin would be between +2 and +11% for 2030, while rainfall over the White Nile basin would be between +2 and +11% over the Blue Nile basin. 

Thus, there are large uncertainties in predicting climatic changes over the Nile basin and their impacts on its flows. This complicates the development of water resources plans in basin countries. By nature, the future is uncertain and this is partly handled via emissions’ scenarios that capture different visions of how the world will develop in the future in terms of population, technology, and energy use. However, the recent scenarios used in the preparation of the 4th assessment report (AR4) of the IPCC (2007) have not yet been assessed for the Blue Nile basin.

This research aims to construct detailed scenarios for precipitation and evapotranspiration over the Blue Nile basin from as many GCMs as possible and to assess their impacts on the flows of the upper Blue Nile at Diem. The rainfall scenarios have been downcaled to the fine resolution required by the hydrological model. Potential (reference crop) evapotranspiration (PET) scenarios have also been developed from GCM output to be consistent with rainfall. Previous studies used simpler methods to calculate future PET; e.g. Conway and Hulme (1996) assumed a 4% increase in PET per degree increase in temperature while Nawaz et al. (2009) used the Thornthwaite temperature method to quantify the change. In this study, by using the physically-based FAO Penman-Monteith method (Allen et al., 1998), the assumption of pure temperature dependence of PET is evaluated. A fine-scale hydrological model of the Nile Basin (NFS, hosted at the Nile Forecast Center of the Ministry of Water Resources and Irrigation, Egypt) has then been used to study the impacts of the developed scenarios on Blue Nile flows.

Fig. 1. Location map of the upper Blue Nile basin (hatched) within the Blue Nile basin (outlined). Circles indicate important cities, triangles indicate river gauges, squares indicate dam locations, and stars indicate raingauges used in the estimation of gridded rainfall fields.

2 Study area

This research focuses on the upper catchment of the Blue Nile (Fig. 1) which extends from 7°45′ to 13°N and 34°30′ and 37°45′E. The Blue Nile is one of the most important tributaries of the Nile as it contributes (at Diem) about 60% of the annual Nile discharge at Dongola. The source of the Blue Nile is the Little Abbay, the largest tributary of Lake Tana, which originates in the Ethiopian Highlands. The Blue Nile leaves Lake Tana at Bahr Dar and flows to the southeast through a series of cataracts. The river then enters a canyon and changes direction to south and then to the west and finally to the northwest forming a large open loop. Along its 940 km journey from Bahr Dar to Diem, near the Ethiopian Sudanese border, the river is joined by several tributaries draining a large area of highlands in western Ethiopia. The elevation of the basin varies greatly from over 4000 m in the headwaters of some tributaries (e.g. the Little Abbay) to 700 m at the foot of the plateau.

The flow analysis has been performed for the Diem station, the outlet of the upper Blue Nile basin which covers an area of about 185 000 km². This sub-basin is characterized by a highly seasonal rainfall pattern with most of the rainfall falling in four months (June to September – JJAS) with a peak in July or August (Fig. 2). The mean annual rainfall for the 1961–1990 period amounts to a little over 1200 mm, of which more than 70% fall during those four months. The interannual variability of basin rainfall is large with a minimum value of 1148 mm and maximum of 1757 mm from an analysis of 11 raingauges over the 1900–1998 period.
(Conway, 2000). There is also considerable decadal variability. From the mid-1960s to the mid-1980s there was a marked decrease in rainfall (Conway, 2005). In the 1990s the rainfall increased again reflected in increasing Blue Nile flows during this period.

Flow lags rainfall by about one month (from a monthly analysis – a daily analysis would provide a more accurate estimate) peaking mostly in September with an average annual volume of about 46 BCM (billion cubic meters) for the 1961-1990 period. PET (calculated using the FAO Penman-Monteith method based on the dataset of the Climatic Research Unit of the University of East Anglia (CRU TS 2.1 – Mitchell and Jones, 2005) as explained in the following sections) is higher during the dry season (December to April) than in the wet season due to increased cloudiness and humidity associated with rainfall. The total annual PET has been estimated at about 1300 mm. The actual evapotranspiration (AET), calculated using the NFS hydrological model, only equals PET during the wet season when there is enough moisture in the soil and thus it only amounts to about 900mm annually (about 70% of PET). The average runoff/rainfall coefficient of the basin is about 20%, which means that a large percentage of rainfall is lost to evaporation indicating the importance of this element in the water balance of the basin. Some of the evaporation also contributes to rainfall in the same region. Mohamed et al. (2005) estimated that the annual moisture recycling over the entire Nile basin is about 11%. Most of the Nile water thus comes from outside the basin boundaries.

Shallow groundwater recharged by rainfall during the rainy season contributes to the baseflow component later during the dry season (Asmerom, 2008). Little is known about deep groundwater recharge and flow to/from adjacent groundwater basins and thus, most previous hydrological studies established the water balance based on rainfall, evaporation and runoff without considering groundwater (e.g. Conway, 1997; Conway, 2000; Sutcliffe and Parks, 1999; Shahin, 1985).

### 3 Datasets and methods

In order to construct detailed climate change scenarios for the upper Blue Nile basin and assess their impacts on river flow, a hydrological model (NFS) is forced using rainfall and PET scenarios based on downscaled and bias corrected GCM simulations. The following sections describe the datasets and methods used to prepare the rainfall and PET scenarios for the hydrological model before it is used to translate these scenarios into flow scenarios at the sub-basin outlet (Diem station). An overview of the datasets and the procedure is given in the flowchart in Fig. 3. Table 1 provides information about the GCMs used in the study.

#### 3.1 Rainfall

Three datasets are used in connection with this analysis: observed rainfall data for present day climate, baseline GCM simulated rainfall and projected rainfall from GCM simulations. Daily GCM rainfall for the baseline and future scenario periods were downloaded from IPCC archive at the PCMDI website (http://www-pcmdi.llnl.gov/). The baseline has been selected to be 1961–1990 as this is the period for which most climate centers have submitted daily data for the 20th century climate experiment (20c3m). For the scenario data, the period 2081–2098 has been selected as to be far enough in the future to have a clear climate change signal. Daily rainfall data from 17 different GCMs using the SRES A1B emission scenario (Nakicenovic and Swart, 2000) have been obtained. The availability of daily data constrained the future period and the number of models selected for the study. The SRES A1B scenario was chosen as it represents a middle of the road estimate of future emissions with CO₂ levels stabilizing at 720 ppmv at the end of the 21st century. Ideally all scenarios could be included, however several researchers (e.g. Prudhomme et al., 2003) have pointed out that the uncertainty across GCMs is generally larger than that across emission scenarios and thus using one emission scenario will encompass a large proportion of the uncertainty. It is acknowledged, though, that the climate system response to different emission scenarios becomes larger near the end of the 21st century compared to earlier times. Therefore, the results presents are indicative of the uncertainty range but do not encompass the whole range.

The observed rainfall dataset has been obtained from the NFS database as gridded daily rainfall fields (at a resolution of 20×20 km²) for the period 1992–2006. Although this period does not overlap with the selected baseline period (1961–1990), it is the historical period available with near-observed daily data, especially satellite data. The effect
of this difference on the long-term mean water balance has been found to be small as there is little difference in the mean monthly rainfall distribution for the two periods. In terms of variability (expressed by the coefficient of variation), the 1961–1990 period rainfall (obtained from monthly records from the NFS database) was found to be more variable than that of the 1992–2006 period although the former encompasses a longer period than the latter. This puts some limits to the representativeness of the observed baseline period, especially in case of estimating changes in variability which was not the focus of this research.

These rainfall fields were created by merging satellite-based and gauge-based gridded fields using a spatially fixed set of monthly weights. These weights were obtained from rainfall analysis over the Blue Nile and thus give more weight to satellite estimates during the rainy season (May–October) over the Ethiopian Plateau. The satellite-based estimate is inferred from half hourly infrared images received from the METEOSAT satellite. These give the cloud-top temperature which is used to delineate rainfall areas using a threshold temperature of $-40^\circ C$. Assuming that very cold temperatures correspond to very high clouds, which in turn are associated with rainfall, the daily rainfall rate over a pixel is calculated as a linear function of the Cold Cloud Duration (CCD). That is the duration for which that pixel is covered by a cold cloud computed by accumulating all the half-hourly counts. Initially, existing satellite estimation methods at the time of NFS development (early 1990s) were investigated for use within the NFS including the GOES Precipitation Index (GPI) (Arkin, 1979), the READING technique (Milford and Dugdale, 1990), the PERMIT technique (Barrett et al., 1989), the Convective-Stratiform Technique (CST) (Adler and Negri, 1988), and the Progressive Refinement Technique (PRT) (Bellerby and Barrett, 1993). Currently satellite rainfall is estimated by the Nile Hybrid technique (Green-Newby, 1992, 1993) which merges strengths of the CST and PRT Techniques to form a strategy designed to combine short-term and long-term satellite rainfall estimates to produce a high resolution (20 km) daily rainfall estimate.

Gauge estimates are obtained using the Nile Inverse Distance (NID) interpolation (Cong and Schaake, 1995) technique based on WMO synoptic gauge data downloaded daily from the Florida State University website (http://www.met.fsu.edu/rawdata/syn/). The NID is a variant of the inverse distance method in which rainfall at an ungauged location is estimated as the weighted average of rainfall at surrounding gauges where weights are the inverse distance between the gauges and the ungauged location (which are selected here to fall on the same grid as the satellite estimate). The final gauge estimate is the weighted sum of the interpolated estimate and the long-term mean rainfall at the ungauged location. Those weights are calculated such that the influence of the long-term mean increases with the distance to the nearest gauge with recorded rainfall on the day. This is similar to the climatologically aided interpolation (CAI) technique of Willmott and Robeson (1995). This method overcomes the sparseness and discontinuity of the records because, on any particular day, usually data from about 40 stations are

---

**Fig. 3.** An overview of the datasets, procedures, and models used in the study.
Table 1. Information about the 17 GCMs from the IPCC 4th assessment report used in the study.

<table>
<thead>
<tr>
<th>Modeling group</th>
<th>Acronym</th>
<th>Atmospheric resolution (approx)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjerkenes Centre for Climate Research, Norway</td>
<td>BCM</td>
<td>2.8° × 2.8° (T63 L31)</td>
</tr>
<tr>
<td>Canadian Centre for Climate Modeling &amp; Analysis, Canada</td>
<td>CGCM</td>
<td>3.8° × 3.7° (T47 L31)</td>
</tr>
<tr>
<td>Canadian Centre for Climate Modeling &amp; Analysis, Canada</td>
<td>CGCM63</td>
<td>2.8° × 2.8° (T63 L45)</td>
</tr>
<tr>
<td>Meteo-France / Centre National de Recherches Mtorologique, France</td>
<td>CNRM</td>
<td>2.8° × 2.8°</td>
</tr>
<tr>
<td>CSIRO Atmospheric Research, Australiaa,b,d</td>
<td>CSIRO30</td>
<td>1.9° x 1.9° (T63 L18)</td>
</tr>
<tr>
<td>CSIRO Atmospheric Research, Australiaa</td>
<td>CSIRO35</td>
<td>1.9° × 1.9°</td>
</tr>
<tr>
<td>NOAA/Geophysical Fluid Dynamics Laboratory, USAa</td>
<td>CM20</td>
<td>2.0° × 2.5° (L24)</td>
</tr>
<tr>
<td>NOAA/Geophysical Fluid Dynamics Laboratory, USAa</td>
<td>CM21</td>
<td>2.0° × 2.5° (L24)</td>
</tr>
<tr>
<td>NASA/Goddard Institute for Space Studies, USA</td>
<td>AOM</td>
<td>3.0° × 4.0° (L12)</td>
</tr>
<tr>
<td>LASG/Institute of Atmospheric Physics, China</td>
<td>GOAL</td>
<td>2.8° × 2.8° (T42 L26)</td>
</tr>
<tr>
<td>Institute for Numerical Mathematics, Russia</td>
<td>INMCM</td>
<td>4.0° × 5.0° (L21)</td>
</tr>
<tr>
<td>Center for Climate System Research, National Institute for Environmental Studies</td>
<td>MIROCH</td>
<td>1.1° × 1.1° (T106 L56)</td>
</tr>
<tr>
<td>Frontier Research Center for Global Change, Japan</td>
<td>MIROC</td>
<td>2.8° × 2.8° (T42 L20)</td>
</tr>
<tr>
<td>Center for Climate System Research, National Institute for Environmental Studies</td>
<td>MIROCM</td>
<td>2.8° × 2.8° (T42 L20)</td>
</tr>
<tr>
<td>Max Planck Institute for Meteorology, Germanya</td>
<td>ECHAM</td>
<td>1.9° × 1.9° (T63 L31)</td>
</tr>
<tr>
<td>Meteorological Research Institute, Japan</td>
<td>MRI</td>
<td>2.8° × 2.8° (T42 L30)</td>
</tr>
<tr>
<td>National Center for Atmospheric Research, USAc</td>
<td>CCSM</td>
<td>1.4° × 1.4° (T85 L26)</td>
</tr>
<tr>
<td>National Center for Atmospheric Research, USAa,c</td>
<td>PCM</td>
<td>2.8° × 2.8° (T42 L26)</td>
</tr>
</tbody>
</table>

a These models did not provide 2 m specific humidity. Instead specific humidity at the lowest model level above orography were used.
a Test was performed using models having both variables (11 models). The effect on PET was absorbed in the calculated correction factors.
b The surface wind for 2081–2098 was missing. Instead the 1961–1990 surface wind was used. Most models showed only small wind changes between the baseline and future periods.
c The surface wind was not provided. The CRU climatological wind fields, spatially averaged to model resolution, were used instead.
d Surface pressure was missing. Sea-level pressure was converted to surface pressure using the available orography.

available. These are not necessarily the same ones received on another day. The observed daily database has 236 entries while only 142 of these lie within or close to the basin (indicated by stars on Fig. 1). In the early stages of NFS development, gauge-based fields were used to calibrate the satellite estimation equations but operationally, the two estimates are calculated separately and then a weighted average is computed as the final rainfall estimate. More details about the NFS rainfall estimation component can be found in the operations manual (Nile Forecast Center, 1999a) and Elshamy (2006). The NFS rainfall estimates are continuously evaluated by the system operators but these evaluations are not published. A comparison of gauge rainfall and NFS rainfall over a small sub-catchment of the Blue Nile showed very good agreement (Moges, personal communication, 2007). In an evaluation study of several satellite rainfall products over the Ethiopian highlands, Dinku et al. (2007) reported that a product called TAMSAT performed generally well. TAMSAT uses the same algorithms and data used by the NFS (except that it has a spatially variable threshold temperature). We therefore regard the NFS rainfall fields to be of adequate quality for the current analysis over the Blue Nile area. Comparing rainfall from the Climatic Research Unit of the University of East Anglia (CRU TS 2.1 – Mitchell and Jones, 2005) dataset to the NFS rainfall revealed that CRU generally had higher rainfall over the region.

3.2 Evapotranspiration

A potential evapotranspiration (PET) climatology was created from the CRU dataset (CRU TS 2.1 – Mitchell and Jones, 2005). This dataset comprises monthly time series of rainfall, temperature, diurnal temperature range, vapor pressure, cloud cover, wet day frequency, and frost day frequency for global land areas gridded at a resolution of 0.5°. PET climatology (reference crop evapotranspiration) for 1961–1990 has been calculated using the FAO Penman-Monteith (P-M) method (Allen et al., 1998) based on temperature, diurnal temperature range, vapour pressure, and cloudiness data. The calculations required the use of wind data and as these were not available as a time series, climatological values from the CRU CL 1.0 (New et al., 1999) were used. To our knowledge, the CRU dataset is the only observed dataset available with all the necessary variables to calculate PET using the
Penman-Monteith method. PET climatologies were also calculated from each of the GCMs for the baseline (1961–1990) and future (2081–2098) periods using the same method.

A monthly time series of PET fields could also have been used, but a test using time series for the CRU dataset and for a selected GCM showed that the mean hydrograph of the Blue Nile flow is not sensitive to interannual variability of PET. This has generally been reported for several hydrological models over different regions (e.g. Oudin et al., 2005).

One of the GCM outputs is actual evapotranspiration (AET) which could have been used, together with rainfall, to study the water balance over the basin. However, this was not done for several reasons. First, GCM rainfall is biased and thus AET will be biased since rainfall is partitioned between AET, infiltration, and runoff by the GCM land-surface scheme. While we could bias-correct the rainfall using observations, no observations of AET are available to attempt to remove the GCM bias. AET has been inferred from satellite images (National Oceanic Atmospheric Administration-Advanced Very High Resolution Radiometer (NOAA-AVHRR) using energy balance methods over parts of the Nile basin (e.g. Mohamed et al., 2004) but these estimates suffer from poor temporal resolution as the satellite overpass frequency is low. Meanwhile, we could bias correct PET calculated from GCM output. Second, GCM AET will have the coarse resolution of the GCM and will need to be downsampled. As the method used for downscaling depends on the resolution of observed data, there is no means to downscale AET. Third, in order to apply a hydrological model, PET is required, rather than AET. Otherwise, we could use GCM generated runoff directly and eliminate the need for a hydrological model. It is well known that GCM runoff suffers from several drawbacks (e.g. it cannot be compared directly to streamflow due to lack of routing within most GCMs in addition to the coarse GCM resolution) as stressed by several authors (e.g. Wood et al., 1991; Kite et al., 1994). Although routing may be less important when studying the annual water balance of relatively large basins, it is still important to obtain a reasonable reproduction of the seasonal cycle (Kuhl and Miller, 1992; Arora and Boer, 1999; Evans, 2003).

3.3 The NFS hydrological model

The NFS is used to translate rainfall and PET scenarios into streamflow scenarios. The NFS is a real-time distributed hydro-meteorological forecast system designed for forecasting Nile flows at designated key points within the Nile. Of major interest is the inflow of the Nile into the High Aswan Dam, Egypt. The system is hosted at the Nile Forecast Center (NFC) of the Ministry of Water Resources and Irrigation (MWRI), Giza, Egypt, which kindly provided a copy of the NFS software (version 5.1 – Nile Forecast Center, 2007) and the available documentation for this research.

The core of the NFS is a conceptual distributed hydrological model of the whole Nile system including soil moisture accounting, hillslope and river routing, lakes, wetlands, and man-made reservoirs within the basin. This model is defined on the quasi-rectangular grid of the METEOSAT satellite from which the system receives imagery to estimate rainfall as explained in Sect. 3.1. Each grid cell (pixel) imitates a small basin with generalized hillslopes and stream channels. Inputs to each grid cell are precipitation and potential evaporation. This input is applied to a two-layer soil moisture accounting (water balance) model of the pixel. The upper layer is thin to represent the short-term detention of storm water. This layer receives precipitation inputs and evapotranspiration occurs at the potential rate. All rainfall is assumed to infiltrate the upper layer. Excess rainfall percolates to the lower layer which has a larger moisture capacity. If the evaporative demand is not met from the upper layer, water is extracted from the lower layer at rate which varies linearly with the current moisture capacity. Subsurface runoff is calculated as a non-linear function of the relative moisture content of the lower layer (a non-linear reservoir). Surface runoff is calculated as a fraction of the excess rainfall which depends on the amount of the lower layer deficit.

Surface and subsurface runoff are subsequently input to the pixel’s hillslope routing model, simulating the transfer of water towards the main channel. Generated runoff is then routed through this channel to the downstream pixel according to a pre-defined connectivity sequence (established via GIS). Sayed and Saad (2002) evaluated the daily performance NFS in daily simulations and found that it could explain 93% of the observed daily variance of the Diem flows over the period 1997–2002. Elshamy (2008) evaluated the long-term performance of the NFS and found that it could explain 79% of the observed monthly variance of Diem flows over the period 1940–1999 with errors related to the quality of rainfall data. For more details about the NFS refer to Nile Forecast Center (1999b) and Elshamy (2006).

To calculate the impacts of the developed climate change scenarios, the simulated streamflow at Diem is compared to observations for the baseline period. Flows started to be recorded at Roseries in 1912 but the gauging station had to be moved about 136 km upstream to Diem in 1961 due to the construction of the Roseries Dam. A daily record is available due to Diem in the NFS database starting May 1965 and this was collated with the record at Roseries to build a monthly time series for Roseries/Diem (referred to as Diem from now on) for the baseline period. The mean annual total flow at Diem for 1961–1990 amounts to 46.17 BCM.

3.4 Bias-correction and spatial downscaling

GCMs can have large biases when it comes to reproducing the regional features of climate. This study uses a distribution mapping approach to bias correct the intensity of daily precipitation. Methods for mapping one distribution onto

another are well established in probabilistic modeling and it has been used to correct bias of both monthly and daily GCM precipitation data (Wood et al., 2002; Ines and Hansen, 2006). The basic procedure follows that of Ines and Hansen (2006) and is outlined below. The flowchart in Fig. 3 gives a summary of the datasets used for the bias-correction and downscaling.

The mean rainfall $\bar{x}$ (mm/day) is the product of the mean intensity on wet days (defined as days when precipitation exceeds 0.1 mm) and the relative frequency of wet days. Correcting any bias of one or both will therefore also correct the total rainfall itself. Ideally one would correct both the frequency and intensity. However, in order to preserve the climate change signal only the intensity was corrected here.

First, a two-parameter gamma distribution (see Eq. 1) is fitted separately to the observed and GCM rainfall on wet days to model the intensity distributions.

$$F_i(x; \alpha, \beta) = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_{x_i}^{\infty} e^{-\frac{x}{\beta}} \frac{1}{\beta} dx \quad (1)$$

The shape and scale parameters ($\alpha$ and $\beta$) for each gamma distribution (for each pixel/month combination) are determined using maximum likelihood estimation. Then the cumulative distribution of the daily GCM rainfall $F_{i,\text{model}}(x)$ is mapped onto the cumulative distribution of the observations $F_{i,\text{obs}}(x)$. The bias corrected modeled rainfall $x_{i,\text{model}}$ on day $i$ can then be calculated as:

$$x_{i,\text{model}} = \begin{cases} F_{i,\text{obs}}^{-1}\left(F_{i,\text{model}}(x_i)\right), & x_i > 0.1 \text{ mm} \\ 0, & x_i < 0.1 \text{ mm} \end{cases} \quad (2)$$

This whole procedure is applied separately for each month in order to correct for errors in the seasonal cycle. By correcting the full distribution, corrections on days with little rainfall will be different than for the days with extreme rainfall, thus accounting for the models poorer ability to simulate extremes. This correction procedure corrects both the mean intensities and extremes. As such it differs from a simple multiplicative shift.

For the scenario data we use a correction factor that takes into account changes in variability between the baseline and scenario simulation. A correction factor $f_i$ for each scenario value $x_{i,\text{scen}}$ is found by finding the ratio between the control and scenario values at the point where the cumulative distribution functions of the control and scenario variables coincide:

$$f_i(x_{i,\text{scen}}) = \frac{F_{i,\text{control}}^{-1}\left(F_{i,\text{control}}(x_{i,\text{scen}})\right)}{F_{i,\text{model}}^{-1}\left(F_{i,\text{model}}(x_{i,\text{scen}})\right)} \quad (3)$$

The value for $F_{i,\text{control}}(x_{i,\text{scen}})$ is found by linear interpolation between the two $F_{i,\text{control}}(x_{\text{control}})$ values closest to $x_{i,\text{scen}}$. The final corrected value for the scenario $x_{i,\text{scen}}$ is then found using Eq. (2) multiplied by the climate change correction factor:

$$x_{i,\text{scen}}'' = F_{i,\text{obs}}^{-1}\left(F_{i,\text{scen}}\left(x_{i,\text{scen}}\right)\right) \times f_i(x_{i,\text{scen}}) \quad (4)$$

So even if the absolute rainfall values of the GCM are corrected, the relative change (scenario/control) is conserved for the mean, variability and extremes. Thus, no artificial relative changes are added to the scenario after the bias correction. The extent to which this method preserves the relative rainfall change signal is assessed by comparing the percentage changes in rainfall before and after the correction as will be shown in Sect. 4.

The observed dataset used in this application is the merged satellite-gauge estimates obtained from the NFS database at a fine resolution of $20 \times 20$ km$^2$. Thus several pixels fall into one large gridcell of any GCM. By adjusting the fitted distributions for each pixel within a GCM gridcell and repeating the procedure for each GCM gridcell, a high resolution (spatially downscaled) rainfall field is obtained.

To bias-correct the PET future scenarios, monthly gridded correction factors have been calculated as ratios of the observed (i.e. CRU) climatology and GCM climatology for the baseline period. These factors were then applied to the future monthly PET climatology. As the factors were calculated using the CRU resolution, the correction implies downscaling to the CRU resolution in the same way as done for precipitation.

It could be argued that, while both GCM rainfall (P) and evapotranspiration (AET) are biased, their difference (P-AET), i.e. the convergence over a large-scale basin, would be much less biased. It is assumed that the upper Blue Nile basin is relatively large to satisfy this criterion. The basin
is characterized by little over-year storage due to the limited extent of lakes and swamps in the basin in addition to the long dry season over which the soil moisture store gets almost completely depleted. Then the mean annual P-AET value can be close to the mean annual basin runoff. Comparing P-AET to calculated runoff from the NFS for the baseline period (1961–1990) using observed data confirms this feature (both are close to 273 mm/yr). However, comparing P-E (where E is the GCM simulated latent heat flux divided by the latent heat of vaporization) versus simulated runoff using the methodology described above over the baseline period (see Sect. 4.3) for the different GCMs revealed that most models have very large biases ranging from about −82% to +465% (Fig. 4 – notice the scale). Only one model (ECHAM) produced a P-E value that is close to its simulated runoff for the baseline period (within about 6%). Most models are positively biased (i.e. produced too much rainfall and too little evaporation). These results justify the outlined methodology for bias-correcting GCM results and downscaling them to be used by the NFS hydrological model. In addition, this approach is required to obtain the monthly distribution of runoff, which cannot be derived from monthly P-E values directly even if their GCM values were not biased.

4 Results and analysis

In the following sections, results based on the output of 17 GCMs for the SRES A1B emissions scenario are used to investigate the impacts of climate change on the upper Blue Nile flow at Diem. Rainfall, PET, and AET are summarized as mean areal averages over the basin area. First, the sensitivity of the basin flow to changes in rainfall and PET is analyzed. An assessment of the ability of the bias correction method to preserve the climate signal is then presented. The quality of baseline simulations using the bias corrected rainfall and PET are then assessed before future changes in the water balance are discussed.

4.1 Sensitivity analysis

The long-term water balance of the basin can be illustrated by plotting the actual evapotranspiration ratio to rainfall (AET/P) against the aridity index (PET/P). This is often referred to as the Budyko framework where the water balance is mainly determined by rainfall and available energy for evaporation (Budyko, 1974). The baseline annual water balance of the upper Blue Nile is shown in Fig. 5. PET can be seen to slightly exceed AET and the basin therefore belongs marginally to a moisture constrained regime. During the wet season (JJAS), on the contrary, rainfall greatly exceeds PET and evapotranspiration is energy constrained.

To test the sensitivity of the water balance to changes in PET and rainfall, the NFS was driven with PET and rainfall climatology (1961–1990) and uniform changes between ±30% for each variable stepping 5% at a time while the other is kept constant. The sensitivity curves are plotted in Fig. 5. Increasing PET leads to a smaller runoff coefficient and moves the basin towards a more moisture constrained regime. The PET sensitivity curve has a smaller runoff coefficient compared to the canonical Budyko curve. It is well known that due to strong rainfall seasonality, vegetation and other catchment characteristics the water balance may deviate somewhat from the general Budyko curve. Instead, the following relationship for estimating mean annual evapotranspiration (AET) from Zhang et al. (2004) was used:

$$\frac{AET}{P} = 1 + \frac{PET}{P} - \left[ 1 + \frac{PET^w}{P} \right]^{\frac{1}{w}}$$

where $w$ is an adjustable model parameter representing different catchment characteristics. A $w$ value of 2.85 was found to fit the PET sensitivity curve well as shown in Fig. 5. Note the steeper slope of the curve and therefore greater sensitivity of runoff to PET when the aridity index is small, as happens during the wet season when evaporation is energy constrained.

The Blue Nile discharge has been shown by several authors to be very sensitive to changes in precipitation (e.g. Conway and Hulme, 1993; Sayed, 2004). Both studies indicated that a 10% change in precipitation could yield more than 30% change in flow. Approximately the same sensitivity is found here. The Blue Nile discharge is also sensitive to PET changes but to a smaller extent (10% change in PET results in 14% change in flow) because the basin is moisture constrained for most of the year. Thus, the sensitivity is greatest during the wet season since evapotranspiration is already close to the limit represented by PET. The precipitation curve (Fig. 5) has a steeper slope than the PET curve and does not follow the Budyko or Zhang curve (several other values

![Fig. 5. The Budyko framework. Ratio of the mean annual evapotranspiration to rainfall (AET/P) versus the aridity index (PET/P). Straight lines show the energy and water limits to evaporation.](image-url)
of rainfall, the discharge becomes more sensitive to reductions in PET than to increases.

During the sensitivity exercise it was found that CRU-based PET had to be increased by 10% to obtain a reasonable agreement between simulated and observed flows at Diem (i.e. flows were overestimated using CRU-based PET). As seen in Fig. 5, PET acts like an upper limit to AET during the wet season and thus calibrating NFS parameters would not lead to flow reductions. Assessing the impacts of climate change based on such overestimated flows would have led to less confidence in the results. The sensitivity of the climate signal to the quality of flow simulations was tested and found to be relatively small for such PET change (see Sect. 4.3 below).

Repeating the rainfall sensitivity exercise for this new baseline PET increases the sensitivity as the changes are calculated relative to a lower flow value but the general observations given above remain unchanged. Results given below are thus based on the 10% inflated CRU-based PET being used as the observed PET climatology for correcting GCM bias. This factor is carried over to PET scenarios as they are bias-corrected against the 10% inflated CRU-based PET. This factor might be accounting for the different vegetation type (as compared to grass which is the reference crop in the FAO Penman-Monteith method) or compensating for the use of standard coefficients in the Angstrom formula which relates solar radiation to extraterrestrial radiation and relative sunshine duration (refer to Allen et al., 1998, for more information).

4.2 Bias corrected rainfall

Figure 6 shows the rainfall climatology of the 17 GCMs for the upper Blue Nile basin for the baseline period (1961–1990) before and after bias correction. It is evident that many of the models either overestimated or underestimated the rainfall while some were even unable to reproduce the seasonal cycle. The correction scheme brings the distributions close to the observed pattern (denoted NFS 92-06) except for CCSM and INMCM, which originally had much too low and somewhat bimodal rainfall patterns. Notably, the scheme also preserves the climate change signal (% change in mean annual rainfall total) for most models/months as shown in Fig. 7. A few large deviations occur in the dry season when the gamma distribution assumption becomes less appropriate and the number of points used to fit it becomes small for some models/months. These deviations have little effect on flow as they occur outside the wet season. Note that biases in wet day frequencies are not corrected. This gives a poorer correction for the baseline period, but more accurate conservation of the model’s original climate signal. For PET the ratio-based bias correction is more exact and the points are concentrated around the 1:1 line (not shown).

4.3 Baseline simulations

Using the bias corrected rainfall from the 17 GCMs and the 10% inflated CRU-based PET, the NFS was used to simulate the flow for the baseline period 1961–1990. Most simulations were satisfactory with $R^2$ (Nash-Sutcliffe Efficiency) for the monthly time series ranging between 0.59 and 0.85 except for GOAL, which had a relatively low $R^2$ of 0.49. RMSE ranged between 1.88 and 3.45 BCM, which amounts to about 4.1–7.5% of the mean observed annual total. In terms of the simulation of the mean annual hydrograph most models had small volume biases of ±10% of the mean annual total. The exceptions were CCSM with a high negative bias of 29% and CNRM with a positive bias of 13%. The rainfall bias correction was less successful for the CCSM and the INMCM models, as mentioned above, and this remaining bias was amplified in the flow simulations with large negative bias for the CCSM. In case of simulated AET, the ensemble members are close to each other especially for the wet season when AET from all models is equal to PET (not shown). The main deviation from the ensemble is observed for the CCSM model again due to its severe underestimation of rainfall both before and after correction. The results from these models should be viewed in light of these problems.

A test was done to see if the climate signal translated into flows is sensitive to the quality of the baseline simulation. The baseline simulations were repeated using correction factors based on the original CRU-based PET (i.e. without the 10% inflation). This resulted in overestimated flows for the...
baseline period from most models. The same PET correction factors were used for the future period. Repeating the flow simulations, it was found that the direction of flow change (i.e. increase or reduction) was preserved and the magnitude (difference between future and baseline) remained similar. To avoid the biases in flow simulations contaminating the climate signal, future changes for each model were calculated with respect to the baseline simulation of that model for this test and for the results reported below.

4.4 Rainfall and PET changes

All models predict an increase in annual surface temperature for the 2081–2098 period compared to the baseline period with values ranging between 2°C and 5°C. In turn, all GCMs report increases in total annual PET varying from +2 to +14%. The PET increase is not uniform throughout the year and the importance of PET changes is dominated by the projected changes during the wet season.

Conway and Hulme (1996) assumed a 4% increase in PET per degree change in temperature based on the results of other earlier studies. Figure 8 indicates that this assumption was a good one as the least-squared fit through the model ensemble has a similar slope (3.75% per °C). There are some outliers, e.g. GOAL predicts too little increase in PET for a very similar temperature change for which CSIRO30 predicts a much higher increase. This may be due to the role of other climatic variables such as wind speed and humidity and how they change in the future. For the wet season the linear assumption seems to remain valid but at a steeper slope (5.6% increase in PET/°C). The steeper slope during the wet season indicates a larger role of the radiation component of PET as the effect of cloudiness on PET is more dominant during the wet season. Models deviate from the least-squared fit due to the effect of other climatic variables on PET and the different parameterizations of surface processes within GCMs. Using a temperature-based method would probably give the general trend of the results in terms of PET change but would not provide the same level of detail when models are analyzed individually.

GCMs disagree on the direction of precipitation change for the 2081–2098 period compared to the baseline (Fig. 9a). Recall that, after correction, the baseline precipitation is very similar for almost all GCMs (Fig. 6b). Changes in total annual precipitation range between −15% to +14% but more models report reductions (10) than those reporting increases (7). Several models (6) report small changes within 5%. The largest reductions are reported by the CGCM and CGCM63 experiments while the largest increases are reported by MIROC and INMCM. The ensemble mean of all models shows almost no change in the annual total rainfall and only a slight reduction (−2.4%) in the wet season total. Without the CCSM and INMCM models, which were not well bias corrected, the ensemble mean change becomes −1.5% annually and −3.5% for the wet season.

4.5 Water balance and flow changes

In terms of changes to runoff (Fig. 9b), more models are predicting flow reductions (11) than those showing increases (6). This is mainly because all models show increases in PET and AET during the wet season. The increase in AET can exceed small increases in rainfall leading to overall flow reductions. The ensemble mean annual flow at Diem is reduced by 15% compared to the baseline. However, the very high sensitivity of flow to rainfall makes such a number highly uncertain. There is some evidence that higher reductions in rainfall may be accompanied by higher increases in PET possibly due to reductions in cloud cover. This further enhances the effect of rainfall reduction on runoff leading to very high reductions in the mean annual flow for some models (up to 60%). Some models predicted large increases as well (up to 45%).

The simulated change to the annual water balance is shown in Fig. 10. The ensemble mean for the 2081–2098 period is shifted towards a higher aridity index compared to the baseline climatology. In this manner, the basin is moving towards a more moisture constrained regime. The ensemble mean runoff coefficient is also reduced by about 3.5%. This is mainly due to evapotranspiration being energy constrained during the wet season. For no change in rainfall, higher PET results in increased AET and a reduced wet season runoff coefficient.

For most models the future water balance is well predicted by the Zhang et al. (2004) curve (ω=2.85) discussed in Sect. 4.1. The exceptions are the four models with the largest increase or reduction in rainfall. This is perhaps to be

Fig. 7. Percentage change in rainfall between 1961–1990 and 2081–2098 for all months and models plotted against the corresponding bias-corrected values.
Fig. 8. Annual (a) and wet season (b) changes in PET versus changes in temperature for the Upper Blue Nile Basin. The straight line indicates the least-square fit through GCM results.

Fig. 9. Mean monthly precipitation and runoff distributions for the upper Blue Nile Basin for 2081–2098 as simulated by the 17 bias-corrected GCMs. The rainfall baseline is the observed 1992–2006 mean and baseline PET is the 10% inflated CRU based 1961–1990 climatology.

Fig. 10. Ratio of the mean annual evapotranspiration to rainfall (AET/P) versus the aridity index (PET/P) for GCM simulations.

As an alternative, linear relationships can also be utilized such as that given in Fig. 8 to estimate PET based on temperature change. A similar relationship was developed relating changes in flow to changes in rainfall (Fig. 11). Figure 11 attempts to separate the effects of future changes in PET and rainfall on flow. There seems to be a linear relationship between rainfall changes as depicted by the different GCMs and flow changes on annual basis and for the wet season (with a one-month lag for flow). The slope of the relationship is steep (3.01) for the annual total and even steeper (3.25) for the wet season which means that changes in rainfall are amplified by a factor of three in flows. The fitted lines are nearly parallel to the sensitivity results where PET was held constant, which indicates that changing PET only changes the slope of the line slightly but it mainly shifts it down (i.e. towards less flow) by 14% for the annual total and by 10% for the wet season. These results can be used to obtain rough estimates for the flow response of the upper Blue Nile basin to future changes in rainfall and PET in conjunction with the Zhang curve mentioned above.
5 Conclusions

This study analyzed the output of 17 general circulation models (GCMs) included in the 4th IPCC assessment report. Downscaled precipitation and potential (reference crop) evapotranspiration (PET) were constructed for the upper Blue Nile basin for a baseline climatology (1961–1990) and the SRES A1B scenario for the 2081–2098 period. These were used to drive a fine-scale hydrological model of the Nile Basin to assess their impacts on the flows of the upper Blue Nile at Diem, which accounts for about 60% of the annual flow of the Nile at Dongola.

The 17 downscaled GCMs have large rainfall biases for the present climate (1961–1990) with two models even failing to simulate the seasonal cycle. The probabilistic bias correction method was capable of correcting most errors reproducing the main features of observed rainfall. The scheme removed the bias nearly completely for 15 out of 17 GCMs. Notably, the bias correction method also preserved the models’ original climate change signal on a monthly time scale.

There is no consensus among the GCMs on the direction of precipitation change for the upper Blue Nile basin. Changes in total annual precipitation range between −15% to +14% but more models report reductions (10) than those reporting increases (7). Several models (6) report small changes within 5%. The largest reductions are reported by the CGCM and CGCM63 experiments while MIROCM and INMCM show the largest increases. The ensemble mean of all models shows almost no change in the annual total rainfall and only a slight reduction (−2.4%) in the wet season total. The numbers change only marginally if the two models with the greatest biases are excluded.

The great inter-model divergence in rainfall is consistent with the summary given by Conway (2000) for rainfall over the Blue Nile region. Conway (2000) showed that there is not yet any evidence, neither in observations nor predicted future rainfall, of any clear anthropogenic signal. This may be because such a signal does not exist or is simply too small compared to the very large natural variability over the upper Blue Nile basin. This highlights the need for using large ensembles, as in this study, for assessing future changes to rainfall over the Blue Nile. Given the large rainfall biases of the current generation of GCMs, it could of course be that they simply fail to capture any weak anthropogenic signal should it exist. Another limitation of this study is the use of relatively short reference and scenario periods (30 and 20 years, respectively). This makes the across model spread a function of both real differences in the models’ response to greenhouse gas changes and differences due to insufficient sampling of the models’ decadal climate variability. The influence of using relatively short reference and scenario periods is discussed in details in Sorteberg and Kvamsto (2006).

Changes to the water balance of the upper Blue Nile were assessed using the Budyko framework. The basin is shown to belong marginally to a moisture constrained regime. However, during the wet season the basin is largely energy constrained. For no change in rainfall, increasing PET thus leads to a reduced wet season runoff coefficient. The sensitivity of the water balance to changes in PET was assessed using the NFS hydrological model. The PET sensitivity curve had a smaller runoff coefficient compared to the canonical Budyko curve. It is well known that due to strong rainfall seasonality, vegetation and other catchment characteristics the water balance may deviate from the general Budyko curve. Instead, an empirical curve was estimated from the extended Budyko framework in Zhang et al. (2004). It was found that a “Zhang” curve with a w value of 2.85 fitted the PET sensitivity curve well where w is an adjustable model parameter representing different catchment characteristics.

All models predict the temperature to increase between 2°C and 5°C and consequently PET to increase by 2–14%. In terms of the water balance the ensemble mean for the
The 2081–2098 period is shifted towards a higher aridity index (PET/P) compared to baseline climatology. In this scenario the basin is therefore moving towards a more moisture constrained regime. For moderate or no changes in rainfall it is shown that this shift for most models is well predicted by the Zhang et al. (2004) curve ($w=2.85$). For the energy constrained wet season, higher PET results in increased AET and a reduced wet season runoff coefficient. The ensemble mean annual runoff coefficient is therefore reduced by about 3.5%.

Predictions of the changes in the flow are more uncertain. The flow of the upper Blue Nile basin is shown to be highly sensitive to both precipitation and evaporation changes. Changes in rainfall and PET are amplified by a factor of 3 and 1.4 respectively when translated to changes in flow. The predicted ensemble mean annual flow at Diem is reduced by 15% compared to baseline. However, due to the very high sensitivity of flow to rainfall this number should be interpreted with care. As already stated, the direction of future rainfall change in simulations is not robust. A few models with large changes in rainfall may have a large effect on the ensemble mean flow. We therefore instead put emphasis on the predicted change in the runoff coefficient. Assuming no change or moderate changes in rainfall, the simulations presented here indicate that the upper Blue Nile basin may become more moisture constrained in the future with a slightly reduced runoff coefficient.

In addition, simple relationships between changes in temperature, PET, precipitation, and flow have been developed. They enable simple estimation of climate change impacts on the study basin and facilitate comparison with other modeling studies. This was achieved through the relatively large ensemble of GCMs used. PET changes were found to have a strong linear relationship with temperature changes. PET increases by 3.75% (annually) and 5.6% (wet season) per °C warming. The increase in PET shifts the rainfall-flow change relationship down (by 14% and 10% for the annual and wet season totals respectively) rather than changing its shape. Discharge changes can thus be estimated roughly based on temperature and precipitation changes using these relationships or the Zhang curve.

It is important to be aware of the many limitations to this study. Only one scenario from the 4th IPCC assessment report was used. The SRES A1B scenario was chosen as it represents a middle of the road estimate of future emissions with CO2 levels stabilizing at 720 ppm at the end of the 21st century. A greater range of temperature and PET changes would be obtained by also including other scenarios. Another limitation is the assumption of stationarity in the bias-correction procedure. It was assumed that the present bias would be the same as the future bias, which may of course not be the case. The bias correction was needed to obtain realistic flows given the large rainfall biases of the current generation of GCMs. Most GCMs have a very crude description of land surface climate feedbacks in addition to their coarse resolution. It would therefore be very interesting to see how the results presented here would compare with simulations of regional climate models with a better representation of the water cycle over the basin.

Acknowledgements. This research was done under the Nile Basin Research Programme (NBRP), hosted at the University of Bergen, Norway during Autumn 2007 and Spring 2008. We thank Semu Moges for valuable comments to a draft of this manuscript. We acknowledge the modeling groups of the Program for Climate Model Diagnosis and Intercomparison (PCMDI) and the WCRP’s Working Group on Coupled Modeling (WGCM) for their roles in making available the WCRP CMIP3 multi-model dataset. Support of this dataset is provided by the Office of Science US Department of Energy. The Nile Forecast Center, Ministry of Water Resources and Irrigation of Egypt, is acknowledged for providing the Nile Forecasting System software and documentation. The manuscript was greatly improved based on the comments of two anonymous reviewers.

References


Wood, A. W., Maurer, E. P., Kumar, A., and Lettenmaier, D. P.: Long-range experimental hydrologic forecasting for the eastern

