

The hydrological response of the Ourthe catchment to climate change as modelled by the HBV model

T. L. A. Driessen^{1,*}, R. T. W. L. Hurkmans^{1,**}, W. Terink¹, P. Hazenberg¹, P. J. J. F. Torfs¹, and R. Uijlenhoet¹

¹Hydrology and Quantitative Water Management Group, Wageningen University, Wageningen, The Netherlands

* now at: Royal Haskoning, Nijmegen, The Netherlands

** now at: Bristol Glaciology Centre, School of Geographical Sciences, University of Bristol, Bristol, UK

Received: 10 November 2009 – Published in Hydrol. Earth Syst. Sci. Discuss.: 19 November 2009

Revised: 8 March 2010 – Accepted: 29 March 2010 – Published: 12 April 2010

Abstract. The Meuse is an important river in Western Europe, which is almost exclusively rain-fed. Projected changes in precipitation characteristics due to climate change, therefore, are expected to have a considerable effect on the hydrological regime of the river Meuse. We focus on an important tributary of the Meuse, the Ourthe, measuring about 1600 km². The well-known hydrological model HBV is forced with three high-resolution (0.088°) regional climate scenarios, each based on one of three different IPCC CO₂ emission scenarios: A1B, A2 and B1. To represent the current climate, a reference model run at the same resolution is used. Prior to running the hydrological model, the biases in the climate model output are investigated and corrected for. Different approaches to correct the distributed climate model output using single-site observations are compared. Correcting the spatially averaged temperature and precipitation is found to give the best results, but still large differences exist between observations and simulations. The bias corrected data are then used to force HBV. Results indicate a small increase in overall discharge, especially for the B1 scenario during the beginning of the 21st century. Towards the end of the century, all scenarios show a decrease in summer discharge, partially because of the diminished buffering effect by the snow pack, and an increased discharge in winter. It should be stressed, however, that we used results from only one GCM (the only one available at such a high resolution). It would be interesting to repeat the analysis with multiple models.

1 Introduction

An important river in Northwestern Europe is the river Meuse. Although its catchment covers only about 33 000 km², six million people in The Netherlands and Belgium depend on it for their water supply. Besides, it is important for navigation, and its catchment is densely inhabited (de Wit et al., 2007). The Meuse is a typical rain-fed river. In dry spells, therefore, discharge almost exclusively originates from groundwater aquifers that are recharged during winter (de Wit et al., 2001). If the summer discharge becomes too low, this has consequences for both water quantity and quality, as well as for example for water supply, navigation, and agriculture. Extremely high discharges, on the other hand, may cause large damage as well: the near floods of 1993 and 1995 for example, caused several hundreds of thousands of people in The Netherlands to be evacuated (Chbab, 1995).

It is widely recognized that the increasing trend in temperature over the past decades is likely to continue during the coming century (IPCC, 2007). With this warming, precipitation characteristics are also expected to change (Trenberth et al., 2003): more precipitation is expected to fall in the form of extreme events. For an adequate management of the water resources in rainfed river basins, such as the Meuse, therefore, it is important to have an idea of how precipitation characteristics will change and how the basin will respond to that. Global Climate Models (GCMs) are widely used tools to create projections of future climate (IPCC, 2007). Because their spatial resolution is too low for hydrological applications, their output should be downscaled to a higher spatial resolution. One way to do this is by nesting a Regional Climate Model (RCM) in the GCM over the domain of interest (Lorenz and Jacob, 2005). In this study, we use data from



Correspondence to:
R. T. W. L. Hurkmans
(ruud.hurkmans@bristol.ac.uk)

the well-known GCM ECHAM5/MPIOM (e.g., Arpe et al., 2005), downscaled with the RCM REMO (Jacob, 2001), provided by the Max Planck Institut für Meteorologie in Hamburg, Germany. The final resolution of the data that is used in this study is very high (0.088 degrees, or about 10 km) compared to other similar studies (e.g., Shabalova et al., 2003; Lenderink et al., 2007; van Pelt et al., 2009). Because of the typical low spatial resolution, even after downscaling, climate change impact assessments are usually carried out over large river basins, such as the Rhine (e.g., Kwadijk and Rotmans, 1995; Hurkmans et al., 2010; Shabalova et al., 2003), or the entire Meuse (e.g., de Wit et al., 2007; van Pelt et al., 2009; Booij, 2005).

The high resolution of the scenarios employed here allows to zoom into a smaller catchment than the previously mentioned studies. In this case, we focus on an important tributary of the Meuse, the Ourthe, covering about 1600 km². The Ourthe catchment is situated on the edge of the Ardennes and is, due to its hydraulic gradient, limited groundwater storage and steep sandstone slopes, a fast responding river that contributes significantly to the discharge peak when a precipitation event forces the Meuse river basin. It has been the subject of hydrological modelling studies in various previous investigations (e.g., Leander and Buishand, 2007; Berne et al., 2005; Hazenberg et al., 2010), mostly involving the hydrological model HBV (Hydrologiska Byråns Vattenbalansavdelning; Bergström and Forsman, 1973). In this study, the HBV version HBV Light 2.0 (Seibert, 2005) is employed, which is explained in more detail in Sect. 2.2.

In order to obtain an idea of the hydrological changes in the Ourthe catchment as a result of climate change, we will use the high-resolution climate model output data to force the HBV model. An important step before running the hydrological model is the correction of any structural errors that are usually present in the climate model output (Lenderink et al., 2007), for example as a result of the coarse resolution of GCMs and the downscaling process. After a description of the study area (Sect. 2.1) and hydrological model (Sect. 2.2), more details about this bias and the correction process are provided in Sect. 3.1. In Sects. 3.2 and 3.3, respectively, the model calibration and the calculation of potential evaporation are described, and in Sect. 4, the climate change effects on the hydrology of the Ourthe will be discussed in terms of average fluxes and storages, as well as extreme peak flows and stream flow droughts. In Sect. 5, finally, the conclusions will be presented.

2 Study area, model and data

2.1 Study area

The Ourthe catchment is situated in the southeastern part of Belgium and is partly adjacent to North-West Luxembourg (Fig. 1). Near the city of Nisramont, the Ourthe Occidentale (western branch) and the Ourthe Orientale (eastern branch)

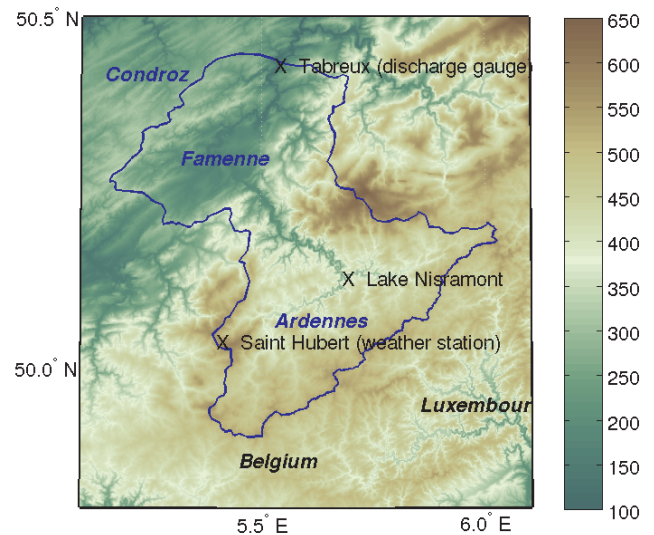


Fig. 1. Digital elevation model of the Ourthe catchment. Also shown are the discharge measuring gauge at Tabreux, the weather station in St. Hubert and Lake Nisramont, indicated by black crosses. The color scale indicates elevation above sea level in meters.

join and form the river Ourthe. The catchment area south of this location is located in the Ardennes mountain range, which mainly consists of sandstone. From this point the river flows in a northwesterly direction through a Middle-Devonian limestone area into the flat Famenne region that is characterized by shale. North-west of the Famenne lies the Condroz region, which has Late-Devonian sandstone anticlines and Early-Carboniferous limestone synclines on top of the earlier mentioned shale. Since the higher Condroz region acts as a natural boundary, the Ourthe flows in a northerly direction, where several smaller tributaries, such as the Vesdre and the Amblève, join the Ourthe river along its way towards Liège, where it eventually joins the river Meuse. This study only addresses the catchment area upstream of Tabreux (Fig. 1). The Vesdre and Amblève sub-catchments are, therefore, not taken into account. There are two main reasons to look at these systems separately: (1) the Amblève and Vesdre are joining the Ourthe just before its outlet in the river Meuse and therefore do not affect much of the Ourthe catchment and (2) there is uncertainty in the discharge measurements just before the confluence point of the Ourthe with the Meuse, because water levels depend on the setting of the weir at Angleur (Velner, 2000).

The Ourthe is a rain-fed river that is situated in a particularly hilly region and therefore has a fast discharge component in its hydrological system. This is especially the case in the upstream part of the catchment, where limited groundwater storage and steep sandstone slopes are present. Altitudes vary roughly between 100 and 650 m a.s.l. With a length of 175 km measured from the source of the Ourthe Occidentale it has an average hydraulic gradient of about 3 m km⁻¹.

The Ourthe catchment upstream of Tabreux, as defined in this study, has a surface area of about 1597 km².

In the period between 1969 and 1998 the average yearly precipitation sum was 970 mm, with a minimum of 680 mm and a maximum of 1230 mm measured at the weather station of St. Hubert. The average yearly evapotranspiration at St. Hubert was 590 mm and the average yearly discharge, as a result, 380 mm.

2.2 HBV model

The hydrological model used in this study to simulate the hydrological response of the Ourthe catchment is the Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström and Forsman, 1973). In this study, HBV Light Version 2.0 (Seibert, 2005) is used. The only two differences between HBV Light and the other versions are in the model initialization, which should be done using a warming-up period in HBV Light, and a routing parameter that can take all real values instead of just integer values (Seibert, 2005). The HBV model is a simple conceptual rainfall-runoff model, which is suitable for different purposes, such as simulation of long streamflow records, streamflow forecasting and hydrological process research. It has been applied in many different catchments, including the Rhine (te Linde et al., 2008; Hundecha and Bárdossy, 2004) and the Meuse (Leander and Buishand, 2007; Booij, 2005). As can be seen schematically in Fig. 2, the HBV model describes the water balance using three storage reservoirs: a soil moisture zone, an upper zone storage (for sub-surface stormflow) and a lower zone storage. Including an algorithm for snow accumulation and melt (based on the degree-day method) and an algorithm accounting for lakes the general water balance equation becomes the following:

$$P - E - Q = \frac{d}{dt} [SP + SM + UZ + LZ + \text{lakes}] \quad (1)$$

where P , E and Q refer to precipitation, evaporation and discharge, respectively, all with dimensions [L T⁻¹]. SP and SM stand for snow pack and soil moisture and UZ and LZ are related to the upper and lower groundwater zone, all with dimensions [L]. The lakes-term [L] refers to the storage in lakes. A subroutine for meteorological interpolation is available to represent the spatial distribution of temperature and precipitation. The model requires precipitation, temperature and potential evapotranspiration as input (Seibert, 2005). For more details about the HBV model, see Bergström and Forsman (1973) and Seibert (2005).

2.3 Datasets

As atmospheric forcing for the HBV model, daily observations of the weather station at St. Hubert (Fig. 1) are available, although all are spanning different periods: precipitation is available for 1968–2005 and temperature and potential

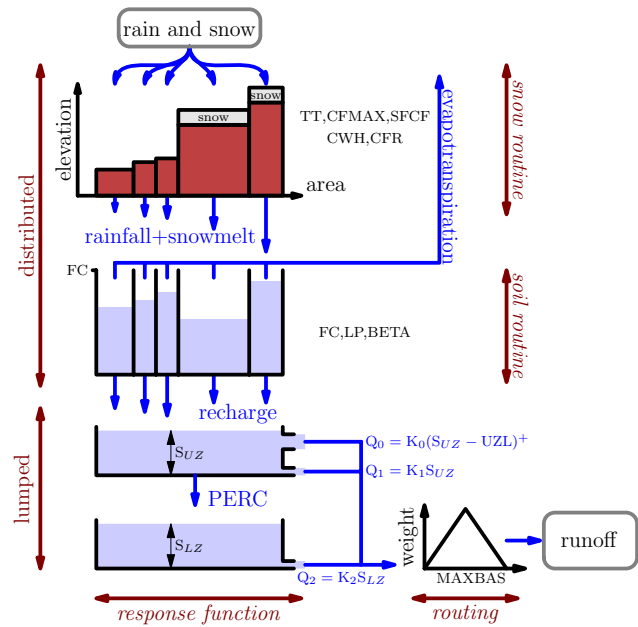


Fig. 2. A simplified schematisation of the HBV Light model (after Seibert (2000)).

evapotranspiration for 1968–1999. In addition, daily observed discharge at Tabreux is available for the period 1968–2005. Beside the observations, climate model output is available from the Max Planck Institut für Meteorologie in Hamburg, Germany. The Regional Climate Model (RCM) REMO (Jacob, 2001) is used to downscale data from either a Global Climate Model (GCM) or a re-analysis dataset, ERA15. For the ERA15 reanalysis dataset, as many global observations as possible have been collected for the period 1979–1993. In areas where the density of observation was sparse, satellite-based observations have been used. A data assimilation scheme and a numerical weather prediction model propagated information about the state of the global atmosphere. This model output together with the observations and forcing fields were used as input for the reanalysis (Gibson et al., 1999). The ERA15 dataset is therefore a mixture of in-situ and satellite observations and modelled data. For this research, a version of the ERA15 dataset extended with operational analyses has been used. This dataset, spanning the period 1979–2003, downscaled using REMO to a horizontal resolution of 0.088° (Jacob et al., 2008), is referred to as ERA hereafter.

Regional climate scenarios are used to force the HBV model to extract the effect of climate change. Three scenarios are available from the GCM ECHAM5/MPIOM, each spanning the period 2001–2100. The three climate scenarios are based on the A1B, A2 and B1 carbon-emission scenarios as defined by the IPCC (IPCC, 2000). The ECHAM5/MPIOM data, with a spatial resolution of about 400 km, is downscaled in two steps by REMO (first to 0.44° and then to 0.088°; Jacob et al., 2008), similar to the ERA

Table 1. Overview of average annual sums of precipitation (P) and potential evaporation (E_p ; derived using the method explained in Sect. 3.3), and average annual mean temperature (T), for the reference period and two 39-year periods in the 21st century according to three IPCC-scenarios.

Variable	1962–2000	2002–2040			2062–2100		
		A2	A1B	B1	A2	A1B	B1
P [mm y^{-1}]	1094	1103	1129	1199	1163	1107	1113
T [$^{\circ}\text{C}$]	7.19	7.69	7.83	7.73	10.29	10.53	9.31
E_p [mm y^{-1}]	674	665	673	666	690	712	680

dataset. In addition, a reference dataset is available spanning the period 1951–2000, which also consists of downscaled ECHAM5/MPIOM data. Hereafter, the scenarios and reference data from ECHAM5/MPIOM, downscaled with the REMO model, will be referred to as the ECHAM5 reference and scenario datasets. Mean annual values of precipitation and temperature for all scenarios and the reference situation are shown in Table 1. In addition, potential evaporation values are shown that are derived using the method explained in Sect. 3.3. Although the ECHAM5 reference run resembles the current climate in a statistical sense, it cannot be used for model calibration because it does not represent the actual time series. The ERA dataset, on the other hand, contains observations and can, therefore, be used for model calibration. In this study, we use only a spatially lumped hydrological model. The observations, therefore, suffice to calibrate the HBV model. The ERA dataset will be used to assess different methods of bias correction. The best method will then be applied to the reference and scenario datasets, which will in turn be used to force the HBV model in order to extract the climate change signal.

3 Methodology

3.1 Bias correction

Because of the very low spatial resolution of a GCM, precipitation cannot be modelled explicitly but needs to be parameterized. This is one of the sources of structural model errors in precipitation as modelled by GCMs (Lenderink et al., 2007). An extra error is added by the RCM in the downscaling process (Leander and Buishand, 2007). Before any hydrological application, the model biases need to be corrected for. However, there is only a long observational time series available for one weather station in the catchment, Saint Hubert (Fig. 1), which is situated at a relatively high altitude and close to the boundary of the catchment. The first step is to compare different ways of correcting for the bias to assess which one results in the best representation of the observations. To this end, an assessment of different bias correction methods is performed using the ERA dataset in order to find the best available method that can be applied to the ECHAM5 datasets. All atmospheric datasets contain three-hourly time

series for temperature, precipitation, atmospheric pressure, vapour pressure, shortwave and longwave incoming radiation and wind speed. However, corrections will only be applied to temperature and precipitation, since there are only observations available for these two parameters.

3.1.1 Bias correction methods

Several approaches to correct for model bias have been proposed (e.g., Leander and Buishand, 2007; Shabalova et al., 2003; Hay et al., 2002). We select the method of Leander and Buishand (2007), which is similar to that of Shabalova et al. (2003), because it has been applied to the Meuse, and also to the Rhine using the same data as the present study (Hurkmans et al., 2010). Leander and Buishand (2007) propose a power transformation, which corrects non-linearly for the coefficient of variation (CV) as well as the mean of the precipitation:

$$P^* = a \times P^b, \quad (2)$$

where P and P^* are the uncorrected and the corrected precipitation, respectively, and a and b are parameters that define the correction. The parameters are iteratively estimated for each five-day period in a window including the 30 days before and after the five-day block over all years of the dataset, resulting in a 65-day window for each block, following Leander and Buishand (2007). The value of b is determined per block by matching the coefficient of variation (CV) of the corrected daily precipitation with the observed daily precipitation. The value of a is then determined such that the mean of the transformed daily precipitation values matches the mean of the observed precipitation values. Thus, parameter b depends only on the CV and its determination is independent of parameter a .

The bias correction of the temperature uses the same 65-day windows as precipitation and involves shifting and scaling to adjust the mean and variance, respectively, according to:

$$T^* = \bar{T}_{\text{obs}} + \frac{\sigma(T_{\text{obs}})}{\sigma(T_{\text{cal}})} (T_{\text{uncor}} - \bar{T}_{\text{obs}}) + (\bar{T}_{\text{obs}} - \bar{T}_{\text{cal}}), \quad (3)$$

where T_{uncor} and T^* represent the uncorrected daily temperature and the corrected daily temperature, respectively. \bar{T}_{obs}

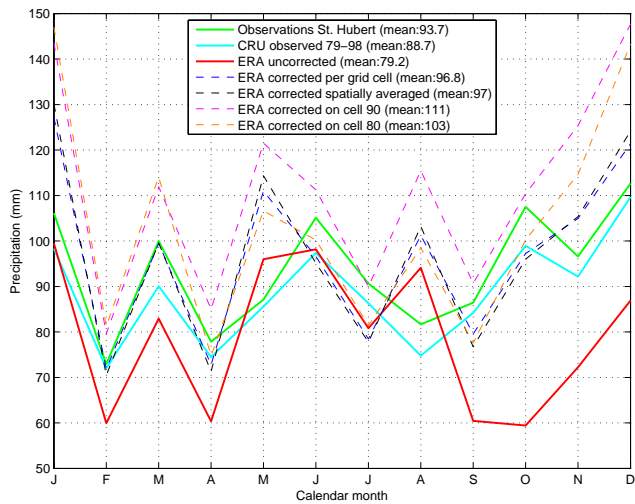


Fig. 3. Mean monthly precipitation sums of the observations and (un)corrected datasets for the period 1979–1996.

and $\sigma(T_{\text{obs}})$ are the mean and standard deviation of the observed daily temperature for the considered five-day period. Similarly, \bar{T}_{cal} and $\sigma(T_{\text{cal}})$ are the mean and standard deviation of the uncorrected temperature for the considered five-day period. The same bias correction methodology was also applied to the Rhine basin by Hurkmans et al. (2010), and was investigated in detail by Terink et al. (2010). They show that after the correction not only the estimation of average precipitation was improved, but also for example that of extreme values, 10-day sums and the first order autocorrelation. In addition, Terink et al. (2010) evaluated the bias-correction method using split-sample tests, where the correction parameters were determined using one part of the dataset and validated using another. They found a reduction of the precipitation bias in both the calibration and validation part. The temperature bias was almost completely removed in these studies.

Because the spatial resolution of the climate model data is 0.088° (about 10 km), and the observations are only available for one location, different approaches are possible to perform the bias correction. Parameters a and b can, for example, be calculated based on the grid cell corresponding to Saint Hubert, or based on the spatial average (assuming that the observation is representative for the entire catchment). Four approaches are compared in this study. The first method calculates different bias correction parameters per grid cell, comparing every individual uncorrected cell data with the Saint Hubert observation. The second method averages the uncorrected data of all the grid cells into one spatially averaged time series and compares it with the Saint Hubert observation. The set of calculated parameters is then applied to all the grid cells. The third method calculates one set of correction parameters by comparing the uncorrected data of the grid cell corresponding to Saint Hubert (denoted as cell 90) and applies this set to all the grid cells. The fourth and last

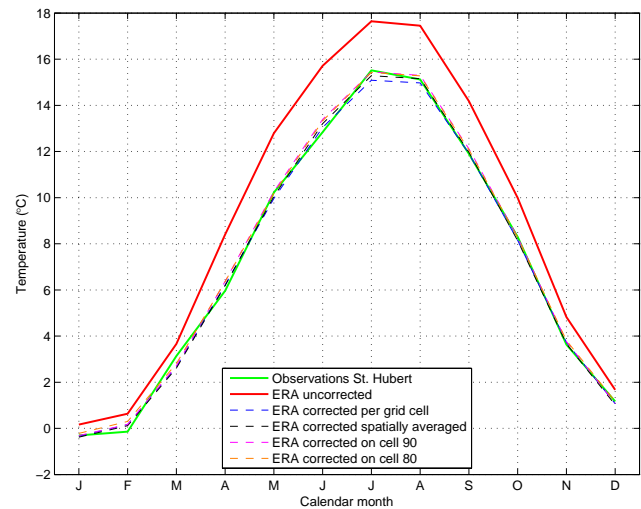


Fig. 4. Mean monthly temperature of the observations and (un)corrected datasets for the period 1979–1996.

method is similar to the third method, but employs a neighbouring grid cell (denoted as cell 80), that better represents the observations at Saint Hubert in terms of monthly precipitation sums than cell 90.

The correction parameters for precipitation are based on the time period 1979–2003, whereas those for temperature are based on the time period 1979–1996. Figure 3 shows the temporal distribution of mean monthly precipitation sums averaged over all the grid cells. To check the representativeness of the meteorological station at St. Hubert for the catchment average value, a precipitation dataset from the Climate Research Unit (CRU) is plotted in Fig. 3 as well. It contains monthly averages of observations, interpolated to a spatial resolution of 0.167° degree, or about 19 km. In Fig. 3, it can be seen that the climatologies of the CRU dataset and the Saint Hubert station look very similar. Therefore, we conclude that the single observation site at Saint Hubert represents the entire catchment at the monthly time scale reasonably well.

3.1.2 Results of the four bias correction methods

In terms of mean monthly precipitation sums, the observations are closest to the correction per grid cell and the spatially averaged correction method. Fig. 4 shows the temporal distribution of mean monthly temperature averaged over all the grid cells. It can be seen directly that the bias corrections for temperature are very solid for practically every method. Furthermore, the error of the uncorrected ERA dataset with respect to the observations appears to be larger in summertime than in wintertime.

The overall mean precipitation of the uncorrected ERA dataset is lower than that of the observations, but after correction it is higher for all four applied methods. The spatially averaged bias correction method produces a mean that is closest to the observational mean. The bias correction on

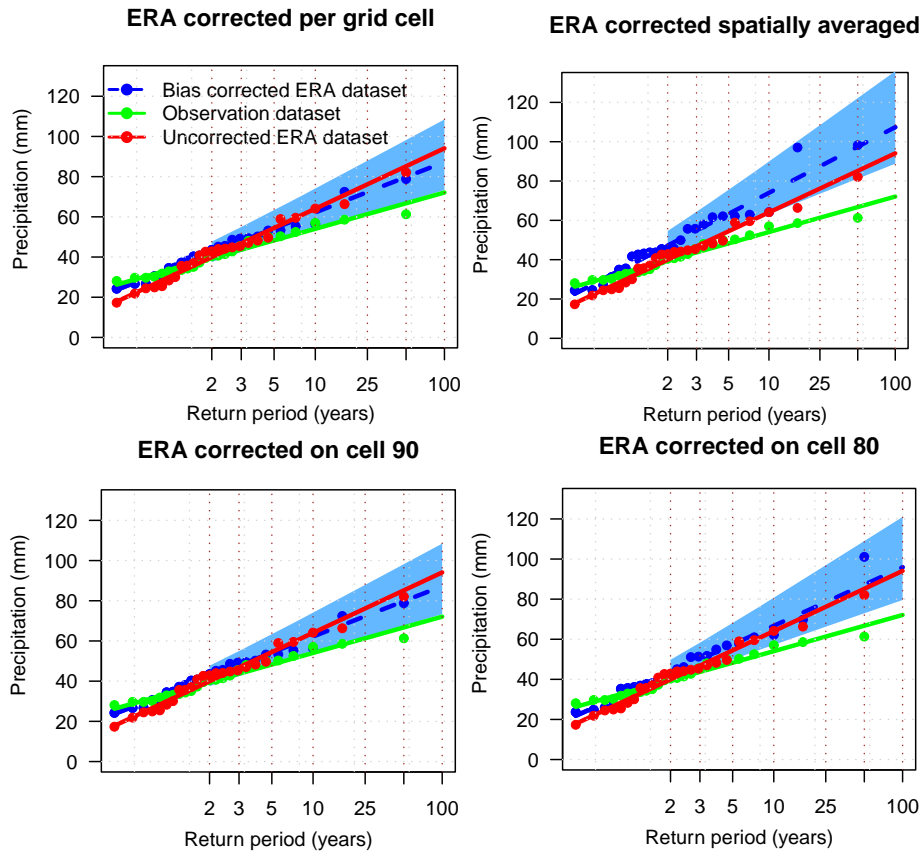


Fig. 5. Gumbel plots of annual maxima of daily precipitation sums of cell 90 for the period 1979–2003. The shaded area is the 95% confidence interval of the bias-corrected dataset. The four panels show results for the four employed bias correction methods.

cell 90 has the highest mean value as well as the highest standard deviation. To investigate extreme values as well, Fig. 5 shows the sorted annual maximum daily precipitation sums of cell 90 versus their return period over 25 years and fits a Gumbel distribution through the data points. Each subplot consists of three datasets of which the observations, the green line, and the uncorrected dataset, the red line, are shown in every plot to put the bias-corrected dataset, that is plotted in blue, in perspective.

Furthermore, the four subplots each contain a corresponding 95%-confidence-interval of the bias-corrected dataset that is produced using the profiling log-likelihood method (Smith, 1985). This is a method to take into account the uncertainty in the fitted Gumbel distribution without making additional assumptions. At every quantile of the dataset of annual precipitation maxima, corresponding to the return periods in Fig. 5 ($q_{(0.50)}$, $q_{(0.33)}$, $q_{(0.20)}$, $q_{(0.10)}$, $q_{(0.04)}$, $q_{(0.02)}$, $q_{(0.01)}$), the Gumbel distribution is reparameterized such that the specific quantile becomes one of the parameters (Coles, 2001). At each quantile (or return period), the 95%-confidence-interval is then estimated from the resulting likelihood function using an iterative method (Venzon and Moolgavkar, 1988).

Figure 5 also shows that the Gumbel fit of the uncorrected dataset is on a higher level than the fits corresponding to the observation dataset. It indicates that the method that corrects per grid cell and the method that corrects on cell 90 both show a fit that is situated between the observed and uncorrected dataset, while the other methods both have a Gumbel fit that is situated above those of the observation dataset as well as the uncorrected dataset.

3.1.3 Selection of the method for further analyses

Based on the performance of the different methods, one should be selected to force the hydrological model. The analysis above shows that there are big differences between the observation dataset and the uncorrected dataset with respect to precipitation sums and climatologies. The bias correction per grid cell and on spatially averaged time series fit best when comparing the mean values of the datasets. A disadvantage of the correction per grid cell is that the values for individual pixels are corrected towards only one observation and that, as a consequence, the spatial variability is reduced. The spatial standard deviation is, therefore, lower than that of the other methods (not shown). In contrast to the correction

of precipitation, the correction of the temperature time series is successful for all the four methods and results in bias-corrected datasets that show the same magnitude and timing as the observed temperature time series.

The bias correction based on a spatially averaged precipitation time series produces a dataset that has a mean which is closest to the observation dataset. The extreme value analysis of the yearly precipitation maxima of cell 90, where Saint Hubert is located, shows not much deviation between the datasets. The largest difference between the datasets at a return period of 100 years based on a Gumbel distribution is about 15 mm. The difference between the uncorrected ERA dataset and the observation dataset is small and their trend lines cross each other at a return period of approximately two years. It is likely that the Gumbel plot representing the Ourthe catchment the best on average should be situated between the green and red trend lines. Fig. 3 shows that the mean of the uncorrected dataset is significantly less than of the observation dataset. Thus, when putting these graphs next to Fig. 5 it can be stated that the uncorrected ERA dataset underestimates the total volume of precipitation, but it slightly overestimates the extremes for return periods larger than 2 years with respect to the observations of Saint Hubert. Based on the interpretations above the bias correction method that corrects spatially averaged data is chosen as the best method to correct the ERA dataset.

Throughout the analyses the bias correction method that corrects per grid cell and the method that corrects spatially averaged times series have shown reasonable results. The strongest argument to choose the latter is based on the preservation of the spatial distribution of precipitation and the fact that less weight is put on the overestimation of extreme precipitation. A strong emphasis should be put on the fact that the method that corrects the spatially averaged time series is the best method of the four assessed methods, but does not guarantee perfectly bias-corrected datasets. The ERA dataset differs quite strongly from the observations on a structural basis and this can never corrected for completely.

Before forcing the HBV model, the ECHAM5 reference and scenario datasets have to be bias-corrected using the observations. Because the bias will most likely be different, new correction parameters are determined by repeating the analysis described above (using the method correcting the spatially averaged data), with data from the reference period and the observations. The reference period is corrected for the period 1961–2000, where the parameter values for temperature are based on the period 1968–1996, while those for precipitation are based on the period 1968–2000. The same parameter values are then applied to correct the three ECHAM5 scenario datasets for the period 2001–2100, assuming that the bias is stationary and will not change under future conditions. This is a common assumption as there is often no information about the stationarity of the bias (e.g., Hurkmans et al., 2010; van Pelt et al., 2009; Leander and Buishand, 2007). However, it may not be completely valid

as the bias tends to increase with increasing temperatures (Christensen et al., 2008). It should be noted that the over-correction of the precipitation that can be seen in Fig. 3 is not present in the correction for the reference period (not shown).

3.2 Model calibration

The HBV model requires precipitation, temperature and potential evaporation as input. As already mentioned, the observations are available for the time period 1968–1996 and are suitable to calibrate the model with. Before the calibration is started, the catchment is divided into five elevation zones, each with its own areal fraction. The elevation zones are used to lapse precipitation and temperature with elevation: precipitation is assumed to increase by 10% with every increase of 100 m and temperature is assumed to decrease with 0.61°C per 100 m increase in elevation. Precipitation is lapsed with 10% per 100 m elevation difference, which is often used in studies using HBV (Seibert, 2005). By this adjustment, the dependency of precipitation on topography, which is present in the atmospheric forcing data, is reintroduced in the lumped model simulation. The fraction of lakes is set to 0.0003, due to the presence of the Lake of Nisramont which has a surface area of 0.47 km², and the elevation of the precipitation and temperature measurements is fixed at 553 m, because this corresponds with the altitude of Saint Hubert.

Fifteen parameters are included in the calibration, see Seibert (2000) for details and parameter ranges. The calibration is carried out using the GAP optimization tool (Genetic Algorithm and Powell), that refers to a genetic algorithm to approximate the solution of the optimization problem (Seibert, 2000) and to a routine using Powell's quadratically convergent method for local optimization of the problem (Press et al., 1992). The genetic algorithm starts with one or more populations of 50 randomly generated parameter sets that are located within the given ranges. These sets are being evaluated by running the model and thus the goodness of fit of each set is determined by the value of the employed objective function. Parameter sets with a good value are given a higher probability to generate new sets than those sets that gave poorer results (Seibert, 2000). The multi-criteria objective function used in this research is a combination of the well-known Nash-Sutcliffe modelling efficiency R_{eff} (Nash and Sutcliffe, 1970), and the absolute value of the volume error, $|VE|$, in mm per year. R_{eff} and $|VE|$ are combined into the combined objective function R_c using (Lindström, 1997):

$$R_c = R_{\text{eff}} - 0.1 \times |VE| \quad (4)$$

Calibration is carried out for the time period 1969–1988, i.e. 20 years, whereas 1968 is used to initialize the system. For validation, the time period 1989–1996 will be used. The calibration period is substantially larger than the validation period, because after the model has been calibrated and validated it has to be able to run large scenario datasets of 100 years. It is expected that calibration on a long time period is

Table 2. Parameters of the HBV model and their optimal values resulting from calibration and validation.

Parameter	Symbol	Unit	Optimal value
Threshold temperature	TT	$^{\circ}\text{C}$	0.7927
Degree-day factor	CFMAX	$\text{mm}^{\circ}\text{C}^{-1}\text{d}^{-1}$	2.724
Snowfall correction factor	SFCF	–	1.2
Refreezing coefficient	CFR	–	0.03417
Water holding capacity	CWH	–	1.695×10^{-5}
Maximum of soil moisture zone	FC	mm	119.91
Threshold for evaporation reduction	LP	–	1.0
Shape coefficient	Beta	–	2.012
Recession coefficient (upper stor.)	K_0	d^{-1}	0.2891
Recession coefficient (upper stor.)	K_1	d^{-1}	0.1563
Recession coefficient (lower stor.)	K_2	d^{-1}	0.0525
Threshold for K_0 to become K_1	UZL	mm	14.30
Maximum percolation	PERC	mm d^{-1}	1.495
Routing parameter	MAXBAS	d	3.565
Correction factor for pot. evaporation	CET	$^{\circ}\text{C}^{-1}$	0.00980

beneficial in order to reach this ability. In this calibration and forecasting process the assumption is made that calibrated parameters stay constant over time. The results of calibration and validation runs showed the best performance for the validation period with a R_{eff} of 0.89, a $|\text{VE}|$ of 8 mm and a correlation coefficient of 0.90, while the calibration period showed a R_{eff} of 0.84, a $|\text{VE}|$ of 34 mm and a correlation coefficient of 0.86. The resulting set of optimal parameters is presented in Table 2. While the higher R_{eff} for the validation period is remarkable, it should be noted that the validation period is relatively short compared to the calibration period (8 versus 20 years). The correlation coefficient was not part of the calibration procedure, but is just an extra measure of model performance

3.3 Calculation of potential evaporation

Potential evaporation is not included in the forcing dataset, therefore it needs to be calculated based on the available forcing parameters to make the reference and scenario datasets suitable as HBV input. For this purpose the simplified Makkink equation is used (Makkink, 1957):

$$L_v \times \text{ET}_{\text{ref}} = 0.65 \times \frac{s}{s + \gamma} \times K^{\downarrow} \quad (5)$$

In this equation L_v stands for the latent heat of vaporization, which has a constant value of $2.451 \times 10^6 \text{ J kg}^{-1}$; ET_{ref} is the reference evapotranspiration in $\text{kg m}^{-2} \text{ s}^{-1}$ and is the parameter that has to be calculated in order to obtain the potential evapotranspiration; s is the slope of the saturated water vapour pressure as a function of temperature curve in kPa K^{-1} ; K^{\downarrow} is the incoming solar radiation in W m^{-2} ; γ the psychrometric constant in kPa K^{-1} . Once the reference

evapotranspiration is calculated the potential evapotranspiration, ET_p , is calculated with use of the crop factor, k_c :

$$\text{ET}_p = k_c \times \text{ET}_{\text{ref}} \quad (6)$$

The crop factor is based on the vegetation cover of the Ourthe catchment, which is provided by the PELCOM database (Mücher et al., 2000). The Ourthe is divided into six land cover types, each with a specific crop factor (given within parentheses; Makkink, 1957): coniferous (1.30), deciduous (1.04) and mixed forest (1.17), grass (1.0), arable rain-fed crops (1.0) and urban area (1.0). Multiplying the crop factor of each land cover type with the corresponding surface fraction results in the average crop factor of 1.06. With this general correction, plausible evapotranspiration values are created using the incoming solar radiation, air temperature and atmospheric pressure of the bias-corrected ECHAM5 dataset.

4 Results and discussion

The calibrated HBV model is then used to simulate discharges at Tabreux for the ECHAM5 reference period and the three ECHAM5 climate scenarios. Time periods of 39 years are selected in order to enable a fair comparison between the reference period and the climate scenarios. The reference period uses one year as warming-up period for initialization and thus simulates the period 1962–2000.

In order to put the different model simulations into perspective, Table 3 shows three streamflow statistics (mean streamflow, mean annual maximum and mean annual minimum), for 1) observed streamflow, 2) streamflow as simulated by the HBV model forced with observed precipitation (Saint Hubert), 3) streamflow as simulated by the HBV

model forced by ERA; and 4) streamflow as simulated by the HBV model forced by the reference data. Because the latter is not constrained by observations, only statistics can be compared, not the actual time series. All values are calculated over the same 17-year period (1980–1996), as this is the period of overlap between all datasets. In Table 3, it can be seen that the HBV model with observed forcing is quite similar to the observations, especially in terms of mean streamflow, but that ERA and the reference forcing datasets produce considerably more streamflow. It should be noted here that all discharges in Table 3 are obtained using the same set of model parameters (Table 2).

The climate scenario datasets contain forcing input for 100 years, while the maximum number of timesteps accepted by HBV Light is 20 000. Therefore two consecutive model runs are needed to capture a 100-year scenario. In the time series that are now created (2002–2100) two periods are isolated, so that clear climate signals can be distinguished: the beginning of the 21st century, 2002–2040, and the end of the 21st century, 2062–2100. The ECHAM5 reference and scenario forcing data are averaged over the catchment and, therefore, the model parameter indicating the elevation of the temperature and precipitation measurement needs to be set at 251 m, which is the average elevation of the Ourthe catchment. All model simulations are carried out using a time step of 1 day.

4.1 Mean outflow

First, the average monthly values of the individual discharge components are investigated. For the beginning and end of the 21st century, Fig. 6 shows the quick upper zone outflux, created by the sum of Q_0 and Q_1 (Fig. 2), the slower outflux from the lower zone, Q_2 , and the total streamflow, i.e. the sum of all fluxes. In the first part of the century the scenarios have monthly values that are similar to those of the reference, but the average total discharge is somewhat higher. Especially the B1 scenario has a high monthly mean of 71 mm and produces the highest discharges in the period from August to January. This increase is strongest in the months January, October and December, when it is ranging between 20% and 54%. The higher discharge of the B1 scenario can be explained due to the increased precipitation for this scenario. In the second part of the century, all the scenario curves have evolved in a more pronounced way. All scenarios indicate less discharge for the months April until September, while in the months December and January discharge is 27% to 38% higher compared to the reference. The overall mean of the scenario discharge has further increased for the A2 scenario, while the discharges for the A1B and B1 scenarios decrease with respect to the beginning of the century and the reference discharge. The total streamflow graphs also show the almost constant discharge of the reference dataset for the months December until April, which can be explained by the buffering effect due to the presence of snow in the catchment. The volume of the snowpack is shown in the upper two plots of

Table 3. Streamflow statistics of four employed sources: 1) observed discharge, 2) simulated discharge where HBV is forced by observations; 3) simulated discharge where HBV is forced by ERA, and 4) simulated discharge where HBV is forced by the reference run. Mean annual average, mean annual maximum and mean annual minimum values are shown.

Dataset	M.A. Average [m ³ s ⁻¹]	M.A. Max. [m ³ s ⁻¹]	M.A. Min. [m ³ s ⁻¹]
Observed	24.0	200.0	2.57
HBV-OBS	21.8	157.3	2.97
HBV-ERA	39.4	305.1	5.10
HBV-REF	35.9	262.3	4.35

Fig. 7, which show very clearly that the storage of snow is diminishing with time. This decrease in volume has an impact on the discharge that logically becomes more dynamic due to the lack of buffering and therefore reacts more strongly to the precipitation forcing. During the second part of the century, the snow storage decreases further and results in a discharge climatology that shows hardly any buffering effect by the snowpack.

The quick flow and baseflow plots in Fig. 6 show similar trends compared the total streamflow of the Ourthe. In winter, the quick flow component is almost twice as large as the baseflow component, while in summer the baseflow is larger than the quick flow component. This effect occurs mainly due to the large variation in the mean monthly quick flow, which can differ 70 mm on a monthly basis throughout the year. The gray shaded areas in Fig. 6 indicate the range between the 25th and 75th percentile of the reference dataset and show the variation of mean discharges in the 39-year reference period. From these ranges, it can be concluded that the mean monthly discharge of total streamflow during winter has a larger variation than the mean monthly discharge during summer.

4.2 Mean storages

The changes in streamflow are in the first place produced by a change in forcing. However, this change is also affecting the storage in the catchment reservoirs. Assessing the different reservoirs gives a further insight why changes in the streamflow occur. On average, the reference period is characterized by snow in the months December until April and a snow storage peak in February of about 44 mm (Fig. 7). All the climate scenarios show a decrease in snow storage in the beginning of the century that extends towards the end of the century. The mean monthly storages in the soil moisture zone are also plotted in Fig. 7 and show that in the winter months December, January, and February the soil moisture zone is nearly saturated until the field capacity value of 120 mm. At the end of the 21st century, the mean storage in the soil moisture zone is clearly less than during the reference period, which

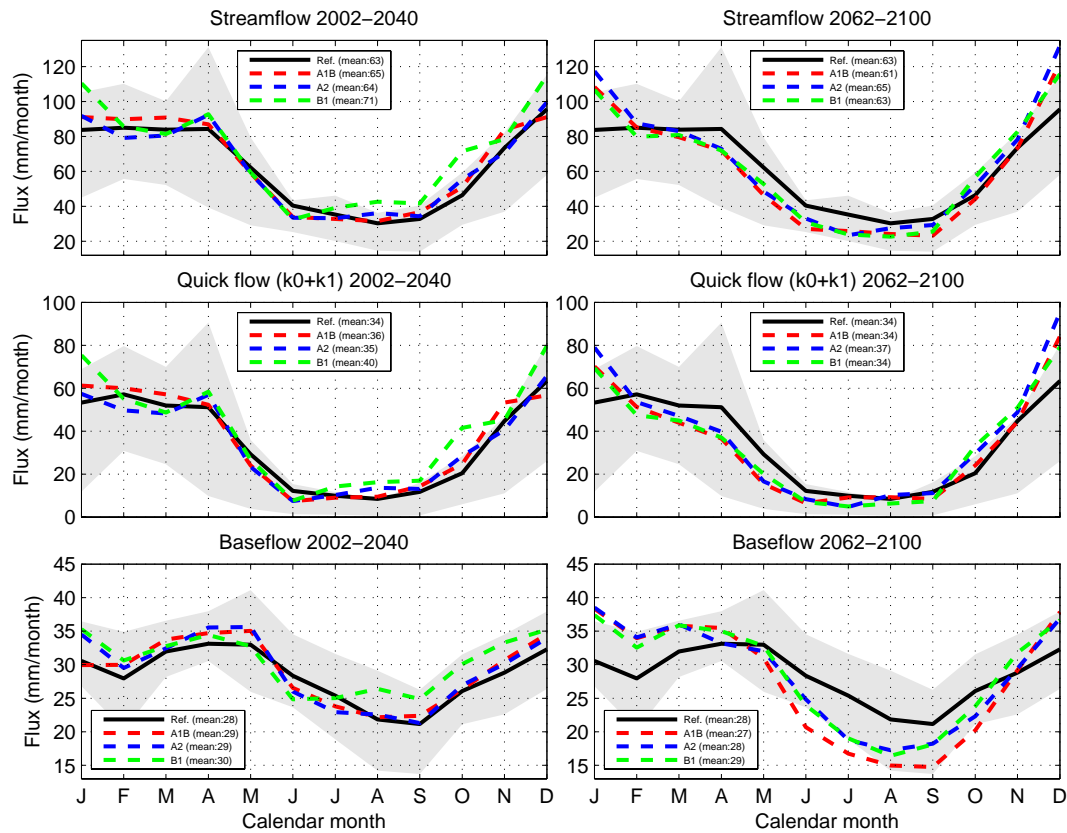


Fig. 6. Mean monthly sums of streamflow, quick flow from the upper zone and baseflow from the lower zone for the reference period and the climate scenarios, shown for the period 2002–2040 (left) and the period 2062–2100 (right). The range between the 25th and 75th percentile of the reference period is plotted as a light gray shaded band.

is mainly due to the decreased water storage in the summer months. This decrease is most pronounced for the A1B scenario that stores on average 5 mm less than the reference. The lower two graphs in Fig. 7 show the monthly mean storage volumes for the lower and the upper zone. On average, the lower zone stores four times more water than the upper zone, which can be explained by the fact that the outflow of the lower reservoir is smaller due to its recession coefficient, K_2 , that is three to five times smaller than the recession coefficients for the upper zone outflow. Structural changes in storage become more distinctive towards the end of the century, especially in the lower zone storage. During the months May until October, less water is stored in this reservoir and the opposite is true for the months November until April. Thus, the seasonal effect in storage becomes stronger for all scenarios.

4.3 Streamflow droughts

In this study, a drought occurs when the discharge drops below a certain threshold. This threshold is arbitrary and is in this case defined as the 75th percentile of the reference period, i.e., the discharge that is exceeded for 75% of the time. Thus, the lower 25% of the daily discharges are considered

as belonging to a drought (Fleig et al., 2006; Hurkmans et al., 2009, 2010). For the Ourthe this comes down to a threshold discharge of $14.1 \text{ m}^3 \text{ s}^{-1}$. In Table 4 drought statistics are shown for the reference period as well as for the several scenarios for the two time periods that have been considered.

The results presented in Table 4 differ strongly throughout the 21st century. In the beginning of the century, all three indicators decrease for all scenarios. The B1 scenario in particular is characterized by an average annual maximum duration of about 32 days, which represents a decrease of 25% with respect to this statistic during the reference period. At the end of the century, the average number of events has decreased even further, but the average of the annual maximum durations shows large increases with respect to the beginning of the century, especially for the A1B and B1 scenarios. The difference of the A1B and B1 scenarios with respect to the reference period is about 25 and 12 days, respectively. Remarkable is also that all scenarios show a significantly higher drought intensity than the reference period towards the end of the century. Especially the A1B scenario has a high drought intensity, which is $1.6 \text{ m}^3 \text{ s}^{-1}$ higher than the reference period.

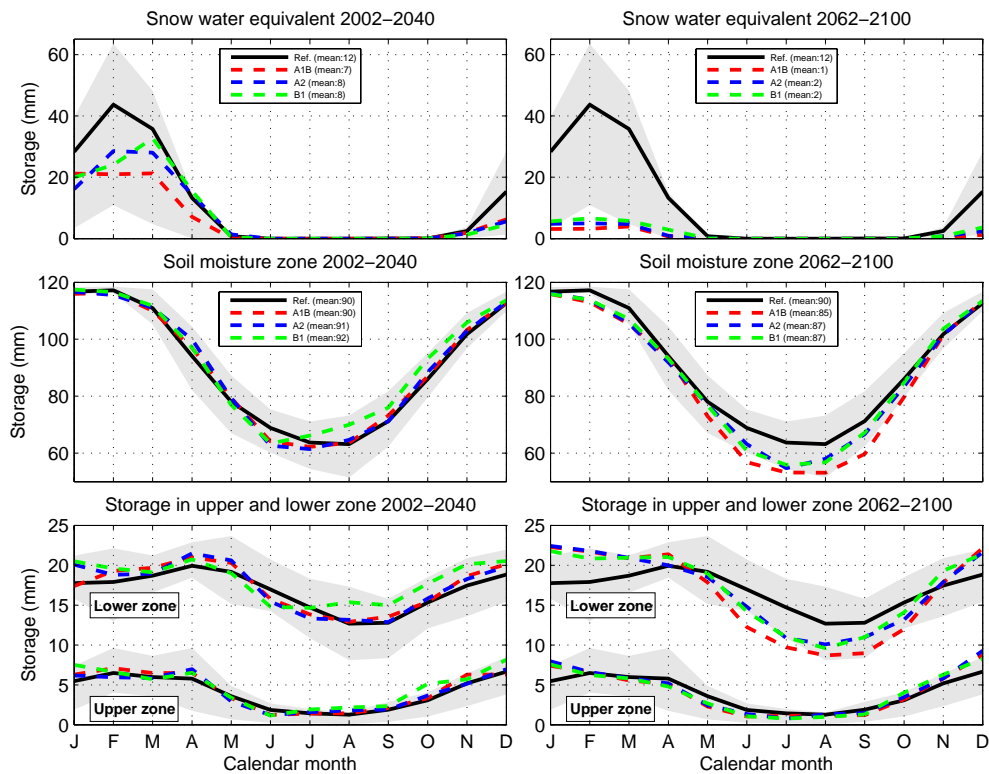


Fig. 7. Mean monthly storages of the snow water equivalent, soil moisture, upper zone and lower zone storage for the reference period and the climate scenarios, shown for the period 2002–2040 (left) and the period 2062–2100 (right). The range between the 25th and 75th percentile of the reference period is plotted as a light gray shaded band.

Table 4. Statistics of drought events for the reference period and the climate scenarios for the period 2002–2040 and 2062–2100. A drought event is calculated with respect to a specific threshold discharge. The statistical parameters are indicated with D1, D2 and D3, where D1 is the average number of events per year, D2 is the average of the yearly maximum durations of a drought in days, and D3 is the average of yearly maximum intensities, where the intensity is defined as the deficit volume divided by duration, in $\text{m}^3 \text{s}^{-1}$.

Period	Reference			A1B scenario			A2 scenario			B1 scenario		
	D1	D2	D3	D1	D2	D3	D1	D2	D3	D1	D2	D3
1962–2000	5.0	41.6	5.1									
2002–2040				4.3	38.5	4.9	4.4	41.2	5.0	4.2	32.3	4.4
2062–2100				3.9	66.9	6.7	4.3	48.7	5.6	4.2	53.5	5.5

Streamflow droughts can also be analyzed by plotting the distribution of the annual maximum cumulative deficit volumes as a function of their return period (Hurkmans et al., 2010), as can be seen in Fig. 8. Here, a deficit volume is defined as the total volume of water during the period that the streamflow remains below the threshold, i.e. the cumulative intensity until the threshold is exceeded again. A Generalized Pareto (GP) distribution is fitted through the data points. For the beginning of the century, not much difference can be seen between the A1B scenario and the reference period. The B1 scenario clearly indicates, through its data points as well as through its fitted distribution, that smaller discharge deficit volumes are expected for the same return period with

respect to the reference period, while the A2 scenario follows the same curve as the reference period until a return period of about six years. For longer return periods, the data points and the fitted distribution are situated above the data points and distribution of the reference. However, it should be stressed that mainly the three highest discharge deficits of the A2 scenario play a role in producing the convex shape of the fitted line. Not too much emphasis should be put on this peculiar shape, because the number of available data points for high return periods is small. At the end of the century, all the data points and fitted distributions of the scenarios indicate significantly larger annual maximum discharge deficit volumes than the reference for the same return periods. The

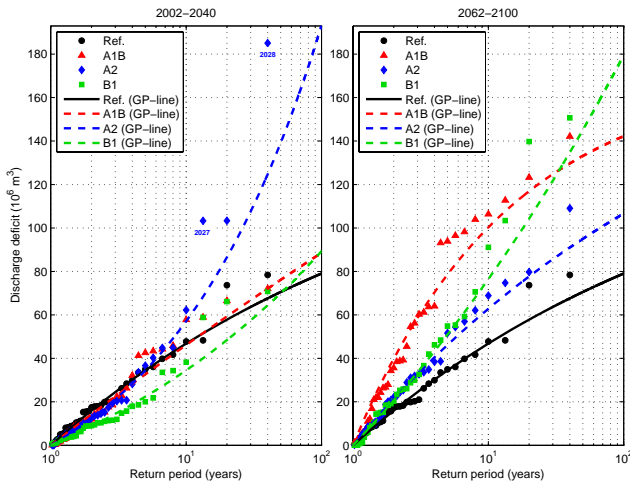


Fig. 8. Annual maximum cumulative deficit of streamflow with respect to the 75th percentile of the reference period versus its return period for the reference period and the climate scenarios, shown for the period 2002–2040 (left) and the period 2062–2100 (right). Corresponding Generalized Pareto (GP) distributions are fitted to the data.

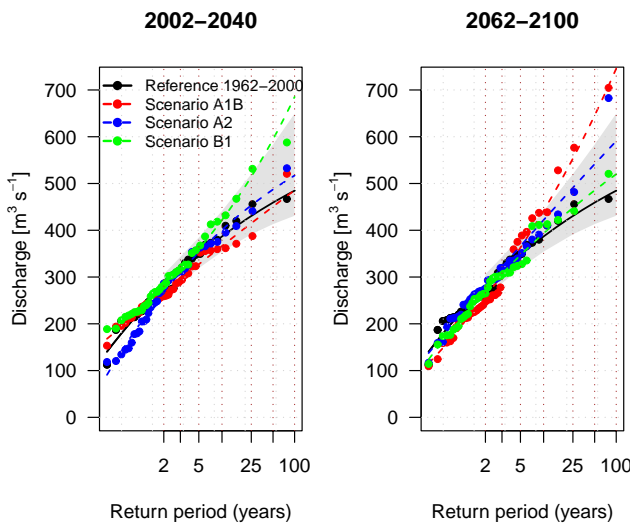


Fig. 9. Annual maximum discharge at Tabreux versus its return period for the reference period and the climate scenarios shown for the period 2002–2040 (left) and the period 2062–2100 (right). Corresponding Generalized Extreme Value (GEV) distributions are fitted to the data. The 95%-confidence interval of the GEV-fit for the reference dataset is plotted as a light gray shaded band.

difference with respect to the reference becomes larger with increasing return periods for the A2 and B1 scenarios. This increase is larger for the B1 scenario than for the A2 scenario. Still, the A1B scenario shows already large differences at low return periods. The increased discharge deficit volumes for the A1B scenario were expected, since Table 4 also showed a considerable increase in annual maximum duration and intensity.

4.4 Peak discharges

For water management purposes, an analysis of the maximum daily discharge is also of importance. In Fig. 9, two graphs show the extreme peak discharges of the scenarios as a function of their return periods for the time periods 2002–2040 and 2062–2100. In the beginning of the century, no strong differences can be distinguished. It is remarkable to notice that the A1B scenario is characterized by a series of five data points that are situated between a return period of 5 and 15 years and have nearly the same discharge level. The A2 scenario has lower peak discharges at low return periods, but follows the reference period quite closely for return periods above two years. In contrast, the B1 scenario has higher peak discharges for return periods above five years with respect to the reference. Towards the end of the century, the A1B scenario shows lower peak discharges than the reference for return periods lower than 3 years, but higher peak discharges than the reference for longer return periods. Furthermore, five of the six highest peak discharges are situated above the 95%-confidence interval of the reference, meaning that under the A1B scenario it is even more likely that higher peak discharges occur.

5 Summary and conclusions

By means of high-resolution regional climate scenarios and a well-known hydrological model, we have obtained projections of the hydrological behaviour of the Ourthe catchment during the 21st century. In addition, an important step in this process, the bias correction of the climate model output has been investigated extensively. Although ERA15 is often used to represent data (e.g., Kotlarski et al., 2005; Hurkmans et al., 2008), for precipitation large differences with respect to the observations appear when considering a relatively small catchment such as the Ourthe. Because the bias correction merely adjusts mean and variability of the precipitation, after bias correction still considerable differences exist between modelled and observed precipitation time series. This should be kept in mind when interpreting our results. Four different bias correction methods have been assessed by looking at the mean monthly values and the spatial distribution for temperature and precipitation, as well as extreme precipitation events. It can be concluded that determination of the bias correction parameters of a spatially averaged dataset after which the same parameters are applied to each grid cell, yields the best results of the methods that were investigated.

By comparing the beginning, 2002–2040, and the end, 2062–2100, of the 21st century with the reference period, 1962–2000, changes in various hydrological variables are investigated. For many variables, the change in the beginning is not so pronounced as it is towards the end of the century. This is true for practically every scenario. In general, the discharge of the Ourthe is characterized by a growing contrast

between seasons, meaning that the winters become wetter and the summers become drier. This holds for all three scenarios that were investigated. The overall annual streamflow increases during the entire century according to all scenarios. Only the A1B scenario projects a small decrease towards the end of the century. All scenarios, A1B being the most extreme, project a decrease in baseflow in the summer months, associated with declining soil moisture and ground-water storages in the summer. Especially in the A2 and A1B scenarios, snow storage is declining throughout the century and practically disappearing towards the end of the century. Also in the B1 scenario it is decreasing, but not as dramatic as in the other two scenarios due to the relatively small temperature increase in the B1 scenario. Streamflow droughts typically decrease during the beginning of the century due to the increase of precipitation and discharge. In contrast, more extreme streamflow droughts are projected to occur towards the end of the century. This holds for all scenarios, again the A1B scenario being the most extreme. Peak flows are not projected to change dramatically or even decrease slightly, except for the B1 scenario in the beginning of the century, where the extreme peak flows are somewhat higher than in the reference situation. The same holds for the A1B scenario at the end of the century.

A drawback of the employed bias correction is the availability of data at only one location. If available, future research should therefore incorporate multiple observation points in order to get a better representation of average precipitation and temperature over the Ourthe catchment. This would enable to use different sets of bias correction parameters and assign them to the time series of the grid cells that they represent. Besides, alternative studies should investigate the use of different hydrological models and, probably more importantly, different climate models to provide the downscaled forcing input. Experiments where GCMs have been compared (e.g., Covey et al., 2003; Reichler and Kim, 2008) indicated a large range even in simulations of the current climate, let alone for a future climate. See Hurkmans et al. (2010) for a more extensive discussion regarding this problem. Using multiple or other models would take into account such a range. Similar comparisons, however, showed that the GCM that was used in this study simulates the current climate very well and is relatively close to the multi-model mean in ensemble projections (IPCC, 2007). In addition, our results are in line with studies using the same models in the Rhine basin (Hurkmans et al., 2010), as well as with studies applying different models in the same region (van Pelt et al., 2009), who all found higher discharges in winter and lower in summer for the end of the 21st century. Also Graham et al. (2007) found lower summer discharges and higher winter discharges, both in the Rhine basin and the Baltic region, using a variety of RCMs and a few GCMs.

Acknowledgements. This research was financially supported by the European Commission through the FP6 Integrated Project NeWater and the BSIK ACER project of the Dutch Climate Changes Spatial Planning Programme. The authors would like to thank Roel Velner and the KMI for providing observed data for the Ourthe catchment. Daniela Jacob from the Max Planck Institut für Meteorologie in Hamburg, Germany is kindly acknowledged for providing the downscaled climate scenarios.

Edited by: G. Blöschl

References

- Arpe, K., Hagemann, S., Jacob, D., and Roeckner, E.: The realism of the ECHAM5 models to simulate the hydrological cycle in the Arctic and North European area, *Nord. Hydrol.*, 36, 349–367, 2005.
- Bergström, S. and Forsman, A.: Development of a conceptual deterministic rainfall-runoff model, *Nord. Hydrol.*, 4, 147–170, 1973.
- Berne, A., ten Heggeler, M., Uijlenhoet, R., Delobbe, L., Dierckx, Ph., and de Wit, M.: A preliminary investigation of radar rainfall estimation in the Ardennes region and a first hydrological application for the Ourthe catchment, *Nat. Hazards Earth Syst. Sci.*, 5, 267–274, 2005, <http://www.nat-hazards-earth-syst-sci.net/5/267/2005/>.
- Booij, M. J.: Impact of climate change on river flooding assessed with different spatial model resolutions, *J. Hydrol.*, 303, 176–198, doi:10.1016/J.Hydrol.2004.07.013, 2005.
- Chbab, E. H.: How extreme were the 1995 flood waves on the rivers Rhine and Meuse?, *Phys. Chem. Earth*, 20, 455–458, 1995.
- Christensen, J. H., Boberg, F., Christensen, O. B., and Lucas-Picher, P.: On the need for bias correction of regional climate change projections of temperature and precipitation, *Geophys. Res. Lett.*, 35, L20709, doi:10.1029/2008GL035694, 2008.
- Coles, S.: An introduction to statistical modeling of extreme values, Springer series in statistics, Springer Verlag, 2001.
- Covey, C., AchutaRao, K. M., Cubasch, U., Jones, P., Lambert, S. J., Mann, M. E., Philips, T. J., and Taylor, K. E.: An overview of results from the Coupled Model Intercomparison Project, *Global Planet. Change*, 37, 103–133, doi:10.1016/S092108181(02)00193-5, 2003.
- de Wit, M., Warmerdam, P., Torfs, P., Uijlenhoet, R., Roulin, E., Cheymol, A., van Deursen, W., van Walsum, P., Ververs, M., Kwadijk, J., and Buiteveld, H.: Effect of climate change on the hydrology of the river Meuse, Tech. Rep. 410 200 090, Dutch National Research Programme on Global Air Pollution and Climate Change, 2001.
- de Wit, M. J. M., van den Hurk, B., Warmerdam, P. M. M., Torfs, P. J. J. F., Roulin, E., and van Deursen, W. P. A.: Impact of climate change on low-flows in the river Meuse, *Clim. Change*, 82, 351–372, doi:10.1007/S10584-006-9195-2, 2007.
- Fleig, A. K., Tallaksen, L. M., Hisdal, H., and Demuth, S.: A global evaluation of streamflow drought characteristics, *Hydrol. Earth Syst. Sci.*, 10, 535–552, 2006, <http://www.hydrol-earth-syst-sci.net/10/535/2006/>.

- Gibson, J., Kallberg, P., Uppala, S., Hernandez, A., Nomura, A., and Serrano, E.: ECMWF Re-Analysis Project Report Series: 1. ERA-15 Description (version 2), Tech. rep., ECMWF, 1999.
- Graham, L. P., Hagemann, S., Jaun, S., and Beniston, M.: On interpreting hydrological change from regional climate models, *Climatic Change*, 81, 97–122, doi:10.1007/s10584-006-9217-0, 2007.
- Hay, L. E., Clark, M. P., Wilby, R. L., Gutowski Jr., W. J., Leavesly, G. H., Pan, Z., Arritt, R. W., and Takle, E. S.: Use of regional climate model output for hydrologic simulations, *J. Hydrometeorol.*, 3, 571–590, 2002.
- Hazenbergh, P., Leijnse, H., and Uijlenhoet, R.: Weather radar estimation of large scale stratiform winter precipitation in a hilly environment and its implications for rainfall runoff modeling, *Water Resour. Res.*, in review, 2010.
- Hundecha, Y. and Bárdossy, A.: Modeling of the effect of land use changes on the runoff generation of a river basin through parameter regionalization of a watershed model, *J. Hydrol.*, 292, 281–295, doi:10.1016/j.jhydrol.2004.01.002, 2004.
- Hurkmans, R. T. W. L., de Moel, H., Aerts, J. C. J. H., and Troch, P. A.: Water balance versus land surface model in the simulation of Rhine river discharges, *Water Resour. Res.*, 44, W01418, doi:10.1029/2007WR006168, 2008.
- Hurkmans, R. T. W. L., Terink, W., Uijlenhoet, R., Moors, E. J., Troch, P. A., and Verburg, P. H.: Effects of land use changes on streamflow generation in the Rhine basin, *Water Resour. Res.*, 45, W06405, doi:10.1029/2008WR007574, 2009.
- Hurkmans, R. T. W. L., Terink, W., Uijlenhoet, R., Torfs, P. J. J. F., Jacob, D., and Troch, P. A.: Changes in streamflow dynamics in the Rhine basin under three high-resolution climate scenarios, *J. Climate*, 23, 679–699, doi:10.1175/2009JCLI3066.1, 2010.
- IPCC: Special report on emissions scenarios - a special report of working group III of the Intergovernmental Panel on Climate Change, <http://www.grida.no/climate/ipcc/emission/index.htm>, (last access: April 2010), 2000.
- IPCC: Fourth assessment report: Climate change 2007: Climate change impacts, adaptation and vulnerability. Summary for policy makers., 2007.
- Jacob, D.: A note to the simulation of the annual and inter-annual variability of the water budget over the Baltic Sea drainage basin, *Meteorol. Atmos. Phys.*, 77, 61–73, 2001.
- Jacob, D., Göttel, H., Kotlarski, S., Lorenz, P., and Sieck, K.: Klimaauswirkungen und Anpassung in Deutschland – Phase 1: Erstellung regionaler Klimaszenarien für Deutschland, Forschungsbericht 204 41 138 UBA-FB 000969, Umweltbundesamt, Dessau, Germany, 2008.
- Kotlarski, S., Block, A., Böhm, U., Jacob, D., Keuler, K., Knoche, R., Rechid, D., and Walter, A.: Regional climate model simulations as input for hydrological applications: evaluation of uncertainties, *Adv. Geosci.*, 5, 119–125, 2005, <http://www.adv-geosci.net/5/119/2005/>.
- Kwadijk, J. and Rotmans, J.: The impact of climate change on the river Rhine: a scenario study, *Clim. Change*, 30, 397–425, 1995.
- Leander, R. and Buishand, T. A.: Resampling of regional climate model output for the simulation of extreme river flows, *J. Hydrol.*, 332, 487–496, doi:10.1016/j.jhydrol.2006.08.006, 2007.
- Lenderink, G., Buishand, A., and van Deursen, W.: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach, *Hydrol. Earth Syst. Sci.*, 11, 1145–1159, 2007, <http://www.hydrol-earth-syst-sci.net/11/1145/2007/>.
- Lindström, G.: A simple automatic calibration routine for the HBV model, *Nord. Hydrol.*, 28, 153–168, 1997.
- Lorenz, P. and Jacob, D.: Influence of regional scale information on the global circulation: a two-way nesting climate simulation, *Geophys. Res. Lett.*, 32, L18706, doi:10.1029/2005GL023351, 2005.
- Makkink, G.: Testing the Penman formula by means of lysimeters, *International Journal of Water Engineering*, 11, 277–288, 1957.
- Mücher, S., Steinnocher, K., Champeaux, J.-L., Griguolo, S., Wester, K., Heunks, C., and van Katwijk, V.: Establishment of a 1-km Pan-European land cover database for environmental monitoring, in: *Proceedings of the Geoinformation for All XIXth Congress of the International Society for Photogrammetry and Remote Sensing (ISPRS)*, edited by: Beek, K. J. and Molenaar, M., *Int. Arch. Photogramm. Remote Sens., GITC*, Amsterdam., 33, 702–709, 2000.
- Nash, J. E. and Sutcliffe, I. V.: River flow forecasting through conceptual models. Part I – A discussion of principles, *J. Hydrol.*, 10, 282–290, 1970.
- Press, W., Flannery, B., Teukolsky, S., and Vetterling, W.: *Numerical recipes in FORTRAN: The art of scientific computing*, Cambridge University Press, Cambridge, Great Britain, 2nd ed., 1992.
- Reichler, T. and Kim, J.: How well do coupled models simulate today's climate?, *Bull. Amer. Meteor. Soc.*, 89, 303–311, doi:10.1175/BAMS-89-3-303, 2008.
- Seibert, J.: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, *Hydrol. Earth Syst. Sci.*, 4, 215–224, 2000, <http://www.hydrol-earth-syst-sci.net/4/215/2000/>.
- Seibert, J.: HBV light version 2, User's manual, Dept. of Physical Geography and Quaternary Geology, Stockholm University, 2005.
- Shabalova, M. V., van Deursen, W. P. A., and Buishand, T. A.: Assessing future discharge of the river Rhine using regional climate model integrations and a hydrological model, *Clim. Res.*, 23, 233–246, 2003.
- Smith, R. L.: Maximum likelihood estimation in a class of nonregular cases, *Biometrika*, 72, 67–90, 1985.
- te Linde, A. H., Aerts, J. C. J. H., Hurkmans, R. T. W. L., and Eberle, M.: Comparing model performance of two rainfall-runoff models in the Rhine basin using different atmospheric forcing data sets, *Hydrol. Earth Syst. Sci.*, 12, 943–957, 2008, <http://www.hydrol-earth-syst-sci.net/12/943/2008/>.
- Terink, W., Hurkmans, R. T. W. L., Torfs, P. J. J. F., and Uijlenhoet, R.: Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin, *Hydrol. Earth Syst. Sci. Discuss.*, 7, 221–267, 2010, <http://www.hydrol-earth-syst-sci-discuss.net/7/221/2010/>.
- Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, *Bull. Amer. Meteor. Soc.*, 84, 1205–1217, doi:10.1175/BAMS-84-9-1205, 2003.
- van Pelt, S. C., Kabat, P., ter Maat, H. W., van den Hurk, B. J. J. M., and Weerts, A. H.: Discharge simulations performed with a hydrological model using bias corrected regional climate model input, *Hydrol. Earth Syst. Sci.*, 13, 2387–2397, 2009, <http://www.hydrol-earth-syst-sci.net/13/2387/2009/>.

Velner, R.: Neerslag-afvoer modellering van het stroomgebied van de Ourthe met het HBV model: een studie ten behoeve van verlenging van de zichttijd van hoogwatervoorspelling op de Maas, Master's thesis, Wageningen University, 2000.

Venzon, D. J. and Moolgavkar, S. H.: A method for computing profile-likelihood-based confidence intervals, *Appl. Statist.*, 37, 89–94, 1988.