



# Uncertainty Analysis of Climate Change Impact on River Flow Extremes Based on a Large Multi-Model Ensemble

Jan De Niel<sup>1</sup>  · E. Van Uytven<sup>1</sup>  · P. Willems<sup>1</sup> 

Received: 16 April 2019 / Accepted: 27 September 2019 /  
Published online: 30 October 2019

© Springer Nature B.V. 2019

## Abstract

Water managers are faced with a changing climate in the decision-making process while adaptation and mitigation strategies need to be developed. The climate change impact towards the end of the century, however, is highly uncertain and coping with this is a great challenge for decision makers. Over the recent years, combined efforts of hydrologists and climatologists have led to many climate change impact studies on water resources. However, most studies only use a limited ensemble size and/or focus on only one contributing source and hence possibly underestimate the total uncertainty. For two Belgian catchments, we simulated daily flow with five different lumped conceptual hydrological models and ten different parameter sets each, forced by the output of 24 global climate models covering four different emission scenarios, combined with 9 different downscaling methods over reference (1961–1990) and future (2071–2100) periods, resulting in a large multi-model ensemble with 41,850 members. Results show that both low and peak flows would become more extreme in the future, and these changes are stronger with increased radiative forcing. The most important uncertainty sources in low-flow projections are the global climate models (explaining 27–36% of the total variance) and the hydrological model structure (34–42%). For peak flow projections, these are global climate models (32–39%) and statistical downscaling methods (21–26%). Also, interaction effects account for a significant part of the uncertainty (24–38%). The results of this study illustrate that one might end up with biased results and overly confident conclusions when only focusing on some of the uncertainty sources in multi-model ensembles.

**Keywords** Hydrological extremes · Rainfall-runoff model · Uncertainty · ANOVA · CMIP5

## Highlights

- Extreme river flows will become more extreme under climate change
- Climate models are the dominating source of uncertainty
- Statistical downscaling is an important source of uncertainty for peak-flow impacts
- Hydrological model structure is an important source of uncertainty for low-flow impacts

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11269-019-02370-0>) contains supplementary material, which is available to authorized users.

✉ Jan De Niel  
jandeniel@gmail.com

<sup>1</sup> Hydraulics Section, Department Civil Engineering, KU Leuven, Leuven, Belgium

## 1 Introduction

The latest report of the Intergovernmental Panel Climate Change [IPCC 2013] states that since the 1950s observed changes in the climate system such as heavy precipitation events are unprecedented over decades to millennia. It is very likely that further changes by the late twenty-first century will take place. For Europe, on average, flood peaks with return periods above 100 years might double in frequency within three decades [Alfieri et al. 2015b]. By the end of this century, the socio-economic impact of river floods could increase on average by 220% due to climate change only [Alfieri et al. 2015a]. More specifically, in Belgium, there is a possible increase of up to 40% for mean precipitation; extreme precipitation changes, however, might increase by up to 50% in winter and 100% in summer [Tabari et al. 2015].

To quantify climate change impacts on water resources, a vertical top-down methodology is typically followed (see e.g. Xu 1999): global climate models (GCMs) with driving greenhouse gas scenarios (GHGs) or representative concentration pathways (RCPs), in combination with downscaling methodologies, dynamically through regional climate models (RCMs) or statistically with statistical downscaling methods (SDMs), first describe the local climate change signal for a given future period. This is further used as input for a (calibrated) hydrological model (HM) in order to quantify changes in hydrological variables, which are often related to water availability and/or flooding conditions. Each consecutive step of the above modelling chain adds uncertainty to the results. Knowledge and understanding of these uncertainties is crucial for water managers to guide them in the decision-making process.

Recent research acknowledges that not adequately addressing all sources of uncertainty may result in biased climate change impact studies and resulting decisions [Chen et al. 2011; Najafi and Hessami Kermani 2017]. Most studies, however, only consider a limited ensemble describing the combined climate and hydrological uncertainty.

Table 1 shows some of the recent literature focusing on climate change impact on water resources and its related uncertainty. Note that this list is a limited sample taken from all available literature. The table shows that, mostly, the research is focusing on only one specific source of uncertainty and, hence, these studies possibly underestimate the total uncertainty involved. This has recently been confirmed by Seiller et al. [2017], who concluded that “[...] *the diagnosis of the impacts of climate change on water resources are quite affected by the hydrologic models selection and calibration metrics*”. Moreover, Sunyer et al. [2015], among others, pointed out that various downscaling methods will result in significant differences in the (magnitude) of impact results.

Here, we apply an ensemble of 93 GCM runs (a combination of 24 climate models with 4 RCPs, see Table 2), 9 SDMs, 5 HMs and 10 parameter sets per HM, for two catchments in Belgium. The resulting balanced ensemble of 41,850 members per catchment, is, to the authors’ knowledge, unseen in literature. It allows us to better investigate the relative importance of each contributing source of uncertainty. This study focuses on the impacts of climate change on low and peak flows, hence the types of impacts of importance for water management: hydrological extremes, floods and water availability.

## 2 Study Area

Two study areas in Belgium have been selected: the Grote Nete catchment and the Dijle catchment (Fig. 1). The Grote Nete catchment in the North-Eastern part of Belgium has an area of 402 km<sup>2</sup> with a sandy soil coverage. The Dijle catchment in central Belgium has an area of

**Table 1** Selected recent climate change impact studies on water resources, sorted by decreasing total ensemble size

Reference	Greenhouse gas scenarios	Climate models	Downscaling methods	Hydrological models	HM pars.	# ensemble members
[Chamorro et al. 2017]	4	5	1	2	1000	40,000
[Seiller et al. 2017]	2	29	2	12	3	4176
[Chen et al. 2011]	2	6	4	3	10	1440
[Teng et al. 2011]	2	15	1	5	1	150
[Hundeicha et al. 2016]	1	15	8	1	1	120
[Teklesadik et al. 2017]	4	4	1	6	1	96
[Vidal et al. 2016]	1	4	3	6	1	72
[Liu et al. 2011]	3	18	1	1	1	54
[Liu et al. 2017]	2	21	1	1	1	42
[Gosling et al. 2017]	1	5	1	8	1	40
[Islam et al. 2017]	2	12	1	1	1	24
[Meaurio et al. 2017]	2	5	2	1	1	20
[Pandey et al. 2019]	2	3	1	1	1	6
[Velázquez et al. 2013]	1	1	1	4	1	4
[Pandey et al. 2017]	3	1	1	1	1	3
[Vansteenkiste et al. 2014b]	3*		1	3	9	81 (*)
[Dams et al. 2015]	3*		1	4	4	36 (*)
[Vansteenkiste et al. 2012]	3*		1	2	1	6 (*)

an asterisk (\*) indicates that a (small) number of scenarios were created based on a larger ensemble of climate model runs

636 km<sup>2</sup> with as main land use: urban (about 20% of the study area), pasture and forests (resp. 16 and 18%) and arable land (silt and silty loam soils). Both catchments have high baseflow during winter and very low discharges during summer. Belgium has a maritime, temperate climate with annual precipitation ranging between 700 and 1000 mm/year. The long term average evapotranspiration amount is about 670 mm/year. These catchments have been subject to several climate change impact analyses previously [Van Steenberg and Willems 2012; Vansteenkiste et al. 2012, 2014a, 2014b; Willems., 2014].

## 3 Materials and Methods

### 3.1 Meteorological Data

#### 3.1.1 Global Climate Models

The ensemble includes 93 global, CMIP5 based climate model runs, hereafter denoted as GCM runs, and comprises 24 different models and 4 RCPs (RCP 2.6, 4.5, 6.0 and 8.5, Table 2). The periods 1961–1990 and 2071–2100 were chosen as reference and future period, respectively.

For all selected climate model runs, data is available for precipitation and other variables required for the various downscaling methods.

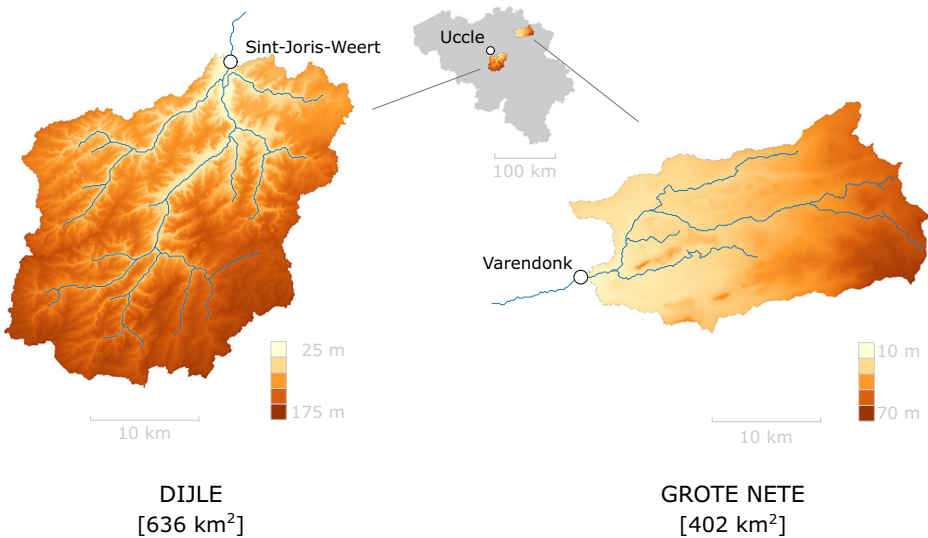
Evapotranspiration (ET<sub>o</sub>) is, here, calculated based on the Penman-Bultot evapotranspiration calculation method [Baguis et al. 2010; Bultot et al. 1983] and requires various meteorological variables such as radiation, heat fluxes, relative humidity and wind speed

**Table 2** Selected global climate models

Modeling left or Group	Institute ID	Model name	RCP			
			2.6	4.5	6.0	8.5
Beijing Climate left, China Meteorological Administration	BCC	BCC-CSM1.1(m)	1	1	1	
CSIRO (Commonwealth Scientific and Industrial Research Organisation, Australia), and BOM (Bureau of Meteorology, Australia)	CSIRO-BOM	ACCESS 1.0				1
Canadian Centre for Climate Modelling and Analysis	CCCma	CanESM2	5	5		5
Centro Euro-Mediterraneo per I Cambiamenti Climatici	CMCC	CMCC-CESM				1
		CMCC-CMS		1		1
Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	CNRM-CERFACS	CNRM-CM5	1*	1*		1*
Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence	CSIRO-QCCCE	CSIRO-Mk3.6.0		1	1	1
College of Global Change and Earth System Science, Beijing Normal University	GCESS	BNU-ESM	1*	1*		
Institute for Numerical Mathematics	INM	INM-CM4		1*		1*
Institut Pierre-Simon Laplace	IPSL	IPSL-CM5A-LR	1*	3*	1*	3*
		IPSL-CM5A-MR	1*	1*	1*	1*
		IPSL-CM5B.LR		1		1
Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	MIROC	MIROC-ESM-CHEM	1	1	1	1
		MIROC-ESM	1	1	1	1
Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	MIROC	MIROC5	3	2	3	3
Met Office Hadley Centre (additional HadGEM2-ES realizations contributed by Instituto Nacional de Pesquisas Espaciais)	MOHC	HadGEM2-A	1	1	1	1
Max Planck Institute for Meteorology (MPI-M)	MPI-M	MPI-ESM-LR	1	1		1
		MPI-ESM-MR	1	1		1
Meteorological Research Institute	MRI	MRI-CGCM3	1*	1*	1*	1*
NASA Goddard Institute for Space Studies	NASA GISS	GISS-E2-R		2		
Norwegian Climate Centre	NCC	NorESM1-M	1		1	1
Geophysical Fluid Dynamics Laboratory	NOAA GFDL	GFDL-CM3			1	1
		GFDL-ESM2G	1*	1*	1*	1*
		GFDL-ESM2M		1*	1*	1*

an asterisk (\*) indicates the availability of all variables needed to calculate evapotranspiration according to Penman-Bultot [Baguis et al. 2010; Bultot et al. 1983]

as input. This puts a limit on the availability of GCM runs: ETo can only be calculated for those GCM runs indicated with an asterisk (\*) in Table 2. More information on the ETo calculation method and how it has been calibrated for Belgium can be found in Bultot et al. [1983] and Baguis et al. [2010].



**Fig. 1** Selected case studies in Belgium

### 3.1.2 Downscaling Methods

The spatial and temporal resolutions of the climate models mismatch the resolutions required for hydrological impact modelling [Xu 1999]. Also, the output of climate models is potentially biased. This impedes the direct application of climate model outputs and requires downscaling. In general, two downscaling approaches exist: dynamical and statistical downscaling. High resolution regional and limited area climate models (RCMs and LAMs) form the basis for dynamical downscaling. However, recent research has shown that high resolution climate models do not appear to give added value to simulation of precipitation extremes and remain biased [Tabari et al. 2016]. Moreover, compared to the GCMs, the availability of RCMs and LAMs is limited and this might create biased uncertainty estimates [Van Uytven and Willems 2018]. Given the research objective to focus on all uncertainty sources and, thus, to design a large multi-ensemble, we have chosen to implement statistical downscaling methods (SDMs) and apply these on an ensemble of global climate models.

The ensemble includes 9 SDMs. Different categorizations exist for these SDMs; here, we consider the categorization of Maraun et al. [2010] which divides the methods into three separate categories:

Model Output Statistics (MOS 1–5): 1 delta change method (method CFM in Sunyer et al. 2015); 3 different types of quantile perturbation methods (the quantile perturbation method accounting for changes in the number of wet days as described in Ntegeka et al. 2014 and, methods SD-A-4 and SD-A-5 in Willems and Vrac 2011) and 1 event based perturbation method (method  $S_B$  in Sørup et al. 2017);

Weather Generator (WG): 1 event based, re-sampling based weather generator [Thorndahl et al. 2017];

Perfect Prognosis (PP 1–3): 3 different types of weather typing methods including 1 resampling method (method SD-B-1 in Willems and Vrac 2011) and two analogue methods (methods SD-B-3 and SD-B-7 in Willems and Vrac 2011). These methods are implemented, based on the ERA-40 re-analysis data [Uppala et al. 2005].

Methods MOS 1–5 are change factor based statistical downscaling methods. Change factors are calculated using the climate model outputs for the future period with respect to the outputs for the reference period. They are thereafter applied to the observed time series. The considered weather generator is also a change factor based statistical downscaling method. The parameters of the weather generator are calibrated using observations and are next modified by change factors. Hence, the modified change factors are representative for the future climate. Change factor based statistical downscaling methods assume that the climate model biases remain time-invariant and are accounted for by considering change factors rather than the direct climate model outputs.

The weather typing methods are perfect prognosis methods. For this type of statistical downscaling methods, the relation between the predictors and the predictand is defined using historical observations. Next, the relation is applied to the climate model output assuming the predictors are accurately and adequately simulated (i.e. perfect prognosis assumption). However, the original implementation of the weather typing methods is not tailored to the biases in the weather type occurrences (Willems and Vrac 2011). This has been identified as a method shortcoming by Van Uytven et al. (2019). Remark that also change factor based statistical downscaling methods have shortcomings, e.g. they are data driven and not physically based, and that the shortcomings of statistical downscaling methods are accounted for by ensembles.

## 3.2 Hydrological Modelling

### 3.2.1 Hydrological Model Structures

The ensemble includes five hydrological model structures: VHM [Willems 2014;], HBV [Lindström et al. 1997], NAM [Nielsen and Hansen 1973; DHI 2009], GR4J [Perrin et al. 2003] and PDM [Moore 1985, 2007]. These are all conceptual hydrological models (in contrast to physically based models) and are implemented in a lumped way (in contrast to spatially distributed modelling – some of these hydrological models have been applied in a distributed form, however, this is not used in this study).

Previous research has shown that conceptual lumped hydrological models are able to show a similar or higher overall performance compared to the more complex distributed and physically based models [Breuer et al. 2009; Viney et al. 2009; Ghavidelfar et al. 2011; Liu et al. 2011; Apip et al. 2012; Lobligois et al. 2014; de Boer-Euser et al. 2017]. On top of that, physically based and distributed hydrological models are more prone to the issue of equifinality and overparameterization [Gupta and Sorooshian 1983; Beven 1989, 1993, 2006; Jakeman and Hornberger 1993; Uhlenbrook et al. 1999; Das et al. 2008; Sivakumar 2008; Andréassian et al. 2012]. The two above reasons, complemented with the low computational cost of lumped conceptual models compared with distributed physically based models, drive us to the choice for the former in the context of this study.

### 3.2.2 Calibration of Hydrological Models

The calibration protocol is detailed below:

- Calibration period: 1 September 2003–31 December 2005, with a warming up period starting on 13 August 2002 for the Grote Nete catchment; for the Dijle catchment this was

1 November 2009–28 February 2013, with a warming up period starting on 1 September 2008.

- Validation period: 1 January 2006–31 December 2008 (the Grote Nete catchment) and 1 March 2013–1 March 2016 (the Dijle catchment), with the calibration period as warming up period
- Algorithm: the MOSCEM-UA algorithm [Vrugt et al. 2003] was used to perform the (automatic) calibration Nash-Sutcliffe efficiency (NSE) and NSE of the log of the flows (NSElog) were chosen as objective functions for the calibration. The MOSCEM-UA algorithm results in a Pareto-front of best parameter sets, of which 10 sets, evenly spread across the objective function space, are selected for the further analysis.
- Daily time step was used.

Resulting hydrographs of the calibrated and validated hydrological models can be found in the [Supplementary material](#).

### 3.3 Impact Calculations

Calibrated hydrological models are forced by time series of 100 years of current and future (projected) meteorological variables (precipitation P and potential evapotranspiration ETo). From the resulting (current and future) discharge series, low and peak-flow extremes are extracted by means of a Peak-over-threshold method [Willems 2009]. After ranking the flow extremes from more extreme to less extreme, empirical return periods are assigned and a climate change impact factor is calculated per return period:

$$IF_x(T) = \frac{X(T)_{future}}{X(T)_{current}}, \quad (1)$$

with X the variable of interest (low-flow extreme, or peak-flow extreme) and T the empirical return period. Impact factors are further averaged over all empirical return periods larger than 1 year (only if the behavior of the impact factors are found to be stable within that range) in order to obtain one impact factor per ensemble member. With Eq. (1), one can see that  $IF < 1$  (for low-flow extremes) and  $IF > 1$  (for peak-flow extremes) would correspond with more extreme conditions in the future.

The forcing data for conceptual hydrological models is precipitation P and potential evapotranspiration ETo. However, as shown by, only a limited number of GCM runs are able to produce projections for ETo according to the Penman-Bultot method. Every downscaled time series of P is therefore combined with every downscaled series of ETo. After propagation through the impact model, impact factors are averaged for the various ETo-projections. This results in one set of impact factors per downscaled time series of P. Such an approach does not bias the impact results, as shown in De Niel et al. [2018].

For the stochastic downscaling methods, impacts are calculated for each of the 100 stochastic simulations. After propagation through the impact model, the impact factors are averaged.

Note that the land use/land cover will – most probably – have changed by the end of this century and these changes will – most probably – further impact river flow extremes. However, as the present study is limited to an uncertainty analysis of climate change impacts only, we do not consider these land use/land cover changes.

### 3.4 Uncertainty Analysis

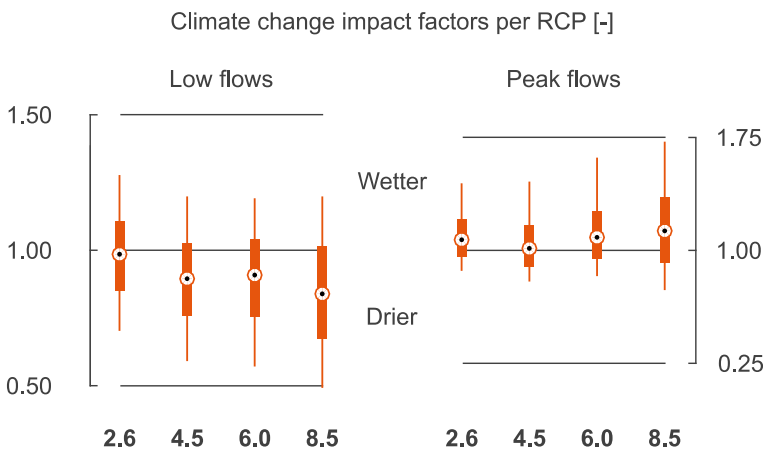
In order to quantify the relative contribution of the various members of the impact modelling chain (GCM runs, SDM, HM and parameter sets) to the total uncertainty in low-flow and peak-flow impact factors, ANOVA (analysis of variance) is applied. This technique is widely applied in hydrology [Dams et al. 2015; Giuntoli et al. 2015; Sunyer et al. 2015; Vidal et al. 2016; Meresa and Romanowicz 2017] and is able to distinguish the main effects of each contributing source of uncertainty, as well as interaction effects between the various factors.

## 4 Results

Figure 2 shows the impact factors for low-flow extremes and peak-flow extremes. These impact factors (one per ensemble member) are grouped based on the representative concentration pathways (RCPs). For both peak-flow and low-flow extremes, more extreme conditions are projected towards the end of the century: the impact factors are generally larger than 1 for peak-flow extremes, and smaller than 1 for low-flow extremes, irrespective of the RCP considered. Furthermore, the impact becomes more extreme with increasing radiative forcing: the mean impact factor for peak-flows considering a radiative forcing of  $8.5 \text{ W/m}^2$  is equal to 1.18, vs. 1.07 with a radiative forcing of  $4.5 \text{ W/m}^2$ . Similar for low-flow impacts: 0.84 for  $8.5 \text{ W/m}^2$  vs. 0.98 for  $4.5 \text{ W/m}^2$ . Note that, for this figure, the results of both catchments are pooled since they were showing almost no difference.

The large spread in impact results, as seen in Fig. 2, is further investigated. The ANOVA results in Fig. 3 show the relative contributions of the different sources of uncertainty to the spread in impact results, both for the impact on peak-flow extremes and low-flow extremes. The results are shown for both case studies and seem to be comparable for both catchments.

The GCM runs (including RCP and initial conditions) have a similar contribution of 30–40% to the spread in peak-flow and low-flow impacts. For low-flow extremes, the downscaling methodology has no no significant influence on the impact results. The



**Fig. 2** Climate change impacts on low-flow (left) and peak-flow (right) extremes become more extreme with increased radiative forcing

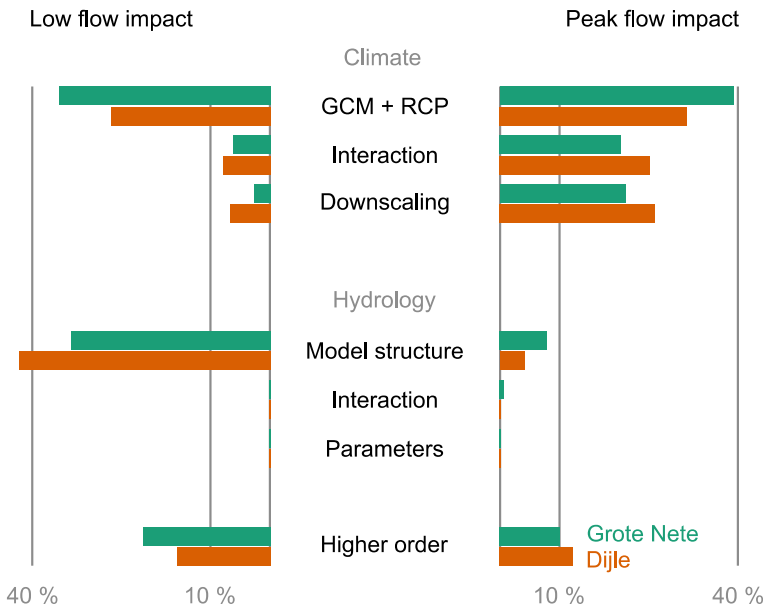


hydrological model structure, on the contrary, explains up to 34% and 42% (for Grote Nete catchment and the Dijle catchment, respectively) of the spread in low-flow impact results. For peak-flow extremes, the downscaling methodology has a significant influence on the impact results (21% and 26% for the Grote Nete and Dijle catchment, respectively).

The ANOVA-technique does not only allow to quantify the main effects as discussed above, but also the interaction effects between the various sources of uncertainty. The interaction effects were found missing in most of the selected climate change analyses from Table 1. Interestingly, Fig. 3 indicates that the interaction effects between GCM runs and SDMs account for about 20% of the total uncertainty range (Dijle catchment) to 25% (Grote Nete catchment) for peak-flow changes, and only 4–8% for low-flow changes. Interactions between hydrological model structure and model parameters are found to be negligible. The remaining second order, and higher order effects, add up to 10–20%.

### 5 Discussion and Conclusions

It is shown that, for our case studies, hydrological extremes will become more extreme with climate change: dry becomes drier and wet becomes wetter; the sign of the changes is quite clear. This is in agreement with previous studies on the same and neighboring catchments [Tavakoli et al. 2014; Vansteenkiste et al. 2014b; Dams et al. 2015]. Furthermore, Hundecha et al. [2016] stated that, for European catchments with a rainfall dominated flood regime (opposed to spring/summer snowmelt floods), extreme flow indices would increase in the future. Lobanova et al. [2018] concluded that peak-flows in winter periods would increase across Central, Northern and Eastern Europe, whereas, in general, no significant changes were



**Fig. 3** Contributions of various sources on the total spread of low-flow (left) and peak-flow impact results (right), for the two case studies, indicate relative importance of climatological and hydrological modelling

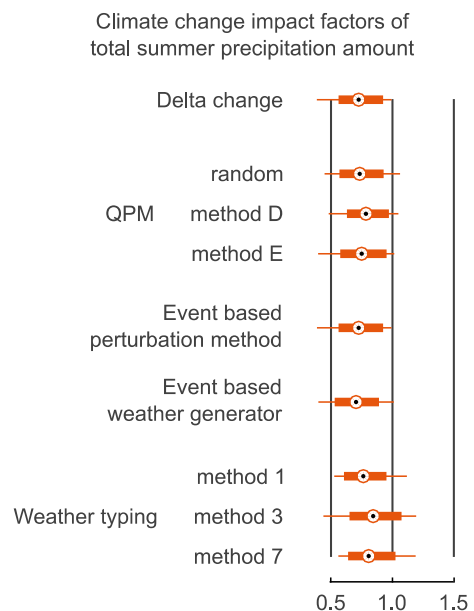
projected for low-flow periods. Gosling et al. [2017] found, for the Rhine river basin, a projected decrease in low-flows, and increase in peak-flows; moreover, the projected changes became stronger with an increased signal of global warming, confirming results from present study. For other river basins in the world, however, the sign of change was unclear. Meaurio et al. [2017] also found a significant increase in durations (days) of low-flows. On the contrary, Seiller et al. [2017] argued that in a multi-model ensemble, the direction of projected change might not be clear.

Despite the clear signal related to the direction of change for our case studies, there is a high uncertainty concerning the magnitude of these changes.

Low-flow conditions in our case studies are generally governed by total seasonal precipitation. We concluded that the contribution of SDMs to the uncertainty in projected low-flow extremes, is rather limited. This can be explained by the total summer precipitation amounts for the various downscaling methodologies, which remains fairly constant (Fig. 4): median values for RCP 8.5 GCM runs vary between 0.70 for the event based weather generator to 0.84 for weather typing method 3 (Fig. 4). This is confirmed by Vidal et al. [2016], who reported only a minor contribution of SDMs on low-flow impacts for two selected catchments in France. In contrast, Chen et al. [2011] found, based on the older CMIP3 GCMs and for a Canadian watershed, a significant contribution of SDMs on the low flow impacts. The ensemble size for SDMs in both studies was, however, rather limited.

On the (high) contribution of uncertainty for future low-flow extremes related to the hydrological models, our findings confirm earlier research on the same case studies [Vansteenkiste et al. 2014b]. This earlier research, however, considered a smaller ensemble of lumped conceptual hydrological models and older generation GCMs. With only one downscaling methodology applied, they concluded that “[...] model choice as well as calibration strategy hence have a critical impact on low flows, more than on peak flows”. Similar conclusions were found by Velázquez et al. [2013], for one Canadian and one German catchment and an ensemble of four hydrological models. Another recent model-inter-comparison study compared 11 hydrological

**Fig. 4** Climate change impact on summer precipitation amount for selected GCM runs with RCP 8.5 ( $n = 30$ )



models for the Meuse river and its sub-catchments [de Boer-Euser et al. 2017]. It was found that, for current conditions, the behavior of the different hydrological models was very different under drier conditions, even though the overall performance of the models was similar.

Peak-flow conditions in our case studies are generally due to extreme precipitation events in winter periods. The largest fraction of the uncertainty in peak-flow impact is explained by the various downscaling methods. Indeed, it is known that different types of downscaling methodologies might result in significantly different quantifications of extreme peak-flows [Chen et al. 2011; Teng et al. 2011; Tavakoli et al. 2014; Hundecha et al. 2016; Meaurio et al. 2017].

Interaction effects between the various levels in the impact modelling chain (2nd and higher order), although often neglected in most climate change impact studies, are found to be responsible for 24–38% of the total uncertainty on future extreme river flows. Here, we acknowledge the importance of these interaction effects, however, we do not dig deeper into the underlying reasons of these interaction terms.

From the literature review (Table 1), one could conclude most studies focus on only one uncertainty source in their impact analysis. With an very large multi-model ensemble, we were able to show that biased results might be obtained in doing so. When considering peak-flow extremes, one should focus on climate modelling and downscaling uncertainty, and it seems somehow justifiable to limit the ensemble of hydrological models; on the contrary, neglecting hydrological model structure uncertainty in low-flow predictions will undoubtedly lead to biased results.

In most climate change impact analyses to date, the various GCMs are considered independent. And, as such, the impact factors resulting from these GCMs are considered equiprobable. However, with the growing number of climate models that become available through public databases, this assumption has been challenged by several studies [Knutti 2010; Pennell and Reichler 2010; Knutti et al. 2013; Sunyer et al. 2013; Hosseinzadehtalaei et al. 2017b]. They all concluded that there is a significant interdependency among GCMs (and RCMs) and only a reduced number of GCMs should be used for impact analyses. Failure to take into account this interdependency would lead to biased results, conservative uncertainty estimations and overly confident predictions. In order to solve this issue of climate model dependency, Knutti et al. [2017] recently proposed a climate model projection weighting scheme. In the present study, impact factors for low-flow and peak-flow extremes were analyzed per RCP (Fig. 2). For each RCP, all GCM runs were considered equiprobable and thus violating the above interdependency issue. We acknowledge this limitation; results hence should be interpreted with caution. When inter-dependency would be taken into account, we expect that (1) the range of results in Fig. 2 would not increase – it might decrease; and (2) the uncertainty contributions of the various factors might shift, most probably towards higher contribution of the GCMs. Moreover, recent research has shown that weighting GCMs would only have a limited impact on hydrological variables [Chen et al. 2017; Das et al. 2018]. Thus, it is expected that the main conclusions of the study (dry gets drier and wet gets wetter, plus the importance of hydrological model structure and downscaling methodology in the uncertainty of impact results) would not change by including the inter-dependency among GCMs.

Also, we acknowledge the debate related to the choice of calculation formula for potential evapotranspiration and its influence on hydrological projections in a climate change context [Kingston et al. 2009; Milly and Dunne 2016, 2017; Seiller and Anctil 2016; Hosseinzadehtalaei et al. 2017a; Paparrizos et al. 2017; Seong et al. 2017]. This source of uncertainty has not been investigated in the present study. In order to add the uncertainty related to potential evapotranspiration calculation formulae, historical evapotranspiration needs to be re-calculated, as well as the future projections. On top of that, each hydrological model should be calibrated separately for each ETo calculation formula, making the problem a lot more complex.

**Acknowledgements** The second author obtained a scholarship from the Fund for the Scientific Research (FWO) - Flanders. This financial support is gratefully acknowledged. We, also, acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modelling groups for producing and making available their model output. Finally, Hossein Tabari is acknowledged for calculating the evapotranspiration for the climate model data. This was completed in a previous study.

The climate model data is available through the website of the Earth System Grid Federation, <https://esgf.llnl.gov/>. The observations at Uccle were made available by the Royal Meteorological Institute of Belgium.

The ERA40 re-analysis data set is available through the public datasets web interface of ECMWF (<http://apps.ecmwf.int/datasets>).

**Author Contributions** JD, EV and PW worked on the conceptualization of the research and jointly developed the methodology. JD mainly worked on the hydrological part of the analysis; EV worked on the climatological part. JD prepared the visualisation and the initial draft, which was critically reviewed and revised by EV and PW.

## Compliance with Ethical Standards

**Conflict of Interest** The authors confirm they have no conflicts of interest to declare.

## References

- Alfieri L, Feyen L, Dottori F, Bianchi A (2015a) Ensemble flood risk assessment in Europe under high end climate scenarios. *Glob Environ Chang* 35:199–212. <https://doi.org/10.1016/j.gloenvcha.2015.09.004>
- Alfieri L, Burek P, Feyen L, Forzieri G (2015b) Global warming increases the frequency of river floods in Europe. *Hydrol Earth Syst Sci* 19:2247–2260. <https://doi.org/10.5194/hess-19-2247-2015>
- Andréassian V, Le Moine N, Perrin C, Ramos MH, Oudin L, Mathevet T, Lerat J, Berthet L (2012) All that glitters is not gold: the case of calibrating hydrological models. *Hydrol Process* 26(14):2206–2210. <https://doi.org/10.1002/hyp.9264>
- Apip TS, Tachikawa Y, Takara K (2012) Spatial lumping of a distributed rainfall-sediment-runoff model and its effective lumping scale. *Hydrol Process* 26(6):855–871. <https://doi.org/10.1002/hyp.8300>
- Baguis P, Roulin E, Willems P, Ntegeka V (2010) Climate change scenarios for precipitation and potential evapotranspiration over central Belgium. *Theor Appl Climatol* 99(3–4):273–286. <https://doi.org/10.1007/s00704-009-0146-5>
- Beven K (1989) Changing ideas in hydrology - the case of physically-based models. *J Hydrol* 105(1–2):157–172. [https://doi.org/10.1016/0022-1694\(89\)90101-7](https://doi.org/10.1016/0022-1694(89)90101-7)
- Beven K (1993) Prophecy, reality and uncertainty in distributed hydrological modelling. *Adv Water Resour* 16: 41–51
- Beven K (2006) A manifesto for the equifinality thesis. *J Hydrol* 320(1–2):18–36. <https://doi.org/10.1016/j.jhydrol.2005.07.007>
- Breuer L et al (2009) Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: model intercomparison with current land use. *Adv Water Resour* 32(2):129–146. <https://doi.org/10.1016/j.advwatres.2008.10.003>
- Bultot F, Coppens A, Dupriez G (1983) Estimation de l'évapotranspiration potentielle en Belgique. Publications/publicaties série/serie A No 112 Royal Meteorol Inst Belg
- Chamorro A, Kraft P, Pauer G, Exbrayat J-F, Breuer L (2017) Effect of (quasi-) optimum model parameter sets and model characteristics on future discharge projection of two basins from Europe and Asia. *Clim Chang* 142:559–573. <https://doi.org/10.1007/s10584-017-1974-4>
- Chen J, Brissette FP, Poulin A, Leconte R (2011) Overall uncertainty study of the hydrological impacts of climate change for a Canadian watershed. *Water Resour Res* 47:1–16. <https://doi.org/10.1029/2011WR010602>
- Chen J, Brissette FP, Lucas-Picher P, Caya D (2017) Impacts of weighting climate models for hydro-meteorological climate change studies. *J Hydrol* 549:534–546. <https://doi.org/10.1016/j.jhydrol.2017.04.025>
- Dams J, Nossent J, Senbeta TB, Willems P, Batelaan O (2015) Multi-model approach to assess the impact of climate change on runoff. *J Hydrol* 529:1601–1616. <https://doi.org/10.1016/j.jhydrol.2015.08.023>

- Das T, Bárdossy A, Zehe E, He Y (2008) Comparison of conceptual model performance using different representations of spatial variability. *J Hydrol* 356(1–2):106–118. <https://doi.org/10.1016/j.jhydrol.2008.04.008>
- Das J, Treesa A, Umamahesh NV (2018) Modelling impacts of climate change on a river basin: analysis of uncertainty using REA and possibilistic approach. *Water Resour Manag* 32:4833–4852. <https://doi.org/10.1007/s11269-018-2046-x>
- de Boer-Euser T et al (2017) Looking beyond general metrics for model comparison - lessons from an international model intercomparison study. *Hydrol Earth Syst Sci* 21(1):423–440. <https://doi.org/10.5194/hess-21-423-2017>
- De Niel J, Van Uytven E, Willems P (2018) On the correlation between precipitation and potential evapotranspiration climate change signals for impact analyses on water resources. *Hydrol Sci J* 64(4):420–433. <https://doi.org/10.1080/02626667.2019.1587615>
- DHI (2009) Mike 11 software manual
- Ghavidelarf S, Alvankar SR, Razmkhah A (2011) Comparison of the lumped and quasi-distributed Clark runoff models in simulating flood hydrographs on a semi-arid watershed. *Water Resour Manag* 25(6):1775–1790. <https://doi.org/10.1007/s11269-011-9774-5>
- Giuntoli I, Vidal J, Prudhomme C, Hannah DM (2015) Future hydrological extremes : the uncertainty from multiple global climate and global hydrological models, *earth Syst. Dynam.* 267–285. <https://doi.org/10.5194/esd-6-267-2015>
- Gosling SN et al (2017) A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1 °C , 2 °C and 3 °C. *Clim Chang* 141:577–595. <https://doi.org/10.1007/s10584-016-1773-3>
- Gupta VK, Sorooshian S (1983) Uniqueness and observability of conceptual rainfall-runoff model parameters: the percolation process examined. *Water Resour Res* 19(1):269–276. <https://doi.org/10.1029/WR019i001p00269>
- Hosseinzadehtalaei P, Tabari H, Willems P (2017a) Quantification of uncertainty in reference evapotranspiration climate change signals in Belgium. *Hydrol Res* 48(3). <https://doi.org/10.2166/nh.2016.243>
- Hosseinzadehtalaei P, Tabari H, Willems P (2017b) Uncertainty assessment for climate change impact on intense precipitation : how many model runs do we need? *Int J Climatol* 37(April):1105–1117. <https://doi.org/10.1002/joc.5069>
- Hundecha Y et al (2016) Inter-comparison of statistical downscaling methods for projection of extreme flow indices across Europe. *J Hydrol* 541:1273–1286. <https://doi.org/10.1016/j.jhydrol.2016.08.033>
- IPCC (2013), Summary for Policymakers, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
- Islam SU, Dery SJ, Werner AT (2017) Future climate change impacts on snow and water resources of the Fraser River basin, British Columbia. *J Hydrometeorol* 18:473–496. <https://doi.org/10.1175/JHM-D-16-0012.1>
- Jakeman AJ, Hornberger GM (1993) How much complexity is warranted in a rainfall-runoff model. *Water Resour Res* 29(8):2637–2649. <https://doi.org/10.1029/93WR00877>
- Kingston DG, Todd MC, Taylor RG, Thompson JR, Arnell NW (2009) Uncertainty in the estimation of potential evapotranspiration under climate change. *Geophys Res Lett* 36:3–8. <https://doi.org/10.1029/2009GL040267>
- Knutti R (2010) The end of model democracy? *Clim Chang* 102(3):395–404. <https://doi.org/10.1007/s10584-010-9800-2>
- Knutti R, Masson D, Gettelman A (2013) Climate model genealogy : generation CMIP5 and how we got there. *Geophys Res Abstr* 40(March):1194–1199. <https://doi.org/10.1002/grl.50256>
- Knutti R, Sedláček J, Sanderson BM, Lorenz R, Fischer EM, Eyring V (2017) A climate model projection weighting scheme accounting for performance and interdependence. *Geophys Res Lett* 44(4):1909–1918. <https://doi.org/10.1002/2016GL072012>
- Lindström G, Johansson B, Persson M, Gardelin M, Bergström S (1997) Development and test of the distributed HBV-96 hydrological model. *J Hydrol* 201(1–4):272–288. [https://doi.org/10.1016/S0022-1694\(97\)00041-3](https://doi.org/10.1016/S0022-1694(97)00041-3)
- Liu T, Willems P, Pan XL, Bao AM, Chen X, Veroustraete F, Dong QH (2011) Climate change impact on water resource extremes in a headwater region of the Tarim basin in China. *Hydrol Earth Syst Sci* 15(11):3511–3527. <https://doi.org/10.5194/hess-15-3511-2011>
- Liu J, Luo M, Liu T, Bao A, De Maeyer P, Feng X (2017) Local climate change and the impacts on hydrological processes in an arid Alpine catchment in Karakoram. *Water* 9(344):1–17. <https://doi.org/10.3390/w9050344>
- Lobanova A et al (2018) Hydrological impacts of moderate and high-end climate change across European river basins. *J Hydrol Reg Stud* 18:15–30. <https://doi.org/10.1016/j.ejrh.2018.05.003>
- Lobligeois F, Andréassian V, Perrin C, Tabary P, Loumagne C (2014) When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrol Earth Syst Sci* 18(2):575–594. <https://doi.org/10.5194/hess-18-575-2014>

- Maraun D et al (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. *Rev Geophys* 48(3):RG3003. <https://doi.org/10.1029/2009RG000314>
- Meaurio M, Zabaleta A, Boithias L, Epelde AM, Sauvage S, Sánchez-Pérez J-M, Srinivasan R, Antiguada I (2017) Assessing the hydrological response from an ensemble of CMIP5 climate projections in the transition zone of the Atlantic region ( Bay of Biscay ). *J Hydrol* 548:46–62. <https://doi.org/10.1016/j.jhydrol.2017.02.029>
- Meresa HK, Romanowicz RJ (2017) The critical role of uncertainty in projections of hydrological extremes. *Hydrol Earth Syst Sci* 21(8):4245–4258. <https://doi.org/10.5194/hess-21-4245-2017>
- Milly PCD, Dunne KA (2016) Potential evapotranspiration and continental drying. *Nat Clim Chang* 6(10):946–949. <https://doi.org/10.1038/nclimate3046>
- Milly PCD, Dunne KA (2017) A hydrologic drying Bias in water-resource impact analyses of anthropogenic climate change. *J Am Water Resour Assoc* 53(4):822–838. <https://doi.org/10.1111/1752-1688.12538>
- Moore RJ (1985) The probability-distributed principle and runoff production at point and basin scales. *Hydrol Sci J* 30(2):273–297. <https://doi.org/10.1080/02626668509490989>
- Moore RJ (2007) The PDM rainfall-runoff model. *Hydrol Earth Syst Sci* 11(1):483–499. <https://doi.org/10.5194/hess-11-483-2007>
- Najafi R, Hessami Kermani MR (2017) Uncertainty modeling of statistical downscaling to assess climate change impacts on temperature and precipitation. *Water Resour Manag* 31(6):1843–1858. <https://doi.org/10.1007/s11269-017-1615-8>
- Nielsen SA, Hansen E (1973) Numerical simulation of the rainfall-runoff process on a daily basis. *Nord Hydrol* 4:171–190
- Ntegeka V, Baguis P, Roulin E, Willems P (2014) Developing tailored climate change scenarios for hydrological impact assessments. *J Hydrol* 508:307–321. <https://doi.org/10.1016/j.jhydrol.2013.11.001>
- Pandey BK, Gosain AK, Paul G, Khare D (2017) Climate change impact assessment on hydrology of a small watershed using semi-distributed model. *Appl Water Sci* 7(4):2029–2041. <https://doi.org/10.1007/s13201-016-0383-6>
- Pandey BK, Deepak K, Kawasaki A, Mishra PK (2019) Climate change impact assessment on blue and green water by coupling of representative CMIP5 climate models with physical based hydrological model. *Water Resour Manag* 33(1):141–158. <https://doi.org/10.1007/s11269-018-2093-3>
- Paparrizos S, Maris F, Matzarakis A (2017) Sensitivity analysis and comparison of various potential evapotranspiration formulae for selected Greek areas with different climate conditions. *Theor Appl Climatol* 745–759. <https://doi.org/10.1007/s00704-015-1728-z>
- Pennell C, Reichler T (2010) On the effective number of climate models. *J Clim* 24:2358–2367. <https://doi.org/10.1175/2010JCLI3814.1>
- Perrin C, Michel C, Andréassian V (2003) Improvement of a parsimonious model for streamflow simulation. *J Hydrol* 279(1–4):275–289. [https://doi.org/10.1016/S0022-1694\(03\)00225-7](https://doi.org/10.1016/S0022-1694(03)00225-7)
- Seiller G, Anctil F (2016) How do potential evapotranspiration formulas influence hydrological projections ? *Hydrol Sci J* 61(12):2249–2266. <https://doi.org/10.1080/02626667.2015.1100302>
- Seiller G, Roy R, Anctil F (2017) Influence of three common calibration metrics on the diagnosis of climate change impacts on water resources. *J Hydrol* 547:280–295. <https://doi.org/10.1016/j.jhydrol.2017.02.004>
- Seong C, Sridhar V, Billah MM (2017) Implications of potential evapotranspiration methods for streamflow estimations under changing climatic conditions. *Int J Climatol*. <https://doi.org/10.1002/joc.5218>
- Sivakumar B (2008) Dominant processes concept, model simplification and classification framework in catchment hydrology. *Stoch Environ Res Risk Assess* 22(6):737–748. <https://doi.org/10.1007/s00477-007-0183-5>
- Sorup HJD, Georgiadis S, Gregersen IB, Arnbjerg-Nielsen K (2017) Formulating and testing a method for perturbing precipitation time series to reflect anticipated climatic changes. *Hydrol Earth Syst Sci* 21(1):345–355. <https://doi.org/10.5194/hess-21-345-2017>
- Sunyer MA, Madsen H, Rosbjerg D, Arnbjerg-Nielsen K, Sunyer MA, Madsen H, Rosbjerg D, Arnbjerg-Nielsen K (2013) Regional interdependency of precipitation indices across Denmark in two ensembles of high-resolution RCMs. *J Clim* 26(20):7912–7928. <https://doi.org/10.1175/JCLI-D-12-00707.1>
- Sunyer MA et al (2015) Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe. *Hydrol Earth Syst Sci* 19(4):1827–1847. <https://doi.org/10.5194/hess-19-1827-2015>
- Tabari H, Taye MT, Willems P (2015) Water availability change in Central Belgium for the late 21st century. *Glob Planet Change* 131(0):115–123. <https://doi.org/10.1016/j.gloplacha.2015.05.012>
- Tabari H, De Troch R, Giot O, Hamdi R, Termonia P, Saeed S (2016) Local impact analysis of climate change on precipitation extremes : are high-resolution climate models needed for realistic simulations ? *Hydrol Earth Syst Sci* 3843–3857. <https://doi.org/10.5194/hess-20-3843-2016>



- Tavakoli M, De Smedt F, Vansteenkiste T, Willems P (2014) Impact of climate change and urban development on extreme flows in the Grote Nete watershed, Belgium, *Nat. Hazards* 71(3):2127–2142. <https://doi.org/10.1007/s11069-013-1001-7>
- Teklesadik AD, Alemayehu T, Van Griensven A (2017) Inter-model comparison of hydrological impacts of climate change on the upper Blue Nile basin using ensemble of hydrological models and global climate models. *Clim Chang* 141:517–532. <https://doi.org/10.1007/s10584-017-1913-4>
- Teng J, Vaze J, Chiew FHS, Wang B, Perraud J-M (2011) Estimating the relative uncertainties sourced from GCMs and hydrological models in modeling climate change impact on runoff. *J Hydrometeorol* 13:122–139. <https://doi.org/10.1175/JHM-D-11-058.1>
- Thorndahl S, Andersen AK, Larsen AB (2017) Event-based stochastic point rainfall resampling for statistical replication and climate projection of historical rainfall series. *Hydrol Earth Syst Sci* 21(9):4433–4448. <https://doi.org/10.5194/hess-21-4433-2017>
- Uhlenbrook S, Seibert J, Leibundgut C, Rodhe A (1999) Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrol Sci J* 44(5):779–797. <https://doi.org/10.1080/02626669909492273>
- Uppala SM et al (2005) The ERA-40 re-analysis. *Q J R Meteorol Soc* 131(612):2961–3012. <https://doi.org/10.1256/qj.04.176>
- Van Steenbergen N, Willems P (2012) Method for testing the accuracy of rainfall-runoff models in predicting peak flow changes due to rainfall changes, in a climate changing context. *J Hydrol* 414–415:425–434. <https://doi.org/10.1016/j.jhydrol.2011.11.017>
- Van Uytven E, Willems P (2018) Greenhouse gas scenario sensitivity and uncertainties in precipitation projections for Central Belgium. *J Hydrol* 558:9–19. <https://doi.org/10.1016/j.jhydrol.2018.01.018>
- Van Uytven, E., J. De Niel, J., and P. Willems (2019), Uncovering the shortcomings of a weather typing based statistical downscaling method, *Hydrol. Earth Syst. Sci. Discuss.*, <https://doi.org/10.5194/hess-2019-40>, 2019
- Vansteenkiste T, Tavakoli M, Ntegeka V, Willems P, De Smedt F, Batelaan O (2012) Climate change impact on river flows and catchment hydrology: a comparison of two spatially distributed models. *Hydrol Process* 3662(August 2012):3649–3662. <https://doi.org/10.1002/hyp.9480>
- Vansteenkiste T, Tavakoli M, Van Steenbergen N, De Smedt F, Batelaan O, Pereira F, Willems P (2014a) Intercomparison of five lumped and distributed models for catchment runoff and extreme flow simulation. *J Hydrol* 511:335–349. <https://doi.org/10.1016/j.jhydrol.2014.01.050>
- Vansteenkiste T, Tavakoli M, Ntegeka V, De Smedt F, Batelaan O, Pereira F, Willems P (2014b) Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections. *J Hydrol* 519:743–755. <https://doi.org/10.1016/j.jhydrol.2014.07.062>
- Velázquez JA, Schmid J, Ricard S, Muerth MJ, Gauvin St-Denis B, Minville M, Chaumont D, Caya D, Ludwig R, Turcotte R (2013) An ensemble approach to assess hydrological models' contribution to uncertainties in the analysis of climate change impact on water resources. *Hydrol Earth Syst Sci* 17:565–578. <https://doi.org/10.5194/hess-17-565-2013>
- Vidal JP, Hingray B, Magand C, Sauquet E, Ducharme A (2016) Hierarchy of climate and hydrological uncertainties in transient low-flow projections. *Hydrol Earth Syst Sci* 20(9):3651–3672. <https://doi.org/10.5194/hess-20-3651-2016>
- Viney NR et al (2009) Assessing the impact of land use change on hydrology by ensemble modelling (LUCHEM) II: ensemble combinations and predictions. *Adv Water Resour* 32(2):147–158. <https://doi.org/10.1016/j.advwatres.2008.05.006>
- Vrugt JA, Gupta HV, Bastidas LA, Bouten W, Sorooshian S (2003) Effective and efficient algorithm for multiobjective optimization of hydrologic models. *Water Resour Res* 39(8):1–19. <https://doi.org/10.1029/2002WR001746>
- Willems P (2009) A time series tool to support the multi-criteria performance evaluation of rainfall-runoff models. *Environ Model Softw* 24(3):311–321. <https://doi.org/10.1016/j.envsoft.2008.09.005>
- Willems P (2014) Parsimonious rainfall-runoff model construction supported by time series processing and validation of hydrological extremes - part 1: step-wise model-structure identification and calibration approach. *J Hydrol* 510:578–590. <https://doi.org/10.1016/j.jhydrol.2014.01.017>
- Willems P, Vrac M (2011) Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change. *J Hydrol* 402(3–4):193–205. <https://doi.org/10.1016/j.jhydrol.2011.02.030>
- Xu C-Y (1999) Climate change and hydrologic models: a review of existing gaps and recent research developments, *water Resour. Manag.* 13(5):369–382. <https://doi.org/10.1023/A:1008190900459>