Journal of Hydrology 550 (2017) 201–217

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Evaluating the robustness of conceptual rainfall-runoff models under climate variability in northern Tunisia

H. Dakhlaoui [a,b,*], D. Ruelland [c,*], Y. Tramblay [d], Z. Bargaoui [a]

[a] LMHE, Ecole Nationale d'Ingénieurs de Tunis, Université Tunis El Manar, BP 37, 1002 Tunis le Belvedère, Tunisia
[b] Ecole Nationale d'Architecture et d'Urbanisme, Université de Carthage, Rue El Quods, 2026 Sidi Bou Said, Tunisia
[c] CNRS, UMR HydroSciences Montpellier, Place E. Bataillon, 34395 Montpellier Cedex 5, France
[d] IRD, UMR HydroSciences Montpellier, Place E. Bataillon, 34395 Montpellier Cedex 5, France

ARTICLE INFO

Article history:
Received 9 January 2017
Received in revised form 24 March 2017
Accepted 13 April 2017
Available online 18 April 2017
This manuscript was handled by A. Bardossy, Editor-in-Chief, with the assistance of Axel Bronstert, Associate Editor

Keywords:
Rainfall-runoff modelling
Hydro-climate variability
Model robustness
Uncertainty
Tunisia

ABSTRACT

To evaluate the impact of climate change on water resources at the catchment scale, not only future projections of climate are necessary but also robust rainfall-runoff models that must be fairly reliable under changing climate conditions. The aim of this study was thus to assess the robustness of three conceptual rainfall-runoff models (GR4j, HBV and IHACRES) on five basins in northern Tunisia under long-term climate variability, in the light of available future climate scenarios for this region. The robustness of the models was evaluated using a differential split sample test based on a climate classification of the observation period that simultaneously accounted for precipitation and temperature conditions. The study catchments include the main hydrographical basins in northern Tunisia, which produce most of the surface water resources in the country. A 30-year period (1970–2000) was used to capture a wide range of hydro-climatic conditions. The calibration was based on the Kling-Gupta Efficiency (KGE) criterion, while model transferability was evaluated based on the Nash-Sutcliffe efficiency criterion and volume error. The three hydrological models were shown to behave similarly under climate variability. The models simulated the runoff pattern better when transferred to wetter and colder conditions than to drier and warmer ones. It was shown that their robustness became unacceptable when climate conditions involved a decrease of more than 25% in annual precipitation and an increase of more than +1.75 °C in annual mean temperatures. The reduction in model robustness may be partly due to the climate dependence of some parameters. When compared to precipitation and temperature projections in the region, the limits of transferability obtained in this study are generally respected for short and middle term. For long term projections under the most pessimistic emission gas scenarios, the limits of transferability are generally not respected, which may hamper the use of conceptual models for hydrological projections in northern Tunisia.

1. Introduction

Several studies have shown that climate change will have severe impacts on available water resources worldwide (e.g. Alcamo et al., 2007; Kundzewicz et al., 2007). In this context, adaptation to climate change is indispensable, which calls for a detailed understanding of local climate change projections and their impacts on water resources. Rainfall-runoff models combined with regional climate change scenarios are widely used to assess the impact of climate change at catchment scale (e.g. Bastola et al., 2011; Chen et al., 2011; Ruelland et al., 2012; 2015). This approach requires realistic future climate projections and robust hydrological models that are fairly reliable under changing climate conditions. Although several studies have shown climate projections to be the main source of uncertainty in climate change impact studies (e.g. Wilby and Dessai, 2010), some authors point out that the hydrological models could considerably increase total uncertainty (e.g. Kay et al., 2009; Poulin et al., 2011).

Abbreviations: CD, Cold Dry period; CMD, Catchment Moisture Deficit; CW, Cold Wet period; DSST, Differential Split Simple Test; GSST, General Split Simple Test; HD, Hot Dry period; HW, Hot Wet period; KGE, Kling-Gupta Efficiency; NSE, Nash Sutcliffe Efficiency; PET, Potential Evapotranspiration; RRM, Rainfall-Runoff models; SCE, Shuffle Complex Evolution; VE, Volume Error.

* Corresponding authors at: LMHE, Ecole Nationale d'Ingénieurs de Tunis, Université Tunis El Manar, BP 37, 1002 Tunis le Belvedère, Tunisia (H. Dakhlaoui).
E-mail addresses: hammouda.dakhlaoui@laposte.net (H. Dakhlaoui), denis.ruelland@um2.fr (D. Ruelland).

http://dx.doi.org/10.1016/j.jhydrol.2017.04.032
0022-1694/

1.1. The issue of hydrological modelling under non-stationary climate conditions

It can be argued that physically-based models have a greater potential to provide predictions beyond the range of conditions used for calibration, but their robustness under long-term climate variability is difficult to evaluate since they are generally based on a large number of input variables, usually not available over long time spans. This is particularly true in developing countries in which catchments are generally poorly gauged. Moreover, Refsgaard and Knudsen (1996) suggested that there is no immediate justification for using an advanced type of model (i.e. a physically-based one) to represent flows following a significant change in rainfall, and concluded that simple models could justifiably be used to assess the impact of climate change on water resources. In a more recent study, Karlsson et al. (2014) investigated the impact of climate change on hydrology in a Danish catchment using a conceptual model (NAM) and a physically-based model (Mike-SHE). The authors found the two different models to provide very similar efficiency in reproducing runoff at the outlet, even though the efficient reproduction of low flows depended to a great extent on the structural representation of the different flow paths in the two models tested. In another study, Coron et al. (2014) found similar behaviour in a sophisticated conceptual distributed model (CEQUEAU) and two parsimonious lumped models (GR4j and the Mouelhi formula, Mouelhi et al., 2006) with respect to water bias under climate variability. A review of the literature showed that conceptual parsimonious models are still considered to be suitable tools to assess the impact of climate change in developing countries (e.g. Tramblay et al. 2013a; Teng et al. 2011; Wilby and Dessai, 2010), since they can be used even under poor data availability. However, a challenging aspect of hydrological conceptual models is the identifiability of their parameters which need to be estimated through calibration procedures to match the behaviour of the model to that of the catchment. The usual sources of uncertainty in hydrological modelling under stationary conditions (concerning the climate and the physical characteristics of the basin) are linked to the structure of the model, the calibration procedures, and erroneous data used for calibration and validation (e.g. Liu and Gupta, 2007; Brigode et al., 2013). Under non-stationary conditions, such as climate change, an additional source of uncertainty is parameter instability due to possible changes in the physical characteristics of the catchment and in the main processes at play (e.g. Coron et al., 2012; Poulin et al., 2011; Thompson et al., 2013).

1.2. Accounting for parameter uncertainty

Model structure uncertainty is usually assessed by testing different models and quantifying the range of their outputs (e.g. Jiang et al., 2004; Hublart et al., 2015). The assessment and reduction of parameter uncertainty is generally performed using Monte-Carlo procedures including the Glue method (Beven and Binley, 1992; Bastola et al., 2011), Bayesian inference methods (Neuman, 2003; Huard and Mailhot, 2008), evolutionary algorithms such as the Shuffled Complex Evolution Metropolis algorithm (Vrugt et al., 2003) or depth functions (Bárdossy and Singh, 2008). Several methods have been proposed to assess and/or reduce uncertainty due to climate variability: the use of the longest available data series (e.g. Poulin et al., 2011), the selection of appropriate calibration criteria (e.g. Hartmann and Bárdossy, 2005) and the use of the differential split sample tests (DSST, Klemeš, 1986; Vaze et al., 2011; Tramblay et al., 2013a; Ruelland et al., 2015).

The DSST is the standard method used to investigate parameter instability under climate change, which consists in calibration and validation exercises of hydrological models using sub-periods with contrasted climate conditions. The idea behind performing a DSST is that the errors made by extrapolation from given observed climate conditions to different observed conditions could correspond to the errors made when using reference data for calibration and extrapolating to future climatic conditions (Seibert, 2003). It makes it possible to evaluate model transferability (or transposability), which can be defined as the ability of a model to perform with the same level of accuracy under conditions that differ from those used for its calibration (Seiller et al., 2012). Model transferability can be considered as part of model robustness classically defined as the insensitivity of the model parameters to calibration data. Recently, a DSST has been used in a number of studies to evaluate model reliability under climate variability. Precipitation, temperature and potential evapotranspiration (PET) are the main climate variables used to define the contrasted climatic conditions of sub-periods. Generally, these climate variables are used separately to cluster different climate periods. Refsgaard and Knudsen (1996) applied a (dry/wet) DSST to three catchments in Zimbabwe with black, gray and white box hydrological models (NAM, WETBALL and Mike-SHE). These authors reported that all the models proved their ability to simulate the runoff pattern in periods with much reduced rainfall and runoff compared to the calibration period. Seibert (2003) used a DSST (period with the highest peak flows *vs.* period with the lowest peak flows) to evaluate the reliability of the HBV model prediction outside the model calibration conditions in four snow-dominated Swedish basins. He demonstrated that a model calibrated in years with lower runoff peaks will not necessarily provide accurate results when tested on years with higher peak runoff. In the Upper Neckar catchment (Germany), Hartmann and Bárdossy (2005) divided a 30-year observation period into three sub-periods, first in terms of mean annual temperature (warm, normal and cold), and second, in terms of annual precipitation (wet, normal, and dry years). They demonstrated the importance of a good choice of calibration criteria to improve robustness under climate variability, but did not express a clear preference for calibrating a model in specific climate conditions. Wu and Johnston (2007) calibrated the SWAT model in a catchment in northern Michigan using different climatic datasets representing drought versus average rainfall conditions, and evaluated the performance using the same validation period. Although the hydrological model was well calibrated in both periods, the drought-calibrated version provided better validation efficiency. Vaze et al. (2011) used annual precipitation to identify the driest and wettest continuous periods in 61 catchments in southwest Australia. They found that transferring parameters to a drier climate was particularly challenging. They suggested that calibration periods of at least 20 years were needed to prove a model's robustness to climate variability but only if the difference in mean rainfall was between -15% (drier climate) or +20% (wetter climate). Ruelland et al. (2015) evaluated the efficiency of the GR4j model using two contrasted (dry and wet) 10-year sub-periods applied to four western Mediterranean catchments (France, Spain and Morocco) and found that extrapolating the model to drier or wetter conditions was not straightforward.

1.3. Motivations of the present study

The absence of a consensus in the aforementioned literature on using a DSST to assess the impact of the climate conditions during the calibration period on model transferability and on which climate conditions should be preferred to ensure optimal model robustness shows that this question requires further investigation. Moreover, most of these studies clustered the observed period into contrasted climate conditions by using only one climatic variable. Few studies attempted to use more than one climate variable for the purpose of classification. Seiller et al. (2012) used a DSST based