Journal of Hydrology 523 (2015) 147–159

Contents lists available at ScienceDirect

Journal of Hydrology

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

Does model performance improve with complexity? A case study with three hydrological models

Rene Orth [a,*], Maria Staudinger [b], Sonia I. Seneviratne [a], Jan Seibert [b], Massimiliano Zappa [c]

[a] Institute for Atmospheric and Climate Science, ETH Zurich, Universitätstrasse 16, CH-8092 Zurich, Switzerland
[b] Department of Geography, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland
[c] Swiss Federal Institute for Forest, Snow and Landscape Research (WSL), Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland

ARTICLE INFO

Article history:
Received 20 June 2014
Received in revised form 24 November 2014
Accepted 17 January 2015
Available online 29 January 2015
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Emmanouil N. Anagnostou, Associate Editor

Keywords:
Hydrological model comparison
Runoff validation
Soil moisture validation
Simple conceptual model as benchmark

SUMMARY

In recent decades considerable progress has been made in climate model development. Following the massive increase in computational power, models became more sophisticated. At the same time also simple conceptual models have advanced. In this study we validate and compare three hydrological models of different complexity to investigate whether their performance varies accordingly. For this purpose we use runoff and also soil moisture measurements, which allow a truly independent validation, from several sites across Switzerland. The models are calibrated in similar ways with the same runoff data. Our results show that the more complex models HBV and PREVAH outperform the simple water balance model (SWBM) in case of runoff but not for soil moisture. Furthermore the most sophisticated PREVAH model shows an added value compared to the HBV model only in case of soil moisture. Focusing on extreme events we find generally improved performance of the SWBM during drought conditions and degraded agreement with observations during wet extremes. For the more complex models we find the opposite behavior, probably because they were primarily developed for prediction of runoff extremes. As expected given their complexity, HBV and PREVAH have more problems with over-fitting. All models show a tendency towards better performance in lower altitudes as opposed to (pre-) alpine sites. The results vary considerably across the investigated sites. In contrast, the different metrics we consider to estimate the agreement between models and observations lead to similar conclusions, indicating that the performance of the considered models is similar at different time scales as well as for anomalies and long-term means. We conclude that added complexity does not necessarily lead to improved performance of hydrological models, and that performance can vary greatly depending on the considered hydrological variable (e.g. runoff vs. soil moisture) or hydrological conditions (floods vs. droughts).

1. Introduction

In recent decades great progress has been made in the understanding of the functioning of the climate system (IPCC, 2013). Following these scientific advances the quality and performance of climate models has significantly improved. Together with an astonishing enhancement of computational power this has led and is still leading to the development of very sophisticated models that represent the system in great detail through the consideration of numerous involved processes (e.g. Gent, 2011). On the other hand, simple conceptual models have evolved rapidly at the same time (e.g. Budyko, 1974; Donohue et al., 2007; Kirchner, 2009; Koster and Mahanama, 2012). Sometimes it is beneficial to have less complex and less computationally demanding models for instance for first-order analyses, or to run a large number of test cases. Also in the (not uncommon) case of uncertain or poorly resolved input data, simple (lumped) models may compete with complex models (Beven, 1989). Moreover for practical applications such as risk analysis or forecasting, the performance of conceptual models may serve as a benchmark for sophisticated models to determine their added value and hence their suitability in a particular case (Gurtz et al., 2003; Perrin et al., 2006; Kobierska et al., 2013), even if any judgment on model performance necessarily depends on the evaluation measure (Andreassian, 2009).

Mostly conceptual models consider specific parts of the climate system and make use of first-order approximations to represent the most important processes. For example in hydrology there is a long history of modeling the response of runoff to a given precipitation event in a given catchment using both simple and sophisti-

* Corresponding author.
E-mail address: rene.orth@env.ethz.ch (R. Orth).

http://dx.doi.org/10.1016/j.jhydrol.2015.01.044
0022-1694/© 2015 The Authors. Published by Elsevier B.V.

cated approaches. The sophisticated models with their many parameters can closely match reproduce measurements over the calibration period, but they tend to suffer from over-parametrization over the validation period (Beven, 1989). In contrast, simple models with their few parameters cannot capture runoff as well during the calibration phase but show a consistent performance in the validation period (Perrin et al., 2001; Holländer, 2009). In other words, a model needs to be both reliable and robust, therefore it is necessary to incorporate the best of both worlds and to develop models with simple structure but adequate complexity.

In this study we compare and evaluate three state-of-the-art hydrological models of different complexity in a collaborative effort between three research groups. This case study will help to determine if higher complexity (necessarily) leads to better model performance, and therefore an improved representation of observed hydrological processes.

Previous studies have mostly focused on various aspects of runoff modeling (e.g. Beven, 1989; Kirchner, 2009; Bosshard et al., 2013; Kobierska et al., 2013). As the runoff data is used for both the model training and its validation, it is common to use different time periods for calibration and validation of the models. We follow a similar methodology, but by using soil moisture measurements we furthermore analyze the models' soil moisture dynamics (Schlosser et al., 2000; Gurtz et al., 2003; Orth and Seneviratne, 2013b). This allows us to perform the validation for an independent variable which is not used for model calibration.

To get a better impression of the models' behaviors under various conditions we consider eight well-observed, near-natural catchments (i.e. with little or no human influence) in different climate regimes, located across Switzerland. Moreover we evaluate the abilities of the models to capture extreme conditions, considering both dry and wet extremes (Zappa and Kan, 2007; Orth and Seneviratne, 2013a). This integrated analysis will allow us to identify particular strengths and weaknesses of each model, which should be considered when selecting a model for a specific application.

2. Models and data

In this section we provide a brief description of the three hydrological models compared in this study (see overview in Table 1). After a description of the common soil moisture routine we present the individual models ordered with respect to their complexity, such that the most simple model is described first and the most complex model is presented last. Furthermore, we introduce the observational data used to calibrate, run and validate the models.

2.1. Common soil moisture routine

All three models applied in this study use a similar approach to compute soil moisture dynamics which is based on the water balance equation:

$$w_{n+\Delta t} = w_n + (P_n + S_n - E_n - Q_n)\Delta t \tag{1}$$

where w_n denotes soil moisture at the beginning of time step n and P_n, S_n, E_n and Q_n refer to accumulated rainfall, snow melt, evapotranspiration (hereafter referred to as ET) and recharge to groundwater, respectively, during time step n. In this study we apply a time step of $\Delta t = 1\,day$.

In order to calculate soil moisture in Eq. (1), the models use precipitation directly from observations and they estimate snow melt with a degree-day approach. To derive runoff for Eq. (1) all models use an approach introduced by Bergström (1976). In this approach, a fraction of the water input to the soil (rainfall and snow melt, $P_n + S_n$) is added to the soil moisture content. The remaining part of $P_n + S_n$ forms the runoff Q_n, which comprises surface (immediate) and sub-surface (delayed) runoff. The models use different approaches to estimate the conversion of the surface- and sub-surface runoff to streamflow. The partitioning of $P_n + S_n$ is a nonlinear function of the soil moisture content scaled with its maximum value:

$$\frac{Q_n}{P_n + S_n} = \left(\frac{w_n}{c_s}\right)^{\beta} \text{with } \beta \geqslant 0 \tag{2}$$

where c_s denotes the water holding capacity of the soil and β is a shape parameter that determines the sensitivity of (normalized) runoff to (relative) soil moisture. To estimate ET the models follow a similar approach such that normalized ET is a function of relative soil moisture content only. However, the exact formulation of this estimation and the quantity used to normalize ET differs across the models.

Finally the estimated runoff, ET and snow melt accumulated during a particular day are used in Eq. (1) along with observed precipitation from that day to yield soil moisture at the beginning of the next day.

2.2. Simple water balance model

The simple water balance model (SWBM) is a conceptual, lumped model initially proposed by Koster and Mahanama (2012), and subsequently adapted by Orth and Seneviratne (2013b) for application on the daily time scale. Compared to the version of Orth and Seneviratne (2013b), we additionally include further implementations in the SWBM, as described hereafter.

Table 1
Overview of conceptual hydrological models applied in this study.

	SWBM	HBV	PREVAH
Full name	Simple Water Balance Model	Hydrologiska Byråns Vattenbalansavdelning model	PREecipitation-Runoff-EVApotranspiration Hydrological response unit model
Reference	Orth et al., 2013	Bergström, 1995	Viviroli et al., 2009
Spatial structure	lumped	semi-distributed	fully distributed
Spatial resolution	Catchment	Several elevation zones, one for every 100 m altitude difference	200 m × 200 m
Number of vertical layers	2	3	3
Objective function	Nash–Sutcliffe efficiency (Eq. (5))	Nash–Sutcliffe efficiency (Eq. (5))	Combination of (i) Nash–Sutcliffe efficiency (Eq. (5)), (ii) logarithm thereof, and (iii) relative runoff error
Number of calibrated parameters	7	16	12 (+2 for Dischma)
Required forcing variables	Precipitation, (net) radiation, temperature	Precipitation, temperature	Precipitation, temperature, relative humidity, (global) radiation, wind speed, sunshine duration
Snow modeling	Degree-day approach with constant threshold temperature	Degree-day approach	Degree-day approach with correction w.r.t. slope and aspect
Spin-up period	5 years	3 years	10 years