Journal of Hydrology 527 (2015) 858–871

Contents lists available at ScienceDirect

Journal of Hydrology

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

How much can we gain with increasing model complexity with the same model concepts?

Hong Li [a], C.-Y. Xu [a,*], Stein Beldring [b]

[a] Department of Geosciences, University of Oslo, Norway
[b] Norwegian Water Resources and Energy Directorate, Norway

ARTICLE INFO

Article history:
Received 17 January 2015
Received in revised form 20 May 2015
Accepted 23 May 2015
Available online 28 May 2015
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Yu Zhang, Associate Editor

Keywords:
Distributed models
HBV
Internal variables
Lumped model
Model comparison

SUMMARY

The main purpose of this study is to test the hypothesis that, with appropriate structures, increasing model complexity with the same model concepts would lead to an increase in the model efficiency in simulating either runoff or internal variables. Five variants of the Hydrologiska Byråns Vattenbalansavdelning (HBV) model, i.e. a lumped model (LWhole), a semi-distributed model (SBand), a grid-model without routing (GRZero), a grid-model with hillslope routing (GROne), and a grid-model with both hillslope and channel routing (GRTwo) are compared in a cold and mountainous catchment in central southern Norway. The five models are compared with respects to (1) runoff simulation at the catchment outlet and the interior points, and (2) simulations of internal variables, i.e. evapotranspiration, snow water equivalent and groundwater depth. The results show that the models with higher complexity can improve the runoff simulation both at the catchment outlet and the interior points. However, there is no superiority of complex grid-models over simple grid-models in reproducing internal variables in this study.

1. Introduction

The fundamental work of hydrologists is to quantify relationship of precipitation over a catchment area and resulting runoff. In the last three decades, a large amount of hydrological models have been published to describe this relationship (Singh and Woolhiser, 2002). These models are suited to simulate runoff at certain spatial and temporal scales (Praskievicz and Chang, 2009). Basically, the differences among the models come from three aspects (Butts et al., 2004): (1) assumptions about factors influencing hydrology or response functions, (2) numerical solutions, and (3) spatial discretisation, such as sub-catchments or square grids.

There has been an evident trend to develop models with high degree of complexity (Mayr et al., 2013; Perrin et al., 2001). The development of hydrological models has gone through three stages, from input–output black-box models, through lumped conceptual models, to physically-based and/or conceptual distributed models. The physically-based distributed models tend to represent physical processes by partial differential equations at a fine spatial resolution. To develop models with high degree of physical dependence and structural complexity is long-sighted and useful in terms of knowledge gained about hydrological processes from catchment studies. However, it would lead to increasing difficulty in estimating parameters and large parameter uncertainty (Butts et al., 2004).

Hydrologists show high interest in comparing hydrological models of varying complexity. The conclusions are not consistent due to many reasons, for example, the models used in the comparisons, data quality and catchment characteristics, etc. However, if the models differ in the assumptions about factors influencing hydrology or the runoff process, results of comparisons are determined by the basins used.

To focus on the model itself, we assume that the responses of runoff to areal mean precipitation are described by the model concepts, which can be described by a lumped model. The model complexity exists in the way that water balance components are presented. For example, the lumped HBV model formulates all basic concepts of the HBV model. The semi-distributed models with elevation bands or sub-catchments and the grid-distributed models are of a higher degree of model complexity. The routing procedure, such as the Muskingum method also adds an additional degree of complexity, because it links the model elements to each other.

This research examines if increasing degree of model complexity improves the model performance and how much. A review of

* Corresponding author.
E-mail address: chongyu.xu@geo.uio.no (C.-Y. Xu).

http://dx.doi.org/10.1016/j.jhydrol.2015.05.044
0022-1694/© 2015 The Authors. Published by Elsevier B.V.

the scientific literature did not provide a clear guidance on this issue. In 2004, the Distributed Model Intercomparison Project (DMIP) compared distributed and lumped versions of twelve models with radar precipitation data at a 4 km spatial resolution and hourly temporal resolution (Smith et al., 2004). Twenty one events in eight catchments with areas ranging from 65 to 2484 km^2 were selected in the comparative studies. This project results showed that overall, lumped models outperformed distributed models in terms of discharge simulation, and some distributed models showed comparable results to lumped models in many basins (Reed et al., 2004). The results depended on the basin's shape, orientation, soil and climatic characteristics (Butts et al., 2004; Reed et al., 2004).

Different results are reported by other studies. Atkinson et al. (2003) demonstrated that the best simulation of runoff was obtained by a fully distributed model in a small catchment at an hourly time step. Han et al. (2014) and Yan and Zhang (2014) respectively examined effects of watershed subdivision on modelling runoff. They found that with a larger number of sub-catchments, the model gave a better performance and there was a threshold level, beyond which no significant improvements could be obtained with increasing number of sub-catchments. This confirms the results by comparing two watershed subdivision schemes by Varado et al. (2006). Michaud and Sorooshian (1994) compared the lumped and distributed Soil Conservation Service (CSC) models on a semi-arid catchment of 150 km^2 and the results showed that neither of the models could accurately simulate the peak flows or runoff volumes from 24 sever thunderstorms at an one-minute time step. Using the same model, Boyle et al. (2001) found improvements were related to spatial distribution of model input and streamflow routing and no additional benefits could be obtained when the number of watersheds were more than three.

A well-known way for a better model comparison is involving internal points and state variables in model evaluation (Lindström et al., 1997; Uhlenbrook et al., 1999; Vaché and McDonnell, 2006). Alley (1984) was probably the first to notice that models were similar in runoff simulation while substantial differences existed in other variables. Furthermore, Jiang et al. (2007) found that models differed least in discharge among simulations of discharge, actual evapotranspiration and soil moisture. Varado et al. (2006) applied a conceptual model on the Donga catchment (Benin), and found that the model was only good at simulating runoff, but not at reproducing groundwater table. Vaché and McDonnell (2006) showed that, among the models they evaluated, only the most complex model successfully reproduced both discharge dynamics and residence time. The results of their research indicate that the best objective function value in discharge are obtained during calibration, but with sacrifice of other hydrological variables, where the error resides. In order to achieve deliberate conclusions, comparisons including other hydrological measurements, such as evapotranspiration and groundwater should be adopted (Bookhagen and Burbank, 2010; Lindström et al., 1997; Wagener et al., 2001). This multi-variable comparison will not only contribute to an improved understanding of hydrology processes, but also provide guidance for developing hydrological models.

The aim of this paper is to test the hypothesis that, with appropriate way of organising, increasing complexity leads to an increase in model efficiency of simulating either runoff or other internal variables. A conceptual rainfall–runoff model, the HBV model is selected and modified to five model variants of different complexities, i.e. a lumped model (LWhole), a semi-distributed model with ten elevation bands (SBand), a simple grid-model (GRZero), a grid-model with hillslope routing (GROne), and a grid-model with both hillslope routing and river channel routing (GRTwo). The selection of the models is threefold. Firstly, LWhole is the most widely used HBV model variant for scientific and operational purposes globally. Secondly, SBand is currently used in the Norwegian Water Resources and Energy Directorate (NVE), which is responsible for flood forecasting and water resources administration in Norway. Thirdly, three distributed models are included in the sense of having a higher level of physical realism. The spatial variability and how the runoff routes to the catchment outlet are described in the three grid distributed models at various degrees of complexity.

Lindström et al. (1997) made the first attempt to make a distributed version of the HBV model. They used a typical resolution of 40 km^2 of sub-catchments and each sub-catchment was further divided into elevation bands. This modification significantly improved model performance. In the late 1990s, Uhlenbrook et al. (1999) and Krysanova et al. (1999) respectively compared the effects of spatial distribution on runoff simulation. Uhlenbrook et al. (1999) compared three model variants with different number of elevation bands and land use zone and various runoff generation conceptualisation, on a small mountainous catchment of 40 km^2 in south western Germany. They concluded that the models considering more spatial variability were better than the lumped models when separately computing of the upper zone storage for each model unit. Krysanova et al. (1999) applied the semi-distributed HBV model of elevation bands to a large German catchment of 96,000 km^2. The model with sub-catchments enabled better runoff simulation than without sub-catchment division. However, Das et al. (2008) compared four versions of the HBV model and found that semi-distributed and semi-lumped (a lumped model for each sub-catchment) outperformed the distributed (1 km regular grid) and lumped model structures. The authors suspected that the input data did not reflect the actual spatial variability. The study by Wrede et al. (2013) in a Swedish lowland catchment of 2000 km^2 using the distributed model (250 m regular grid) also showed that the quality of input data was a limitation factor for model performance. For modelling other variables, Mayr et al. (2013) reported that involving glacier mass balance in calibration gave a better prediction in the glacier mass but a slightly worsened discharge prediction. Moreover, they revealed that incorrect snow and ice simulations did not necessarily affect the quality of the runoff simulation.

2. Study area and data

2.1. Study area

The Norsfoss catchment is located in upstream of the longest Norwegian river, the Glomma River in central southern Norway (Fig. 1). This catchment covers an area of 18,932 km^2. The mean altitude is 732 m above the mean sea level (m amsl) ranging from 147 to 2170 m amsl, and approximately 26% of the area is above the potential tree level. The mean slope is 6.7° with a range from 0.0° to 73.2°. Climate varies along the river from upper mountain regions in north to lowlands in south. Additionally, the north-western part is characterised by lower temperature, lower precipitation and longer snow-cover period than the lowland area. Annual precipitation is 849 mm/year, and yearly mean air temperature is −0.62 °C based on the period from 1961 to 1990 with 10.68 °C in July and −11.48 °C in January. More than 60% of the catchment area is covered by forest and marsh, and approximately 20% is covered by bedrock. Floods are usually associated with snow melt, heavy rainfall or their combination (L'Abée-Lund et al., 2009).

2.2. Geography data and processing

Data of elevation, slope and land covers are provided by the Norwegian Mapping Authority, which bears nationwide