HYDROLOGICAL PROCESSES
Hydrol. Process. (2016)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1002/hyp.10962

Spatial GR4J conceptualization of the Tamor glaciated alpine catchment in Eastern Nepal: evaluation of GR4JSG against streamflow and MODIS snow extent

Santosh Nepal,[1]* Jie Chen,[2,3] David J. Penton,[2] Luis E. Neumann,[2] Hongxing Zheng[2] and Shahriar Wahid[1]

[1] *International Centre for Integrated Mountain Development, Kathmandu, Nepal*
[2] *CSIRO Land and Water Flagship, Canberra, Australia*
[3] *University of South Australia, Australia*

Abstract:

Snow and glacial melt processes are an important part of the Himalayan water balance. Correct quantification of melt runoff processes is necessary to understand the region's vulnerability to climate change. This paper describes in detail an application of conceptual GR4J hydrological model in the Tamor catchment in Eastern Nepal using typical elevation band and degree-day factor approaches to model Himalayan snow and glacial melt processes. The model aims to provide a simple model that meets most water planning applications. The paper contributes a model conceptualization (GR4JSG) that enables coarse evaluation of modelled snow extents against remotely sensed Moderate Resolution Imaging Spectroradiometer snow extent. Novel aspects include the glacial store in GR4JSG and examination of how the parameters controlling snow and glacial stores correlate with existing parameters of GR4J. The model is calibrated using a Bayesian Monte Carlo Markov Chain method against observed streamflow for one glaciated catchment with reliable data. Evaluation of the modelled streamflow with observed streamflow gave Nash Sutcliffe Efficiency of 0.88 and Percent Bias of $<4\%$. Comparison of the modelled snow extents with Moderate Resolution Imaging Spectroradiometer gave R^2 of 0.46, with calibration against streamflow only. The contribution of melt runoff to total discharge from the catchment is 14–16% across different experiments. The model is highly sensitive to rainfall and temperature data, which suffer from known problems and biases, for example because of stations being located predominantly in valleys and at lower elevations. Testing of the model in other Himalayan catchments may reveal additional limitations. © 2016 The Authors. *Hydrological Processes* published by John Wiley & Sons Ltd.

KEY WORDS hydrological modelling; glacier catchment processes; Himalayan water budgets; snow and glacier melt; Tamor catchment

Received 1 January 2016; Accepted 14 July 2016

INTRODUCTION

The Himalayan region has been described as the 'Third Pole' because the amount of snow and ice stored is lesser only to the Polar Regions. In the context of global climate change, understanding snow and ice processes is vital to estimating the impact of climate change on the hydrological regime (Barnett *et al*., 2005; Immerzeel *et al*., 2010; Nepal *et al*., 2014). However, the likely impact on water availability in downstream areas is less clear because of: (1) questions of the appropriate model structure to capture Himalayan dynamics, (2) uncertainty in precipitation and temperature estimates and projections (Immerzeel *et al*., 2010; Lutz *et al*., Nepal, 2016), and (3) uncertainty in the volume of water contained in shrinking glaciers.

Hydrologists have used hydrological models to understand hydrological dynamics at a watershed scale for many decades (Singh and Frevert, 2002). For this, different types of hydrological models are proposed depending upon the model research questions, data availability and understanding of related hydrological processes (Nepal *et al*., 2014). Many hydrological models have been applied in the Himalayan region to understand hydrological system dynamics such as SRM model (Immerzeel *et al*., 2010; Panday *et al*., 2013; Khadka *et al*., 2014), SPHY model (Lutz *et al*., 2014) and J2000 model (Gao *et al*., 2012; Nepal *et al*., 2014; Nepal, 2016). Many of the models have adjustments for snow processes; SPHY and J2000 also simulate glacier melt processes. They typically use an index or degree-day factor (DDF)—see Shea *et al*. (2015), Thayyen and Gergan (2010) and Immerzeel *et al*. (2013) for reference values. Models that explicitly represent physical processes require significantly more detailed data,

*Correspondence to: Santosh Nepal, International Centre for Integrated Mountain Development, Kathmandu, Nepal.
E-mail: Santosh.Nepal@icimod.org

than more abstract lumped conceptual models (Daniel *et al*., 2011). In the case of the Himalayan region, snow and glacier melt processes dominate the hydrology of some high-altitude areas (in contrast to mid-hills) (Lutz *et al*., 2014) and may require explicit representation to capture dynamics expected in climate change studies. For example, the high-altitude areas provide streamflow during the dry season through melt runoff (Immerzeel *et al*., 2012; Nepal, 2016). A further constraint is the limited climate observations because of harsh terrain—many of the high-altitude areas are remote and inaccessible (Nepal, 2012; Pellicciotti *et al.*, 2012). In such cases, selecting a model that balances model complexity and data availability is a challenging task (Knoche *et al*., 2014). Because of the limited amount of available data in such environments, it may not be possible to adequately constrain the model parameters, leading to well-known equifinality problems (Beven and Freer, 2001).

The spatial distribution of precipitation for hydrological models continues to cause difficulties for modellers (Pellicciotti *et al.*, 2012). Several authors (e.g. Andermann *et al*., 2012; Krakauer *et al*., 2013) have assessed the accuracy of ground network surfaces (e.g. APHRODITE by Yatagai *et al*., 2012) and satellite radar (e.g. TRMM described in Huffman *et al*., 2007) against independent ground measurements using the normal observation network, high elevation stations (i.e. EV-K2-CNR stations examined by Diodato *et al*., 2010) and short expeditions (Immerzeel *et al*., 2014). In general, many consider APHRODITE the best of the freely available interpolated products (e.g. Andermann *et al*., 2012), although it underestimates precipitation in the Nepal Himalayas (see notes in Supporting Information of Lutz *et al*., 2014 for more details). Müller and Thompson (2013) also demonstrate that methods for assimilating the satellite and ground network datasets improve rainfall accuracy compared to either network. For this study, the ground network observations are denser than most locations in the Himalayas, providing an opportunity to examine sensitivities to model input data.

Recent advances in model development and regional calibration of GR4J for the snow-covered areas by Valéry *et al*. (2014) and SPHY for Himalayan areas by Lutz *et al*. (2014) demonstrate methods to improve the robustness of estimates of the Himalayan water balance. Valéry *et al*. (2014) removed glaciated catchments from the suite of catchments used for calibration because of their unique properties. Lutz *et al*. (2014) evaluated the SPHY model against streamflow using hydrological stations that were interior to the calibration catchments because of limitations on data availability.

This paper evaluates an adaptation of the GR4J hydrological model (Perrin *et al*., 2003) to include snow and glacier melt processes, hereafter called 'GR4JSG'. The model is proposed as a combination of parsimonious approaches to (1) catchment runoff process, (2) snow-melt processes; and (3) glacial snow stores in this study. The model is evaluated against streamflow and compared to Moderate Resolution Imaging Spectroradiometer (MODIS) snow extent (as motivated by Valéry *et al*., 2014) in the Tamor sub-catchment, an eastern tributary of the Koshi River basin. The model's parameter sensitivity is explored using Markov Chain Monte Carlo (MCMC) sensitivity as implemented in DiffeRential Evolution Adaptive Metropolis (DREAM) (Vrugt *et al*., 2009). This paper explores the model uncertainty because of different assumptions about model structure and biases in the input data, which informs our understanding of the likely limitations in the application of the model. The application of snow and glacier melt processes into the conceptual GR4J model provides melt runoff estimation from the alpine catchment with minimum data requirements (only precipitation, temperature and potET), but may not be accurate, as the parameters were not adequately identified from the stream-flow signal.

STUDY AREA

The Tamor catchment is a tributary of the Koshi River basin located in eastern Nepal as shown in Figure 1. The basin is characterized by steep topography (average slope of 28°) and high mountains, which are geologically active (rising and eroding). Figure 2 shows the river profile, with a landscape of sparsely covered forest in the lower catchment and permanently snow-covered mountain in the upper catchment (10% glaciated). The total area of the catchment is about 4005 km^2 and comprises the mountain Kanchenjunga, third highest mountain in the world (8200 masl). The glacier area covers about 407 km^2 of the basin area in altitude ranging from 4000 masl to 8200 masl according to Bajracharya and Shrestha (2011). The basin shows temperate climate in the lower elevation areas and sub-alpine to alpine climate in high altitude areas. The mean annual rainfall at the highest reference station, Taplejung (1732 masl), is 2.2 m/year with about 74% of the annual rainfall occurs during the summer monsoon between June to September (Nepal, 2012). The basin suffers from multiple water related hazards such as floods and flash floods, including glacial lake outburst, landslides, flooding and erosion (Chen *et al*., 2013).

METHODS AND MATERIALS

This section describes the GR4JSG model, conceptual structure of the catchment, the input data used and the methods of calibration and evaluation.

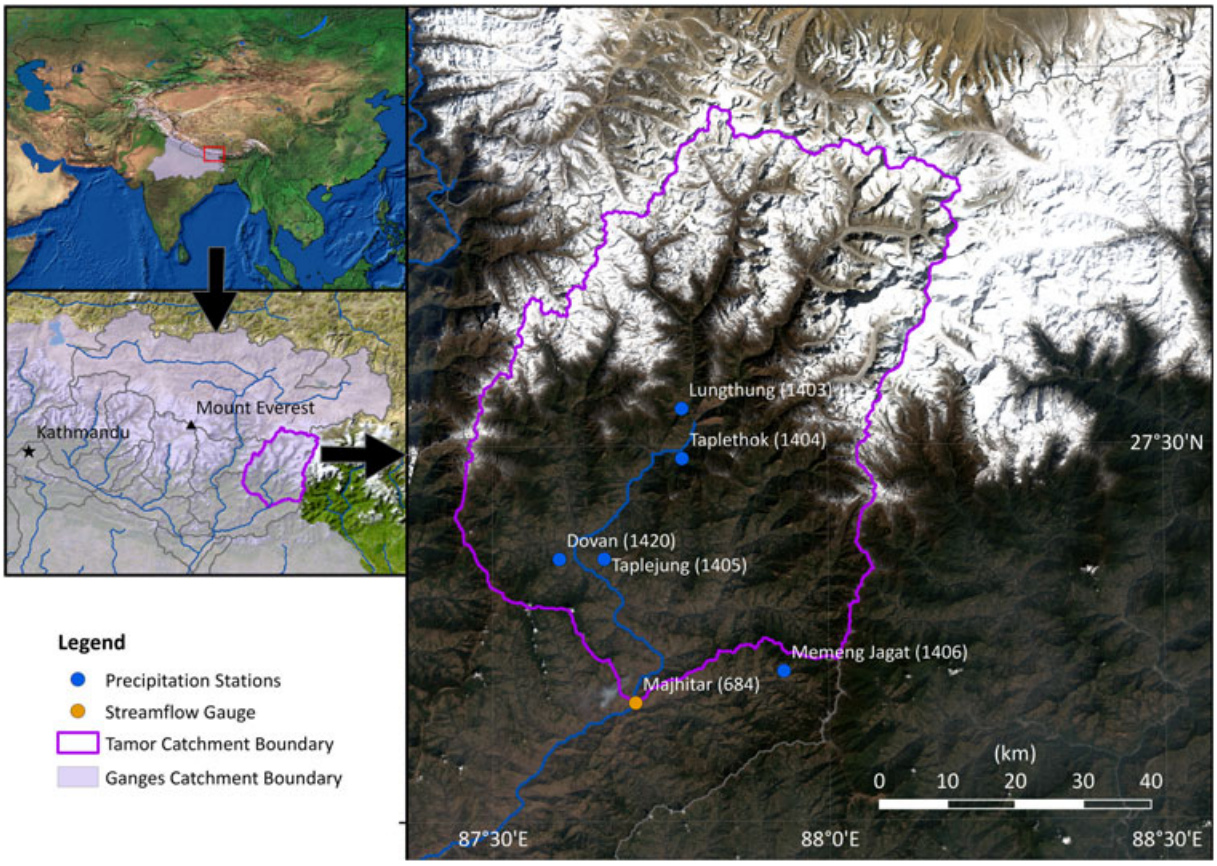

Figure 1. The Tamor catchment is a headwater of the Koshi, which is part of Ganges. It is located on eastern border of Nepal

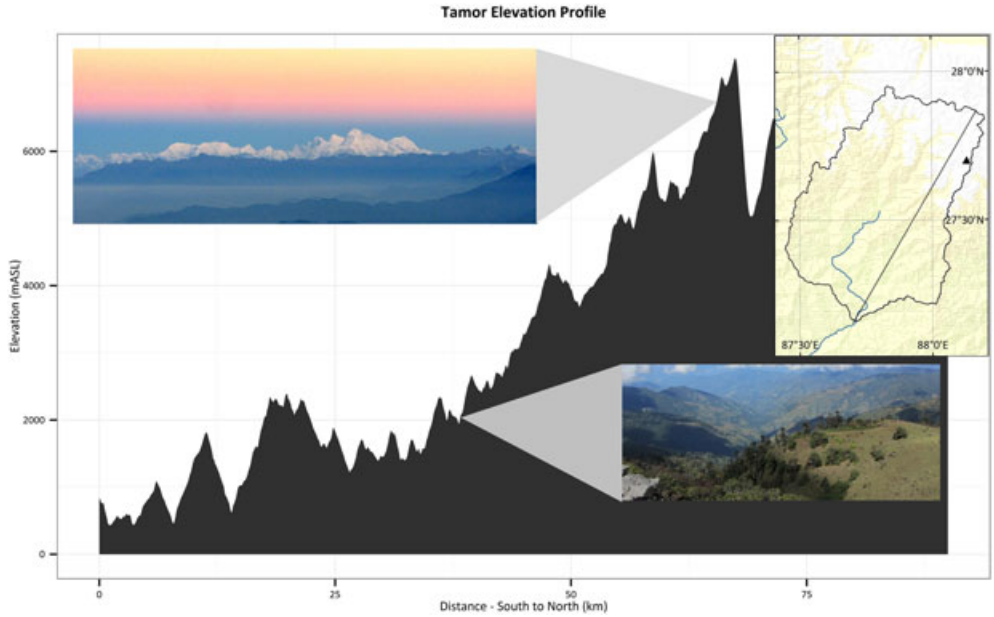

Figure 2. The elevation from south to north of the Tamor catchment. The Tamor catchment has a large variation in elevation from the mid-hills of Nepal to the top of mount Kanchenjunga (8200 masl). Images courtesy of Dr. Mac Kirby

GR4J with snow and glacier (GR4JSG)

GR4J has been widely used in Europe and around the world including in the Himalayas. The GR4J model developed by Perrin *et al.* (2003) is a daily lumped rainfall–runoff model with four parameters (x1, x2, x3 and x4). GR4J consists of two main stores: the production store, and the routing store as shown in the conceptual structure of Figure 3. The inputs to the model are rainfall depth (P) in mm and potential evapotranspiration (E) in mm. As shown in Figure 3, x1 controls the size of the production store (mm), x2 controls the flux to groundwater (mm), x3 controls the size of the routing store (mm) and x4 controls the recession of the unit hydrograph (days). The eWater Source (Welsh *et al.*, 2013) implementation of GR4J was used as the base for this paper. For more details on GR4J, please refer to Perrin *et al.* (2003).

In addition to Perrin *et al.* (2003)'s GR4J model, the GR4JSG model contains stores for snow accumulation (Snow in mm) and glacial water content (Ice in mm). The model distributes precipitation between rain and snow, depending upon the average temperature. The daily partition of precipitation into rainfall and snow is:

$$P_{snow} = \mathrm{P} \,.\, \mathrm{P}_s \tag{1}$$

$$P_{rain} = \mathrm{P} \,.\, (1 \,-\, \mathrm{P}_s) \tag{2}$$

where $\mathrm{P}_{\mathrm{snow}}$ is the amount of precipitation falling as snow (mm), $\mathrm{P}_{\mathrm{rain}}$ is the fraction of snow falling as liquid precipitation (mm) and P_{s} is the proportion snow as defined in the J2000 hydrological model (Krause, 2002; Nepal, 2012):

$$Ps = \frac{TRS + \mathrm{TRANS} - T_{avg}}{2.\,\mathrm{TRANS}} \tag{3}$$

where TRS (°C) is the base temperature which differentiate rain (above TRS) and snow (below TRS). The range for Ps is fixed in between 0 and 1. However, parameter TRANS (°C) provides the temperature range of above and below the TRS in which the rain-snow mixture is calculated and defined in Equation 3. For this model, TRS is considered as 0 °C and TRS is 2 °C which means above +2 °C is rain, −2 °C is snow and mixture of rain-snow in between +2 and −2 °C. T_{avg} is the daily average temperature which is defined from the daily maximum and minimum temperature.

The snowmelt leaving the snow store is estimated by applying temperature index degree-day factor, $\mathrm{DDF}_{\mathrm{Snow}}$ (mm/°C/day) as shown in Equation 4. In this approach, temperature is considered as a heat transfer process for melting of snowpack (Hock, 2003). The amount of snowmelt (mm) is:

$$\mathrm{Snowmelt} \;=\; \mathrm{DDF}_{snow} \,.\, (\mathrm{T}_{melt} - \mathrm{T}_{base}) \tag{4}$$

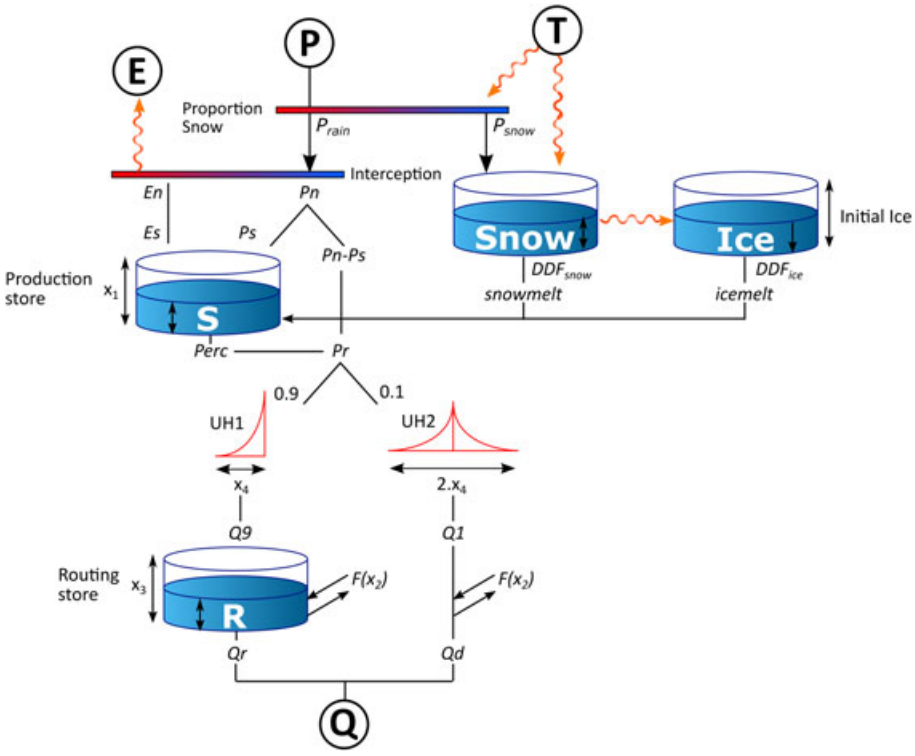

Figure 3. From Perrin *et al.* (2003), the conceptual structure of the GR4J model is shown with modifications for snow and glaciers. When temperature is below zero precipitation falls as snow. In warm conditions, the snow melts (or thermal energy causes ice to melt). Snow and ice melt enter to Production Store (S)

where T_{base} (°C) is the threshold temperature for snowmelt and T_{melt} is the average value of maximum and average temperature.

Likewise, the model simulates the glacier icemelt by applying temperature index degree-day factor, DDF_{Ice} (mm/°C/day). In the glacier area, when the surface is covered with seasonal snow, first the snowpack melts as explained above (Equation 4), then once the snow storage is empty, glacier icemelt begins:

$$\text{Icemelt} = \text{DDF}_{ice} \,.\, (\text{T}_{melt} - \text{T}_{base}). \quad (5)$$

The output from snowmelt and glacier melt are allowed to enter into the production store as shown in Figure 3.

In this study, a large initial volume of water in the ice store makes the glacier area behave in a time invariant manner. In the real environment, the glacier area shrinks and expands. This dynamic process is not included in the GR4JSG model. However, the change in glacier area is a gradual process and reflected at a decadal scale, and might not affect the model results for a short simulation periods of a couple of decades (Jóhannesson *et al*., 1989; Pelto and Hedlund, 2001). Similarly, this study only uses glacier area without differentiating into clean and debris covered glaciers. However, in recent modelling applications, glaciers are distributed into clear and debris-covered in which melt runoff from debris-covered glaciers is reduced by some extent (Nepal *et al*., 2014; Lutz *et al*., 2014). Because of the conceptual nature of the model in which the catchment is distributed in elevation bands, only glacier area is used.

Conceptual structure of the catchment

The conceptual structure of the catchment aimed to: (1) retain a simple parsimonious model, (2) account for differing characteristics of glacier and non-glacier areas, (3) account for significant variation in the catchment's temperature profile, and (4) provide a spatial structure that would allow evaluation against observations. GR4J is a simple parsimonious model which required each extra free parameter to demonstrate a significant improvement in performance Perrin *et al*. (2003). We afforded an additional two free parameters to account for snow and glacial properties, but no extra free parameters.

One of the characteristics of mountainous region is elevation differences where the temperature varies significantly. Valéry (2010) demonstrated the need to account for significant variation in the catchment's temperature profile. Accounting for variation in temperature profile by dividing the catchment into elevation bands improved model performance in highly mountainous catchments by dividing the catchment into elevation bands. Each elevation band had different forcing data based on a temperature lapse rate from a reference temperature station. Following suit, we partitioned the catchment into bands based on elevation from the Shuttle Radar Thematic Mission 3-s digital elevation model with processing by CGIAR to mosaic and fill voids (Jarvis *et al*., 2008).

The combination of elevation bands at 200-m intervals and glacier extent led to 44 functional units (FUs) (Figure 4). A FU is a conceptual portion of a catchment that functions in a similar way for hydrological processes as defined by Argent *et al*. (2009) in similar fashion to the concept of hydrological response units defined by Leavesley *et al*. (1995). Principally, a FU in a catchment can have different model parameters or inputs to another FU in the catchment. For this study, we used identical precipitation data, but different temperature and potential evapotranspiration data for each FU; however, the parameters were changed for different experiments. The same parameter values were applied for all FUs. The runoff from rainfall runoff models in all the FUs are summed at the catchment outlet (i.e. there is no explicit routing scheme between FUs).

Input hydroclimate data

GR4JSG requires precipitation, temperature and evapotranspiration climate data to simulate streamflow, snow extent and the catchment's water budget. The precipitation input was the daily simple mean of the five nearby precipitation stations (as shown in Figure 1). Table I provides a detailed description of these stations and their characteristic features.

In meteorological contexts, Brunt (1933) and many others have investigated the rate that air temperature falls when it ascends, either as a fully saturated air mass, or as a fully dry air mass—this is the adiabatic lapse rate for dry and saturated air. These lapse rates are often used to extrapolate temperatures at particular elevations based on ground observations. Rolland (2003) examined the reliability of extrapolating temperature based on 269 sites, finding that the strongest correlation for maximum temperatures was in summer with additional topographic features such as aspect and slope affecting the interpolation reliability. In the local region: (1) Kattel *et al.* (2013) examined 56 temperature stations (72 to 3920 masl) in Nepal to produce monthly lapse rates varying from 0.43 °C/100 m to 0.61 °C/100 m for mean temperature, (2) Pokhrel *et al.* (2014) used a constant lapse rate of 0.46 °C/100 m for Dudh Koshi catchment (100 km west of Tamor) based on local observations in their analysis of snowmelt modelling approaches, (3) Panday *et al.* (2013) calibrated annual lapse rates (mean 0.51 °C/100 m to 0.68 °C/100 m) in their application of Snowmelt Runoff Model to Tamor catchment, (4) Normand *et al*. (2010) used a second-order polynomial regression function in

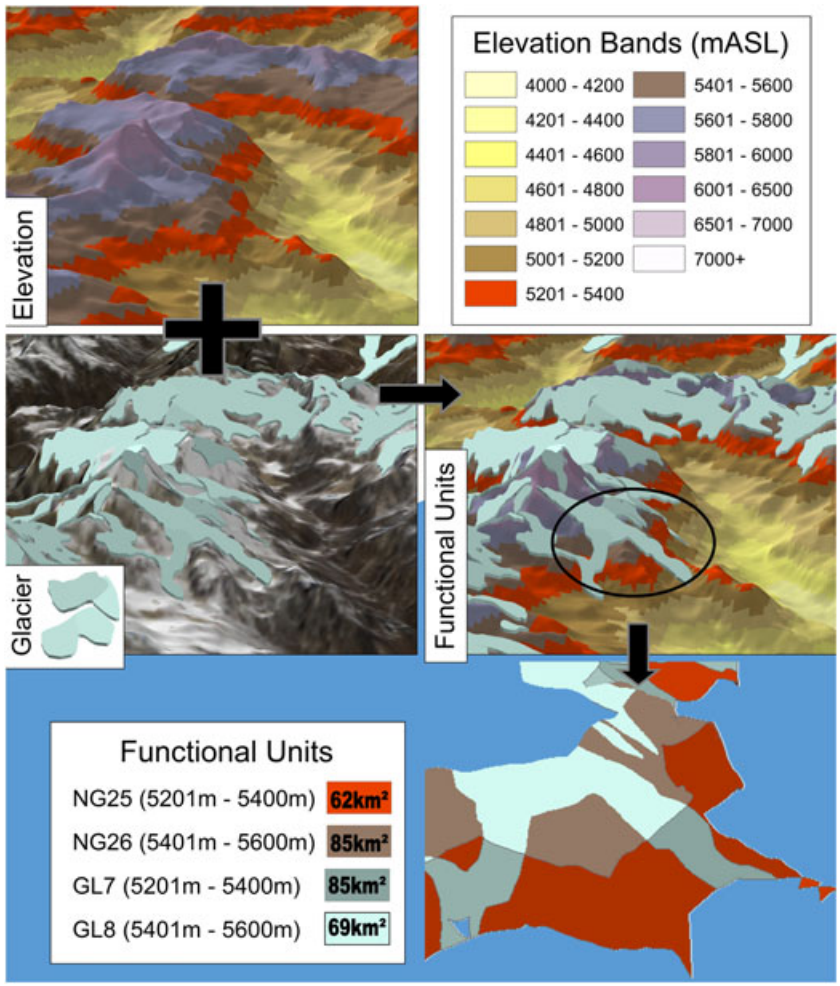

Figure 4. Top, the elevation of the catchment is divided into elevation bands using the Shuttle Radar Digital Elevation Model. Below left, the elevation bands are combined with glacial coverages (based on Landsat and field observation). Right, the combination of glacial coverage and elevation band form functional units. Bottom, four functional units are shown in detail with the total area covered. There are 30 functional units representing the area without glaciers at elevations from 400 m to 8600 m. There are 12 functional units representing the area with glaciers at elevations from 4000 m to 8600 m

their application of HBV to the Tamor catchment (no further details provided); and (5) Nepal *et al.* (2014) derived seasonal lapse rates of 0.55 °C/100 m for monsoon (June to September) and 0.6 °C/100 m for the dry season (October to May) based on observations of the Dudh Koshi catchment. Immerzeel *et al.* (2014) have suggested a more physically based model of temperature lapse rate that includes humidity. The temperature for each functional unit in the Tamor catchment model was corrected using the seasonal lapse rate for summer and winter as suggested by Nepal (2012) for the neighbouring catchment of the Dudh Koshi catchment. For Tamor, the difference in elevation between the median elevation of the functional unit and the elevation at the Taplejung meteorological station was adapted to calculate the temperature of each functional units.

Table I. Hydro-meteorological stations used in the Tamor river sub-basin. Each measurement station measures parameters such as precipitation (P), maximum temperature (Tmax), minimum temperature (Tmin), sunshine hours (SH), wind-speed and discharge (D).

Station ID	Station name	Elevation (mASL)	Parameters
1403	Lungthung	1780	P
1404	Taplethok	1383	P
1405	Taplejung	1732	P, Tmax, Tmin, SH, RH, WS
1406	Memenjagat	1830	P
1420	Dovan	763	P
684	Majhitar	533	D

The potential evapotranspiration data for the Tamor catchment model was from Nepal (2012), which was derived by using Penman–Monteith equation. The model's streamflow predictions were evaluated against the Majhitar (station id 684) streamflow gauge measure-

ments (Table I). The Majhitar streamflow gauge was the outlet of the Tamor catchment model.

Calibration and evaluation method

GR4JSG parameters x1, x2, x3, x4, DDF_{Snow} and DDF_{Ice} are subject to local catchment conditions; that is, the actual parameter values need to be inferred by calibrating to local conditions (see Vaze *et al*., 2011 for further discussion on calibration techniques, and Bennett *et al*. (2013) for discussion on evaluation methods). The melt runoff estimated through calibration parameters of DDF for snow and ice is shown in Equations 4 and 5 which is directly affected by temperatures of the functional units. The calibration techniques involved: (1) split-sample calibration and validation of six parameters, (2) Box-Cox transform (Box and Cox, 1964) on streamflow (for calibration only), and (3) Joseph and Guillaume (2013) likelihood function to compare modelled and observed time-series. The results were evaluated against streamflow and compared to MODIS snow extent.

The split-sample calibration involved training the model from 1 January 2001 to 31 December 2004 then testing the model from 1 January 2005 to 31 December 2009. The availability of good quality rainfall data dictated the length of the calibration and validation periods. The snow stores were slow to reach an equilibrium state. That is, if the snow stores were at zero at the start of a simulation it would take 20 years for some of them to warm up. We seeded the initial value of the snow stores by running a complete simulation with calibrated parameters. From these initial snow store states, running the model in the warm-up period (1994–2000) was sufficient to remove any trends in the snow store states. Even noting that inter-annual variability of the region is relatively low, the short calibration and evaluation period meant the model is unlikely to capture all climatic variability. The performance of the model was evaluated against streamflow by applying Nash–Sutcliff Efficiency (Nash and Sutcliffe, 1970—NSE), Percent Bias (PBIAS) and visual inspection.

The model was also calibrated using the DREAM MCMC algorithm by Vrugt *et al*. (2009) with the likelihood function defined by Joseph and Guillaume (2013). Through Bayesian inference, the adaptive MCMC algorithm, running multiple interacting chains in DREAM, calculates the posterior probability density function of parameters through sampling (i.e. the likelihood that the observed series was generated from particular parameter sets). For accurate description of the posterior probability density functions, many chains are required. Twelve chains provided a quantitative understanding of the parameter variation and the model's sensitivity to particular parameter values. The R-based implementation of DREAM described by Guillaume and Andrews (2012) with parallelism described by Joseph and Guillaume (2013) was used for calibration in which the authors used eight chains. According to Laloy and Vrugt (2012), the number of chains for the original implementation of DREAM should be double the number of dimensions of the problem (in this case 12 chains were used to match six dimensions that were presumed to be orthogonal).

In order to avoid skew in the likelihood function towards high flows, we transformed the streamflow using a Box–Cox Transform (Box and Cox, 1964), with power 0.5 as shown in Equation 6 (note evaluation used untransformed values). A value of 0.5 provided the effect of required, so no investigation of other lambda ranges was necessary.

$$X_t = \begin{cases} \dfrac{(Q_t + \lambda_2)^{\lambda_1} - 1}{\lambda_1}, & \lambda_1 \neq 0 \\ \log(Q_t + \lambda_2), & \lambda_1 = 0, \end{cases} \tag{6}$$

where Q_t is stream flow and λ_1 and λ_2 are transformation parameters (λ_1 is set to 0.5, and $\lambda_2 = 0$). X_t is the transformed value at time t.

The comparison against snow extent involved producing a spatial total for the Tamor catchment from the MODIS 8-day composite snow extent product (Hall *et al*., 2002). For the assessment of snow cover areas, different filters (temporal, spatial and altitude-based filters) were used to reduce the influence of cloud pixels in the snow products, as described by Gurung *et al*. (2011), as shown for some dates in Figure 5. Note the accuracy of the product improves from 5 July 2002 when both Aqua and Terra satellites are providing measurements, and the accuracy is less during continuous periods of cloudy weather during monsoon. MODIS snow products have been widely used for understanding snow extent in different parts of the world (Klein and Barnett, 2003, Parajka and Blöschl, 2006; Wang et al. 2008). Figure 5 shows the MODIS pixels (500-m resolution) overlapping the Tamor catchment. The snow extent for the catchment was the sum of the areas covered by MODIS pixels classified as snow covered. The equivalent modelled snow extent was compared by using coefficient of determination (R^2) and visual inspection after the model was calibrated using streamflow.

Uncertainty and sensitivity analysis

Sources of uncertainty in the catchment runoff modelling include observed data (precipitation, temperature, potential evaporation and streamflow), model assumptions and simplifications, and parameterization (Kuczera *et al*., 2006; Renard *et al*., 2010;

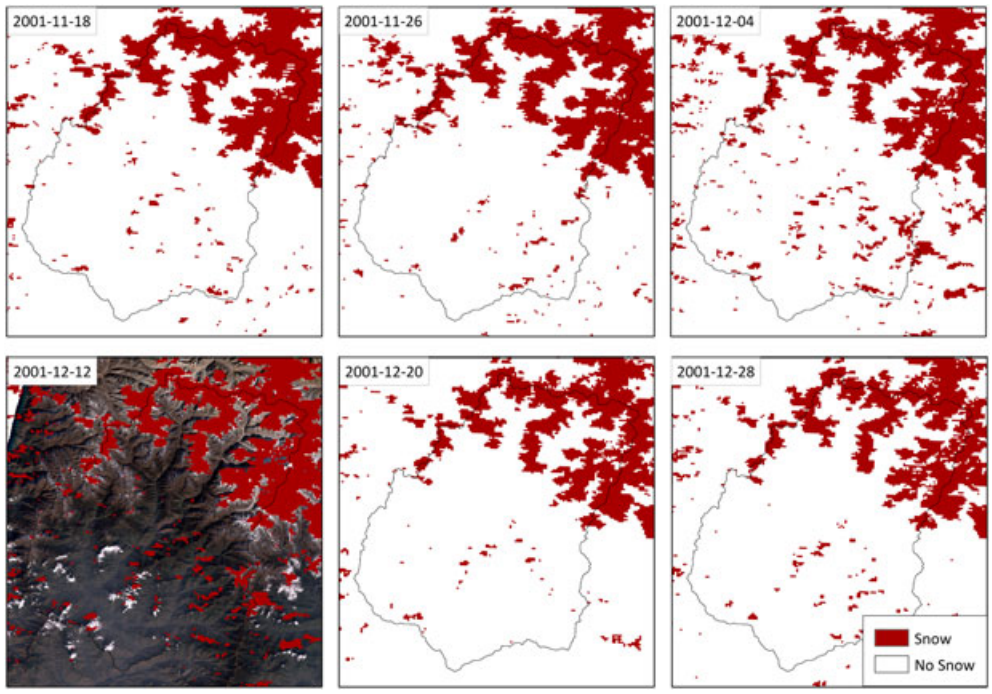

Figure 5. MODIS 8-day snow extent classification, shown in red, across the Tamor catchment. Top row images were for dates 18 November 2001, 26 November 2001 and 4 December 2001. Bottom row images were for dates 12 December 2001 (with Landsat image background of the same date), 20 December 2001 and 28 December 2001

Vaze *et al*., 2011). Andermann *et al*. (2012) discuss difficulties understanding precipitation in the Himalayas including: (1) inaccessible terrain that makes monitoring difficult, and (2) the orographic effects resulting from elevation and climate processes. In many cases, the location of stations around river valleys may cause under-reporting of precipitation (Nepal, 2012). The ground network is sparse and typically limited to elevations under 2000 m (1830 m in the case of Tamor; Table I).

In general, sensitivity analysis techniques provide a method to test the robustness of models to input or parameter uncertainty. Common schemes for conducting sensitivity analysis include variance-based schemes such as Monte Carlo analysis; however, Peeters *et al*. (2014) and Plischke *et al.* (2013) have highlighted the complexity of conducting sensitivity analysis when the solution space is complex and non-linear (which it often is). Nonetheless, the robustness of the model to uncertainties was tested by varying precipitation data and ground-water parameter bounds (one-at-a-time) and assessing the Monte Carlo parameter distributions returned by DREAM.

For testing the sensitivity to precipitation, three additional sets of plausible precipitation inputs were tested: (1) average observed precipitation, (2) 110% of average observed precipitation above 2000 m, and (3) 110% of average observed precipitation between the elevation bands of 2000 m and 3000 m, 100% of precipitation between 3000 m and 4000 m, 70% precipitation between elevation bands 4000 m and 5000 m, 40% precipitation between 5000 m and 6000 m and 10% precipitation above 6000 m. Set (1) represented the extrapolation method of many hydrological studies (e.g. Panday *et al*., 2013). Set (2) represented calibrated multiplicative factors such as that of Lutz *et al*. (2014)—note, Lutz *et al*. (2014) add 17% to Aphrodite rainfall. Set (3) represented a complicated pattern of precipitation variation as depicted by Diodato *et al*. (2010)—note the level of extrapolation is large because of the lack of reference stations.

Summary of experiments

Ten experiments were designed to examine the sensitivity of the model to input data and parameterization (Table II). Experiment A was set up as a baseline using GR4J without any conceptual changes. Experiment B added snow and glacial stores to see the performance improvement of conceptual changes. Experiment C and D tested the sensitivity of the model to different precipitation series. The ground-water flux component (x2) is designed to allow inter-catchment flows of groundwater (Le Moine *et al*., 2007). It is unknown whether such a flux exists in this catchment. However, because x2 can introduce and remove water from the catchment water-balance, it may compensate for other processes (e.g. DDFs may rise to compensate). To understand the role of x2 in this catchment, Experiment E, F and G examined the variation of the parameters for GR4J ground-water flux and ice and

Table II. Calibration performance scores against different assumptions of precipitation driving data, ground-water constraints (×2) and degree-day-factor values. Precipitation set (A) average observed precipitation at 5 stations, (B) 110% of average observed precipitation above 2000 m and (C) 110% of average observed precipitation between the elevation bands of 2000 m and 3000 m, 100% of precipitation between 3000 m and 4000 m, 70% precipitation between elevation bands 4000 m and 5000 m, 40% precipitation between 5000 m and 6000 m and 10% precipitation above 6000 m. NSE is calculated daily.

Exp	Description	Param bound	Glacier area	Precip. set	Calib NSE	Calib PBIAS	Eval NSE	Eval PBIAS	All NSE	All PBIAS
A	Baseline GR4J	None	407 km^2	A	0.867	1.93%	0.848	−4.91%	0.840	−7.18%
B	Baseline GR4JSG	None	407 km^2	A	0.892	2.40%	0.883	0.24%	0.86	−5.45%
C	Sensitivity to precipitation	None	407 km^2	B	0.894	0.14%	0.883	−2.10%	0.856	−7.56%
D	Sensitivity to precipitation	None	407 km^2	C	0.888	−1.40%	0.878	0.58%	0.840	−8.00%
E	Sensitivity to g/w (x2)	x2 = 0	407 km^2	B	0.895	−1.40%	0.880	−3.70%	0.850	−9.00%
F	Sensitivity to precipitation and g/w (x2)	$\|x2\| < 0.5$	407 km^2	B	0.895	−0.14%	0.880	−2.40%	0.856	−7.80%
G	Sensitivity to precipitation and g/w (x2)	$\|x2\| < 0.5$	407 km^2	C	0.875	−6.20%	0.872	−4.20%	0.820	−12.0%
H	Sensitivity to glacial area	x2 = 0	536 km^2	B	0.895	−1.40%	0.880	−3.70%	0.850	−9.00%
J	Performance using tight reference values for ddf	3.0 < iceddf 9.0 > snowddf x2 = 0	407 km^2	B	0.894	1.10%	0.880	0.29%	0.860	−6.20%
K	Sensitivity to ice-snow discrimination	iceddf = snowddf, $\|x2\| < 0.5$	407 km^2	B	0.891	1.10%	0.880	0.80%	0.858	−6.20%

snow stores by constraining the ground-water flux (x2 = 0). Experiment H examined the sensitivity of the model to glacial area (a historic glacial area map was used for this experiment). Experiment J and K investigated the performance of the model when using reference values to constrain the ice and snow degree day factors. In each experiment, DREAM was used to calibrate the model, and the reported best parameter set was the one with the highest likelihood through calibration.

RESULTS AND DISCUSSION

The design and application of GR4JSG to the Tamor catchment demonstrated: (1) the ability to adequately represent streamflow using GR4J and GR4JSG models, (2) the ability to compare hydrological models against MODIS snow extents, (3) the difficulty establishing robust degree-day factors, and (4) interaction between different GR4JSG parameters.

Water balance (snow and glacier melt)

Figure 6 shows the median monthly precipitation and discharge of the catchment during 2001 to 2009 according to station observations (with inter-annual variation of one standard deviation shown as bands). In the period 2001–2009 the average precipitation and discharge are 2135 mm and 1734 mm, respectively. Most of the precipitation (about 72%) falls during the monsoon season and most of the discharge (74%) flows during the same period. During the latter part of the monsoon and post monsoon period, discharge is greater than precipitation indicating a streamflow contribution from ground-water, snowmelt, permafrost melt and/or glacial melt.

The integration of snow and glacier melt plugins to the standard GR4J model made it possible to understand the contribution of snow and glacier melt runoff to the total streamflow as shown in Figure 7. The model estimates that snow and glacier melt runoff from the catchment amount to about 14 to 16% of the total annual runoff across various experiment. The contribution of melt runoff during the pre-monsoon season (March–May) ranges from 23 to 27% of the total runoff, while it is 16 to 18% during the monsoon season (June–September) when the melt runoff coincides with rainfall–runoff. A few other studies in the Tamor basin (Nepal, 2012 and Panday *et al.*, 2013) estimate that the snowmelt contribution to total annual runoff to be about 27%, which is higher than estimated from this study. The lower amount of glacier melt from this study is primarily because of unrealistic

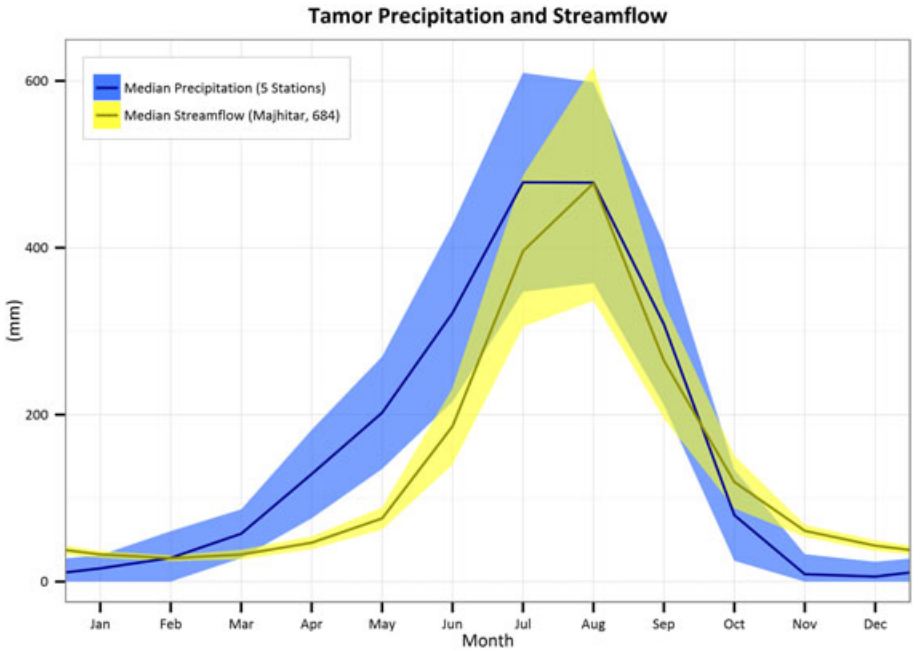

Figure 6. The median monthly precipitation and discharge of the catchment during 2001 to 2009 according to station observations (with inter-annual variation of one standard deviation shown as bands)

lower values of DDFIce. This is also explained in the next section with reference to different experiments (Table III). Similarly, the reason for the difference is also due to the of different conceptualizations of snowfall and snowmelt. In Nepal (2012), runoff from rain-on-snow (because of rainfall on snow buckets in the days following a snow event) is treated as part of the snowmelt, whereas in this model it is considered as runoff generated by rainfall. As suggested by Nepal *et al*. (2014), the contribution of rain-on-snow event is quite high in low elevation areas and gradually decrease in the high-altitude areas. The studies from the nearby western catchment (Dudh Koshi) suggest that the glacier melt contribution is 19% (Lutz *et al*., 2014) between 1998 and 2007 and 17% (Nepal *et al*., 2014) between 1986 and 1997. Nepal *et al*., 2015 compared the GR4JSG and J2000 in the Dudh Koshi

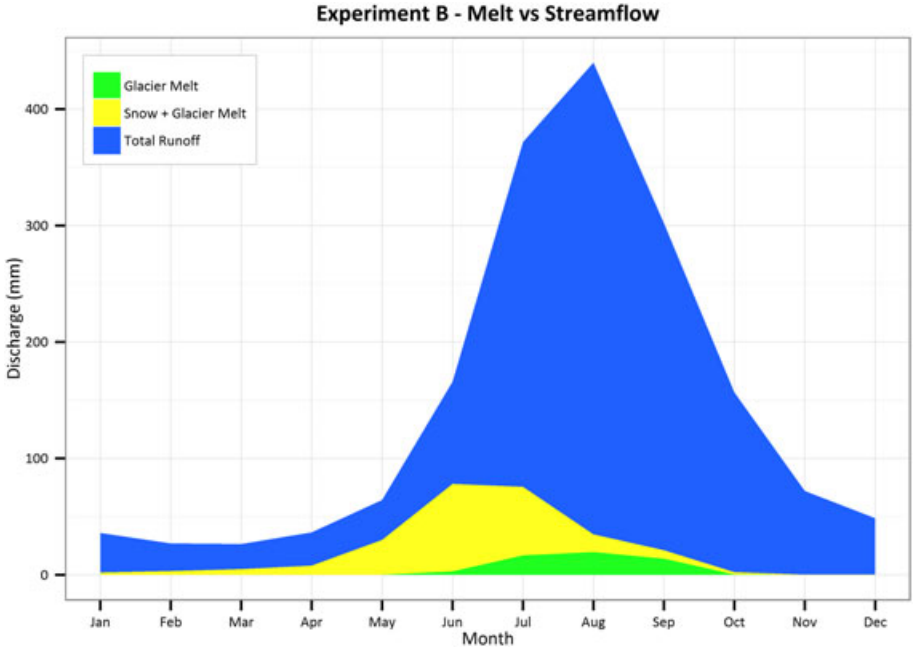

Figure 7. Monthly contribution of different melt runoff (snow and glacier) to modelled streamflow

Table III. Best parameter sets for each experiment

Exp #	Description	x1 (mm)	x2 (mm)	x3 (mm)	x4 (day)	Snowddf mm °C^{-1} day^{-1}	Iceddf mm °C^{-1} day^{-1}
A	Baseline GR4J	1143.243	3.452	146.904	0.800	N/A	N/A
B	Baseline GR4JSG	1498.572	0.522	142.402	0.517	7.801	0.006
C	Sensitivity to precipitation	1498.616	0.240	170.733	0.592	7.776	0.006
D	Sensitivity to precipitation	1313.566	1.303	128.167	0.507	9.669	9.537
E	Sensitivity to g/w (x2)	1499.893	0.000	185.961	0.524	7.807	0.018
F	Sensitivity to precipitation and g/w (x2)	1498.258	0.196	172.444	0.524	7.831	0.009
G	Sensitivity to precipitation and g/w (x2)	1499.432	0.499	99.525	0.632	9.959	9.988
H	Sensitivity to glacial area	1499.882	0.000	188.600	0.532	7.782	0.014
J	Performance using tight reference values for ddf	1499.817	0.000	175.914	0.605	7.754	3.000
K	Sensitivity to ice-snow discrimination	1497.991	0.247	153.613	0.608	3.801	3.801

catchment results similar melt contribution of about 13% (GR4JSG) and 17% (J2000) and variation was attributed to different conceptualization of hydrological processes in both models.

Evaluation against streamflow

The performance of standard GR4J model through the evaluation period was NSE daily (0.85) and PBIAS (−4.9%) for streamflow (Experiment A, Table II and Table III). Note that GR4J by itself provides no insight into the snow and glacier processes. The addition of snow and glacier melt approach in GR4JSG model (Experiment B) led to improvements in the representation of streamflow with daily NSE (0.88) and PBIAS (0.24%) during the evaluation period (Figure 8). The modelled flow values demonstrate a high level of visual agreement with observations through both wet and dry seasons; however, the high peak flows are underestimated in some occasions during both the calibration and validation periods. Most values are represented quite well although there are deviations in theoretical quantile *versus* model quantiles as shown in Figure 8d.

Increasing precipitation in different elevation zones as explained earlier did not produce a significant change to model performance in Experiment C, resulting in similar calibrated parameter values to Experiment B. The increase in precipitation did lead to a smaller value of x2, i.e. less water was imported via the groundwater transfer. In Experiment D, decreasing precipitation above 4000 m led to significant compensation through the x2 parameter, unrealistic degree-day-factors and poor performance (higher biases). Further constraining the x2 parameter led to poorer performance (higher biases)—Experiment E, F and G. Constraining snow and glacial degree-day-factor to narrower theoretical bounds did not change the model behaviour with daily NSE (0.88) and PBIAS (0.3%)—Experiment J.

The snow processes dominated the glacial processes in all the experiments with unconstrained snow and ice degree-day-factors. When the ice degree-day-factors were unconstrained, the calibrated values were far lower than reference values from other models (i.e. 0.006 mm°C^{-1} day^{-1} in Experiment B where Lutz *et al*., 2014 calibrated values of 3.0 mm°C^{-1} day^{-1} for debris covered and 6.0 mm°C^{-1} day^{-1} for debris free glaciers), and the difference between snow and ice degree-day factors for Experiment B was not realistic. Degree-day-factor to the ice degree-day-factor led to a value of 3.8 mm °C^{-1} day^{-1} for the degree-day-factor (Experiment K), which is near other studies (Lutz *et al*., 2014 calibrated snow degree day factor of 4.80 mm°C^{-1} day^{-1}). Therefore, calibration across a range of similar alpine glaciated catchments may be necessary to constrain the degree-day-factors to representative values. However, the current lack of precipitation and temperature data across several catchment makes this task difficult to achieve practically.

Comparison against MODIS snow extent

When the snow bucket in any FU contained more snow than a threshold of 10 mm, it was assumed that the complete area of the FU was covered in snow. Without the use of a low snow threshold to consider the area covered in snow, the model produced unrealistic snow area calculations, presumably because of an oversensitivity to smaller precipitation events. The agreement between MODIS 8-day snow extent and modelled snow extent for GR4JSG Experiment B was $R^2 = 0.46$ for the period 2002 to 2009 (Figure 9) and $R^2 = 0.42$ for Experiment K. From Figure 9, which compares the snow cover extent between the model (panel b) and the MODIS

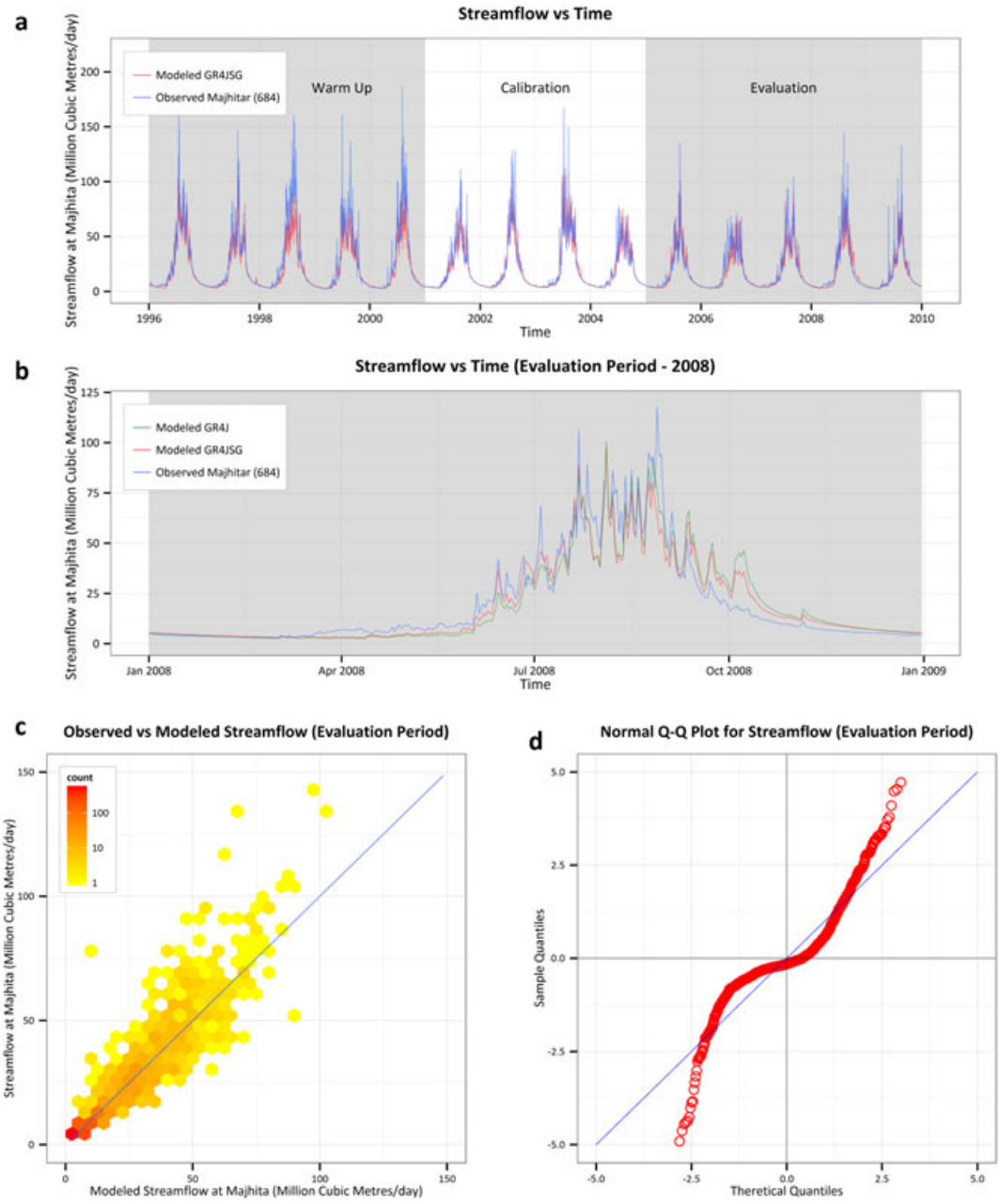

Figure 8. a) Observed *versus* modelled streamflow through the entire period. b) Observed *versus* modelled streamflow through the evaluation period. c) Scatter plot (hexagon density) showing observed *versus* modelled discharge of Experiment B. The efficiency is NSE daily (0.88) and PBIAS (0.24%) through the test period from 1 January 2005 to 31 December 2009. d) Sample quantiles *versus* theoretical quantiles for the evaluation period

snow extent (panel a), it is clear that the model can reproduce the seasonal pattern of snow accumulation and melting.

In Experiment K, with lower degree-day-factors (snow = ice = 3.801), the R^2 was 0.42. In all the experiments, the modelled snow extent during summer was consistently lower than the snow extent from MODIS possibly because of combined effect of overestimation of snow extent in MODIS for cloud cover and/or excessive melt because of temperature input and conceptualization. Because the relationship to snow extent was not part of the calibration objective function, there is a significant possibility that other equally good solutions might produce different levels of agreement.

The correlation between modelled snow extent and MODIS snow extent enables identification of internal model behaviours that are consistent with the conceptual model of catchment behaviour. That is, if the snow pack is growing when the MODIS snow extent is increasing, the model is behaving as expected. Conversely, if the

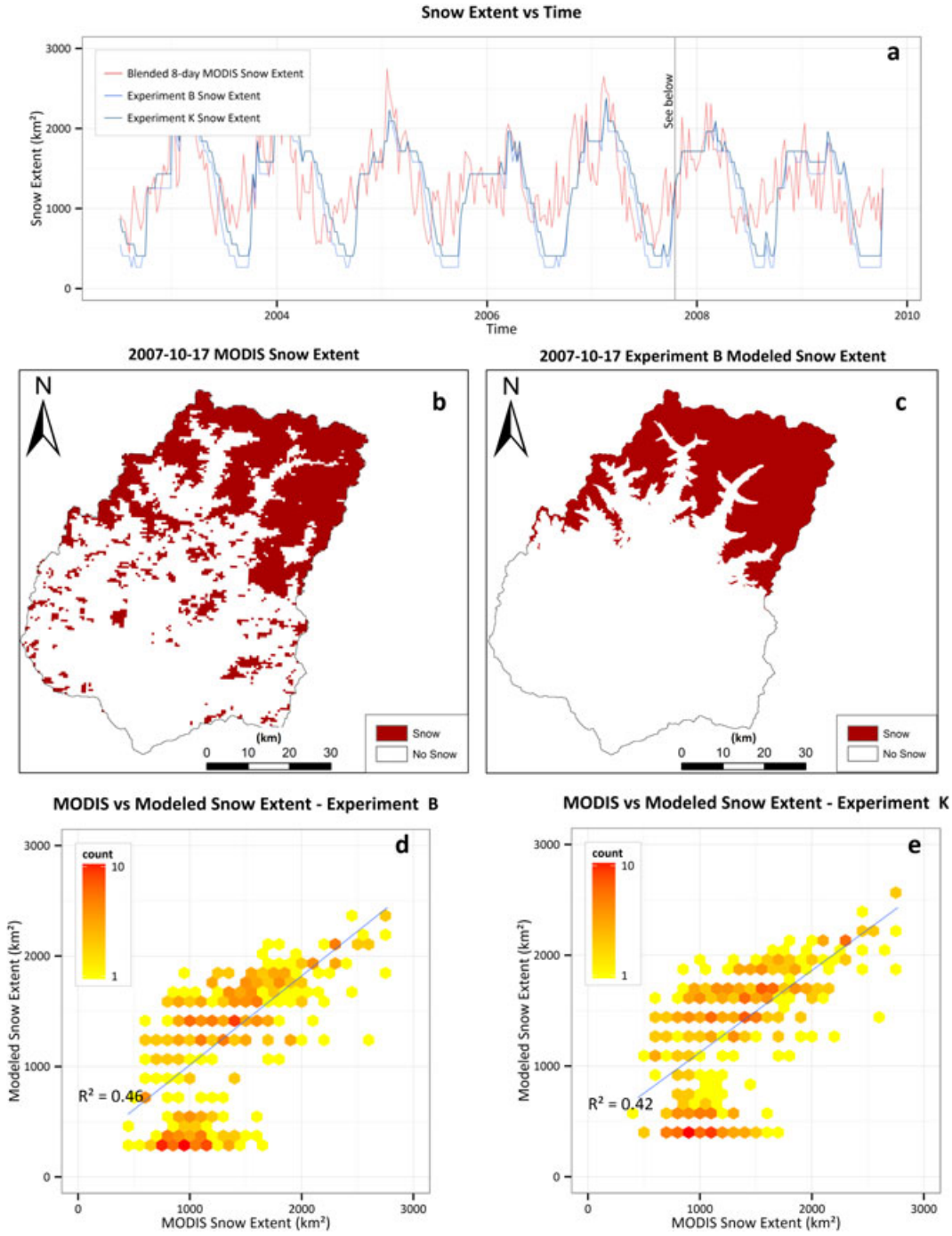

Figure 9. a) MODIS *versus* modelled snow extent for GR4JSG Experiment B and K through the period where data was available (including from both Aqua and Terra). b) Snow extent according to MODIS 8 day, c) attributed snow extent according to snow buckets containing snow, d) scatter plot (hexagon density) of MODIS *versus* modelled snow extent for experiment B. e) Scatter plot (hexagon density) of MODIS *versus* modelled snow extent for experiment K

snow pack is shrinking when the MODIS snow extent is increasing, the model might be responding spuriously to the input data, and it may not respond as expected to new input data (e.g. climate change scenarios). Because snow volume is a significant portion of the Himalayan water balance, any improvement in the estimation of snow volume should lead to hydrological predictions that are more robust. In particular, if snow extent is calibrated as in Finger *et al*. (2015), we would expect improved performance in alpine catchments.

Modellers using GR4JSG and similar models with degree-day-factors should beware of slow trends in the

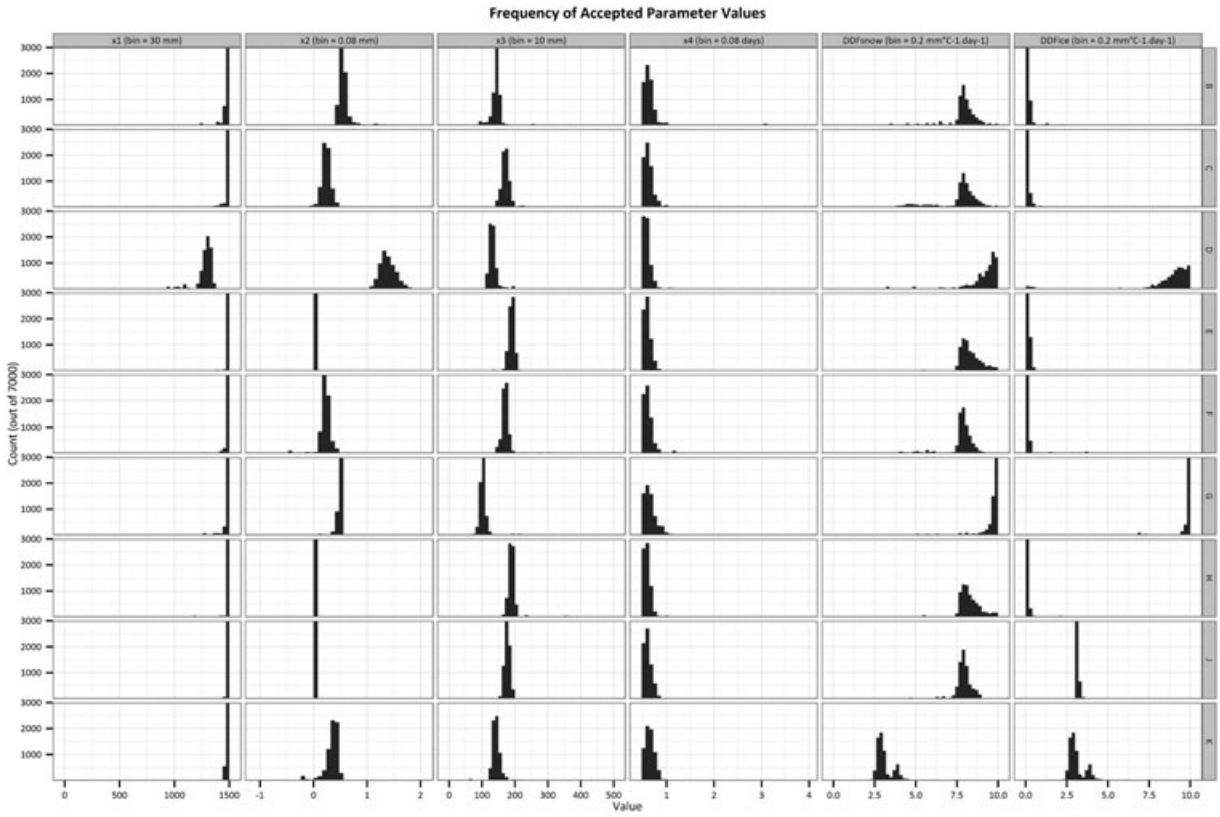

Figure 10. Histogram of acceptable parameter values from DREAM calibration for each experiment—each parameter's range is split into 50 buckets (where acceptable solutions have Gelman diagnostic statistic < 1.2 for convergence and are part of the second half of solutions). Parameters x1, x4 and DDFice were regularly on or near parameter boundaries

internal states of the models (i.e. the model will behave differently depending on the initial states of the snow and ice stores, which may take 20 years to converge to sensible values). However, the analysis of model sensitivities in this study showed that precipitation inaccuracies affect parameter values, which then affect hydrological estimates of ground-water fluxes and snow extent estimates.

Parameter sensitivity

Perrin *et al*. (2003) reported 80% confidence intervals for each parameter based on a large variety of catchments as 100 – 1200 mm for x1, −5 – 3 mm for x2, 20 – 300 mm for x3 and 1.1 – 2.9 days for x4. The calibration algorithms ranged over slightly broader parameter ranges: 0 – 1500 mm for x1, −10 – 5 mm for x2, 0 – 500 mm for x3 and 0.5 – 4 days for x4 to identify unusual parameter values. For all the GR4JSG experiments, the x1 parameter was close to the upper parameter bound (1500 mm) indicating that the production store for this catchment was larger than expected from the original GR4J model application and may be compensating for processes particular to this catchment (e.g. snow and glacier).

Figure 10 shows the parameter frequency derived from the posterior parameter density for the experiments using DREAM. Ideally, parameters of the model would be identifiable and independent, features of a parsimonious model. That is, identifiable in that there would be a single best value for each parameter when compared to the observed data for validation, and hence the posterior parameter distributions would be narrow; and independent, where each parameter value would not influence the best parameter value for another parameter (i.e. not exhibit cross-correlation). For most of the experiments, the model was identifiable, there was a narrow range of acceptable parameter values; however, DDF_{Snow} varied across a range of around 3 mm°C^{-1} day^{-1} in all the scenarios. Most of the parameter values were independent of each other, with minor relationships between DDF_{Snow} and other parameters as shown in Figure 11, which shows the correlation between parameters of acceptable solutions for Experiment B (where acceptable solutions have Gelman diagnostic (Gelman *et al*., 2000) statistic < 1.2 for convergence and are part of the second half of solutions). For Experiment B, Figure 12 shows the relationship between DDF_{Snow} and x3. For differing values of DDF_{Snow}, x3 varied around 50 mm, generally with a linear trend (note DDF_{Snow} converged to the

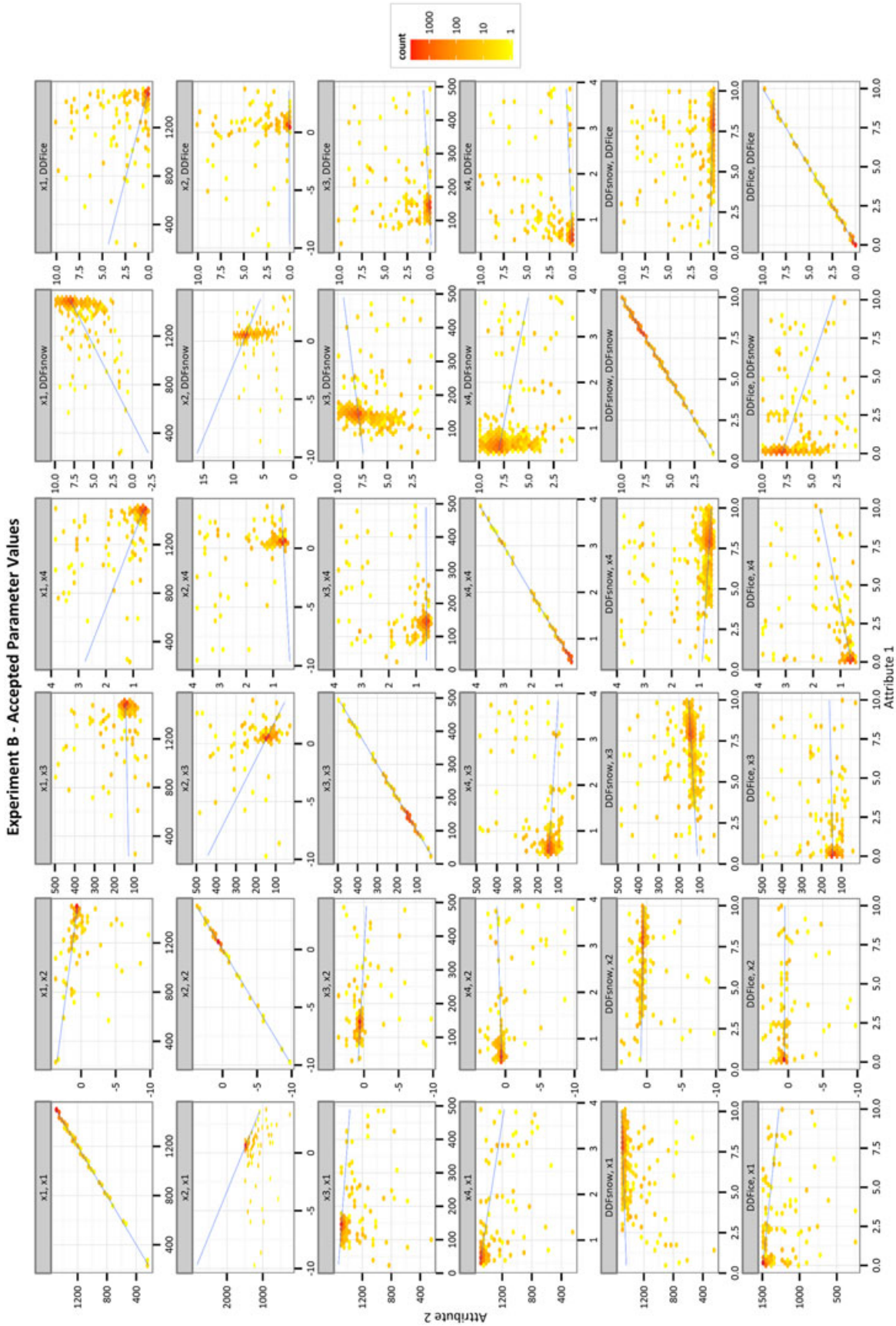

Figure 11. Relationship between pairs of parameters for Experiment B (hexagon density on log scale). Most parameter pairs showed little cross-correlation, with the exception of DDFsnow and other parameters where there was a linear relationship observed (R^2 of 0.3 for DDFsnow and x3)

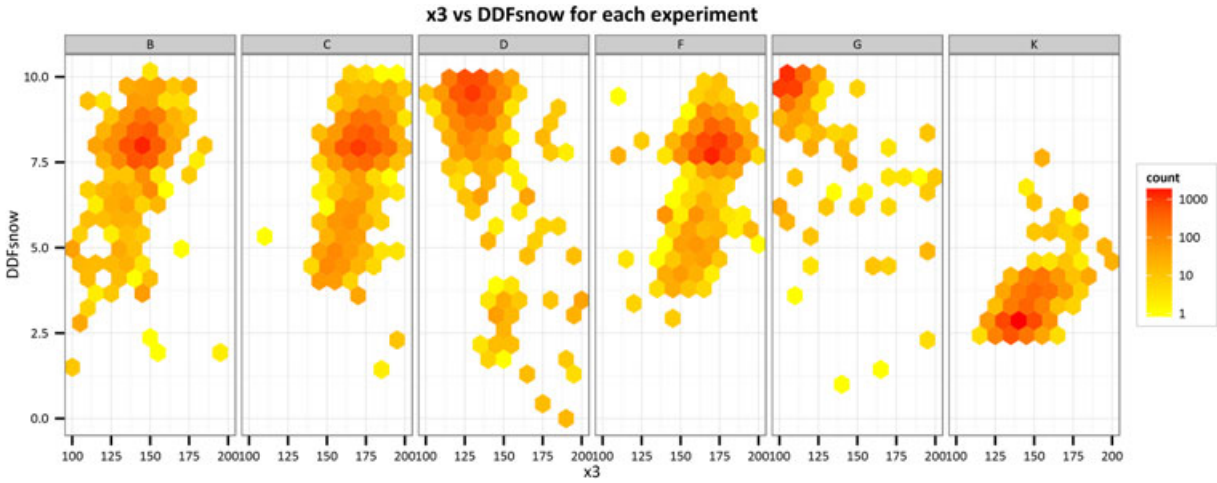

Figure 12. Comparison of DDFSnow and x3 for Experiments B, C, D, F, G and K. Experiments B, C and K have similar characteristics. DDFSnow was on the parameter boundary for Experiments D and G, which may explain the differing relationship in that case. Note, the boundary of Experiment K for DDFsnow was 2.5 mm°C^{-1} day^{-1}, hence the reduced range

parameter boundary in experiments D and K). This would be consistent with the idea that the routing store parameter (x3) compensates for faster runoff generated by a higher DDF_{Snow} value.

In Bayesian inference frameworks, the appropriateness of the likelihood function depends on the model structure and data, and the statistical form of the errors that cause the differences between them. Further information about the structure of errors in the observation series (e.g. rating curve, precipitation) would assist in disaggregating data errors from problems with the conceptual model. Nevertheless, DREAM identified reasonable parameters using likelihood functions designed for other, quite different catchments. Note, producing better posterior parameter densities would involve examination of the residuals in terms of possible autocorrelation and heteroscedasticity (Schoups and Vrugt, 2010). Improved understanding of the prior probability distribution of rainfall and streamflow could lead to improved understanding of parameters and better estimates of the region's water balance.

CONCLUSIONS

This paper described a spatial GR4JSG conceptualization of the glaciated alpine Tamor catchment using elevation bands and degree-day factors. The model conceptualization enabled evaluation of the modelled streamflow and comparison to remotely sensed MODIS snow extent. The agreement of these snow extents to MODIS snow extent for the catchment was R^2 (0.46), with calibration against streamflow only. The evaluation with observed streamflow was NSE (0.88), PBIAS (<4%), using a split-sample calibration and validation method. The model application also provided information about the melt runoff from glacier and non-glacier areas. This study showed the successful implementation of the GR4JSG conceptual model in the data scarce region of the Himalayan region.

However, the application also highlighted a few limitations of the model: (1) sensitivity to initial conditions, (2) slow trends in the internal states, (3) parameters collapsing to boundaries of GR4J's parameter range, (4) insufficient signal to constrain the ice degree-day-factors based on streamflow alone, and (5) interplay of GR4J parameters such as x2 with the degree-day-factors. The model is also highly sensitive to rainfall and temperature driving data, which suffer from known problems and biases. Testing of the model in other Himalayan catchments may reveal additional limitations.

ACKNOWLEDGEMENTS

The Australian Government through the Sustainable Development Investment Portfolio (SDIP) run by the Department of Foreign Affairs and Trade (DFAT) funded this work to ICIMOD and CSIRO. In ICIMOD, this study was undertaken as a part of the Koshi Basin Programme (KBP). This study was partially supported by core funds of ICIMOD contributed by the governments of Afghanistan, Australia, Austria, Bangladesh, Bhutan, China, India, Myanmar, Nepal, Norway, Pakistan, Switzerland and the United Kingdom.

We are grateful for provision of the DREAM algorithm for calibration and uncertainty analysis (Guillaume and Andrews, 2012). We acknowledge the assistance of the Nepal Department of Hydrology and Meteorology and thank the Government of Nepal for allowing us access to hydroclimate data. Geoff Davis and Geoff Adams from eWater assisted with technical aspects associated with the

incorporation of hydrological models into the eWater Source framework. We acknowledge NOAA for MODIS data and the NASA/Glovis programme for providing LANDSAT imagery. We thank the many individuals and reviewers at ICIMOD and CSIRO that assisted with various parts of the project including data preparation and reviewing of manuscripts. We thank the anonymous reviewers for their constructive comments and feedback. The views and interpretations in this publication are those of the author's and they are not necessarily attributable to their organizations.

REFERENCES

Andermann C, Longuevergne L, Bonnet S, Crave A, Davy P, Gloaguen R. 2012. Impact of transient groundwater storage on the discharge of Himalayan rivers. *Nature Geoscience* **5**(2): 127–132.

Argent RM, Perraud JM, Rahman JM, Grayson RB, Podger GM. 2009. A new approach to water quality modelling and environmental decision support systems. *Environmental Modelling & Software* **24**(7): 809–818.

Bajracharya SR, Shrestha B (Eds). 2011. *The status of glaciers in the Hindu Kush-Himalayan region*. ICIMOD: Kathmandu.

Barnett TP, Adam JC, Lettenmaier DP. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature* **438**(7066): 303–309.

Bennett ND, Croke BF, Guariso G, Guillaume JH, Hamilton SH, Jakeman AJ, Marsili-Libelli S, Newham LT, Norton JP, Perrin C, Pierce SA. 2013. Characterising performance of environmental models. *Environmental Modelling & Software* **40**: 1-20. Vancouver.

Beven K, Freer J. 2001. Equifinality, data assimilation, and data uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology* **249**: 11–29.

Box GEP, Cox DR. 1964. An analysis of transformations. *Journal of the Royal Statistical Society: Series B* **26**: 211–252.

Brunt D. 1933. The adiabatic lapse-rate for dry and saturated air. *Quarterly Journal of the Royal Meteorological Society* **59**(252): 351–360.

Chen NS, Hu GS, Deng W, Khanal N, Zhu YH, Han D. 2013. On the water hazards in the trans-boundary Kosi River basin. *Natural Hazards and Earth System Science* **13**(3): 795–808.

Daniel EB, Camp JV, LeBoeuf EJ, Penrod JR, Dobbins JP, Abkowitz MD. 2011. Watershed modeling and its applications: a state-of-the-art review. *Open Hydrology Journal* **5**: 26–50.

Diodato N, Tartari G, Belocchi G. 2010. Geospatial rainfall modeling at eastern Nepalese highland from ground environmental data. *Water Resources Management* **24**: 2703–2720.

Finger D, Vis M, Huss M, Seibert J. 2015. The value of multiple data set calibration versus model complexity for improving the performance of hydrological models in mountain catchments. *Water Resources Research* **51**(4): 1939–1958.

Gao T, Kang S, Krause P, Cuo L, Nepal S. 2012. A test of J2000 model in a glacierized catchment in the central Tibetan Plateau. *Environmental Earth Sciences* **65**(6): 1651–1659.

Gelman A, Goegebeur Y, Tuerlinckx F, Van Mechelen I. 2000. Diagnostic checks for discrete data regression models using posterior predictive simulations. *Applied Statistics* **49**(2): 247–268.

Guillaume J, Andrews F. 2012. dream: DiffeRential Evolution Adaptive Metropolis. R package version 0.4-2. URL http://CRAN.R-project.org/package=dream

Gurung DR, Giriraj A, Aung KS, Shrestha B, Kulkarni AV. 2011. *Snow-cover mapping and monitoring in the Hindu Kush-Himalayas*. ICIMOD: Kathmandu.

Hall DK, Riggs GA, Salomonson VV, DiGirolamo NE, Bayr KJ. 2002. MODIS snow-cover products. *Remote Sensing of Environment* **83**(1): 181–194.

Hock R. 2003. Temperature index melt modelling in mountain areas. *Journal of Hydrology* **282**(1): 104–115.

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, …, Stocker EF. 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *Journal of Hydrometeorology* **8**(1): 38-55.

Immerzeel WW, Petersen L, Ragettli S, Pellicciotti F. 2014. The importance of observed gradients of air temperature and precipitation for modeling runoff from a glacierized watershed in the Nepalese Himalayas. *Water Resources Research* **50**: DOI:10.1002/2013WR014506

Immerzeel WW, Beek LPHV, Konz M, Shrestha AB, Bierkens MFP. 2012. Hydrological response to climate change in a glacierized catchment in the Himalayas. *Climate Change* **110**(3-4): 721–736.

Immerzeel WW, Pellicciotti F, Bierkens MFP. 2013. Rising river flows throughout the twenty-first century in two Himalayan glacierized watersheds. *Nature Geoscience* 6. DOI:10.1038/NGEO1896

Immerzeel WW, van Beek LPH, Bierkens MFP. 2010. Climate change will affect the Asian water towers. *Science* **328**(5984): 1382–5.

Jarvis A, Reuter HI, Nelson A, Guevara E. 2008. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90 m Database (http://srtm.csi.cgiar.org).

Jóhannesson T, Raymond C, Waddington ED. 1989. Time-scale for adjustment of glaciers to changes in mass balance. *Journal of Glaciology* **35**(121): 355–369.

Joseph JF, Guillaume JHA. 2013. Using a parallelized MCMC algorithm in R to identify appropriate likelihood functions for SWAT. *Environmental Modelling & Software* **46**: 292–298.

Kattel DB, Yao T, Yang K, Tian L, Yang G, Joswiak D. 2013. Temperature lapse rate in complex mountain terrain on the southern slope of the central Himalayas. *Theoretical and applied climatology* **113**(3-4): 671–682.

Khadka D, Babel MS, Shrestha S, Tripathi NK. 2014. Climate change impact on glacier and snow melt and runoff in Tamakoshi basin in the Hindu Kush Himalayan (HKH) region. *Journal of Hydrology* **511**: 49–60.

Klein AG, Barnett AC. 2003. Validation of daily MODIS snow cover maps of the Upper Rio Grande River Basin for the 2000–2001 snow year. *Remote Sensing of Environment* **86**(2): 162–176.

Krakauer NY, Pradhanang SM, Lakhankar T, Jha AK. 2013. Evaluating satellite products for precipitation estimation in mountain regions: a case study for Nepal. *Remote Sensing* **5**(8): 4107–4123.

Krause P. 2002. Quantifying the impact of land use changes on the water balance of large catchments using the J2000 model. *Physics and Chemistry of the Earth* **27**: 663–673.

Knoche M, Fischer C, Pohl E, Krause P, Merz R. 2014. Combined uncertainty of hydrological model complexity and satellite-based forcing data evaluated in two data-scarce semi-arid catchments in Ethiopia. *Journal of Hydrology* DOI:10.1016/j.jhydrol.2014.10.003

Kuczera G, Kavetski D, Franks S, Thyer M. 2006. Towards a Bayesian total error analysis of conceptual rainfall–runoff models: characterising model error using storm-dependent parameters. *Journal of Hydrology* **331**(1–2): 161–177. DOI:10.1016/j.jhydrol.2006.05.010

Laloy E, Vrugt JA. 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Research* **48**, W01526. DOI:10.1029/2011WR010608

Le Moine N, Andréassian V, Perrin C, Michel C. 2007. How can rainfall-runoff models handle intercatchment groundwater flows? Theoretical study based on 1040 French catchments. *Water Resources Research* **43**, W06428. DOI:10.1029/2006WR005608

Leavesley GH, Stannard LG. 1995. The precipitation–runoff modeling system — PRMS. In *Computer Models of Watershed Hydrology*, Singh VP (ed.). Water Resources Publications; 281–310.

Lutz AF, Immerzeel WW, Shrestha AB, Bierkens MFP. 2014. Consistent increase in High Asia's runoff due to increasing glacier melt and precipitation. *Nature Climate Change*. DOI:10.1038/nclimate2237

Müller MF, Thompson SE. 2013. Bias adjustment of satellite rainfall data through stochastic modeling: methods development and application to Nepal. *Advances in Water Resources* **60**: 121–134.

Nash J, Sutcliffe JV. 1970. River flow forecasting through conceptual models part I—a discussion of principles. *Journal of Hydrology* **10**(3): 282–290.

Nepal S. 2012. Evaluating upstream–downstream linkages of hydrological dynamics in the Himalayan region. PhD. Thesis. Friedrich Schiller University of Jena, Jena.

Nepal S, Krause P, Flügel W-A, Fink M, Fischer C. 2014. Understanding the hydrological system dynamics of a glaciated alpine catchment in the Himalayan region using the J2000 hydrological model. *Hydrological Processes* **28**(3): 1329–1344.

Nepal S. 2016. Impacts of climate change on the hydrological regime of the Koshi river basin in the Himalayan region. *J. Hydro-environment Res.* **10**: 76–89.

Nepal S, Zheng H, Penton D, Neumann L. 2015. Comparative performance of GR4JSG and J2000 hydrological models in the Dudh Koshi catchment of the Himalayan region. In *MODSIM2015, 21st International Congress on Modelling and Simulation*, Weber T, McPhee MJ, Anderssen RS (eds). Modelling and Simulation Society of Australia and New Zealand: Gold Coast; 2395–2401. ISBN:978-0-9872143-5-5.

Normand S, Konz M, Merz J. 2010. An application of the HBV model to the Tamor Basin in Eastern Nepal. *Journal of Hydrology and Meteorology* **7**(1): 49–58.

Panday PK, Williams CA, Frey KE, Brown ME. 2013. Application and evaluation of a snowmelt runoff model in the Tamor River basin, Eastern Himalaya using a Markov Chain Monte Carlo (MCMC) data assimilation approach. *Hydrological Processes*. DOI:10.1002/hyp.10005

Parajka J, Blöschl G. 2006. Validation of MODIS snow cover images over Austria. *Hydrology and Earth System Sciences Discussions* **3**(4): 1569–1601.

Peeters LJM, Podger GM, Smith T, Pickett T, Bark RH, Cuddy SM. 2014. Robust global sensitivity analysis of a river management model to assess nonlinear and interaction effects. *Hydrology and Earth System Sciences* **18**(9): 3777–3785.

Pellicciotti F, Buergi C, Immerzeel WW, Konz M, Shrestha AB. 2012. Challenges and uncertainties in hydrological modeling of remote Hindu Kush-Karakoram-Himalayan (HKH) basins: suggestions for calibration strategies. *Mountain Research and Development* **32**(1): 39–50.

Pelto MS, Hedlund C. 2001. Terminus behavior and response time of North Cascade glaciers, Washington, USA. *Journal of Glaciology* **47**(158): 497–506.

Perrin C, Michael C, Andreassian V. 2003. Improvement of a parsimonious model for streamflow simulations. *Journal of Hydrology* **279**: 275–289.

Plischke E, Borgonovo E, Smith CL. 2013. Global sensitivity measures from given data. *European Journal of Operational Research* **226**(3): 536–550.

Pokhrel BK, Chevallier P, Andréassian V, Tahir AA, Arnaud Y, Neppel L, Bajracharya OR, Budhathoki KP. 2014. Comparison of two snowmelt modelling approaches in the Dudh Koshi basin (eastern Himalayas, Nepal). *Hydrological Sciences Journal* **59**(8): 1507–1518.

Renard B, Kavetski D, Kuczera G, Thyer MA, Franks SW. 2010. Understanding predictive uncertainty in hydrologic modeling: the challenge of identifying input and structural errors. *Water Resources Research* **46**: W05521. DOI:10.1029/2009WR008328

Rolland C. 2003. Spatial and seasonal variations of air temperature lapse rates in Alpine regions. *Journal of Climate* **16**(7): 1032–1046.

Schoups G, Vrugt JA. 2010. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research* **46**: W10531.

Shea JM, Immerzeel WW, Wagnon P, Vincent C, Bajracharya S. 2015. Modelling glacier change in the Everest region, Nepal Himalaya. *The Cryosphere* **9**(3): 1105–1128.

Singh VP, Frevert DK (Eds). 2002. *Mathematical models of small watershed hydrology and applications*. Water Resources Publications: Highlands Ranch, Colorado.

Thayyen RJ, Gergan JT. 2010. Role of glaciers in watershed hydrology: a preliminary study of a "Himalayan catchment". *The Cryosphere* **4**(1): 115–128.

Valéry A. 2010. *Modélisation précipitations–débit sous influence nivale. Élaboration d'un module neige et évaluation sur 380 bassins versants*. Agro Paris Tech: Paris, France.

Valéry A, Andréassian V, Perrin C. 2014. 'As simple as possible but not simpler': what is useful in a temperature-based snow-accounting routine? Part 1—Comparison of six snow accounting routines on 380 catchments. *Journal of Hydrology* **517**: 1166–1175.

Vaze J, Jordan P, Beecham R, Frost A, Summerell G. 2011. Guidelines for rainfall–runoff modelling: towards best practice model application. *eWater Cooperative Research Centre* http://ewater.org.au/uploads/files/eWater-Guidelines-RRM-(v1_0-Interim-Dec-2011).pdf

Vrugt JA, ter Braak CJF, Diks CGH, Robinson BA, Hyman JM, Higdon D. 2009. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation* **10**(3): 273–290. DOI:10.1515/IJNSNS.2009.10.3.273

Wang W, Liang S, Meyers T. 2008. Validating MODIS land surface temperature products using long-term nighttime ground measurements. *Remote Sensing of Environment* **112**(3): 623–635.

Welsh WD, Vaze J, Dutta D, Rassam D, Rahman JM, Jolly ID, …, Lerat J. 2013. An integrated modelling framework for regulated river systems. *Environmental Modelling & Software* **39**: 81-102.

Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A. 2012. APHRODITE: constructing a long-term daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *Bulletin of American Meteorological Society*. DOI:10.1175/BAMS-D-11-00122.1

