Research papers

Runoff prediction in a poorly gauged basin using isotope-calibrated models

Tsutomu Yamanaka [a,*], Wenchao Ma [b]

[a] Faculty of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8577, Japan
[b] Center for Research in Isotopes and Environmental Dynamics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8577, Japan

ABSTRACT

Predictions in ungauged basins have been a major challenge in hydrologic sciences, and there is still much work needed to achieve robust and reliable predictions for such basins. Here, we propose and test a novel approach for predicting runoff from poorly gauged basins using a minimum complex model calibrated with isotope data alone (i.e., without observed discharge data). The model is composed of two water-stores (soil water and groundwater) and considers their connectivity to runoff in terms of both water and isotope budgets. In a meso-scale basin in which riverbed deformations frequently occur, making automatic observation of river discharge difficult, we measured hydrogen and oxygen isotope composition (δ^2H and δ^{18}O) of precipitation and river water twice-weekly for one year. Runoff predicted by the model agreed well with that observed monthly or bimonthly. Monte Carlo simulation revealed a strong coherence between model performance in isotope simulation and runoff prediction, demonstrating that the use of isotopes as dynamic proxies of calibration targets helps reliably constrain model parameters. Our results indicate that this approach can serve as a powerful tool for prediction of runoff hydrographs, particularly for basins in which the stage-discharge relationship is highly variable.

© 2016 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

River gauging (i.e., measuring water level and, by inference, flow) has been conducted since the foundation of human civilization, as indicated by the Nilometer in ancient Egypt (Dooge, 2004). However, many drainage basins worldwide are ungauged or poorly gauged, and in some cases existing gauging stations are even declining, so that prediction of runoff in such basins is considered to be a challenge (Sivapalan, 2003). The Prediction in Ungauged Basins (PUB) initiative (Sivapalan et al., 2003) launched by the International Association of Hydrological Sciences has resulted in considerable advances in hydrology (Blöschl et al., 2013), although much of this success has been in gauged rather than in ungauged basins; thus, there is still a long way to go to achieve robust and reliable predictions without river gauging (Hrachowitz et al., 2013).

Prediction of runoff hydrographs has been made using a wide spectrum of rainfall-runoff models (Beven, 2012). However, these models all must be calibrated by comparing predicted runoff with observed one (Wagener et al., 2004). Accordingly, if sufficient high

quality records of river discharge are not available for a basin, reliable runoff predictions cannot be made. Measured water levels can usually be converted to discharge using known stage-discharge relations (also known as H-Q curves, rating curves). However, in some basins in which riverbed morphology and stream geometry are frequently changed because of high sediment yield, the stage-discharge relation is variable and the discharge cannot be determined from water level measurements (Herschy, 2009). In this case, it is difficult to calibrate models and thus reliably predict runoff hydrographs.

There are three methods of predicting runoff hydrographs in ungauged basins (Parajka et al., 2013). The first is to estimate model parameters from basin characteristics *a priori*. Although some parameters of completely physics-based models can be well constrained, their performance is generally worse than that of calibrated models (Duan et al., 2006). The second method is to calibrate model parameters for well gauged basins and then apply them to an ungauged basin of interest by extrapolation of the model parameters along with basin characteristics while assuming hydrological similarity between basins. Extrapolations tend to be more reliable if process realisms are held (Tetzlaff et al., 2013); however, it is not easy to confirm the realisms without observation. The third method is to calibrate models using observed proxy data other than (but intrinsically associated with) discharge. For

---

instance, McGuire et al. (2007) calibrated a hillslope model with artificially applied tracer (Br⁻), then obtained reasonable runoff predictions. Tracer data can yield integral fingerprints of flow paths and storage (Leibundgut et al., 2009; McDonnell and Beven, 2014), and provide a powerful test of whether the model is right for the right reasons (Kirchner, 2006). Therefore, the third approach is expected to help assure process realisms in the models, although its feasibility has not been assessed in detail.

Here, we propose a novel approach for predicting river runoff from poorly gauged basins using a model calibrated with environmental stable isotopes of water alone. We selected environmental isotopes as a proxy of calibration targets because they are ideal tracers for water flow/storage and involve information integrated over the basin as well as discharge (Genereux and Hooper, 1998; McDonnell and Beven, 2014). Stable isotope tracers have been conventionally used for hydrograph separation (see Klaus and McDonnell (2013) for overview) and/or mean transit (or residence) time estimation (see McGuire and McDonnell (2006) for overview), and such "soft data" can be used for calibrate/validate runoff models (Seibert and McDonnell, 2002; Stadnyk et al., 2005; Vaché and McDonnell, 2006; Sayama and McDonnell, 2009). However, the present study directly incorporate stable isotopes into runoff model. The objectives of this study were to test the proposed approach for a basin in central Japan, to validate its usefulness and to clarify its limitations.

2. Materials and methods

2.1. Study basin

The study basin was the upstream part of the Midai basin (35.6°N, 138.4°E, 55.2 km²) in Yamanashi Prefecture, central Japan (Fig. 1). The Midai River is a tributary of the Kamanashi River, which is a northwestern branch of the Fuji River. The Midai River is steep, with a bed slope of 1/10–1/50, and is known to have caused severe floods since prehistoric times (Takeuchi, 2004). The Midai River carries a large amount of sediments because it originates from fragile areas in mountains (Ministry of Land, Infrastructure and Transport, Kofu Work Office, 2005). Although the Yamanashi Prefectural Office has a gauging station that routinely measures the water level of Midai River, frequent bed deformations due to strong flow and large sediment yield prevent conversion of the measured water levels to discharge rates. Therefore, it is necessary to establish a novel approach for runoff prediction without discharge data, which is the principal reason we selected this basin.

The study basin has a temperate inland climate with warm rainy seasons and cold dry seasons. At Kofu station of the Japan Meteorological Agency (JMA), which is adjacent to Midai basin, the long-term mean annual precipitation is 1135 mm and the annual mean temperature (mean annual range) is 14.7 °C (23.8 °C). Relatively light snow accumulation is common. The vegetation is dominated by natural *Quercus crispula* stands and *Pinus densiflora* plantations.

Topography is characterized by a large elevation range exceeding 2000 m and steep slopes with a mean of 32.4° (Table 1). Geology is dominated by andesite and mudstone (including shale/slate). Forest, the dominant land cover, occupies 73.5% of the total area. The study basin is not an experimental one, and no further information about runoff-related characteristics are available.

2.2. Observation

For isotopic measurements, we collected river water and precipitation samples twice-weekly (i.e., 3 or 4 days intervals) for one year from 1 May 2011 to 30 April 2012. River water sampling was conducted at a water purification plant that receives water from the Midai River. Precipitation was sampled at an open space within the plant using a precipitation collector with a polyethylene funnel that was designed to prevent evaporation of stored water (Yamanaka et al., 2015). Hydrogen and oxygen stable isotope ratios (^{2}H/^{1}H and ^{18}O/^{16}O) in all samples were measured by tunable-diode laser spectroscopy using a liquid water isotope analyzer (L1102-i, Picarro Inc., Santa Clara, CA, USA) at the Center for Research in Isotopes and Environmental Dynamics, University of Tsukuba. Measurement results are expressed using δ notation (δ^2H and δ^{18}O) relative to the Vienna Standard Mean Ocean Water (V-SMOW). Measurement errors for the analyzer were 1.0‰ for δ^2H and 0.1‰ for δ^{18}O (Yamanaka and Onda, 2011).

River discharge was observed monthly (or bimonthly) by the velocity-area method (Herschy, 2009) at a site 1.8 km downstream from the sampling site. Flow velocity was measured by an electromagnetic flowmeter (AEM1-D, JFE Advantech, Nishinomiya, Japan) at 0.6 of the depth from the water surface for each vertical. The river stream usually branched off within a channel, and the number and location of small branches within the channel was not fixed owing to the riverbed deformation. Thus, we measured discharge for all the branches and then summed up. On 9 Sep. 2011, flow velocity could not be measured because the flow was too fast for safe wading, and thus discharge data at that time are not available.

To confirm the possibility of reconstructing continuous discharge data, we referred to water level data observed in close proximity to our site at 10 min intervals by the Yamanashi Prefectural Office. However, the stage-discharge relationship was not fixed month by month.

2.3. Model

The exponential tank model originally developed by Kondo et al. (1992) was modified and employed in the present study. The model is composed of two tanks, A and B (Fig. 2). Tank A represents soil water storage and tank B represents groundwater storage. Water storage in the tank (S_A, mm) and flow from the tank (F_A, mm d⁻¹) at a time step j were calculated as follows:

$$S_A^j = S_A^{j-1} + d_A^j - E_T^j \Delta t, \tag{1}$$

$$F_A^j = (1 - f_I) P_G^j - d_A^j / \Delta t, \tag{2}$$

where E_T is the evapotranspiration (mm d⁻¹), f_I is the interception fraction, P_G is the gross rainfall (mm d⁻¹), Δt is the time interval (=1 d), and superscripts j and $j - 1$ denote time steps. The remaining portion of precipitation in tank A, d_A (mm), is represented as:

$$d_A^j = (S_{A\max} - S_A^{j-1})\left\{ 1 - \mathrm{Exp}\left[-(1 - f_I) P_G^j \Delta t / S_{A\max}\right] \right\}, \tag{3}$$

where $S_{A\max}$ is a model parameter representing the maximum storage of tank A (mm).

Similarly, the water budget in tank B can be given as follows:

$$S_B^j = S_B^{j-1} + d_B^j - F_C^j \Delta t, \tag{4}$$

$$F_B^j = F_A^j - d_B^j / \Delta t, \tag{5}$$

$$d_B^j = (S_{B\max} - S_B^{j-1})\left[1 - \mathrm{Exp}(-F_A^j \Delta t / S_{B\max}) \right], \tag{6}$$

where S_B is the water storage in tank B (mm), d_B is the remaining portion of F_a in tank B (mm), and $S_{B\max}$ is the maximum storage of tank B (mm). A base flow component, F_C (mm d⁻¹), is given as:

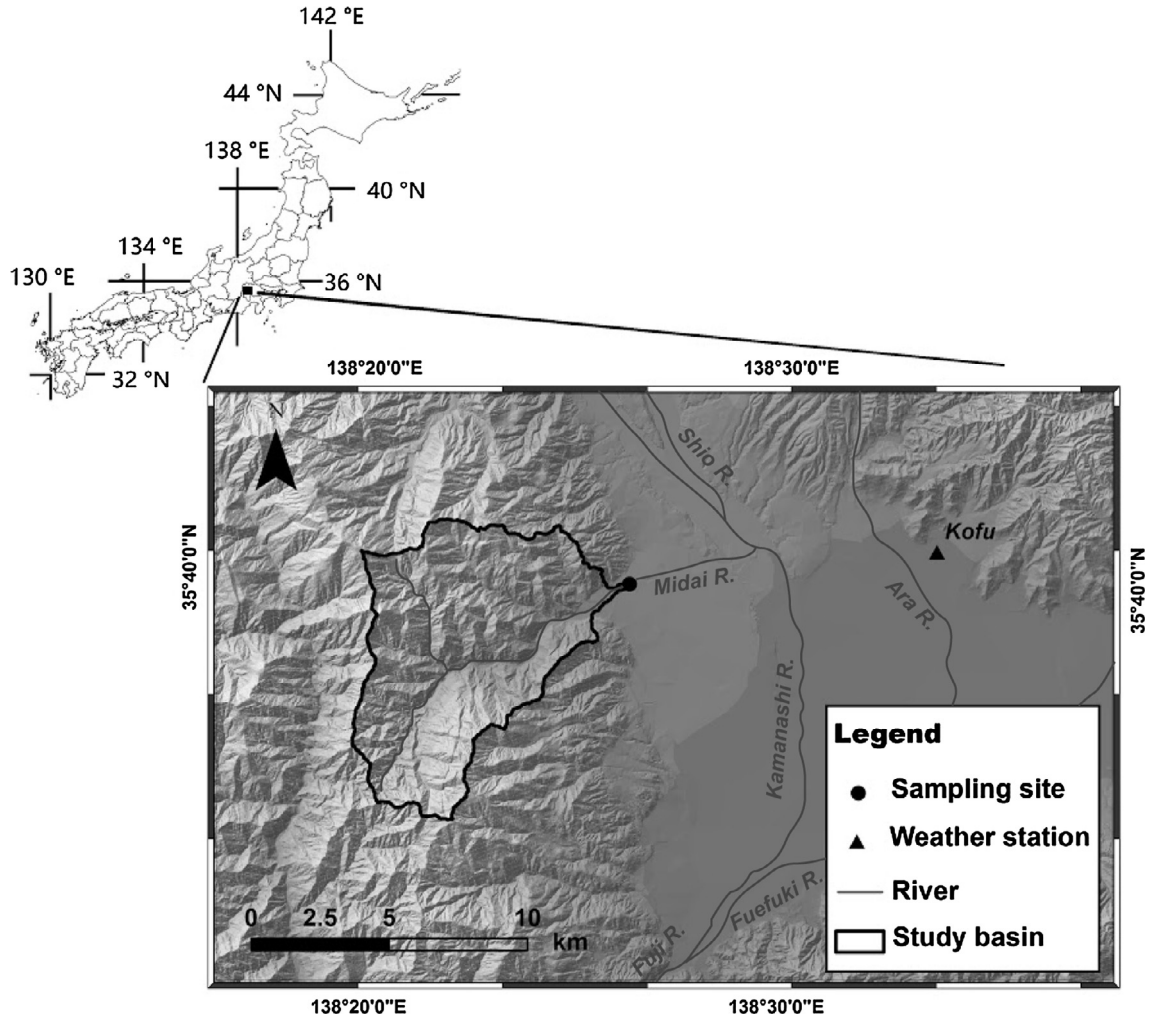$$F_C^j = k_C (S_B^{j-1})^2, \tag{7}$$

Fig. 1. Map of the study basin and locations of sampling site and weather station.

Table 1
Summary of characteristics of the study basin.

Item	Unit	Value
Area	km^2	55.2
Topography		
Mean elevation	m	1362.7
Minimum elevation	m	476.0
Maximum elevation	m	2492.4
Mean slope		32.4
Maximum slope		82.6
Geology		
Andesite	%	56.6
Mudstone (shale, slate)	%	29.7
Porphyrite	%	7.5
Diluvial gravel	%	4.6
Others	%	1.6
Land use/land cover		
Forest	%	73.5
Residential area	%	10.1
Agricultural land	%	9.0
Others	%	7.4

where k_C (mm^{-1} d^{-1}) is a factor assuming that base flow is proportional to the square of storage height, and represents drainage ability of the basin. Finally, river discharge, Q (mm d^{-1}), can be given as the sum of quick flow and base flow as follows:
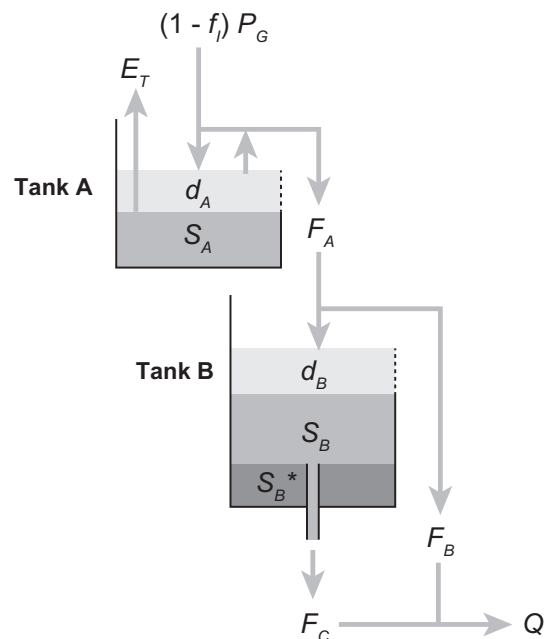
$$Q^j = F_B^j + F_C^j. \tag{8}$$

Fig. 2. Schematic illustration of the model (see text for definition of symbols).

We introduced additional procedures into the tank model for computing isotope mass balance, as explained below. The δ values of water in tank A (δ_A, ‰), tank B (δ_B, ‰) and river water (δ_R, ‰) were computed as follows:

$$\delta_A^j = \left\lfloor \delta_A^{j-1} S_A^{j-1} + \delta_P^j (1 - f_I) P_G^j \Delta t - \delta_{FA}^j F_A^j \Delta t - \delta_A^{j-1} E_T^j \Delta t \right\rfloor \Big/ S_A^j, \tag{9}$$

$$\delta_B^j = \left\lfloor \delta_B^{j-1}(S_B^{j-1} + S_B^*) + \delta_{FA}^j d_B^j \Delta t - \delta_B^{j-1} F_C^j \Delta t \right\rfloor \Big/ \left(S_B^j + S_B^* \right), \tag{10}$$

$$\delta_R^j = \left(\delta_{FA}^j F_B^j + \delta_B^j F_C^j \right) \Big/ Q^j, \tag{11}$$

where δ_P is the isotopic composition of precipitation (‰), and S_B^* is a model parameter (mm) representing the water storage portion that does not affect the base flow-storage relationship but contributes to isotopic variations in groundwater storage (Barnes and Bonell, 1996). We assumed the isotopic composition of F_A was as follows:

$$\delta_{FA}^j = (1 - f_{hc})\delta_P^j + f_{hc}\delta_A^j \tag{12}$$

where f_{hc} is a parameter representing the hydrologic connectivity of soil water to runoff. This treatment is a reflection of incomplete mixing of precipitation and soil water during groundwater recharge and runoff generation. Recently, Good et al. (2015) suggested that partial disconnection between mobile and immobile soil water should be considered in isotope budget studies. Brooks et al. (2009) and Evaristo et al. (2015) showed isotopic evidence of such a disconnection. At hillslope scale, density of macro pores for preferential flow can control the disconnection. At the basin scale, we may have to consider connectivity between landscape units (Burt, 2005) such as hillslope-riparian-stream connectivity (Jencso et al., 2010; Tetzlaff et al., 2014). At or near the riparian zones, saturation-excess overland flow tends to occur easily, as described by well-known variable source area concept (Hewlett and Hibbert, 1967; Ward and Robinson, 1990); thus, mixing of soil water with runoff water is limited (Birkel et al., 2015). This is also true if infiltration-excess overland flow occurs. In the present model, f_{hc} implicitly represents all such factors causing incomplete mixing of precipitation and soil water.

2.4. Input data and calibration

The JMA Radar-AMeDAS (Automated Meteorological Data Acquisition System) precipitation data (e.g., Makihara, 1996) with a spatial resolution of 1 km were employed for computation of basin-mean daily precipitation as P_G. The f_I was estimated to be 0.128 as an area-weighted average of typical f_I for every land use/vegetation types within the study basin (Ma and Yamanaka, 2016). The E_T was estimated by the FAO Penman–Monteith method (Allen et al., 1998) using daily meteorological data (solar radiation, air temperature, relative humidity, and wind speed) observed at Kofu station (Fig. 1) of JMA. We assumed that the crop coefficient is unity. Validity of the estimated E_T for several different basins adjacent to the present study basin was confirmed by Ma and Yamanaka (2016). Daily δ_P data were reconstructed from data observed twice a week under an assumption that the precipitation δ value does not change during each sampling period (i.e., 3 or 4 days). In addition, δ_P was corrected considering isotopic lapse rates (−11.66‰ km^{-1} for δ^{2}H and −1.724‰ km^{-1} for δ^{18}O; Yamanaka et al., 2015), and the difference between the basin-mean elevation (1362.7 m) and the sampling site elevation (476 m). Consequently, basin-mean δ_P values are 10.34‰ (1.529‰) lower in δ^{2}H (δ^{18}O) than the observed at the sampling site.

We performed a simulation for a period from 1 January 2010 to 30 September 2012, using the first 1-yr period for spin-up. Initial

values of S_A and S_B were set to 500 mm and 100 mm, respectively. Initial values of both δ_A and δ_B were set to −95‰ for δ^{2}H and −13‰ for δ^{18}O. The δ_P values for periods other than the observation period (i.e., 1 May 2011 to 30 April 2012) were complemented by assuming the same annual cycle. In other words, we assumed the same δ_P value for the same day of year (DOY). Though this assumption may not be always true, biases in simulated δ_R were expected to be the largest in May 2011 and we confirmed no such a bias in that period.

After this, five unknown parameters remained, S_{Amax}, S_{Bmax}, k_C, S_B^*, and f_{hc}. The values of these parameters were fixed through calibration by stochastic hill climbing optimization (Brownlee, 2012), with a measure of the root mean square error (RMSE) of δ_R for δ^{2}H and δ^{18}O. In this optimization, multiple restarts with different initial parameter-sets were adopted to confirm whether the result does not become trapped in local optima. To avoid unrealistic drifts in water budget, we added the following constraint.

$$\sum (f_I P_G - E_T - Q) < 100, \tag{13}$$

where Σ denotes the annual total; although the constraint value of 100 does not have any physical meanings, we confirmed that optimum values of unknown parameters (especially S_{Amax} and S_{Bmax}) tend to unrealistically diverge when we used the larger values. In addition, f_{hc} should range from 0 to 1, and the other 4 parameters as well as S_A and S_B should have positive signs.

The authors of the present study have proposed an isotope-calibrated five-layer tank model for estimating time-variant transit time (Ma and Yamanaka, 2013, 2016). However, the model contains too many parameters to calibrate without discharge data. The reason we developed and employed the minimum complex model comprising only two tanks with five unknown parameters was to reduce the equifinality problem (Beven and Binley, 1992). Reducing model layers implies that accuracy of transit time (if we estimate) will be reduced. However, it is expected that reproducibility of δ_R is less affected, because isotopic contents in deeper layers are not very different generally.

2.5. Monte Carlo simulation

Simulation results from calibrated (i.e., optimized) models still involve uncertainties in both isotope and runoff predictions. In addition, higher performance in isotope prediction is not always associated with that in runoff prediction. To confirm the relationship between reproducibility of predicted isotopes and runoff, we performed Monte Carlo simulations after calibrating unknown parameters. A total of 1000 runs were made with different randomly selected values of S_{Amax}, S_{Bmax}, and k_C, within the range of optimum values ±100%. Because S_B^* and f_{hc} do not affect runoff prediction (i.e., no appearance in Eqs. (1)–(8)), they were not considered.

In addition, three sets of 100 runs were made with randomly selected values of one out of S_{Amax}, S_{Bmax}, and k_C to reveal the relationship of uncertainty in each parameters with model performance.

3. Results

3.1. Isotope reproducibility

Fig. 3 shows the daily precipitation amount, elevation-corrected δ_P, and observed δ_R. The corrected δ_P ranged from −31.8 to −153.6‰ for δ^{2}H and −2.57 to −20.44‰ for δ^{18}O, with lower values in winter and higher values in summer. Variations in observed δ_R were less than 10% of those, and the annual cycle corresponding to those for precipitation was not clear. However, moderate
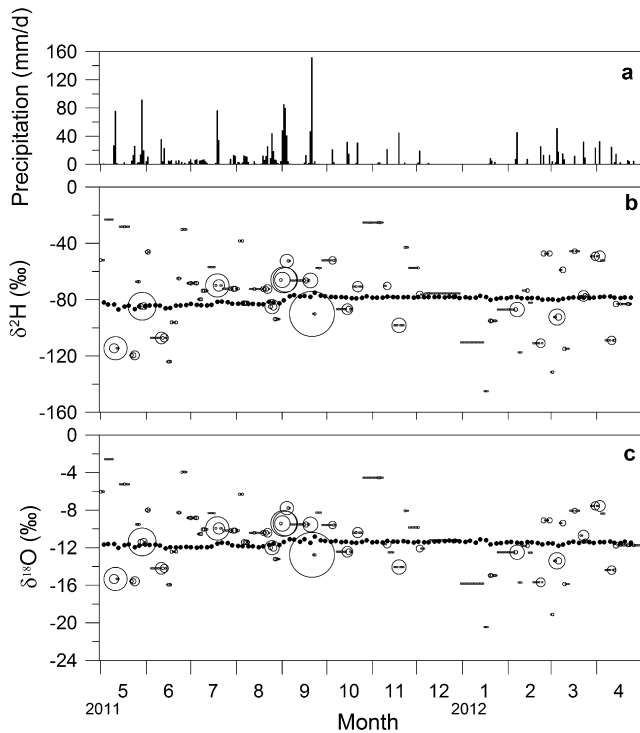
Fig. 3. Temporal variations in (a) daily precipitation amount and daily values of (b) δ^2H and (c) $\delta^{18}O$ for precipitation (closed circle) and river water (open circle, scaled by precipitation amount).
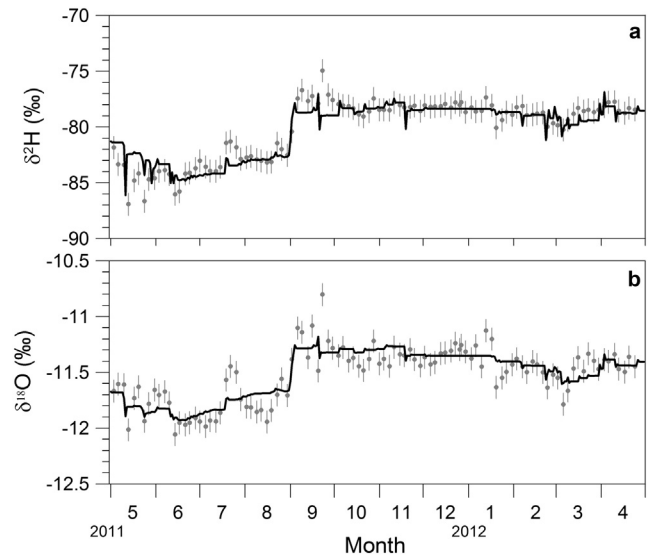

Fig. 4. Comparison of observed (grey closed circle with error bars indicating measurement accuracy) and predicted (black thick line) values of (a) δ^2H and (b) $\delta^{18}O$ for river water.

increases in summer and short-term changes throughout the year were found in observed δ_R.

These tendencies were captured well by our simulations (Fig. 4). The coefficient of determination (R^2) between observed and simulated δ_R was 0.84 for δ^2H and 0.73 for $\delta^{18}O$. The RMSE of δ_R was 1.1‰ for δ^2H and 0.13‰ for $\delta^{18}O$, being almost equivalent to (but slightly worse than) the measurement accuracy of the isotope analyzer used. The ratio of RMSE to the standard deviation of observed data (RSR) was 0.40 for δ^2H and 0.52 for $\delta^{18}O$. According to a performance rating proposed by Moriasi et al. (2007), these figures correspond to 'very good' ($0.00 \leqslant RSR \leqslant 0.50$) or 'good' ($0.50 < RSR \leqslant 0.60$).

Errors in simulated δ_R can be attributed to input data of δ_P with coarse temporal resolution and spatial heterogeneity, E_T with low accuracy, and f_I as a constant first approximation, as well as model structure and parameterizations. Although isotopic enrichment caused by soil surface evaporation, which was not considered in this model, could potentially introduce errors in simulated δ_R, especially for $\delta^{18}O$, our preliminary assessment considering this effect (see Ma and Yamanaka (2016) for inclusion into tank models) showed no improvements, suggesting that it is a minor factor, at least for the study basin, or was cancelled by errors in assumed δ_P and/or initial values of δ_A and δ_B.

Calibrated values of unknown model parameters are summarized in Table 2. For S_{Amax}, S_{Bmax}, and k_C, the difference between the best values in δ^2H- and $\delta^{18}O$-calibarion were relatively large, indicating uncertainties in model structure and/or parameter optimization. However, those were nearly the same for S_B^* and f_{hc}, suggesting robustness of the calibrated values.

3.2. Runoff prediction

Runoff predicted by the δ^2H ($\delta^{18}O$)-calibrated model agreed well with the observed river discharge (Fig. 5), with an R^2 of 0.96

(0.93). The Nash-Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970) of the δ^2H ($\delta^{18}O$)-calibrated model was 0.86 (0.92), corresponding to 'very good' ($0.75 \leqslant NSE \leqslant 1.00$) performance (Moriasi et al., 2007). Because the number of available data (=9) is not sufficient, absolute performance rating may not be appropriate. However, relative significance of calibrated models can be evaluated using NSE as shown below.

Fig. 6 shows the relationships between RSR in isotope simulation (RSR$_i$) and NSE in runoff prediction (NSE$_r$) from 1000 Monte Carlo runs. Strong coherence between them was observed for the entire domain (Fig. 6a), whereas data plots were scattered in a smaller domain around the best results (Fig. 6b). For runoff prediction, slightly better results than the isotopically optimized prediction were obtained, even if isotope reproducibility was worse. This means that isotopically optimized model is not always the best one for runoff prediction. However, all 'very good' runs for isotope simulation were also 'very good' for runoff prediction, demonstrating that better isotopic calibration deduced acceptable better model for runoff prediction.

Both RSR$_i$ and NSE$_r$ were insensitive to changes in S_{Amax}, although the best values were slightly different between isotope simulation and runoff prediction (Fig. 7a). Changes in S_{Bmax} introduced large changes in RSR$_i$, but smaller changes in NSE$_r$ (Fig. 7b). Changes in k_C had greater effects on both isotope simulation and runoff prediction than the other two parameters (Fig. 7c). Sensitivity of NSE$_r$ to k_C changes was greater around the best k_C value for isotope simulation, suggesting the importance of local optimization for improving runoff prediction. These results suggest that isotope simulation is highly sensitive to base flow, but less sensitive to soil water storage associated with quick flow. The lower sensitivity of the model to soil water was likely caused by lower temporal resolution (i.e., twice-weekly) of input isotope data. In contrast, runoff prediction appears to be sensitive to both components, but less sensitive to the groundwater storage associated with quick flow. For a model with the best parameter values in δ^2H ($\delta^{18}O$) calibration, base flow (i.e., F_C) accounted for more than 76% (93%) of the total runoff even under peak flow conditions. This explains why both isotope simulation and runoff prediction are highly sensitive to a base flow related parameter, k_C.

Table 2
Calibrated values of unknown model parameters.

Target	S_{Amax}	S_{Bmax}	k_C	$S_B{}^*$	f_{hc}
δ^2H	344.02	8868.01	3.895×10^{-4}	314.93	0.2787
$\delta^{18}O$	644.66	72076.01	9.025×10^{-5}	343.17	0.2282
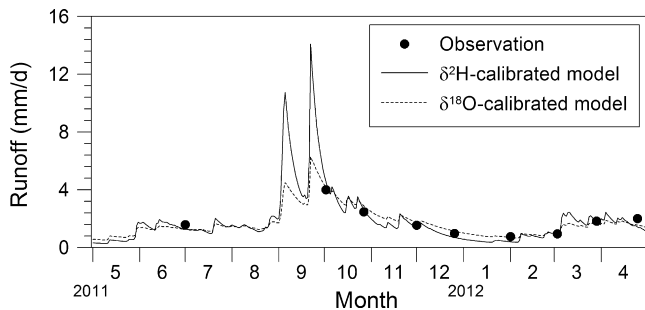
Fig. 5. Comparison of observed (closed circle) and predicted (solid line, δ^2H-calibrated model; dashed line, $\delta^{18}O$-calibrated model) runoff.

4. Discussion

4.1. Usefulness of isotope calibration for runoff prediction

Some studies have partially demonstrated that incorporation of isotopes into runoff models helped reduce the uncertainty of model parameters (Dunn et al., 2008; Fenicia et al., 2008; Birkel et al., 2010a, 2010b, 2011; Ma and Yamanaka, 2013, 2016). However, it has also been argued that incorporating isotope (or other tracer) data may not always improve models owing to uncertainties associated with additional parameters required in isotope simulations (Seibert et al., 2003; Wissmeier and Uhlenbrook, 2007; McDonnell et al., 2010). As shown in Fig. 6, better models for isotope simulation can provide better prediction of runoff, confirming that isotope calibration is useful in practice for runoff prediction.

Isotope calibration can constrain more accurately the base flow prediction than the quick flow prediction. Thus, this approach is expected to be useful for basins where base flow generally dominates, like the present study basin (base flow contribution > 76%). However, in quick flow dominated basins, inconsistencies among the best parameter values between isotope simulation and runoff prediction may result in worse runoff prediction. One potential factor causing the uncertainty of quick-flow-related parameters is the

temporal resolution of isotope data. Twice-weekly data cannot capture quick flow components exactly. Isotope data with higher temporal resolution (e.g., daily or sub-daily) have the potential to improve runoff prediction by isotope-calibrated models (Birkel et al., 2010a). Advanced tunable diode laser spectroscopy could enable isotope analysis at low cost for a large number of samples (Lyon et al., 2009) or even high frequency (e.g., hourly) monitoring in situ, increasing the usefulness of our approach.

According to McDonnell and Beven (2014), the flow velocities in the system control the tracer response and the celerities control the hydrograph. In addition, velocities are controlled by the water storage filled in the system, while celerities are controlled by the storage deficit. The results that model calibration with single source of information (not discharge but isotopes) was useful suggest a constant (i.e., temporally stable) relationship between velocities and celerities in the study basin. It is also indicated that our model structure helped successfully represent the above two mechanisms. In other words, it should be noted that usefulness of isotope calibration for runoff prediction may depend on constancy of the velocity-celerity relationship in the basin and the selection of model structure.

4.2. Limitations and ways forward

Use of dual isotopes (δ^2H and $\delta^{18}O$) allows us to know the degrees of uncertainties associated with the calibrated model. However, it is still difficult to estimate the reliability of runoff prediction. For instance, if the model structure as presented herein was not suitable for a basin of interest, predicted runoff may include non-negligible errors despite there being good reproducibility in isotope simulation. Therefore, several or more observations of discharge in different stage conditions are needed to confirm the validity of the model. Once this is accomplished, predicted runoff is expected to be reliable, regardless of changes in riverbed morphology.

However, as shown in Fig. 5, uncertainty in peak runoff prediction is relatively large. Although it should be noted that hydrometric measurements (Herschy, 2009) and other runoff models (Beven,
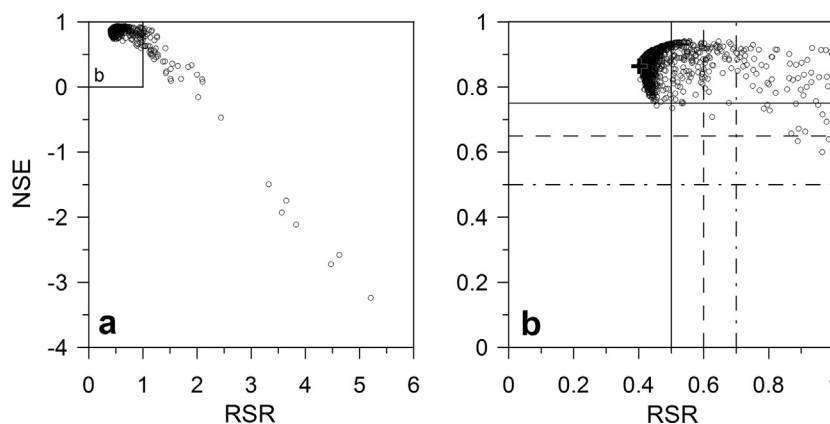

Fig. 6. Relationship between the ratio of RMSE to the standard deviation of observed data (RSR) for isotope simulation and Nash-Sutcliffe efficiency (NSE) for runoff prediction in 1000 Monte Carlo runs using a δ^2H-calibrated model. In the enlarged panel (b), the solid line denotes the boundary between very good and good, the dashed line denotes that between good and satisfactory, and the dot-dashed line denotes that between satisfactory and unsatisfactory.
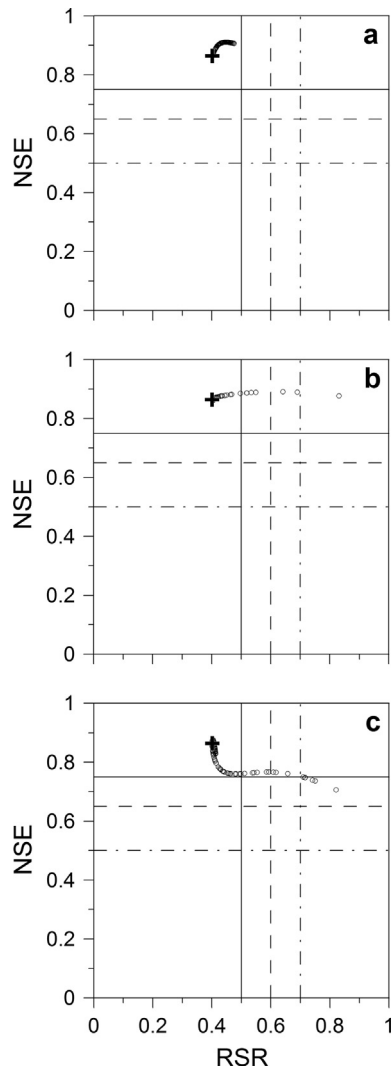
Fig. 7. The same data as shown in Fig. 6, but based on 100 Monte Carlo runs with variable parameters (a) S_{Amax}, (b) S_{Bmax}, and (c) k_C.

2012) tend to include large errors in peak discharge, intrinsic errors in isotope-calibrated models should be carefully examined. Quick flow (i.e., F_B) contributions in the model became large before the peak for every event, after which base flow dominated, suggesting that errors in predicted peak runoff are associated with both components rather than either one alone. Therefore, determination of k_C appears to be the most important aspect, because this parameter is more strongly influenced by slight changes in isotopic reproducibility (see Section 3.2). In other words, more data are needed for calibration to ensure better parameter optimization. Long-term isotope records that include large storm events will be useful for improving the accuracy of peak runoff prediction.

Calibration resulted in f_{hc} values of 0.2–0.3 (Table 2). Interestingly, these values are comparable to a global estimate (0.38 ± 0.28) by Good et al. (2015). Of course, this consistency is physically less meaningful, and the f_{hc} is not always constant. Nevertheless, these findings indicate the need for more attention to the causes and mechanisms of the weak connectivity between soil water and runoff. Further investigations in both headwater small catchments and meso-scale basins with diverse topographical/geological settings will help improve our understanding of catchment hydrology and conceptualization of model structure for isotope calibration.

5. Conclusions

From a practical point of view, we developed a novel approach that calibrate a minimum complex runoff model using isotopic data alone, and showed its ability to predict runoff without hydrometric calibration. Model performance in isotope simulation was closely related to that in runoff prediction, demonstrating the model is right for the right reasons. Thus, we conclude that isotope calibration helps reliably constrain model parameters, and that this approach is useful for runoff prediction in poorly gauged basins such as the study basin, where riverbed deformations frequently occur.

Uncertainties still remain, especially with respect to constraining quick-flow-related parameters and predicting peak runoff. Higher frequency and/or longer-term isotope data are expected to improve these uncertainties. The applicability of the model structure in this study (i.e., two-layer tank representing soil water and groundwater storage) should be further investigated for other basins (including well gauged ones), with a focus on hydrological connectivity on a basin scale.

References

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop Evapotranspiration: Guidelines for Computing Crop Requirements. FAO Irrigation and Drainage Paper No. 56, FAO, Rome.

Barnes, C.J., Bonell, M., 1996. Application of unit hydrograph techniques to solute transport in catchments. Hydorol. Process. 10, 793–802. http://dx.doi.org/10.1002/(SICI)1099-1085(199606)10:6<793::AID-HYP372>3.0.CO;2-K.

Beven, K., 2012. Rainfall-Runoff Modelling: The Primer. Willey-Blackwell, Chichester.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrol. Process. 6, 279–298. http://dx.doi.org/10.1002/hyp.3360060305.

Birkel, C., Dunn, S.M., Tetzlaff, D., Soulsby, C., 2010a. Assessing the value of high-resolution isotope tracer data in the stepwise development of a lumped conceptual rainfall–runoff model. Hydrol. Process. 24, 2335–2348. http://dx.doi.org/10.1002/hyp.7763.

Birkel, C., Tetzlaff, D., Dunn, S.M., Soulsby, C., 2010b. Towards a simple dynamic process conceptualization in rainfall–runoff models using multi-criteria calibration and tracers in temperate, upland catchments. Hydrol. Process. 24, 260–275. http://dx.doi.org/10.1002/hyp.7478.

Birkel, C., Soulsby, C., Tetzlaff, D., 2011. Modelling catchment-scale water storage dynamics: reconciling dynamic storage with tracer-inferred passive storage. Hydrol. Process. 25, 3924–3936. http://dx.doi.org/10.1002/hyp.8201.

Birkel, C., Soulsby, C., Tetzlaff, D., 2015. Conceptual modelling to assess how the interplay of hydrological connectivity, catchment storage and tracer dynamics controls non-stationary water age estimates. Hydorl. Process. 29, 2956–2969. http://dx.doi.org/10.1002/hyp.10414.

Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., Savenije, H., 2013. Runoff Prediction in Ungauged Basins: Synthesis across Processes, Places and Scales. Cambridge University Press, Cambridge.

Brooks, J.R., Barnard, H.R., Coulombe, R., McDonnell, J.J., 2009. Ecohydrologic separation of water between trees and streams in a Mediterranean climate. Nat. Geosci. 3, 100–104. http://dx.doi.org/10.1038/ngeo722.

Brownlee, J., 2012. Clever Algorithms: Nature-Inspired Programming Recipes. Lulu, Raleigh.

Burt, T.P., 2005. A third paradox in catchment hydrology and biogeochemistry: decoupling in the riparian zone. Hydrol. Process. 19, 2087–2089. http://dx.doi.org/10.1002/hyp.5904.

Dooge, J.C.I., 2004. Background to modern hydrology. In: Rodda, J.C., Ubertini, L. (Eds.), The Basis of Civilization: Water Science? International Association of Hydrological Sciences Publication No. 286, pp. 3–12.

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H.V., Gusev, Y.M., Habets, F., Hall, A., Hay, L., Hogue, T., Huang, M., Leavesley, G., Liang, X., Nasonova, O.N., Noilhan, J., Oudin, L., Sorooshian, S., Wagener, T., Wood, E.F., 2006. Model parameter estimation experiment (MOPEX): an overview of science strategy and major results of the second and third workshops. J. Hydrol. 320, 3–17. http://dx.doi.org/10.1016/j.jhydrol.2005.07.031.

Dunn, S.M., Bacon, J.R., Soulsby, C., Tetzlaff, D., Stutter, M.I., Waldron, S., Malcolm, I. A., 2008. Interpretation of homogeneity in δ^{18}O signatures of stream water in a nested sub-catchment system in north-east Scotland. Hydrol. Process. 22, 4767–4782. http://dx.doi.org/10.1002/hyp.7088.

Evaristo, J., Jasechko, S., McDonnell, J.J., 2015. Global separation of plant transpiration from groundwater and streamflow. Nature 525, 91–94. http://dx.doi.org/10.1038/nature14983.

Fenicia, F., McDonnell, J.J., Savenije, H.H.G., 2008. Learning from model improvement: on the contribution of complementary data to process understanding. Water Resour. Res. 44, W06419. http://dx.doi.org/10.1029/2007WR006386.

Genereux, D.P., Hooper, R.P., 1998. Oxygen and hydrogen isotopes in rainfall-runoff studies. In: Kendall, C., McDonnell, J.J. (Eds.), Isotope Tracers in Catchment Hydrology. Elsevier Science B.V., Amsterdam, pp. 319–346.

Good, S.P., Noone, D., Bowen, G., 2015. Hydrologic connectivity constrains partitioning of global terrestrial water fluxes. Science 349, 175–177. http://dx.doi.org/10.1126/science.aaa5931.

Herschy, R.W., 2009. Streamflow Measurement. CRC Press, Boca Raton.

Hewlett, J.D., Hibbert, A.R., 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In: Sopper, W.E., Lull, H.W. (Eds.), Forest Hydrology. Pergamon Press, New York, pp. 275–290.

Hrachowitz, M., Savenije, H.H.G., Blöschl, G., McDonnell, J.J., Sivapalan, M., Pomeroy, J.W., Arheimer, B., Blume, T., Clark, M.P., Ehret, U., Fenicia, F., Freer, J.E., Gelfan, A., Gupta, H.V., Hughes, D.A., Hut, R.W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P.A., Uhlenbrook, S., Wagener, T., Winsemius, H.C., Woods, R.A., Zehe, E., Cudennec, C., 2013. A decade of Predictions in Ungauged Basins (PUB): a review. Hydrol. Sci. J. 58, 1198–1255. http://dx.doi.org/10.1080/02626667.2013.803183.

Jencso, K.G., McGlynn, B.L., Gooseff, M.N., Bencala, K.E., Wondzell, S.M., 2010. Hillslope hydrologic connectivity controls riparian groundwater turnover: implications of catchment structure for riparian buffering and stream water sources. Water Resour. Res. 46, W10524. http://dx.doi.org/10.1029/2009WR008818.

Kirchner, J.W., 2006. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. Water Resour. Res. 42, W03S04. http://dx.doi.org/10.1029/2005WR004362.

Klaus, J., McDonnell, J.J., 2013. Hydrograph separation using stable isotopes: review and evaluation. J. Hydrol. 505, 47–64. http://dx.doi.org/10.1016/j.jhydrol.2013.09.006.

Kondo, J., Watanabe, T., Nakazono, M., 1992. Heat budget evaluation of forest evapotranspiration in Japan. Tenki 39, 685–695.

Leibundgut, C., Maloszewski, P., Külls, C., 2009. Tracers in Hydrology. Willey-Blackwell, Chichester.

Lyon, S.W., Desilets, S.L.E., Troch, P.A., 2009. A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus δ^{18}O. Hydrol. Process. 23, 2095–2101. http://dx.doi.org/10.1002/hyp.7326.

Ma, W., Yamanaka, T., 2013. Temporal variability of mean transit time and transit time distribution assessed by a tracer-aided tank model for a meso-scale catchment. Hydrol. Res. Lett. 7 (4), 104–109. http://dx.doi.org/10.3178/hrl.7.104.

Ma, W., Yamanaka, T., 2016. Factors controlling inter-catchment variation of mean transit time with consideration of temporal variability. J. Hydrol. 534, 193–204. http://dx.doi.org/10.1016/j.jhydrol.2015.12.061.

Makihara, Y., 1996. A method for improving radar estimates of precipitation by comparing data from radars and raingauges. J. Meteorol. Soc. Jpn. 74, 459–480.

McDonnell, J.J., Beven, K., 2014. Debates-The future of hydrological sciences: a (common) path forward? A call to action aimed at understanding velocities, celerities and residence time distributions of the headwater hydrograph. Water Resour. Res. 50, 5342–5350. http://dx.doi.org/10.1002/2013WR015141.

McDonnell, J.J., McGuire, K., Aggarwal, P., Beven, K.J., Biondi, D., Destouni, G., Dunn, S., James, A., Kirchner, J., Kraft, P., Lyon, S., Maloszewski, P., Newman, B., Pfister, L., Rinaldo, A., Rodhe, A., Sayama, T., Seibert, J., Solomon, K., Soulsby, C., Stewart, M., Tetzlaff, D., Tobin, C., Troch, P., Weiler, M., Western, A., Wörman, A., Wrede, S., 2010. How old is stream water? Open questions in catchment transit time

conceptualization, modeling and analysis. Hydrol. Process. 24, 1745–1754. http://dx.doi.org/10.1002/hyp.7796.

McGuire, K.J., McDonnell, J.J., 2006. A review and evaluation of catchment transit time modeling. J. Hydrol. 330, 543–563. http://dx.doi.org/10.1016/j.jhydrol.2006.04.020.

McGuire, K.J., Weiler, M., McDonnell, J.J., 2007. Integrating tracer experiments with modeling to assess runoff processes and water transit times. Adv. Water Resour. 30, 824–837. http://dx.doi.org/10.1016/j.advwatres.2006.07.004.

Ministry of Land, Infrastructure and Transport, Kofu Work Office, 2005. Information about Fuji River Flood Control. Sannichi Printing, Fujiyoshida.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. J. Hydrol. 10 (3), 282–290. http://dx.doi.org/10.1016/0022-1694(70)90255-6.

Parajka, J., Andréassian, V., Archfield, S.A., Bárdossy, A., Blöschl, G., Chiew, F., Duan, L., Gelfan, A., Hlavčová, K., Merz, R., McIntyre, N., Oudin, L., Perrin, C., Rogger, M., Salinas, J.L., Savenije, H.G., Skϕien, J.O., Wagener, T., Zehe, E., Zhang, Y., 2013. Prediction of runoff hydrographs in ungagged basins. In: Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., Savenije, H. (Eds.), Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales. Cambridge University Press, Cambridge, pp. 227–269.

Sayama, T., McDonnell, J.J., 2009. A new time-space accounting scheme to predict stream water residence time and hydrograph source components at the watershed scale. Water Resour. Res. 45, W07401. http://dx.doi.org/10.1029/2008WR007549.

Seibert, J., McDonnell, J.J., 2002. On the dialog between experimentalist and modeler in catchment hydrology: use of soft data for multicriteria model calibration. Water Resour. Res. 38, 1241. http://dx.doi.org/10.1029/2001WR000978.

Seibert, J., Rodhe, A., Bishop, K., 2003. Simulating interactions between saturated and unsaturated storage in a conceptual runoff model. Hydrol. Process. 17, 379–390. http://dx.doi.org/10.1002/hyp.1130.

Sivapalan, M., 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. Hydrol. Process. 17, 3163–3170. http://dx.doi.org/10.1002/hyp.5155.

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiondo, E.M., O'Connell, P.E., Oki, T., Pomeroy, J.W., Schertzer, D., Uhlenbrook, S., Zehe, E., 2003. IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. Hydrol. Sci. J. 48, 857–880. http://dx.doi.org/10.1623/hysj.48.6.857.51421.

Stadnyk, T., Amour, N.St., Kouwen, N., Edwards, T.W.D., Pietroniro, A., Gibson, J.J., 2005. A groundwater separation study in boreal wetland terrain: the WATFLOOD hydrological model compared with stable isotope tracers. Isot. Environ. Health Stud. 41, 49–68. http://dx.doi.org/10.1080/10256010500053730.

Takeuchi, K., 2004. Dujangyan and Shingenteh; lessons from the world's longest living flood control works. In: Rodda, J.C., Ubertini, L. (Eds.), The Basis of Civilization: Water Science? International Association of Hydrological Sciences Publication No. 286, pp. 195–200.

Tetzlaff, D., Al-Rawas, G., Blöschl, G., Carey, S.K., Fan, Y., Hrachowitz, M., Kirnbauer, R., Jewitt, G., Laudon, H., McGuire, K.J., Sayama, T., Soulsby, C., Zehe, E., Wagener, T., 2013. Process realism: flow paths and storage. In: Blöschl, G., Sivapalan, M., Wagener, T., Viglione, A., Savenije, H. (Eds.), Runoff Prediction in Ungauged Basins: Synthesis Across Processes, Places and Scales. Cambridge University Press, Cambridge, pp. 53–69.

Tetzlaff, D., Birkel, C., Dick, J., Geris, J., Soulsby, C., 2014. Storage dynamics in hydropedological units control hillslope connectivity, runoff generation, and the evolution of catchment transit time distributions. Water Resour. Res. 50, 969–985. http://dx.doi.org/10.1002/2013WR014147.

Vaché, K.B., McDonnell, J.J., 2006. A process-based rejectionist framework for evaluating catchment runoff model structure. Water Resour. Res. 42, W02409. http://dx.doi.org/10.1029/2005WR004247.

Wagener, T., Sivapalan, M., McDonnell, J., Hooper, R., Lakshmi, V., Liang, X., Kumar, P., 2004. Predictions in ungauged basins as a catalyst for multidisciplinary hydrology. Eos, Trans. Am. Geophys. Union 85, 451. http://dx.doi.org/10.1029/2004EO440003.

Ward, R.C., Robinson, M., 1990. Principles of Hydrology. McGraw-Hill, London.

Wissmeier, L., Uhlenbrook, S., 2007. Distributed, high-resolution modelling of ^{18}O signals in a meso-scale catchment. J. Hydrol. 332, 497–510. http://dx.doi.org/10.1016/j.jhydrol.2006.08.003.

Yamanaka, T., Onda, Y., 2011. On measurement accuracy of liquid water isotope analyzer based on wavelength-scanned cavity ring-down spectroscopy (WSCRDS). Bull. Terrestrial Environ. Res. Cent., Univ. Tsukuba 12, 31–40.

Yamanaka, T., Makino, Y., Wakiyama, Y., Kishi, K., Maruyama, K., Kano, M., Ma, W., Suzuki, K., 2015. How reliable are modeled precipitation isoscapes over a high-relief mountainous region? Hydrol. Res. Lett. 9, 118–124. http://dx.doi.org/10.3178/hrl.9.118.