Identifying dominant runoff mechanisms and their lumped modeling: a data-based modeling approach

Yoshiyuki Yokoo1,2*, Takahiko Chiba1, Yudai Shikano1 and Chris Leong1
1Graduate School of Symbiotic Systems Science, Fukushima University, Japan
2Institute of Environmental Radioactivity, Fukushima University, Japan

Abstract:
The authors developed a methodology for identifying dominant runoff mechanisms of a watershed and their lumped modeling as a data-based modeling approach with precipitation and runoff data which would contribute to the reduction of uncertainties in both the model structure and the parameter. We firstly separated a hydrograph into several runoff components by a recession analysis of runoff data and a filter separation method. Secondly, we estimated storage as a function of runoff for each component. Finally, we constructed a single Tank model for each component, where both the runoff component and the estimated storage were used as constraint conditions in identifying coefficients of runoff and infiltration. By applying this approach, we found that (1) the constructed Tank model perfectly traced the runoff components separated by the filter separation method, (2) there are almost no uncertainties in the model structure and the parameter if the result of filter separation can be assumed to be reliable, and (3) we can even estimate effective rainfall with our approach. These results imply our methodology allows identifying and modeling dominant rainfall-runoff mechanisms with minimal uncertainties in model structure and parameter, using hourly precipitation and runoff data alone.

KEYWORDS hydrograph separation; uncertainty; Tank model; storage; data-based; inverse estimation

INTRODUCTION

Runoff models always suffer from uncertainties in structure, parameters and simulated runoff, regardless of whether they are lumped or distributed models. Is it possible to construct a runoff model based on dominant runoff processes (Grayson and Blöschl, 2000; Sivakumar, 2004) of a watershed in the framework of data-based modeling approach (Young and Beven, 1994)? At present, such a methodology has not been developed, hence the problem of uncertainties remain unsolved. This should be one of the most critical problems of current runoff modeling studies.

How can we realize the modeling based on dominant runoff processes identified with observed data? The authors think that runoff components separated by the filter separation method developed by Hino and Hasebe (1984) could be recognized as data-based dominant runoff processes because the method creates a numerical filter with recession time constant directly identified by the recession curve of an hourly hydrograph. By repeating the separations, this method can separate a hydrograph into multiple components. Recently, Kobayashi and Yokoo (2013) suggested a methodology to identify dominant runoff processes and their corresponding watershed-scale storage by integrating the filter separation method by Hino and Hasebe (1984) and the data-based watershed-scale storage estimation method by Kirchner (2009). Later, Yokoo et al. (2014) applied the methodology for interpreting the relationship between slope failure and watershed-scale storage in Thailand. Chiba and Yokoo (2015) then improved the methodology by linearizing the relationship between runoff and storage for theoretical consistency in integrating the method by Kirchner (2009) with the filter separation method by Hino and Hasebe (1984). If we can assume that the methodology by Chiba and Yokoo (2015) is widely applicable in any watershed, the last problem becomes the runoff model construction.

To solve the above problems, the present study aims to develop and demonstrate a methodology that firstly identifies dominant runoff mechanisms within a watershed based on observed data by using the method by Chiba and Yokoo (2015) and then constructing a runoff model reflecting the dominant runoff mechanisms without calibration of model structures and parameters. Here, we employed the Tank model of Sugawara (1995) because equations used in Chiba and Yokoo (2015) resemble those of the Tank model. More details will be explained in the next section.

STUDY AREA AND METHODS

Study area and data

We conducted this study in the catchment area of Kami-masaki water-level monitoring station along Sendai River located at Ebino city in Miyazaki prefecture, Japan, as in Figure 1. The details of the watershed are summarized in Table SI. Precipitation data used in this study is measured at Hachigamine monitoring station plotted in Figure 1.
IDENTIFYING AND MODELING RUNOFF MECHANISMS

Figure 1. Watershed of Kamimasaki monitoring station, Sendai River, Japan. The blue pentagon and red circle indicate locations of Hachigamine precipitation monitoring station and Kamimaki water-level monitoring station.

Hydrograph separation

As demonstrated by Kobayashi and Yokoo (2013), we separated a hydrograph into several components by using the filter separation method by Hino and Hasebe (1984). The method firstly estimates a characteristic recession time constant by applying an exponential decay function to the slowest recession part of runoff data plotted on a semilogarithmic axis where the recession curve is found as a recession line segment. The recession time constant of a component \( T_{c,i} \) (with unique component number \( i \)) is identified as the inverse of the recession decay factor of the exponential decay function. The authors then separated a hydrograph into several components by using the filter separation method by Hino and Hasebe (1984).

As demonstrated by Kobayashi and Yokoo (2013), who modified the method by Kirchner (2009), we firstly explored the relationships in Equation (3) for slower components were explored without the original Kirchner’s (2009) conditions for selecting data in rainless night-time periods to minimize the effects of evapotranspiration and precipitation in the water balance equation, because we assumed that the relationships in Equation (3) for slower components should be affected less by evapotranspiration and precipitation. Hence we employed the original Kirchner’s (2009) conditions only for the fastest component.

Secondly we calculated storage \( s^*_i - s_0 \) as a function of runoff \( q \), by the following equation,

\[

s^*_i - s_0 = \frac{1}{a_i} \cdot q^*_i

\]

where \( s_0 \) is a storage when \( q^*_i \) becomes 0 and the actual magnitude of \( s_0 \) stays unknown at this stage. The relationship between \( q^*_i \) and \( s^*_i - s_0 \) in Equation (4) is explored for all the runoff components separated by numerical filtering as in Figure S1(f).

Construction of Tank model

In constructing a Tank model suggested by Sugawara (1995) as in Figure 2(a), we must specify the number of vertically serial tanks and parameter values. In the present study, we assumed that the total number of vertically serial tanks \( i = n_{\text{max}} \) is equal to that of the runoff components separated by the numerical filtering with their unique tank numbers \( i \) explained above. The parameter values are identified by using the time series data of runoff components and their corresponding storage, which is a newly developed method in the present study.

This method firstly identifies the runoff coefficient of the bottom tank \( i = n_{\text{max}} \) that generates the slowest runoff component by taking \( a_i \) in Equation (3). Here \( i \) is the same as defined above and it denotes the order of the tank counted-up from the top tank \( i = 1 \). As we assume there is no infiltration from the bottom tank, we can estimate infiltration from the \((i_{\text{max}} - 1)\)th tank to the bottom tank, \( P_{i_{\text{max}}-1} \), by cal-
Figure 2. The Tank model. $p_0$, $s_i$, $s_{0,i}$, $p$, and $q_i$ are respectively effective precipitation, storage height from the infiltration hole, height of runoff hole, infiltration and runoff of $i$th tank. $a_i$, $b_i$ are runoff and infiltration coefficients, respectively: (a) Tank model with full structure in the case of $s_{0,i} \geq 0$, (b) structure of a specific tank in the case of $s_{0,i} < 0$.

Calculating the water balance of the bottom tank as Equation (5):

$$s_{n+1}^0 - s_{n}^0 = \left( \frac{s_{\text{max}}^0 - s_{\text{min}}^0}{\text{K/K}} \right) - q_{n+1}^0 + p_{n+1}^0$$  \hspace{1cm} (5)

where $n$ denotes time step. In the case of middle tanks between the top and the bottom tanks, we can estimate infiltration from upper tanks $p_i^{n+1}$, by calculating the water balance as Equation (6) successively from the lower tanks to upper tanks.

$$s_{n+1}^i - s_{n}^i = \left( \frac{s_{\text{max}}^i - s_{\text{min}}^i}{\text{K/K}} \right) - q_{n+1}^i + p_{n+1}^i$$  \hspace{1cm} (6)

Similarly, we can estimate effective precipitation $pe^{n+1}$ of the time step $(n+1)$ from the water balance of the top tank as Equation (7).

$$s_{n}^i - s_{n-1}^i = \left( \frac{s_{\text{max}}^i - s_{\text{min}}^i}{\text{K/K}} \right) - q_{n}^i + p_{n}^i + pe^{n+1}$$  \hspace{1cm} (7)

Note that infiltrations $p_{i,n-1}^0$ in Equation (5) and $p_{i,n}^i$ in Equation (6) can be negative values, indicating return flows from slower components to faster components. Also, the effective rainfall $pe^{n+1}$ calculated by Equation (7) can have a negative value, which can be interpreted as instantaneous evapotranspiration from a watershed. These negative fluxes would be regarded as the difference between our method and the original Tank model by Sugawara (1995).

Secondly, we identified the magnitudes of the infiltration coefficients $b_i$ and the initial storage $s_{0,i}$ simultaneously by applying linear regression lines as Equation (8) to the scatter diagrams between $(s_{n}^i - s_{0,i})$ and $p_{i,n}^n$ for all the runoff components except for the slowest component that has no infiltration as in Figure S2.

$$s_{n}^i - s_{0,i} = b_{i,n}^{-1} \cdot p_{i,n}^n - s_{0,i}$$  \hspace{1cm} (8)

These equations (4) to (8) make it possible to identify all the Tank model coefficients together with the estimated infiltration (or return flow) between tanks and effective precipitation (or evapotranspiration) $pe$ at a watershed-scale that is input to the top tank with equation (7). The difference between precipitation at watershed-scale $p$ and $pe$ can also be regarded as evapotranspiration and outflow to the outside of the watershed area, which requires rigorous discussions to be explored in a separate study.

### RESULTS

#### Hydrograph separation

Figure 3 shows the result of hydrographs separated by the numerical filter by Hino and Hasebe (1984). The recession time constants, $T_n$, for each components were estimated using the criteria set by Chiba and Yokoo (2015) as in Table SII. We can see that the estimated runoff components show different recession rates, indicating that they could be regarded as dominant runoff components from the viewpoint of numerical filtering.

#### Storage estimation

Figure 4 shows the changes in estimated storages of $(s_{n}^i - s_{0,i})$ for all the runoff components, which is based on the results in Figure S1(a–f). Unlike the result of the hydrograph separation in Figure 3, storage in the slowest component dominated among all the storage components. These results show lower storage capacities in faster components and higher capacities in slower components, which is often found in bucket-type models such as the Tank model by Sugawara (1995) and is similar to the previous researches (e.g. Kobayashi and Yokoo, 2013; Yokoo et al., 2014; Chiba and Yokoo, 2015).

#### Construction of the Tank model

Figure S2 shows the relationships between infiltration $p_{i,n}^n$ and storage $(s_{n}^i - s_{0,i})$ for each tank that were illustrated for estimating the magnitudes of infiltration coefficients $b_i$. The scatter plots were widely spread around the linear regression lines estimated by the least square method, which originates from the differences in calculation methods between the original Tank model by Sugawara (1995) and this methodology. Although we assumed linear relationships between $p_{i,n}^n$ and $(s_{n}^i - s_{0,i})$, these results suggest further exploration of the relationships is required.

The Tank model parameters identified by our methodology is summarized in Table SII. The negative value of $s_{0,i}$ in Table SII indicates that the runoff hole is located lower than the infiltration hole, hence the infiltration hole sinks into the tank as in Figure 2(b). This is because our runoff storage-discharge relationship is modeled by Equation (4). If $s_{0,i}$ is positive and less than $s$, as in Figure 2(a), runoff is calculated as the product of runoff coefficient $a_i$ and the storage $(s - s_{0,i})$ whereas infiltration is calculated as $s_i/b_i$. If $s_{0,i}$ is negative as in Figure 2(b), the product of runoff coefficient $a_i$ and the storage $(s - s_{0,i})$ becomes runoff that stops when the tank becomes empty, where infiltration, calcu-
lated as $s_i b_i$ stops when storage height from infiltration hole $s_i$ becomes zero.

**How does the Tank model work?**

Figure 5 shows the time series data of the Tank model identified by the method explained in the previous section. Panel (a) illustrates the observed precipitation $p$ measured at Hachigamine monitoring station together with the effective precipitation $pe$ estimated by our method. We can see that the effective precipitation $pe$ is less than the observed precipitation $p$ and also $pe$ becomes even more negative in the rainy season.

One of the potential reasons for $pe$ being less than $p$ could be explained by evapotranspiration if $p$, measured at Hachigamine monitoring station in Figure 1, accurately represented the watershed-scale precipitation. The negative $pe$ after rainfall events in the rainy season should indicate intensive evapotranspiration after rainfall events at the watershed-scale. However, reasons for $pe$ less than $p$ and negative $pe$ must be carefully investigated in a well gauged watershed as a separate study.

The panels (b) to (f) show the temporal changes of storage $(s - s_0)$, runoff $(q)$, and infiltration $(pi)$ for the Tank number $i$ ($i = 1 – 5$)

The reason for faster infiltration peaks is explained by Equations (5–6) which calculates infiltration from the...
upper tank $p^i$, ... In the original method by Sugawara (1995), it calculates runoff as a linear function of storage, but our calculations using Equations (5–6). In other words, infiltration terms calculated in Equations (5–6) work to connect different calculation methods between the original Sugawara (1995) Tank model and our method based on Chiba and Yokoo (2015) that estimated multiple storage-discharge relationships independently of each other. Therefore, our method is compatible with Sugawara’s tank model however we should recognize that our Tank model is not same as the Tank model by Sugawara (1995).

Figure 5 is just a summary of inputs, state variables and outputs, hence there is nothing new, but we would claim that these are calculated only from observed time series of precipitation and runoff without any complex calibration processes nor a priori assumptions for model structures. Benefits of our methodology and some ways forward will be discussed in the following section.

CONCLUDING DISCUSSION

This study developed a methodology for constructing a Tank model (Sugawara, 1995) without complex calibrations nor a priori assumptions for model structures. Using the watershed-scale storage estimated by the method by Chiba and Yokoo (2015) which integrated the filter separation method by Hino and Hasebe (1984) with the watershed-scale storage estimation method by Kirchner (2009), we succeeded in constructing a Tank model only with the observed precipitation and runoff data.

The outstanding benefit of this methodology is that we can construct a Tank model uniquely to a watershed without complex calibration processes nor a priori assumptions for model structures, unlike Yokoo et al. (2001) and Yokoo and Kazama (2012). With such calibrations or assumptions, we cannot be sure of the model parameters nor structure because of their uncertainties. Along the goal of the initiative of prediction in ungauged basins (PUB; Sivapalan, 2003; Sivapalan et al., 2003), the authors could considerably decrease the dependency on calibrations and increase dependency on knowledge, although our methodology demands continuous hourly data of precipitation and runoff.

With such a Tank model, estimations of dominant runoff processes and their relationship between storage and runoff derived from our approach, we can estimate how precipitation is stored in and released from a watershed only from observed hourly precipitation and runoff data. As far as lumped modeling approach is concerned, the residual problem will be how the watershed-scale precipitation is converted to effective precipitation estimated with our approach in using a Tank model. Considerations of the conversion is beyond the scope of this study and it should be the next target of this approach, where spatial distribution of precipitation and evapotranspiration must be deeply investigated using the effective precipitation estimated by our approach as reference data.

Another benefit is that we can compare our outputs or state variables with field studies as demonstrated by previous studies (e.g. Beven, 2006; Birkel et al., 2011; Krakauer and Temimi, 2011; Sayama et al., 2011). Number of dominant processes, amount of storage, residence time of precipitation, infiltration, as well as runoff for each dominant process are always the target of field-based hydrology. Among them, residence time will be a clue to verify the estimated nature of dominant processes. For example, we estimated the recession time constants $T_r$ in hydrograph separations. The magnitude of $T_r$ becomes an approximation of residence time calculated as annual mean storage divided by annual mean flow, because 99.3% of precipitation becomes discharge after a precipitation event within the period of $T_r$. As discussed by Seibert and McDonnell (2002), a Tank model constructed in our methodology must be carefully tested against soft data or process knowledge from experimentalists. Inversely, variables estimated by our approach might suggest a new way of field experiments to verify the variables. Therefore, the next task of this study will be testing our approach in intensively gauged watersheds such as experimental watersheds in collaboration among modelers and experimentalists.

ACKNOWLEDGMENTS

This study was supported by JSPS KAKENHI Grant Numbers JP16K06501 and JP16KK0142. We also thank the hydrological data on Water Information System and geographic data provided by the MLIT, Japan.

SUPPLEMENTS

Table SI. Description of study area
Table SII. Identified parameters of the Tank model
Figure S1. Relationships between $-d q / d t$ and $q_i$, for all the components: (a) component 1, (b) component 2, (c) component 3, (d) component 4, (e) component 5; as well as the estimated relationships between $(s_i - s_n)$ and $q_i$ in panel (f)
Figure S2. Relationship between estimated infiltration from each tank $p_i$ and storage $(s_i - s_n)$: (a) 1st tank, (b) 2nd tank, (c) 3rd tank and (d) 4th tank

REFERENCES

Beven K. 2006. Searching for the Holy Grail of scientific hydrology: $Q = H(\Delta, S, P, \Delta)$ as closure. Hydrology and Earth System Sciences 10: 609–618. DOI: 10.5194/hess-10-609-2006.

Birkel C, Soulsby C, Tetzlaff D. 2011. Modelling catchment-scale water storage dynamics: reconciling dynamic storage with tracer-inferred passive storage. Hydrological Processes 25: 3924–3936. DOI: 10.1002/hyp.8201.

Chiba T, Yokoo Y. 2015. Revision of estimation method of watershed-scale rainwater storage and its effect. Journal of JSCE B1 (Hydraulic Engineering) 71: 1-289–1-294 (in Japanese).

Grayson RB, Blöschl G. 2000. Summary of pattern comparison and concluding remarks. In Spatial Patterns in Catchment Hydrology: Observations and Modeling, Grayson RB, Blöschl G (eds). Cambridge University Press, Cambridge, UK; 355–367.
Hino M, Hasebe M. 1984. Identification and prediction of non-linear hydrologic systems by the filter-separation autoregressive (AR) method: Extension to hourly hydrologic data. *Journal of Hydrology* 68: 181–210. DOI: 10.1016/0022-1694(84)90211-7.

Kirchner JW. 2009. Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research* 45: W02429. DOI: 10.1029/2008WR006912.

Kobayashi S, Yokoo Y. 2013. Estimating watershed-scale storage changes from hourly discharge data in mountainous humid watersheds: toward a new way of dominant process modeling. *Hydrological Research Letters* 7: 97–103. DOI: 10.3178/hrl.7.97.

Krakauer NY, Temimi M. 2011. Stream recession curves and storage variability in small watersheds. *Hydrology and Earth System Sciences* 15: 2377–2389. DOI: 10.5194/hess-15-2377-2011.

Sayama T, McDonnell JJ, Dhakal A, Sullivan K. 2011. How much water can a watershed store? *Hydrological Processes* 25: 3899–3908. DOI: 10.1002/hyp.8288.

Seibert J, McDonnell JJ. 2002. On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration. *Water Resources Research* 38: 23-1–23-14. DOI: 10.1029/2001WR000978.

Sivakumar B. 2004. Dominant processes concept in hydrology: moving forward. *Hydrological Processes* 18: 2349–2353. DOI: 10.1002/hyp.5606.

Sivapalan M. 2003. Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes* 17: 3163–3170. DOI: 10.1002/hyp.5155.

Sivapalan M, Takeuchi K, Franks SW, Gupta VK, Karambiri H, Lakshimi V, Liang X, McDonnell JJ, Mendiondo EM, Connell O, Oki T, Pomeroy JW, 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. *Hydrological Sciences Journal* 48: 857–880. DOI: 10.1623/hysj.48.6.857.51421.

Sugawara M. 1995. Tank Model. In *Computer Models of Watershed Hydrology*, Singh VJ (ed). Water Resources Publications, Highlands Ranch; 165–214.

Yokoo Y, Kazama S. 2012. Numerical investigations on the relationships between watershed characteristics and water balance model parameters: searching for universal relationships among regional relationships. *Hydrological Processes* 26: 843–854. DOI: 10.1002/hyp.8299.

Yokoo Y, Kazama S, Sawamoto M, Nishimura H. 2001. Regionalization of lumped water balance model parameters based on multiple regression. *Journal of Hydrology* 246: 209–222. DOI: 10.1016/S0022-1694(01)00372-9.

Yokoo Y, Wattanakarn C, Wattanakarn S, Semcharoen V, Promasakha na Sakolnakhon K, Soralump S. 2014. Storage under the 2011 Chao Phraya river flood: An interpretation of watershed-scale storage changes at two neighboring mountainous watersheds in northern Thailand. *Hydrological Research Letters* 8: 1–8. DOI: 10.3178/hrl.8.1.

Young PC, Beven KJ. 1994. Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics* 5: 335–363. DOI: 10.1002/env.3170050311.