Time-variability of flow recession dynamics: Application of machine learning and learning from the machine

Minseok Kim[1], Hannes H Bauser[1], Keith J Beven[2], and Peter A. Troch[1]

[1]University of Arizona
[2]Lancaster University

November 24, 2022

Abstract

Flow recession analysis, relating discharge Q and its time rate of change -dQ/dt, has been widely used to understand catchment scale flow dynamics. However, data points in the recession plot, the plot of -dQ/dt versus Q, typically form a wide point cloud due to noise and hysteresis in the storage-discharge relationship, and it is still unclear what information we can extract from the plot and how to understand the information. There seem to be two contrasting approaches to interpret the plot. One emphasizes the importance of the ensembles of many recessions (i.e., the lower envelope or a measure of central tendency), and the other highlights the importance of the event scale analysis and questions the meaning of the ensemble characteristics. In this study, we examine if those approaches can be reconciled. We utilize a machine learning tool to capture the point cloud using the past trajectory of discharge. Our results show that most of the data points can be captured using 5 days of past discharge. We show that we can learn the catchment scale flow recession dynamics from what the machine learned. We analyze patterns learned by the machine and explain and hypothesize why the machine learned those characteristics. The hysteresis in the plot mainly occurs during the early time dynamics, and the flow recession dynamics eventually converge to an attractor in the plot, which represents the master recession curve. We also illustrate that a hysteretic storage-discharge relationship can be estimated based on the attractor.

Time-variability of flow recession dynamics: Application of machine learning and learning from the machine

Minseok Kim[1], Hannes H. Bauser[2], Keith Beven[3], Peter A. Troch[2]

[1]Biosphere 2, University of Arizona, Tucson, AZ, USA
[2]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA
[3]Lancaster Environment Centre, Lancaster University, Lancaster, UK

Key Points:

- A machine learning tool captures time-variable flow recession dynamics that identify scanning curves of the storage-discharge relationship.
- Machine learned individual flow recession curves converge to a common attractor in the recession plot, revealing the master recession curve.
- It leads to a novel way of analyzing the recession plot, unifying the event-based analysis and the analysis of ensemble characteristics.

Corresponding author: Minseok Kim, `minseok.h.kim@gmail.com`

Abstract

Flow recession analysis, relating discharge Q and its time rate of change $-dQ/dt$, has been widely used to understand catchment scale flow dynamics. However, data points in the recession plot, the plot of $-dQ/dt$ versus Q, typically form a wide point cloud due to noise and hysteresis in the storage-discharge relationship, and it is still unclear what information we can extract from the plot and how to understand the information. There seem to be two contrasting approaches to interpret the plot. One emphasizes the importance of the ensembles of many recessions (i.e., the lower envelope or a measure of central tendency), and the other highlights the importance of the event scale analysis and questions the meaning of the ensemble characteristics. In this study, we examine if those approaches can be reconciled. We utilize a machine learning tool to capture the point cloud using the past trajectory of discharge. Our results show that most of the data points can be captured using 5 days of past discharge. We show that we can learn the catchment scale flow recession dynamics from what the machine learned. We analyze patterns learned by the machine and explain and hypothesize why the machine learned those characteristics. The hysteresis in the plot mainly occurs during the early time dynamics, and the flow recession dynamics eventually converge to an attractor in the plot, which represents the master recession curve. We also illustrate that a hysteretic storage-discharge relationship can be estimated based on the attractor.

1 Introduction

Flow recession analysis (e.g., Barnes, 1939; Hall, 1968; Anderson & Burt, 1980; Brutsaert & Nieber, 1977) has been extensively utilized to understand flow dynamics at the catchment scale (e.g., Vogel & Kroll, 1992; Clark et al., 2009; Jachens et al., 2020). Flow recession is a "data-based" catchment scale signature that encapsulates information about catchment characteristics and dynamics (e.g., Troch et al., 2013). The flow recession analysis also provides ways to estimate a type of the storage-discharge relationship (e.g., Kirchner, 2009; Dralle et al., 2018). Typically, the recession plot is constructed by plotting the rate of change in discharge $-dQ/dt$ versus discharge Q in log-log scale, and patterns in the plot have been analyzed and linked to catchment scale processes and properties (e.g., Brutsaert & Nieber, 1977; Troch et al., 2013).

Brutsaert and Nieber (1977) showed that some patterns of data points in the flow recession plot can be explained by a hydraulic groundwater model, viz. the Boussinesq model. The explanatory power of the model implies that catchment scale properties, such as the saturated hydraulic conductivity and the drainable porosity, can be estimated through the flow recession analysis (Brutsaert & Nieber, 1977; Troch et al., 2013). Other studies showed that the data points can also be explained by other mechanisms and models, such as a two parallel bucket model and a model using superposition of multiple linear reservoirs (e.g., Clark et al., 2009; Harman et al., 2009; Gao et al., 2017). Biswal and Marani (2010) showed that the geometry of drainage network also can explain some patterns. While which model represents reality better probably varies from site to site, it is clear that the recession analysis helps hydrologists develop hypotheses about catchment scale flow dynamics.

However, there still remains a fundamental issue on what is the “right” information we can extract from the signature. The data points in the recession plot usually form a wide point cloud due to the measurement noise in Q (e.g., Rupp & Selker, 2006), the auto-correlation in observation errors, and time-varying catchment dynamics and external forcings (e.g., Harman et al., 2009; Shaw & Riha, 2012; Jachens et al., 2020). Before proposing hypotheses about catchment scale dynamics, we need to decide how to interpret the wide point cloud.

Brutsaert and Nieber (1977) suggested using the lower envelope of a point cloud. They used the lower envelope to capture the ensemble characteristics of many recessions (Brutsaert, 2005) and suggested determining the slope of the lower envelope b among the values that can be explained by the Boussinesq model instead of estimating the slope directly using data. The Boussinesq model used in their original study predicts two slopes ($b = 1.5$ for the late time recession and $b = 3.0$ for the early time recession), and the predicted lower envelope has a lower slope in the lower discharge range. Alternatively, Vogel and Kroll (1992) performed an ordinary regression analysis to fit a line to the data as a measure of the central tendency (centrality). Similarly, Kirchner (2009) suggested binning the data and performed a weighted linear regression to account for the uncertainty associated with each bin.

However, recent studies have questioned the use of the lower envelope and the measure of central tendency and have emphasized the importance of analyzing the slope b

of each recession event (e.g., Shaw & Riha, 2012; Tashie et al., 2020; Jachens et al., 2020). The slope fitted to the data points of each event is event-specific, and it seems that the lower envelope does not represent an ensemble of recession dynamics but is a collection of endpoints of each event (Tashie et al., 2020; Jachens et al., 2020). Such event-to-event differences are often attributed to catchment memory effects (e.g., Harman et al., 2009; Tashie et al., 2020; Jachens et al., 2020) or to seasonal dynamics (Shaw & Riha, 2012). Spatial and temporal pattern of external forcings, such as evapotranspiration and precipitation, may also affect the event-to-event variability (Wang & Cai, 2010; Szilagyi et al., 2007). Besides, the slope of each event is in general much steeper than the slope estimated as a central tendency or derived from the Boussinesq model (e.g., Tashie et al., 2020; Jachens et al., 2020). Tashie et al. (2020) further argued that many of the trajectories of each event in the recession plot have a higher slope at the lower discharge range, except for some dry and flat catchments, casting doubt on the applicability of the Boussinesq model.

There seem to be two contrasting approaches. One emphasizes the importance of analyzing the ensembles of many recessions (i.e., the lower envelope or a measure of central tendency), and the other highlights the importance of the event scale analysis and questions the meaning of the ensemble characteristics that are represented by the lower envelope or the measure of central tendency. In this study, we examine if those approaches can be reconciled. We utilize a machine learning tool to capture dynamics represented in the recession plot using the past trajectory of flow. We anticipate that the tool can learn both the time-variability (i.e., the event-by-event variability) and the ensemble of recession dynamics, if both exist. We report the machine learning model results and explain some patterns that the machine learning tool exposed. We finally show that the contrasting approaches can be combined into a single one. While the focus of our study is not on examining underlying hydrological processes in detail, we also infer and hypothesize underlying hydrological processes. In addition, we illustrate that a hysteretic storage-discharge relationship can be estimated using a characteristic trajectory that appears in the recession plot. In the discussion section, we treat the recession plot as a phase space plot, and links to other phase space plots are also discussed.

2 Theoretical background, methods, and study site

2.1 Flow recession analysis

Originally, flow recession analysis used a plot of $-dQ(t)/dt$ versus $Q(t)$. In this study, we use an alternative function:

$$g(t) = -\frac{dQ(t)}{dt}/Q(t) \tag{1}$$

The function $g(t)$, instead of $-dQ/dt$, is plotted versus $Q(t)$. The function g is identical to the catchment sensitivity function of Kirchner (2009). (Note that the catchment sensitivity function expresses the sensitivity of discharge to changes in storage S; i.e., $g = dQ/dS = (dQ/dt)/(dS/dt)$ (Kirchner, 2009). The formulation in (1) is a simplified form for the case of negligible precipitation and evapotranspiration during recession periods that has been utilized predominantly instead of fully considering dS/dt.) We will use the term recession plot interchangeably for either the g vs. Q plot or the $-dQ/dt$ vs. Q plot. When a power function is used to characterize the original recession plot (i.e., $-dQ/dt = aQ^b$), the power function still holds in the g vs. Q plot with the exponent decreased by 1: $g(Q) = aQ^{b-1}$ (Kirchner, 2009).

The catchment sensitivity function can be used to characterize flow recession dynamics and estimate a type of storage-discharge relationship. The inverse of g, $1/g$, is a time scale of the flow recession. When the flow recession over time is approximated using an exponential function as $Q = Q_0 e^{-t/t_c}$, where t_c is the e-folding time of the exponential decay, $1/g$ is constant and is the e-folding time; i.e. $t_c = 1/g$. Otherwise, the decay rate $1/g$ depends on time. Also, assuming there is a one-to-one and invertible function that relates g to Q, the function $g(Q)$ can be utilized to estimate a relationship between the active storage and discharge using: $S_a(Q) = \int_{Q_0}^{Q} (1/g(Q))dQ$, where S_a is the "active" storage (relative to a certain storage at Q_0) which is the portion of the storage that drives discharge (e.g., Kirchner, 2009; Troch et al., 2013). (Note that the active storage is sometimes referred to as "dynamic" storage (Staudinger et al., 2017), "direct" storage (Dralle et al., 2018), or "hydraulically-connected" storage (Carrer et al., 2019).)

Several methods have been suggested to estimate $dQ(t)/dt$ using the discrete time series of Q. One simple way is to estimate it at a constant time step (CTS): $dQ(t+\Delta t/2)/dt = (Q(t + \Delta t) - Q(t))/\Delta t$, where Δt is the time step and $Q(t + \Delta t/2) = (Q(t + \Delta t) +$

$Q(t))/2$ (Brutsaert & Nieber, 1977). However, the method is sensitive to discharge measurement resolution and noise, especially at low flow (Rupp & Selker, 2006). Roques et al. (2017) suggested the exponential time step (ETS) method, where the time step increases in each recession event and an exponential function is fitted to discharge, which is then used to estimate its (smoothed) time derivative.

Also, several criteria to determine recession periods have been suggested. Brutsaert and Nieber (1977) originally proposed using data for periods of $dQ/dt < 0$ and at least 5 days after any precipitation event, with the expectation that it would eliminate as much as possible direct surface recession flow. Recent studies have refined the criteria. For example, in the event-by-event analysis, a sufficient number of samples is required for each event to fit a statistically meaningful (power) function. Dralle et al. (2017) suggested using events that have strictly decreasing Q for more than four days (when one uses daily time step data). The start and end times of each event can be determined using a time series of precipitation J (Lamb & Beven, 1997; Dralle et al., 2017) or based on the transition from decreasing discharge to increasing discharge and vice versa (Dralle et al., 2017; Jachens et al., 2020). Those event-based studies either do not exclude any periods after peak flow (Dralle et al., 2017; Tashie et al., 2020) or exclude only one day after the peak flow (Jachens et al., 2020). In addition, Lamb and Beven (1997) suggested filtering out periods with significant (potential) evapotranspiration. For the catchment sensitivity function, Kirchner (2009) proposed using the $Q >> J$ and $Q >> ET$ criteria, where ET is the evapotranspiration rate, to rule out the effects of those climate forcings.

As mentioned earlier, the function $g(Q)$ (or $-dQ/dt$) has been parameterized using single discharge values Q. However, according to some studies that explain the event-to-event time-variability as memory effects (e.g., Harman et al., 2009; Jachens et al., 2020; Tashie et al., 2020), it seems more natural to parameterize g using the past trajectory of measurable variables. In this study, we use the past trajectory of discharge to better characterize g, rather than using single discharge values.

By doing so, we capture a type of hysteresis in the flow dynamics that can be *observed* during flow recession periods. One way to define hysteresis in hydrology is to define it as a phenomenon where the output of a system depends not only on the current state of the system but also on the past trajectory of system states or inputs (Davies &

Beven, 2015). For catchment scale flow dynamics, discharge is the output, and storage can be used to represent the state of a catchment. The hysteresis in the catchment scale flow dynamics then manifest as a hysteretic storage-discharge relationship. The wide point cloud in the recession plot illustrates the hysteresis between the "active" storage S_a and discharge Q during flow recession periods if the spread is not due to measurement errors (see Figure 1). Theoretically, if there is no hysteresis between the active storage and discharge, the data points in the recession plot should align on a single curve (see Figure 1A). Earlier we introduced that parameterizing g using Q, i.e., $g(Q)$, leads to a non-hysteretic active storage-discharge relationship. Its inverse is also true; if the active storage discharge relationship is non-hysteretic, g only depends on Q (see appendix A1). Thus, capturing the point cloud is identical to capturing the hysteretic flow dynamics during flow recession periods (see Figure 1B). Taking the well-known hysteresis in the soil water retention curve as an example, what we do in this study as to which part of the hysteresis we are looking at is similar to looking at only the drying part of the hysteresis in the soil water retention curve (i.e., the drying scanning curves). We should expect hysteresis in the catchment scale storage-discharge dynamics as a result of differences in the celerity and velocity responses to inputs. This also suggests that the hysteresis should be scale dependent (Beven & Davies, 2015; Beven, 2020b).

While the complete picture of the hysteresis cannot be examined, it is still meaningful as the recession part of the hysteresis can be seen mainly based on discharge data, which is arguably much less uncertain at the catchment scale than other fluxes (e.g., J and ET) (Kirchner, 2009). Other fluxes become much more important if we look at the complete picture of the hysteresis. Nevertheless, in the later discussion, we will also briefly show a possibility of estimating the (relative) total storage-discharge relationship using a modified catchment sensitivity function, assuming the evapotranspiration rate is reliable.

The model to estimate g using the past trajectory of discharge can be written as:

$$g = H(\overleftarrow{Q}) \tag{2}$$

where H is a non-linear hysteretic function, and $\overleftarrow{Q}$ is the past trajectory of discharge. Specifically, we configure the model to estimate the half-step ahead g, $g(t+\Delta t/2)$, using $Q(t)$, $Q(t-\Delta t)$, $\cdots$, $Q(t-m\Delta t)$, where $m+1$ is the length of the past trajec-

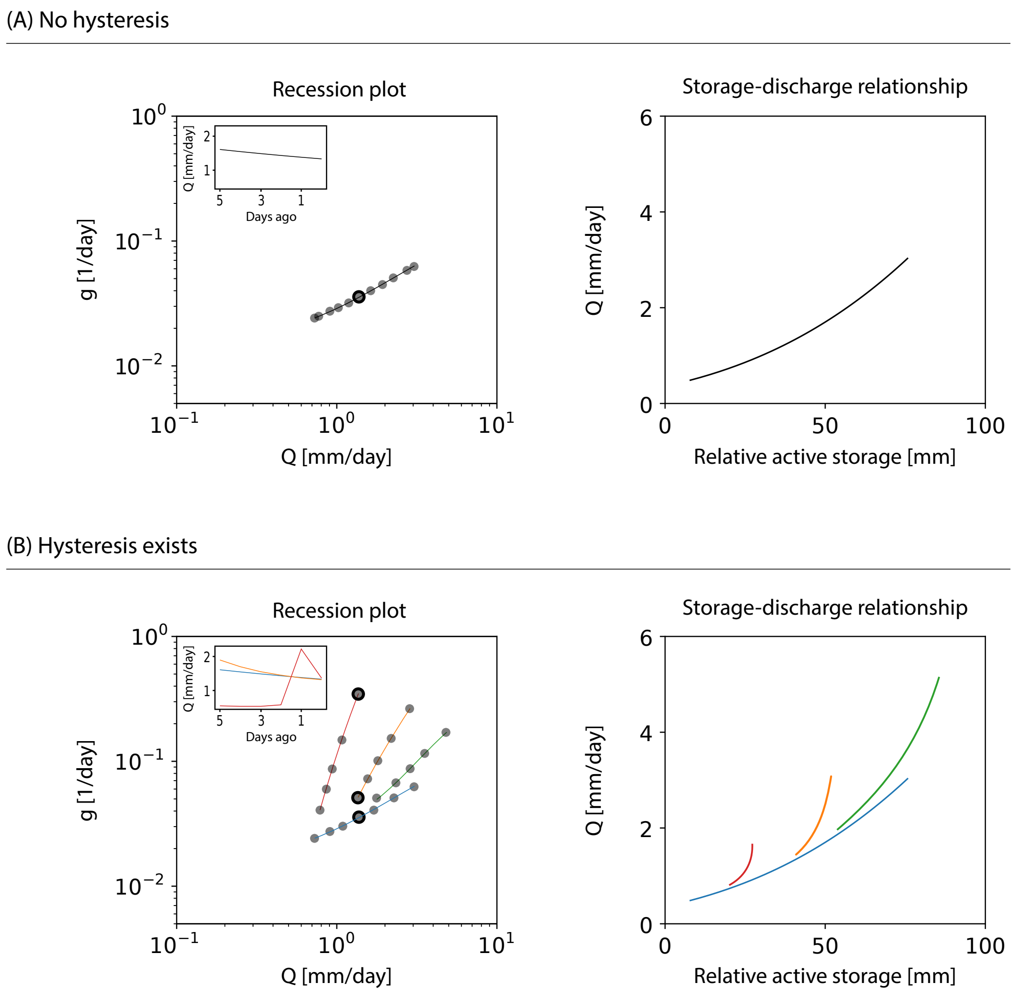

Figure 1. Illustration of the recession plot and the corresponding storage-discharge relationship (A) without hysteresis and (B) with hysteresis. The dots in the recession plot are few selected data points. The lines in the recession plot shows the trajectory of each event. The line color in (B) distinguishes events. When there is no hysteresis between the active storage and discharge, the data points in the recession plot align on a single curve. Otherwise, the hysteresis between the active storage and discharge leads to the scattered data points in the recession plot. The subset figures in both recession plots illustrate the past trajectories of discharge for events at the timings indicated by the black circles in the recession plot. The timing for each event was chosen when discharge is similar at about 1.5 mm/day. (The green event was excluded since discharge did not decrease to the value during the event.) We anticipate that the difference in g at similar discharge can be characterized by the past trajectory of discharge as shown in the subset figure. Note that the subset figure includes the rising limb of discharge for the red event because it includes the trajectory of discharge before the recession starts.

tory of discharge. (Note that while g is estimated for the flow recession periods, the past trajectory of discharge can include rising limbs). During the flow recession periods, the model can estimate the one-step ahead discharge $Q(t{+}\Delta t)$ using $g(t{+}\Delta t/2)$ as: $Q(t+\Delta t) = \frac{2-g(t+\Delta t/2)\Delta t}{2+g(t+\Delta t/2)\Delta t}Q(t)$, assuming that dQ/dt is constant between the two time steps.

The functional form is similar to Beven's Holy Grail problem (Beven, 2006b), that is to find a scale dependent hysteretic function for estimating discharge using the past trajectory of precipitation J and other relevant inputs at the scale of interest. In this study, we use the past trajectory of Q rather than J. One reason is that, often, discharge data is more accurate than catchment scale estimation of J. Also, it is more consistent with the previous studies where Q is used to characterize the function g (or $-dQ/dt$).

Following Young and Beven (1994), model (2) is a "data-based" model in a sense that the model structure is not determined *a priori* as opposed to models in which those structure is determined, for example, the multiple reservoir models (e.g., Clark et al., 2009; Harman et al., 2009; Gao et al., 2017) or spatially-resolved continuum equation based models such as the Richards equation based models. *A priori* determined model structure may adversely affect interpretation of hydrologic dynamics based on model result due mostly to the uncertain model structures (e.g., Beven, 2006a; Kirchner, 2006; Kim & Troch, 2020). The "data-based" modeling approach utilizes the transfer function model (which is originally introduced in control theory; $O(t) = (\sum_{i=0}^{j} b_i z^{-i}/(1{+}\sum_{i=1}^{k} a_i z^{-i}))I(t)$, where O is the output time series, I is the input time series, z is the backward operator, a_i, b_i, j, k are the model parameters), as it has a general form that relates input and output time series (e.g., Young, 2011). As we focus only on capturing the dynamics during recession periods, the transfer function model may reduce to the auto-regressive (AR) model where the output time series $O(t)$ is modeled using its past history (i.e., $(1{+}\sum_{i=1}^{k} a_i z^{-i})O(t) = 0$). Model (2) is similar to the auto-regressive model in that the model utilizes the past history of output time series. While model (2) estimates $g(t+\Delta t/2)$ not $Q(t+\Delta t)$ as our interest is on g, we showed above that $Q(t{+}\Delta t)$ can be estimated using $g(t{+}\Delta t/2)$. The wide point cloud in the recession plot implies that the parameters of the AR model might need to vary over time to account for the non-linearity of the flow recession dynamics. Instead of estimating the time-variable parameters in classic ways, we utilize a machine-learning tool to consider the non-linearity (see the next section for more details).

Also, model (2) can be thought of as a generalization of the model developed by Fleming (2007). Fleming (2007) developed a machine-learning based model that predicts one step ahead discharge, $Q(t{+}\Delta t)$, using $Q(t)$, where the relationship between the two variables depends on $Q(t)$. Again, while model (2) estimates $g(t{+}\Delta t/2)$ not $Q(t{+}\Delta t)$, $Q(t + \Delta t)$ can be estimated using $g(t + \Delta t/2)$ by linear interpolation. The important difference between model (2) and that of Fleming is the use of the past trajectory of Q in model (2) to capture the hysteretic flow recession dynamics. Fleming's model uses only $Q(t)$ to estimate $Q(t + \Delta t)$, while our model utilizes longer past trajectory of Q.

2.2 A machine learning tool: Long Short-Term Memory model

Given the variability in $g(t)$ we use a machine learning tool to learn the function H using data. Machine learning tools have been applied to model several hysteretic hydrologic dynamics that are represented in, for example, the rating curve (Tawfik et al., 1997) and the soil water retention curve (Jain et al., 2004). We choose the LSTM model as a machine learning tool. The LSTM model is a supervised learning algorithm and a type of recurrent neural network, that has been applied successfully to reproduce catchment scale flow dynamics (e.g., Kratzert et al., 2018; Shen et al., 2018). Compared to the classic (or vanilla) recurrent neural networks, the LSTM model has several advantages. The most well known advantage is the improved ability of the LSTM model in remembering past information in memory (Greff et al., 2017).

A LSTM model can be configured with multiple layers such as the recurrent LSTM layer, the dropout layer, and the dense layer (see Figure 2). The recurrent LSTM layer consists of multiple LSTM cells, and a LSTM cell processes an internal state h and a cell state (or a cell memory) c using input data I and three gates: a forget gate f, an input gate i, and an output gate o. The states h and c are vectors of length n, where $n \geq 1$ is referred to as the number of LSTM units. A set of forward operations in a LSTM cell can be written as:

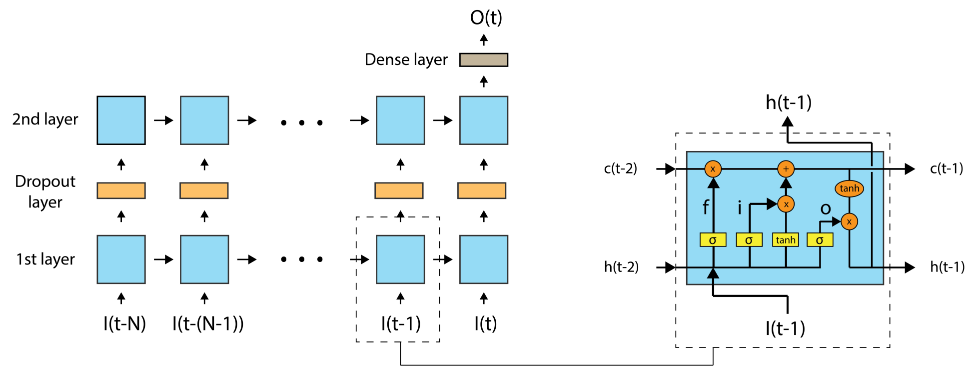

Figure 2. (Left) An example of a LSTM model structure with the dropout layer and the dense layer. The model has two layers of the recurrent LSTM layer with the dropout layer in between. Input time series I_t is fed into the first LSTM layer. The output of the second LSTM layer is fed into the dense layer, which estimates an output O_t of the model. (Right) A detailed structure inside a LSTM cell. h_t is the internal state and c_t is the cell state at time t. f, i, and o denote the forget gate, the input gate, and the output gate, respectively. $\tilde{c}$ is the cell input (modified from Greff et al. (2017)).

$$
\begin{aligned}
f_t &= \sigma(W_f I_t + U_f h_{t-1} + b_f) \\
i_t &= \sigma(W_i I_t + U_i h_{t-1} + b_i) \\
o_t &= \sigma(W_o I_t + U_o h_{t-1} + b_o) \\
\tilde{c}_t &= \tanh(W_c I_t + U_c h_{t-1} + b_c) \\
c_t &= f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \\
h_t &= o_t \circ \tanh(c_t)
\end{aligned}
\tag{3}
$$

where f_t, i_t, o_t, and $\tilde{c}_t$ are activation vectors (of length n) of the forget gate, the input gate, the output gate, and the cell input at time t, respectively, c_t is the cell state vector of length n, h_t is the internal state vector of length n, σ is the sigmoid function, the operator $\circ$ denotes the Hadamard product (element-wise product), I_t is the input feature vector of size m at time t, where m is the number of input features (or variables), W matrices (W_f, W_i, W_o, and W_c) are $n \times m$ weight matrices, U are $n \times n$ weight matrices, and b vectors are the bias vector of length n. The W and U matrices and the b vectors need to be learned using a dataset.

The dropout layer is to prevent the weights from co-adapting too much and reduce overfitting (e.g., Hochreiter & Schmidhuber, 1997). The layer randomly sets a fraction of variables (e.g., input sequence, output sequence, or the recurrent state of the previous time step) to zero at each iteration during training. The dropout rate, a hyperparameter associated with the layer, determines the fraction. The dense layer is a deeply connected neural network layer, and it estimates: $O_t = k(W_d \circ x_t + b_d)$, where O_t is an output sequence of length q, x_t is a length q input sequence to the layer, W_d is a $p \times q$ weight matrix, b_d is a bias vector of length q, and k is an activation function such as the linear function $k(x) = x$.

For example, the model shown in Figure 2 has two layers of the recurrent LSTM layer with the dropout layer in between. The dense layer receives the output of the second LSTM layer as an input sequence. The illustrated model uses $N+1$ days (or time steps) of input data (discharge Q) to estimate an output g, i.e., $I_t = Q(t)$ and $m = 1$ for the first layer, and $O_t = g(t)$ with $q = 1$. Again, while g is estimated for recession periods, the model input I_t can include discharge data in the rising limb. The number of LSTM units for the first and the second layers are hyperparameters that need to be determined by the modeler, and p is equal to the number of LSTM units of the second LSTM layer.

The model needs to be trained using data to estimate the W and U weight matrices and the bias vectors b. Usually, a neural network model is trained over the whole data many times, where the number of iteration over the whole dataset is referred to as the number of epochs. One epoch includes the whole dataset, and an epoch consist of several batches that are a fraction of the dataset. For each batch, the forward pass and the backward pass are performed to train the model using a loss function. The forward pass is what Figure 2 and equation (3) describe; that is the update of the cell state forward in time and according to the direction illustrated in the figure. The backward pass, also called backpropagation, operates in the reverse manner compared to the forward pass. It determines the gradient of the weights in those matrices and the vectors to improve the model performance, and those weights are updated based on the gradient and a gradient descent optimization algorithm. The learning rate, a hyperparameter, determines the step size of the update at each iteration.

Compared to the usual data based approach where the transfer function (or the auto-regressive model) is used, our approach using the ML model is different in a way that non-linearity is considered in the model. While several methods, such as estimating time-variable or state-dependent parameters of the model, were developed to impose non-linearity in the transfer function model (or the AR model), those methods have their own disadvantages. For example, the time-variable parameter estimation method cannot track true time-variability of the parameters if that time-variability is rapid (Young, 2011). In our model, the non-linearity of the model is determined by the LSTM model structure and the associated parameters.

2.3 Study Site and Data

We use discharge data measured at the Calawah River near Forks, WA, USA (latitude 47°57'30", longitude 124°23'30", USGS gauge 12043000). The 334 km^2 catchment is illustrated in Figure 3A. The elevation ranges from about 40 m to 1200 m, and the average topographic slope of this catchment is 0.36 (Falcone, 2011). The catchment is listed in USGS Gages-II as one of the reference catchments, the least-disturbed catchments within the framework of broad regions (Falcone, 2011). Over 90% of the land cover is forest and about 5% is grass/shrub (see Figure 3B). The town of Forks, located near the catchment outlet, has a development area covering about 1% of the catchment area. The land cover of some parts of the catchment has changed over time. From the three land cover maps shown in Figure 3B, a transition from grass/shrub to forest is observed at the east side of the upslope between 1985 and 2015 as forest recovers from clearing. The recovered area accounts for about 5% of the catchment area. Another logging started in the north side of the catchment around 2000 and continued, with about 5% of the catchment having been cleared in 2015. Correspondingly, the land cover in that region has changed from forest to grass/shrub. Regardless of the land cover change, the proportion of each land cover did not change much during 1985-2015; grass/shrub covers 4-6% of the total area, and forests cover 91-93% of the total area during the period.

The CAMELS dataset (Addor et al., 2017) provides daily precipitation and potential evapotranspiration rates for this catchment, derived from the 1 km resolution Daymet data (Thornton et al., 2016). The CAMELS dataset also provides an estimated actual evapotranspiration rate using the Sacramento Soil Moisture Accounting (SAC-SMA) model (Newman et al., 2015). The period of data provided in the CAMELS dataset is between

1980 and 2014, but some portions of data is missing if, for example, discharge is not measured. As a significant portion of discharge data before March 1984 is missing for this catchment, our study period is from March 1984 to December 2014. For the study period, the average precipitation rate is 3,005 mm/year and the mean discharge rate is 2,819 mm/year. The estimated actual evapotranspiration rate is 476 mm/year. Figure 3C shows the precipitation, the discharge, and the actual evapotranspiration rates. This catchment is wet with an aridity index of 0.25. The mass-balance does not close due to an overestimation of the actual evaporation rate in the SAC-SMA model (as the model underestimated discharge), but note that the recession analysis does not rely on the mass-balance and the quality of the actual evaporation time series. Also, while the actual evaporation rate is similar to that is reported for the region (Sanford & Selnick, 2013), the actual evaporation rate might be low for a forest catchment due perhaps to the underestimated precipitation rate in the Daymet data. Note that the amount of precipitation does not affect the flow recession analysis, whereas the timing of precipitation may have a limited influence on the analysis. Missing precipitation event would result in including misplaced data points in the recession plot if the precipitation event was significant enough to affect the flow recession dynamics. Nevertheless, by only using periods with decreasing discharge in the analysis, the effect of missing precipitation events that are significant enough to increase discharge can be eliminated. Note also that many studies do not use the precipitation data and rely only on discharge data when determining recession periods (e.g., Shaw & Riha, 2012; Jachens et al., 2020).

We use daily data in this study, as daily datasets are more commonly available than higher temporal resolution datasets. However, when using a daily dataset, applying the criterion $Q >> ET$, that is used to estimate the catchment sensitivity function in Kirchner (2009), can exclude a lot of low flow data. Thus we do not use that criterion, and in terms of the catchment sensitivity function, our analysis can be seen as analyzing the function in which the effect of evapotranspiration is included implicitly.

2.4 Applied methods and model setup

We used the precipitation time series and the criterion of $dQ/dt <= 0$ to determine the recession period. Periods with $dQ/dt = 0$ were included since actual decreases in discharge might not be recorded due to the measurement resolution. We have not applied the recession event length-based criterion as we do not perform statistical analy-

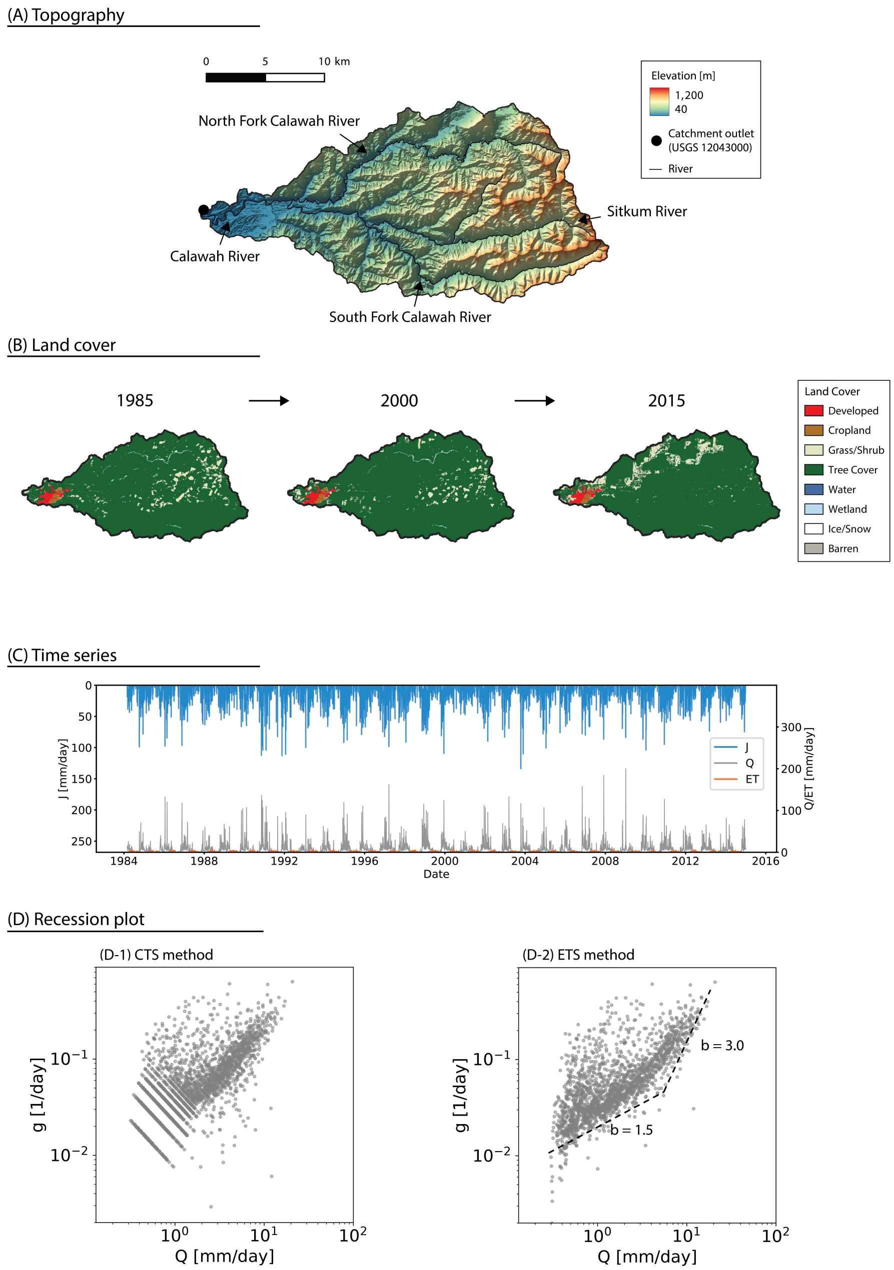

Figure 3. Catchment topography, land cover maps, time series, and the flow recession plots. (A) Topography of the Calawah catchment. This digital elevation map is available through the 3D Elevation Program (3DEP) managed by USGS, and its resolution is 1/3 arc-second. The color illustrates elevation and is shaded with the position of light source at altitude 45° and azimuth 315°. (B) Land cover maps. The LCMAP (Land Change Monitoring, Assessment, and Projection) products managed by USGS were used. (C) Time series of the precipitation J, the discharge Q, and the actual evapotranspiration ET. (D) The recession plots that were estimated using (D-1) the CTS method and (D-2) the ETS method. Note that data points with $dQ/dt = 0$ are not shown in these log-log scale plots. The dotted lines in (D-2) are the lower envelope of (Brutsaert & Nieber, 1977) that were placed close to the lower envelope of the major data points by visual inspection.

sis for each recession event separately. We applied both CTS and ETS methods to estimate the function g. We focus on analyzing the CTS-based estimation since our purpose is analyzing data and because the ETS method involves data smoothing. Nevertheless, we present the ETS-based estimation for comparison.

The LSTM model was constructed with the same structure as described in Figure 2. The model has two recurrent LSTM layers and the dropout layer in the middle. There is also the dense layer after the second recurrent LSTM layer. The mean absolute error (MAE) was used as the loss function. The training period was from March 1984 to December 2000, and the validation period was from January 2001 to December 2014. About 55% of the estimated g values were included in the training period, and another 45% of the values were included in the validation period. The number of LSTM units in each cell in the first layer $n_{u,1}$ and the second layer $n_{u,2}$ were determined using the grid search, and the hyperparameters that minimize the MAE in the validation period were chosen. The values tested in the first grid search were 1, 3, 5, 10, 15, 20, 30, 40, and 50. Additional values of 60, 70, 80, 90, and 100 were tested when the minimum MAE was found at the maximum value explored in the first search. The grid search was performed for several lengths of the past discharge trajectory: 1, 2, 3, 5, 7, and 10 days. The number of trainable parameters n_p is determined by the model structure, $n_{u,1}$, and $n_{u,2}$ as: $n_p = 4(n_{u,1}^2 + n_{u,2}^2 + 2n_{u,1} + n_{u,2}(n_{u,1}+1)) + n_{u,2} + 1$. The Adam solver (Kingma & Ba, 2017) was used for training, and the learning rate was 0.001. The iteration was set to stop if the loss function of the validation set did not improve over 200 iterations. The dropout rate was 0.5. The use of early stopping criteria and the high dropout rate are to reduce overfitting. Also, the model performance during the validation period was checked to ensure that the model performs reasonably well outside of the training period. TensorFlow (Abadi et al., 2015) was used to implement the model.

Here we note that assessing the advantages of the LSTM model over simpler ML models, such as the vanilla recurrent neural network model, is not the focus of our study. A well-known advantage of the LSTM model over simpler neural network models is that the LSTM can utilize longer past data without causing a problem in the parameter estimation, but our application might not take full advantage of the LSTM model if the past trajectory that we need to consider to model flow recession dynamics is relatively short. While how long the past trajectory should be to take the advantage of the LSTM model depends on the properties of the data, there is a chance that we may implement a more complex model than is necessary to efficiently capture the flow recession dynamics. Nevertheless, there are not many benefits to using simpler ML models, especially when our focus is on capturing patterns in data (and not on prediction) and when our focus is not on interpreting the model parameters directly.

3 Results

This section reports the estimated function g and the function learned using the LSTM model. We also show the results of using the central tendency for comparison. Figure 3D shows the recession plots. The catchment sensitivity function g was estimated for 3498 time steps for the CTS method and 2556 time steps for the ETS method. The number of data estimated using the ETS method is less because of the increase of the time step. As expected, the data points are widely scattered. The CTS method-based estimates show a diagonal pattern with its slope of -1 in the low discharge range due to the measurement resolution. The estimation based on the ETS method does not display the pattern as the discharge data was smoothed out. The lower envelope of Brutsaert and Nieber (1977) appears to approximate the lower envelope of the data cloud, with $b = 3$ for high flow and $b = 1.5$ for low flow.

Figure 4A shows the fitted power functions as a measure of central tendency using the binned data. The binned data was estimated using the method suggested in Kirchner (2009) using the CTS method-based estimation. The slope of the fitted line is close to the slope of the lower envelope at low flow and is much lower than the trajectories of each event that are indicated by the gray lines connecting the data points of each event. The coefficient of determination r^2 between the data points and the modeled values using the fitted line is -0.00. Figure 4C shows that there is a structure in the model error. In the modeled value versus the observed value plots, many dots are densely located right above the 1:1 line, and the other dots are very sparsely located under the line. This pattern in the plot, along with the low r^2 values, means that the fitted lines do not represent the data well.

The half-step ahead prediction results of the LSTM model (i.e., $g(t + \Delta t/2)$ estimated using $Q(t)$ and its past values) are shown in Figure 4B. The model results are shown for different lengths of discharge trajectories (1 day and 5 days) that were used in the function H. The LSTM model performance was similar for both training and val-

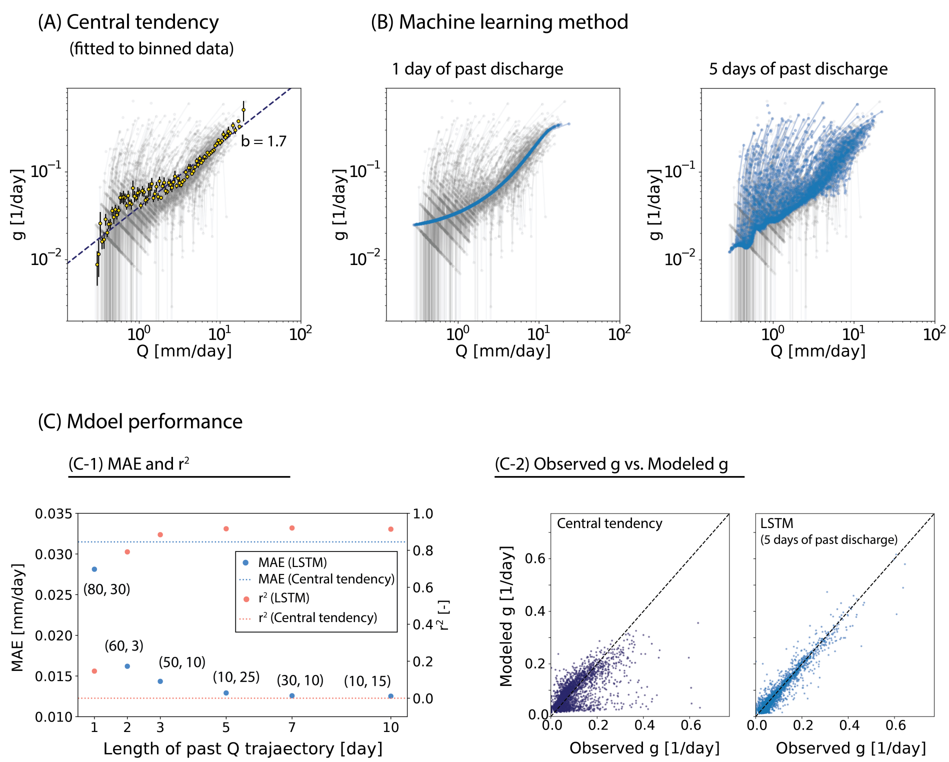

Figure 4. Estimated flow recession dynamics and model performance. The top panels show the estimated flow recession dynamics using (A) the central tendency and (B) the LSTM model. The grey dots are the observed data points, and the grey lines connect the points of each recession event. The yellow circles in (A) are the binned data with the error bar indicating the standard deviation of each bin. The dotted line is the power function fitted to the binned data. In (B), the blue dots are the ML model estimation and the blue lines connect the blue dots of each event. Panel (C-1) illustrates the MAE and the r^2 between the CTS method-based estimation of $g(Q)$ and the modeled $g(Q)$ using the central tendency model and the LSTM model with several lengths of past discharge trajectory as the model input. The chosen values of $n_{u,1}$ and $n_{u,2}$ are displayed in the format $(n_{u,1}, n_{u,2})$. (C-2) illustrates the modeled g and the observed g. The dotted black lines are 1:1 lines.

idation periods (e.g., with the mean absolute error of 0.013 day^{-1} for both periods when 5 days of discharge was used), and the illustrated LSTM results are for both periods. The similar mean absolute error for both periods indicates that overparameterization is unlikely. The plot of the mean absolute error at each iteration during the LSTM model training also does not show any consequences of overparameterization (see Figure S1 for the LSTM model using 5 days of the past discharge). We only show the results estimated using the past trajectory of discharge up to 5 days since there was no significant improvement when we increased the number of days to more than 5 days (see Figure 4C). The chosen number of LSTM units that minimizes the MAE for each case are also illustrated in Figure 4C.

The model results are similar to the pattern of the binned data when only a single discharge value is used, but the model improves significantly as longer past trajectories of discharge are used. The coefficient of determination r^2 is 0.14, 0.88, and 0.92 for the model using 1 day, 3 days, and 5 days of discharge, respectively. Figure 4C shows that the model results are significantly improved compared to the central tendency model. In the modeled value versus the observed value plots, the dots are distributed close to the 1:1 lines.

Figure 5 shows the simulated flow recession dynamics for 16 events. (Note again that the LSTM model can simulate one-step ahead discharge $Q(t{+}\Delta t)$ using the half-step ahead g, $g(t{+}\Delta t/2)$, as described earlier.) In this analysis, we chose events longer than 30 days so that we can see enough recession dynamics for each event. We select events if the condition of $dQ/dt < 0.025$ mm/day^2 holds for more than 30 days, assuming that the discharge increase of 0.025 mm/day over one day is insignificant. Also, the precipitation-based criterion was not applied. As expected, the one-step ahead prediction of Q is close to the observed discharge. When the model is used as a forward model (updating the model input with the modeled Q as it becomes available), the model performance degrades when the first few estimations are biased because the LSTM model was trained for the prediction of the half-step ahead g. Nevertheless, the model tracks patterns of the event trajectories in the recession plotwell as they vary from event to event. (Also, see Figure S2 that illustrates the event-to-event variation more clearly.)

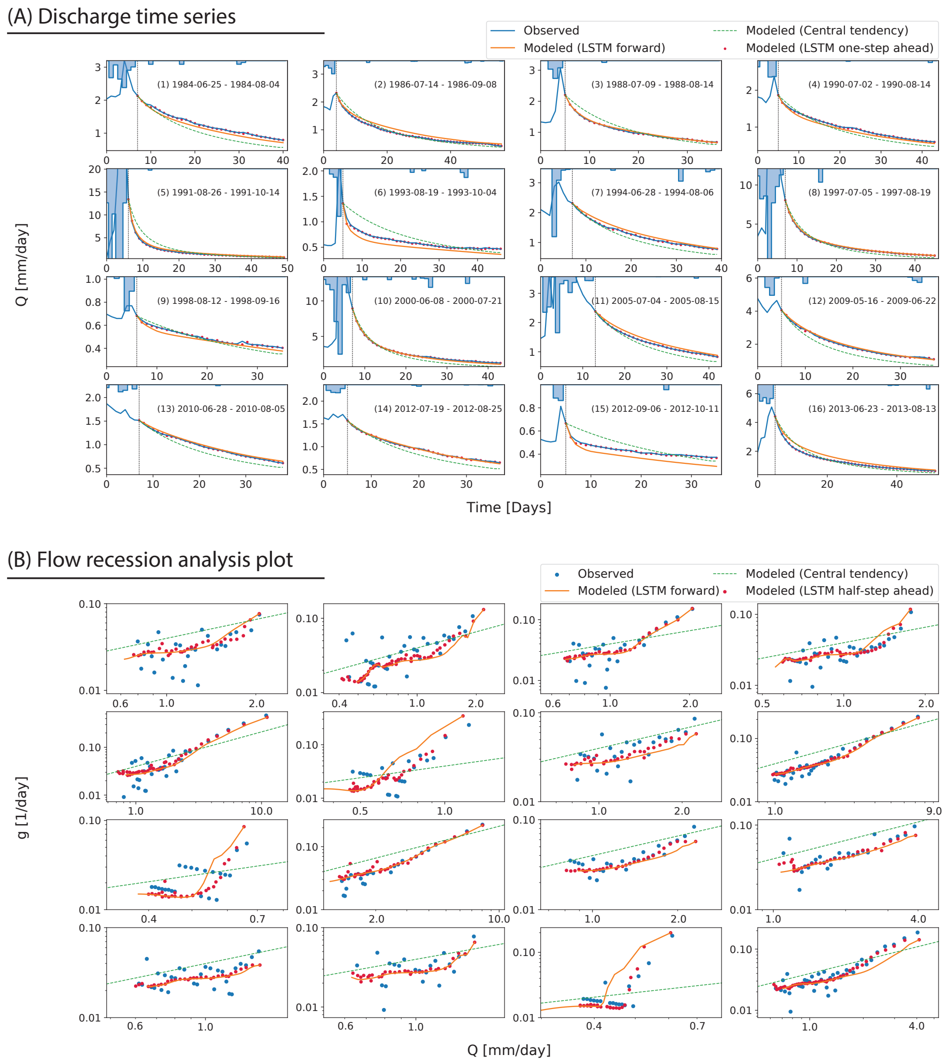

Figure 5. Forward modeling result of the LSTM model for the 16 events. (A) The simulated discharge time series, and (B) the simulated trajectory in the recession plot. The forward model was run after the largest rain event (see the vertical dotted lines in (A)). The red dots represent the one-step ahead or the half-step ahead predictions, and the orange lines illustrate the forward model predictions.

4 Discussion: Learning from the machine

The results indicate that the machine has learned the non-linear hysteretic function H during the flow recession periods. But converting the machine-learned function into a human-readable format is currently a daunting task (e.g., Nearing et al., 2021). It is not easy to interpret the U and W matrices and the b vectors in a physically meaningful way. Nonetheless, our results indicate that the hysteretic recession dynamics can be determined by the last few days of discharge (about 5 days to get $r^2 \approx 0.9$). We can also investigate some machine-learned characteristics and deduce why the machine learned those features. In this study, we focus on analyzing machine-learned characteristics that we observe in the recession plot including the trajectories of each recession event illustrated in the plot as the machine-learned trajectories display patterns that were not clear in the data (see Figure 4). For the discussion, we will treat the recession plot as a phase space plot so that we can leverage terminology and methods developed to explain the trajectory of system dynamics. The result of the LSTM model using 5 days of discharge and the CTS method-based estimation is used for the following analysis. We focus on analyzing the half-step ahead estimation of g instead of the forward model estimation because the half-step ahead estimation is closer to the data (see Figure 5B). Nevertheless, most of the analysis presented in this section are still valid with the forward modeling estimation.

4.1 Recession plot as a phase space plot

In this discussion, we treat the recession plot as "phase space plot". Phase space plots show dynamics of a set of state variables that describe the system state. In other words, phase space plots show (a part of) the phase space where every degree of freedom is represented as an axis in a multidimensional space. The set of state variables of a system is projected as a point in the phase space plot, and its time evolution is represented as a trajectory. Analyzing the trajectory helps understand system dynamics. For example, some systems show an "attractor" in the phase space plot, toward which a system tends to evolve (e.g., Ruelle & Takens, 1971). The phase space plot has been utilized to describe system dynamics in many fields. In classical mechanics, the position and momentum of a particle are used as state variables (e.g., Goldstein, 1980). In thermodynamics and statistical mechanics, macroscopic variables such as pressure and temperature are used as state variables, as considering states of every single particle, i.e., microstates, in a system is not feasible. For example, pressure-volume diagrams have been viewed as describing parts of the phase space. Phase space plots have also been employed occasionally in hydrology. While system-scale variables (macroscopic variables) of hydrologic system that we can measure are limited, several studies utilized phase space plots to analyze catchment dynamics based on measurable or estimable variables (Porporato & Ridolfi, 1997; Duffy & Cusumano, 1998; Beven & Davies, 2015). For example, Duffy and Cusumano (1998) used the concentration-discharge plot as a phase space plot. Beven and Davies (2015) utilized variables such as storage, discharge, and water residence time and transit time, to construct phase space plots.

Discharge is a *measurable* surrogate variable that indicates a state of a catchment. Its time rate of change (or the rate of change divided by discharge; i.e., g) indicates how fast the state evolves. Thus, the recession plot can be viewed as the phase space plot that illustrates a part of the phase space of a catchment. In addition, the recession plot can be thought of as a plot showing a part of the reconstructed phase space based on the method of time delay embedding. Takens' delay embedding theorem asserts that information about the hidden states (unobservable states) of a dynamical system can be contained in a time series of an output and that the phase space can be reconstructed using multiple delayed time series of the output (Takens, 1981). For example, if discharge $Q(t)$ is used as the output, $Q(t)$ and its time delayed variables, $Q(t-\Delta\tau)$, $Q(t-2\Delta\tau)$, $\cdots$, $Q(t-(k-1)\Delta\tau)$, where k is the embedding dimension and $\Delta\tau$ is the time delay, can be used as state variables to reconstruct the phase space (Porporato & Ridolfi, 1997). When k and Δt are chosen appropriately, the reconstruction preserves the properties of the dynamical system that do not change under smooth coordinate changes (i.e., diffeomorphisms); For example, the attractor in the reconstructed phase space is topologically equivalent to the actual attractor, meaning that the trajectory shown in the reconstructed phase space can be used to understand system dynamics. Plotting the dynamics of $Q(t)$ and $dQ(t)/dt$ or $Q(t)$ and $g(t)$ is similar to that of $Q(t{+}\Delta t/2)$ and $Q(t{-}\Delta t/2)$ since $Q(t)$ and $dQ(t)/dt$ (or $g(t)$) have all the information necessary to estimate $Q(t{+}\Delta t/2)$ and $Q(t{-}\Delta t/2)$. In that sense, the recession plot is similar to a reconstructed phase space with the embedding dimension of two. In our case, the time delay is one day.

We note here that suggesting what variables to use to construct (or reconstruct) a phase space that fully describe the system state is not a topic of this study. We only argue that the recession plot, that has been utilized very frequently in hydrology, has a certain similarity to the phase space plot and that we may be benefited by analyzing the recession plot using methods and concepts developed to explain system dynamics using the phase space plots. The embedding dimension of two provides a convenient way of visualization perhaps without losing too much information. While the LSTM model result showed that 5 days of past discharge is needed to capture the flow recession dynamics in the study catchment, $Q(t-\Delta t/2)$ would be a dominant component in determining $Q(t+\Delta t/2)$ or $g(t)$. In what follows we will focus on analyzing the system dynamics shown in the recession plot.

4.2 The attractor and hysteresis in the phase space plot

A characteristic that we observe in the phase space plot is that there is an area where the LSTM estimated points are densely located. Figure 6A shows the Gaussian kernel density estimation $\hat{f}_h(Q, g)$ (e.g., Silverman, 1986) illustrated by the color of each point. Scott's method (Scott, 1992) was used to calculate the bandwidth of the kernel. The yellow and green area is where the points are densely located. The dense area is a region where the catchment has spent a significant amount of time, meaning that the flow dynamics of the dense area are slow or that the flow dynamics associated with that area are repeated frequently. The dense area can be divided into two parts according to its slope in the plot: the lower dense area with low slope (mainly the yellow area) and the upper dense area with high slope (mainly the green area).

An event trajectory shows that the flow dynamics in the lower dense area ($\hat{f}_h >$ 0.35) is slow. The red line in Figure 6A is the LSTM model learned trajectory of an event from Sep. 1, 1991 to Oct. 14, 1991, which ended up in the yellow area. The event spent about half of its time in the yellow area (see the discharge time series in Figure 6B), while the line length of trajectory in the recession plot is much shorter inside the yellow area than the line length of trajectory of the earlier period. Note that the event trajectory in the yellow area also can be estimated using the ETS method, but is not easy to estimate using the CTS method-based estimation due to greater noise (see Figure 6C). Note also that several parts of the trajectory (the red line) are indicated by dashed lines when the associated period is not determined as a recession period. According to the criteria for determining recession periods that we applied, this event was divided into three recession events due to a very small precipitation event (0.83 mm/day) and two small increases in discharge (about 0.02 mm/day increase over one day; see Figure 6B). However, looking at the discharge time series, it makes sense to treat the entire event as a single recession event. The precipitation event appears to be too small to affect the flow dynamics. The increases at two times are very small, and since the cause of the small increases is not clear, it seems better not to use the two small increases to determine the recession period.

The yellow area is not only the area where the flow dynamics are slow but also the area that is often explored. Figure 6E shows that all 16 recession events over 30 days, which were selected previously, converge to the yellow area and then move along that area towards the lower-left corner. The same pattern is also observable in the forward model result (see Figure S2). It means that the yellow area behaves like an "attractor", where all dynamics converge to that area and then move within that area, unless those dynamics are pushed away from it by external forcings. (See Beven and Davies (2015) for more discussion on the attractor in catchment hydrology.) This attractor will be called the "catchment flow attractor" because the attractor is a signature of catchment scale flow dynamics. (Note again that we only focus on the flow recession dynamics in this study and that exploring the potential existence of the catchment flow attractor in the rising limb of discharge is left for future research.) The dynamics in the catchment flow attractor are expected to be equilibrated at a fixed point of zero flow as a point of "maximum entropy" (Beven & Davies, 2015). This state was not explored in this catchment because external forcing (e.g. precipitation) constantly pushes the system away from the point of maximum entropy.

The dense area is where the most characteristic information about catchment scale recession dynamics exist. The area is a better representation of the ensemble of many recessions than the measure of central tendency and the lower envelope of Brutsaert and Nieber (1977). While the binned data captures the pattern of the dense area (see Figure 4A), the binned data places above the dense area because it considers all data points. The situation is similar for the ML result with insufficient length of the past discharge. The full consideration results in the structure of the errors in the modeled g versus observed g plot (Figure 4C), and the error in the forward simulation using the central tendency model (Figure 5). While the performance of the central tendency model can be improved when some data points are filtered out before fitting the line (e.g., filtering out the first few days of data after each rain event and thus focusing more on the late time dynamics and the attractor), it certainly reduces the information content in data and neglects the hysteretic dynamics. The method of Brutsaert and Nieber (1977) seems to fit the dense area to some extent (see Figure 3). However, we lack a method to fit the lower envelop objectively (e.g. Jachens et al., 2020). Furthermore, the upper part of the lower envelop with $b = 3$, which is predicted by the Boussinesq model, is much steeper than the slope of the upper dense area.

The dense area can be parameterized to describe the flow recession dynamics within the area. A function consisting of two linear lines (in log-log space) can be fitted to the data points located in the dense area ($\hat{f}_h > 0.2$). The function can be written as: $\ln g = \max(a_1 + (b_1 - 1)\ln Q, a_2 + (b_2 - 1)\ln Q)$. The crossover between the two lines occurs at $Q^* = (a_2 - a_1)/(b_2 - b_1)$. The lower line fits the catchment flow attractor with $b = 1.69 \pm 0.00$ up to $Q = 3.29$ mm/day (see the black dotted line in Figure 6A). The value is similar to that of the late time dynamics of the Boussinesq model ($b = 1.5$). The slope of the upper line is $b = 2.10 \pm 0.02$. This value is smaller than the value of early time recession of the Boussinesq model ($b = 3$). The slope $b = 2.10$ is similar to the median value of 2.0 which is derived from the event-based analysis for 39 catchments in USA that are not affected by anthropogenic activities (Biswal & Marani, 2010). (Note that more objective or sophisticated parameterization schemes to fit the dense area, such as using the modal linear regression (Yao & Li, 2014), applying a variable threshold for $\hat{f}_h$ over Q, or using a higher-order polynomial in the log-log space, might be applicable but are not employed in this study.)

Based on the trajectory of each event in the recession plot and the attractor, we can define "early time" recession and "late time" recession for each event. The early time recession is until the trajectory converges to the attractor, and the late time recession is after the trajectory converges to the attractor (see Figure 6A). Figure 6G shows that the attractor consists of the late time recession in that there were no increases in the recent past discharge data, while Figure 6F shows that above the attractor, there were increases in the discharge data of the recent past. This definition has an important difference from the original definition based on the Boussinesq model result discussed earlier. Our definition is based on data and the characteristic extracted from data and applies for each event trajectory, whereas the original definition is based on the process-based model and when describing data, it applies to the lower envelope of data.

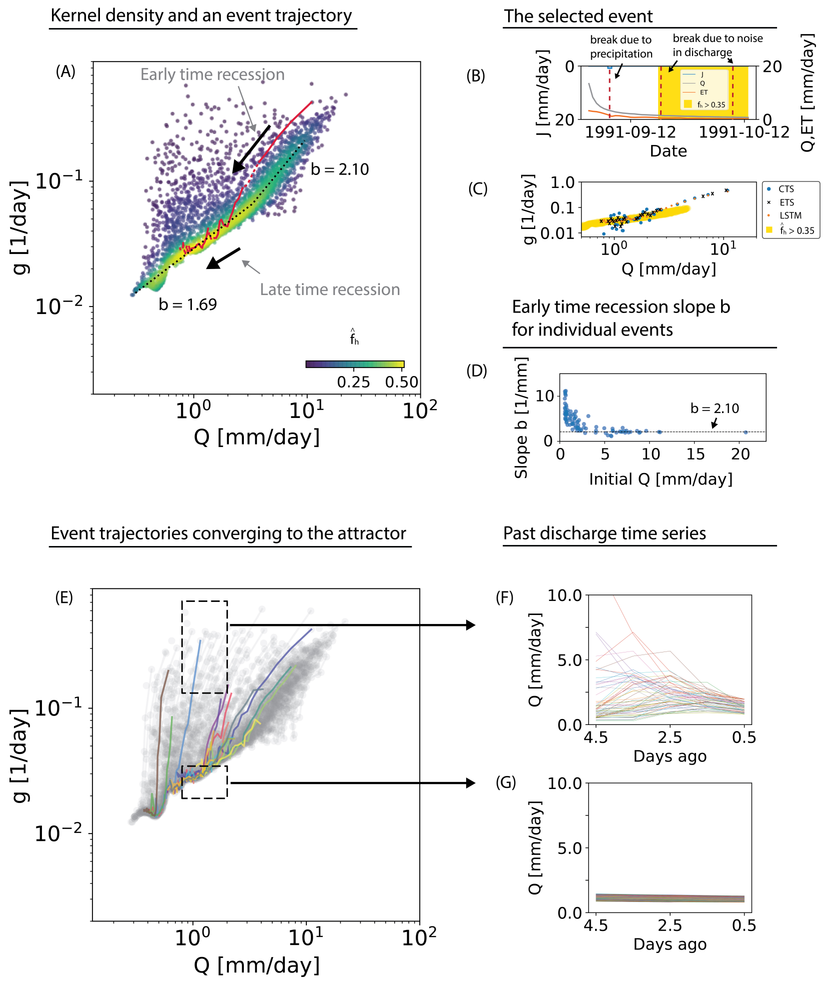

Figure 6. Learning from what the machine learned. (A) Kernel density estimation at each data point. Density is displayed in colors from yellow (dense) to blue (sparse). The red line is the trajectory of the events from September 1, 1991 to October 14, 1991. The line is a solid line during the periods that are determined as a recession period. Otherwise, it is a dashed line. The black arrows indicate the direction of the flow recession dynamics in the plot, and the black dashed lines are the power functions that are fitted to the dense area ($\hat{f}_h > 0.2$). (B) Time series of the precipitation, the discharge, and the actual evapotranspiration during the event. If we use the recession period determination criteria discussed in the text, this event is divided into three events, and the vertical dotted lines show the timing of the division. The yellow area represents the period during which the event moves within the yellow area ($\hat{f}_h > 0.35$) shown in (A). (C) Data points of the event that are estimated using several methods. (D) The fitted early time recession slope b for the events that converged to the attractor as a function of the initial discharge at recession. (Note that only events that have more than three data points in the early time recession were used.) (E) LSTM model-learned trajectories of all events longer than 30 days. (F) 5 days of the past discharge for the data contained in the upper box in (E), and (G) 5 days of the past discharge for the events contained in the lower box in (E).

The early time recession dynamics are, in general, similar to the dynamics reported in the event-based analysis. There is large event-to-event variability. As Jachens et al. (2020) reported, the recession event with lower initial discharge tends to have higher b values (see Figure 6D). At which discharge values a recession event converges to the attractor tends to vary from event to event but with a trend: recession events with lower initial discharge tend to converge to the attractor at lower discharge values. Furthermore, the early time trajectories are mostly convex similar to the result of Tashie et al. (2020). An important difference is that, as opposed to the claim of Tashie et al. (2020), only the early time trajectory is convex, but the entire recession event trajectory is concave unless the event trajectory is forced away prior to its convergence to the catchment flow attractor by external forcings. The event-to-event variability of the early time recession dynamics are reduced for the events with high peak discharge values, resulting in the upper dense area.

These results would suggest that the analysis of the curvature of event trajectory is sensitive to two factors. First, it is sensitive to the $-dQ/dt$ estimation method and the recession event determination criteria. Tashie et al. (2020) used the CTS method to estimate $-dQ/dt$ and used the criteria of decreasing both Q and $-dQ/dt$ for more than 5–7 consecutive days to determine recession periods. Thus, it is possible that the early time dynamics is treated as one event, and the late time dynamics is treated as another event (which is mostly linear in the plot) or not considered as a recession event due to the noisy CTS method-based estimation (e.g., see the previous discussion about the September 1991 - October 1991 event). Second, it is sensitive to precipitation events. As we described earlier, precipitation events can push the dynamics away from the catchment flow attractor before a trajectory converges to the catchment flow attractor. When this happens frequently (e.g., in wet catchments), usual event-based analysis can place more weight on the early time dynamics than the late time dynamics.

4.3 Attractor and the master recession curve

The existence of the catchment flow attractor implies that, at some point in recession, multiple time scale dynamics may reduce to slower dynamics that are similar for all events. The slow dynamics in the catchment flow attractor can be described using the fitted line. The function g decreases with decreasing Q approximately following the power function $g = aQ^{b-1}$, where $b = 1.69$ in this case. When g is the power func-

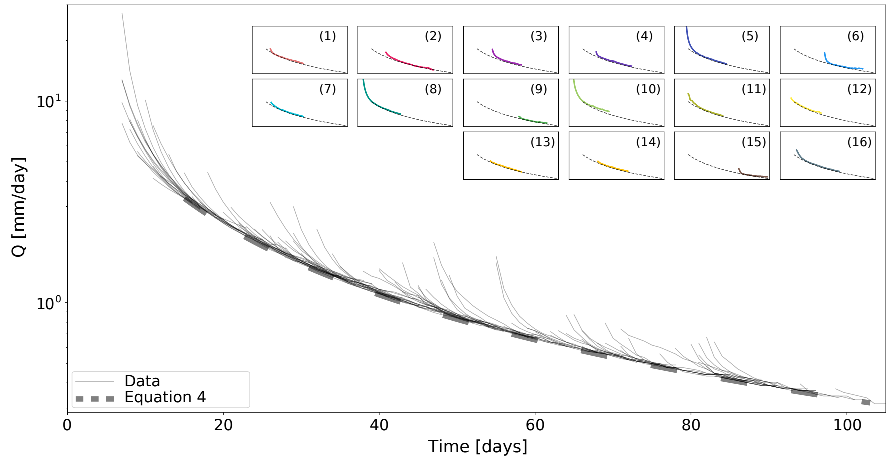

Figure 7. The attractor as the master recession curve. The thin lines illustrate the discharge time series of all recession events longer than 8 days. The thin lines are shifted over time from right to left until it meets the parameterized master recession curve (the thick dashed line).The parameterized master recession curve was determined using Equation 4 with the parameters that are estimated based on the CTS method estimation and the LSTM model using the past 5 days of discharge. The subset figure shows the parameterized master recession curve (the dotted line) and the time-shifted discharge time series of the previously selected 16 events (the solid line).

tion of Q (i.e., $g = aQ^{b-1}$ and $-dQ/dt = aQ^b$), the flow recession in the catchment flow attractor can be written as (e.g., Rupp & Woods, 2008):

$$Q(t) = (Q_0^{1-b} + a(b-1)t)^{1/(1-b)} \tag{4}$$

where Q_0 can be chosen as discharge at a time when the system dynamics converge to the catchment flow attractor, and t is the time lapse since the system converges to the catchment flow attractor. When $b \to 1$, $Q(t) = Q_0 e^{-a/t}$, and the catchment behaves like a linear reservoir. When $b > 1$, the tail of the discharge time series is heavier than the exponential decay. Figure 7 illustrates that equation (4) with the estimated parameters captures the late time flow recession dynamics of each recession event that is longer than 8 days during the study period. The duration criterion was used to filter out as many events as possible that did not converge to the attractor due to precipitation events that occur before the trajectory of the recession event converges to it.

The curve represented by equation (4), that we estimated based on the parameterized catchment flow attractor, is indeed the master recession curve. The term "master recession curve" was coined in Nathan and McMahon (1990) and introduced as a catchment characteristic that represents the most frequent low flow recession dynamics. The master recession curve estimated using the LSTM model result and the additional fitting is an ensemble of the late recession dynamics that can be thought of as a central tendency of the dynamics. The master recession curve has been estimated and discussed in many catchments (e.g., Nathan & McMahon, 1990; Lamb & Beven, 1997; Fiorotto & Caroni, 2013), but this is the first to derive a representation as an attractor that overcomes the variations from event to event without convergence (e.g., Tashie et al., 2020). Our results suggest that this may be the result of combination of the noise in data and the criteria for defining the recession periods that make it hard to recognize the master recession curve from plots of individual recessions. For the catchments where the master recession curve exists, we can expect that the attractor would exist in the recession plot and that each event trajectory converges to the attractor unless pushed away from the attractor due to external forcings. Thus, care must be taken when analyzing the recession plot especially for the low discharge range.

4.4 Process-based interpretation

While did not attempt to provide detailed physical interpretation based on process-based models, we can still infer some processes that might have resulted in the dynamics we observed in the phase space plot and Figure 7. The event to event difference of the early time recession, i.e., the dynamics before those converge to the attractor, might exist due to the difference in the initial condition and boundary conditions (e.g., external forcings, including precipitation and snowmelt, and consequent patterns of storage in the catchment). Difference in the initial condition and the boundary condition would also result in different hydrologic connectivity that could affect to flow recession dynamics. Those conditions would show seasonality, meaning that the event-to-event variability of the early time dynamics may also be dependent on evapotranspiration rates and seasonality. For example, most of the hysteresis is observed in dry seasons (see Figure S3). Sometime later the dynamics of each event converge to the attractor, as the effects of those conditions and forcings vanish. When the effects vanish, the spatial distribution of celerity that controls discharge could be uniquely characterized by the value of discharge at that time.

The contrasting slope of each recession trajectory in the recession plot where it is high during the early time and is low during the late time may also indicate some processes. During the early recession, the discharge decreases at a faster rate. This may be due to the continuous deactivation of some fast flow pathways, such as overland flow and macropore flow, rapid reduction in transmissivity with lower storage, and rapid contraction of variable source areas. For the late time dynamics, we hypothesize that most of the fast flow paths were already deactivated, the contraction of the variable source area is slow, and the flow dynamics are largely dominated by subsurface flow and perennial stream flow, resulting in low g values. During the late time, the Boussinesq model result of the slope of the trajectory is not far from the slope of the parameterized attractor, indicating that the Boussinesq model may be applicable to explain the late time dynamics.

It might be worth noting here that the attractor we defined does not seem to change over time during the study period, while the land cover has changed to some extent mostly as a result of logging. The effect on the shape of the recession part of hydrograph seems negligible based on our analysis as the late time recession dynamics of each recession event can be approximated by the estimated master recession curve (see Figure 7 where all available data are plotted; see also Figure S4 where data for three periods, 1984-1993, 1994-2003, and 2004-2014, were plotted in the same way used to construct Figure 7). Furthermore, the model performance is similar for the training period (March 1984 to December 2000) and the validation period (January 2001 to December 2014) with low mean absolute error for both periods. While there are many studies that show that logging significantly impacts discharge (e.g., Moore & Wondzell, 2005), their implication to our study is unclear. Those studies mostly focus on analyzing small headwater catchments where a significant portion of the catchment is logged, and those studies mostly focused on analyzing the quantity of discharge (e.g., increasing discharge and decreasing evaporation at much larger time scale such as month or annual) and the peak flow rather than the shape of the recession part of hydrograph. Nevertheless, based on studies that show the effect of evapotranspiration on the low flow dynamics (e.g., Szilagyi et al., 2007), one may expect to observe the effect of logging on the shape of the attractor. As potential reasons for no significant temporal variation of the attractor in our study catchment, we can hypothesize that: 1) the scale of this catchment is large enough compared to the area of the land use change; Only about 5% of the catchment has been recovered during the period, and about another 5% of the catchment has been logged during about the latter half of the study period according to the LCMAP data, 2) in this catchment, the shape of the recession hydrograph might not be affected much by the land cover change and the associated change in water partitioning, but other factors such as geomorphologic structure, soil hydraulic properties, and geology that were not changed significantly during the study period dominantly determine the shape.

4.5 Estimating storage-discharge relationship using the attractor

The catchment flow attractor can be utilized to estimate the hysteretic active storage-discharge relationship. In previous studies, the catchment sensitivity function that is estimated as a central tendency has been used to estimate the storage-discharge relationship (e.g., Kirchner, 2009; Dralle et al., 2018), neglecting the hysteresis in the storage-discharge relationship. The existence of the attractor implies that the hysteresis in the storage-discharge relationship is not detectable from the discharge data after each recession event converges to the attractor, while the hysteresis is detectable before the system dynamics converge to the attractor. It means that a non-hysteretic storage-discharge relationship would sufficiently capture the catchment dynamics during recession periods inside the attractor. Using the non-hysteretic part of the relationship, the hysteretic storage-discharge relationship can be estimated in terms of drying scanning curves if we calculate the storage using the mass-balance backward in time starting from the attractor.

Figure 8 shows the (relative) active storage-discharge relationship for the two events (the 1998 July - September event and the 2013 June - August event that are shown in Figures 5 and 6) estimated considering the discharge time series; i.e., $dS/dt = -Q$. The relative active storage was estimated from the point marked by 'X' with the initial condition of zero relative storage. The storage-discharge relationship in Figure 8 shows that the event trajectories overlap at a low flow condition, when the system flow dynamics moves inside the attractor. The overlapped trajectory can be captured by the storage-discharge relationship that is estimated using the parameterized $g(Q)$ for the attractor (see Figure S5). While we estimated the storage from the certain point in the example, it is straightforward to generalize it by estimating the storage-discharge relationship associated with the attractor first and then calculate the storage backward in time from

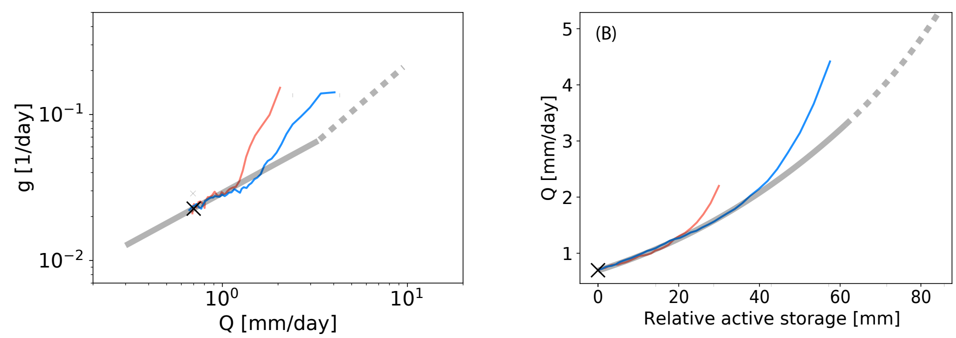

Figure 8. The recession plot and the corresponding storage-discharge relationship. (A) Two event trajectories in the recession plot are illustrated by the red and the blue lines. The solid grey line shows the parameterized attractor, and the dashed grey line shows the parameterized upper dense area. (B) The corresponding active storage-discharge relationship. The marker 'X' in both (A) and (B) indicates g and the active storage at a low flow condition at which the storage is set to zero. The solid grey line and the dashed grey line illustrate the relationship estimated using the parameterized attractor and the parameterized upper dense area.

the attractor. The storage-discharge relationship associated with the upper dense area can also be used to estimate the hysteretic storage-discharge relationship at high flow conditions.

It is also possible to estimate the relative "total" storage considering ET from an initial condition; see Figure S5. The figure implies that another attractor may be found using $g = (dQ/dt)/(-Q - ET)$ (instead of using $g = (dQ/dt)/(-Q)$) and that the attractor may be utilized to estimate the hysteretic (relative) total storage-discharge relationship. Note again that the denominator of g is dS/dt in its full formulation, and the form used in (1) neglects the effect of ET in the storage variation. While this method is, in part, based on the mass-balance, it is different from the traditional mass-balance approach that estimates the relative total storage starting from a fixed initial time. The traditional method can result in the drift of storage over time when the mass-balance is not closed and the uncertainty in the estimated storage accumulates over time. In the method using the attractor, the initial time of storage calculation is the most recent time when the system dynamics is in the attractor, reducing the uncertainty. It is more likely that this attractor may exist under the water-limited condition where ET is limited by

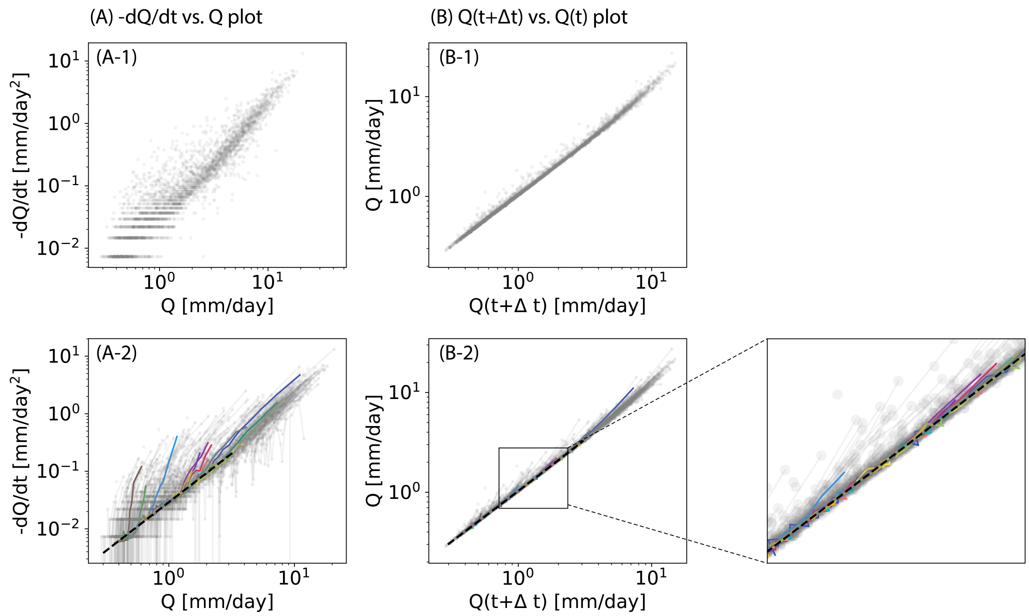

Figure 9. Other phase space plots. (A) $-dQ/dt$ vs. Q plot and (B) $Q(t + \Delta t)$ vs. $Q(t)$ plot. The top figures show the data during the flow recession periods. The bottom figures show the trajectory of each recession event (the grey lines) and the trajectory of the sixteen events that were selected previously (the lines colored in a same way to Figures 7 and 8). The black dashed lines illustrate the parameterized attractor from the g vs. Q plot using the LSTM model result.

water availability. We leave a further discussion about the effect of ET on the catchment sensitivity function and the total storage-discharge relationship for future study.

4.6 Attractor in other phase space plots

In this last subsection of discussion, we briefly introduce other phase space plots and how the attractor and the trajectory of system dynamics appear in the plots. Some plots that we previously discussed can be treated as a phase space plot. For example, the storage-discharge plot is a phase space plot, and where the event trajectories overlap (e.g., the storage-discharge relationship approximated using the parameterized $g(Q)$ for the attractor in the recession plot) is the attractor in the storage-discharge relationship (see Figure 8 and Figure S5). Figure 7 can also be thought of as a phase space plot, while unlike other phase space plots, a reconstruction of data (i.e., shifting event discharge time series) is required to produce the plot. In the phase space plot, the master recession curve is the attractor.

In addition, the plot of $Q(t+\Delta t)$ vs. $Q(t)$ and the other type of the recession plot, $-dQ/dt$ vs. Q plot, are phase space plots. As expected, the $-dQ/dt$ vs. Q plot shows information in a similar way to the g vs Q plot (see Figure 9A), while the hysteretic dynamics are displayed more clearly in the g vs Q plot. All findings that we draw from the g vs Q plot can be observed in the $-dQ(t)/dt$ vs Q plot. For example, while the attractor is not clearly visible in the $-dQ(t)/dt$ vs Q plot generated by data (Figure 9A-1), the machine-learned trajectories show the attractor clearly (Figure 9A-2).

Figure 9B shows the plot of $Q(t+\Delta t)$ vs. $Q(t)$. This plot was introduced in Langbein (1938) and discussed in Linsley et al. (1958) and Brutsaert (2005). They described that the lower envelope and the upper envelope can be used to characterize the slowest recession and the fastest recession dynamics, respectively. We discussed in section 4.1 that, this plot can be used as a phase space plot and, in theory, would show the same information compared to the other phase space plots we discussed. However, there is a notable difference in the pattern of data shown in this plot compared to the pattern in other phase space plots: The area where the data points are densely located is clearly displayed over the entire discharge range. This is because the derivative of discharge, that induced the noise in the recession plot, is not involved. There is noise in the discharge data that creates a zigzag pattern in the event trajectory (see Figure 9B-3), but the noise does not appear large enough to obscure the lower envelope. The parameterized attractor from the g vs. Q plot can be displayed in this plot using: $Q(t+\Delta t) = (Q(t)^{1-b} + a(b-1))^{1/(1-b)}$, that is derived using equation (4), and the parameterized attractor fits the dense area for the low discharge range (< 3.29 mm/day; see Figure 9B-2). It might be tempting to parameterize the attractor from the $Q(t+\Delta t)$ vs. Q plot as the dense area displays clearly unlike the data represented in the recession plot, but care must be taken. While the shape of the dense area for the low discharge range looks almost linear, implying that $b = 1$ in $-dQ/dt = aQ^b$, the b value estimated from the g vs Q plot is 1.69. The parameterized attractor with $b = 1.69$ looks almost linear in this plot, illustrating that the non-linearity is not clearly visible. (See Figure S6 for the consequence of using the $b = 1$ to characterize the low flow dynamics). In addition, the degree of hysteresis is suppressed in this plot compared to other plots that we discussed; however, it does not mean that the hysteresis is negligible as we discussed in preceding sections (e.g., see Figures 6 and 7).

In summary, those plots show the same information in a different way, and some information is displayed more clearly in one plot than the others. The hysteretic flow recession dynamics are shown more clearly in the g vs. Q plot or in the $-dQ/dt$ vs Q plot than the plot of $Q(t{+}\Delta t)$ vs. $Q(t)$. The existence of the attractor can be more clearly inferred from the $Q(t{+}\Delta t)$ vs $Q(t)$ plot. It might be worthwhile to examine other phase space plots, e.g., Poincaré section of Porporato and Ridolfi (1997, 2003) (see Figure S7), if there is additional information about catchment dynamics that we could learn. For example, some phase space plots, such as the Poincaré section, include the rising limb of discharge data. What we have described in this study can be used to explain the flow recession dynamics in the plot (as described in Figure S7), and there may be room for better understanding the rising limb of discharge by utilizing such a plot. Also, there still might be unexplained patterns in the recession data which may be displayed more clearly in other phase space plots.

5 Conclusions

The flow recession analysis has been presented as a tool to understand catchment scale flow dynamics and catchment properties (e.g., Troch et al., 2013). However, there are seemingly contrasting methods of extracting information from the flow recession plot (Q versus $-dQ/dt$ or $(-dQ/dt)/Q$). Traditional methods use the lower envelope to capture the ensemble characteristics of many recessions (Brutsaert & Nieber, 1977), or use a fitted function to entire data points as a measure of centrality (Vogel & Kroll, 1992; Kirchner, 2009). In contrast, recent studies highlight the importance of the event scale analysis and have questioned the use of the lower envelope and the measure of centrality (Jachens et al., 2020; Tashie et al., 2020).

Based on the machine learning model results, we emphasize the importance of analyzing both the ensemble characteristics and the event scale dynamics. The machine learning model, the Long Short-Term Memory (LSTM) model using 5 days of past discharge, captures both the ensemble characteristics and the event scale dynamics of the Calawah catchment. The LSTM model results indicate that the early time dynamics, which are sensitive to initial conditions, lead to the hysteretic trajectories of system dynamics that appears in the recession plot. Analyzing such hysteretic trajectories (event scale trajectories of the early time dynamics) is the focus of previous event scale analysis studies (Jachens et al., 2020; Tashie et al., 2020). The model results further show that the trajectories of system dynamics converge to an attractor, the catchment flow attractor, unless pushed away from the attractor due to external forcings. The early time recession dynamics of large events also share similar trajectories (i.e., the upper dense area determined in the Gaussian kernel density analysis), perhaps because those dynamics for larger events are less sensitive to initial conditions. The catchment flow attractor and the upper dense area represent ensemble characteristics of many recessions. We also briefly illustrated that the catchment flow attractor can be utilized to estimate the the drying part of a hysteretic storage-discharge relationship. The active storage estimated in this study might be affected by evapotranspiration because we did not apply the criterion of $Q >> ET$. The criterion was not applied mainly because we used daily dataset and since applying the criterion would remove significant portion of the dry season data. However, the shape of the attractor does not show seasonal variation, implying that the effect of evapotranspiration on the attractor is not significant in the study catchment. (Note that the effect of evapotranspiration during the recession period may appear more clearly at low flow conditions; see Szilagyi et al. (2007).) One way to confirm the effect of evapotranspiration on the active storage and the catchment sensitivity function would be using hourly data. Applying the condition of $Q >> ET$ to hourly data could avoid filtering out too much of dry season data. For the hourly application, the advantage of the LSTM model will become pronounced, as a larger number of past values would be required to capture the hysteretic flow recession dynamics.

It might be worth noting that our findings are based on an effort to find patterns in the "data-based" modeling result (i.e., the LSTM model results) and to explain these patterns. There are hydrologic models, such as the two bucket model operating in parallel, that could reproduce these patterns. When the dynamics of the faster bucket become negligible to the discharge flowing out of the entire system (i.e., during the late time recession), the slower bucket dominates the flow dynamics, which would then determine the attractor. Depending on the relative contribution of each bucket, the early time recession dynamics can show event-to-event variability. While such a model has been applied to explain the dynamics shown in the recession plot (e.g., Gao et al., 2017), the existence of the attractor in the recession plot and its relation to the master recession curve has not been discussed. That is because such a model application is, in general, limited to explain already *known* patterns (e.g., the time-variability of the early time recession dynamics). In terms of finding new patterns out of noisy data, we believe that applying a “data-based” model is preferable as it is relatively free from model structure error compared to models in which their structure is determined *a priori*. Here, a potential advantage of using a ML model over the traditional data-based model where transfer functions or autoregressive models are used is how flexibly the non-linearity of the model can be considered as we described in section 2.2.

While we focused on analyzing one catchment, we believe that the ML model designed to capture the flow recession dynamics and the developed analysis tool can be generalized in several ways to improve our understanding of catchment scale flow dynamics. This analysis can easily be extended to the continental scale or to the global scale by analyzing many catchments. Analyzing more catchments will allow us to examine if catchment attributes (e.g., area, aridity index, topographical, geological, and ecological properties) can explain some patterns, such as the existence of the dense area (including the attractor) and its slope, concavity, and extent.

Machine learning tools are powerful in that the model structure is flexible. Rather than using only discharge Q, other variables can be used in the function H to examine if there is a better surrogate variable for the function or depending on a purpose of analysis. For example, the past trajectory of precipitation J can be used in the H function when the prediction of an ungauged basin is of interest. Also, both J and Q (and also ET) can be used to better capture the flow recession dynamics and the rising limbs. For a better forecasting, the model can also be trained while continuously updating the modeled Q as the input. Furthermore, the model can also easily be modified to estimate Q instead of g. We showed that the machine learning model result provide a convenient way to extract information out of the noisy catchment scale signature, the recession plot. Following the discussion in Beven (2020a), we hope the approach we applied in this study, making inferences from what the machine learned and what it needed to learn, will be useful for understanding more catchment scale dynamics when such inferences are well guided by scientific knowledge.

Appendix A Non-hysteretic active storage-discharge relationship and one-to-one relationship in the recession plot

Let us assume that there is an invertible, one-to-one relation p so that $S_a = p(Q)$. We also assume that p is differentiable. The temporal fluctuation of the active storage during flow recession periods can be estimated as: $dS_a/dt = -Q$ if we assume negligible water exchange between the active storage and other compartments such as the inactive storage and negligible evapotranspiration loss from the active storage. (This active storage is identical to the "dynamics" storage in (Staudinger et al., 2017).) The ordinary differential equation can then be rewritten as $dS_a/dt = (dp(Q)/dQ)(dQ/dt) = -Q$ and then $-(dQ/dt)/Q = 1/p'(Q)$, and the right hand side term is, by definition, one-to-one relation. $1/p'(Q)$ is indeed $g(Q)$.

Acknowledgments

We gratefully acknowledge the support of National Science Foundation grants EAR-2120113 and GCR-2121155. We also acknowledge support from the Philecology Foundation of Fort Worth Texas. Additional funding support was provided by the Office of the Vice President of Research at the University of Arizona and by the Technology and Research Initiative Fund (TRIF) Water, Environmental, and Energy Solutions (WEES) initiative at the University of Arizona (Shared Equipment Enhancement Funds). The CAMELS dataset is available from https://ral.ucar.edu/solutions/products/camels (Newman et al., 2015).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., ... Xiaoqiang Zheng (2015). *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.*

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. (2017). The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrology and Earth System Sciences*, *21*(10), 5293–5313. doi: 10.5194/hess-21-5293-2017

Anderson, M., & Burt, T. (1980). Interpretation of recession flow. *Journal of Hydrology*, *46*(1), 89-101. doi: https://doi.org/10.1016/0022-1694(80)90037-2

Barnes, B. S. (1939). The structure of discharge-recession curves. *Eos, Transactions American Geophysical Union*, *20*(4), 721-725. doi: https://doi.org/10.1029/TR020i004p00721

Beven, K. (2006a). A manifesto for the equifinality thesis. *Journal of Hydrology*, *320*(1-2), 18–36. doi: 10.1016/j.jhydrol.2005.07.007

Beven, K. (2006b). Searching for the Holy Grail of scientific hydrology: Qt =

(S,R,Delta t)A as closure. *Hydrology and Earth System Sciences*, *10*(5), 609–618. doi: 10.5194/hess-10-609-2006

Beven, K. (2020a). Deep learning, hydrological processes and the uniqueness of place. *Hydrological Processes*, *34*(16), 3608-3613. doi: 10.1002/hyp.13805

Beven, K. (2020b). A history of the concept of time of concentration. *Hydrology and Earth System Sciences*, *24*(5), 2655–2670. doi: 10.5194/hess-24-2655-2020

Beven, K., & Davies, J. (2015). Velocities, celerities and the basin of attraction in catchment response. *Hydrological Processes*, *29*(25), 5214–5226. doi: 10.1002/hyp.10699

Biswal, B., & Marani, M. (2010). Geomorphological origin of recession curves. *Geophysical Research Letters*, *37*(24). doi: 10.1029/2010GL045415

Brutsaert, W. (2005). *Hydrology: An Introduction.* Cambridge University Press. doi: 10.1017/CBO9780511808470

Brutsaert, W., & Nieber, J. L. (1977). Regionalized drought flow hydrographs from a mature glaciated plateau. *Water Resources Research*, *13*(3), 637–643. doi: 10.1029/WR013i003p00637

Carrer, G. E., Klaus, J., & Pfister, L. (2019). Assessing the catchment storage function through a dual-storage concept. *Water Resources Research*, *55*(1), 476-494. doi: 10.1029/2018WR022856

Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H. J., Peters, N. E., & Freer, J. E. (2009). Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale. *Hydrological Processes*, *23*(2), 311–319. doi: 10.1002/hyp.7154

Davies, J. A. C., & Beven, K. (2015). Hysteresis and scale in catchment storage, flow and transport. *Hydrological Processes*, *29*(16), 3604–3615. doi: 10.1002/hyp.10511

Dralle, D. N., Hahm, W. J., Rempe, D. M., Karst, N. J., Thompson, S. E., & Dietrich, W. E. (2018). Quantification of the seasonal hillslope water storage that does not drive streamflow. *Hydrological Processes*, *32*(13), 1978–1992. doi: 10.1002/hyp.11627

Dralle, D. N., Karst, N. J., Charalampous, K., Veenstra, A., & Thompson, S. E. (2017). Event-scale power law recession analysis: Quantifying methodological uncertainty. *Hydrology and Earth System Sciences*, *21*(1). doi: 10.5194/hess-21-65-2017

Duffy, C. J., & Cusumano, J. (1998). A low-dimensional model for concentration-discharge dynamics in groundwater stream systems. *Water Resources Research*, *34*(9), 2235–2247. doi: 10.1029/98WR01705

Falcone, J. (2011). *GAGES-II: Geospatial Attributes of Gages for Evaluating Streamflow* (Tech. Rep.). Reston, Virginia: U.S. Geological Survey. doi: 10.3133/70046617

Fiorotto, V., & Caroni, E. (2013). A new approach to master recession curve analysis. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, *58*(5), 966–975. doi: Doi10.1080/02626667.2013.788248

Fleming, S. W. (2007). Artificial neural network forecasting of nonlinear Markov processes. *Canadian Journal of Physics*, *85*(3), 279–294. doi: 10.1139/P07-037

Gao, M., Chen, X., Liu, J., Zhang, Z., & Cheng, Q. B. (2017). Using two parallel linear reservoirs to express multiple relations of power-law recession curves. *Journal of Hydrologic Engineering*, *22*(7). doi: 10.1061/(ASCE)HE.1943-5584.0001518

Goldstein, H. (1980). *Classical mechanics*. Addison-Wesley.

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2017). Lstm: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232. doi: 10.1109/tnnls.2016.2582924

Hall, F. R. (1968). Base-flow recessions—a review. *Water Resources Research*, *4*(5), 973-983. doi: https://doi.org/10.1029/WR004i005p00973

Harman, C. J., Sivapalan, M., & Kumar, P. (2009). Power law catchment-scale recessions arising from heterogeneous linear small-scale dynamics. *Water Resources Research*, *45*(9), 1–13. doi: 10.1029/2008WR007392

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput.*, *9*(8), 1735–1780. doi: 10.1162/neco.1997.9.8.1735

Jachens, E. R., Rupp, D. E., Roques, C., & Selker, J. S. (2020). Recession analysis revisited: Impacts of climate on parameter estimation. *Hydrology and Earth System Sciences*, *24*(3), 1159–1170. doi: 10.5194/hess-24-1159-2020

Jain, S. K., Singh, V. P., & van Genuchten, M. T. (2004). Analysis of Soil Water

Retention Data Using Artificial Neural Networks. *Journal of Hydrologic Engineering*, *9*(5), 415–420. doi: 10.1061/(asce)1084-0699(2004)9:5(415)

Kim, M., & Troch, P. A. (2020). Transit time distributions estimation exploiting flow-weighted time: Theory and proof-of-concept. *Water Resources Research*, *56*, e2020WR027186. doi: 10.1029/2020WR027186

Kingma, D. P., & Ba, J. (2017). *Adam: A method for stochastic optimization.* arXiv. doi: arXiv:1412.6980

Kirchner, J. W. (2006). Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. *Water Resources Research*, *42*(3), 1–5. doi: 10.1029/2005WR004362

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, *45*(2), 1–34. doi: 10.1029/2008WR006912

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. doi: 10.5194/hess-22-6005-2018

Lamb, R., & Beven, K. (1997). Using interactive recession curve analysis to specify a general catchment storage model. *Hydrological and Earth System Sciences*, *1*, 101–113.

Langbein, W. B. (1938). Some channel-storage studies and their application to the determination of infiltration. *Eos, Transactions American Geophysical Union*, *38*, 435–445.

Linsley, R. K., Kohler, M. A., & Paulhus, J. L. H. (1958). *Hydrology for engineers.* New York: McGraw-Hill.

Moore, R., & Wondzell, S. (2005). Physical Hydrology and the Effects of Forest Harvesting in the Pacific Northwest: a Review. *Journal of the American Water Resources Association*, *41*(4), 763–784. doi: 10.1111/j.1752-1688.2005.tb03770.x

Nathan, R. J., & McMahon, T. A. (1990). Evaluation of automated techniques for base flow and recession analyses. *Water Resources Research*, *26*(7), 1465-1473. doi: 10.1029/WR026i007p01465

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame,

J. M., ... Gupta, H. V. (2021). What role does hydrological science play in the age of machine learning? *Water Resources Research*, *57*(3), e2020WR028091. doi: 10.1029/2020WR028091

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., ... Duan, Q. (2015). Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: Data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrology and Earth System Sciences*, *19*(1), 209–223. doi: 10.5194/hess-19-209-2015

Porporato, A., & Ridolfi, L. (1997). Nonlinear analysis of river flow time sequences. *Water Resources Research*, *33*(6), 1353-1367. doi: 10.1029/96WR03535

Porporato, A., & Ridolfi, L. (2003). Detecting determinism and nonlinearity in river- flow time series flow time series. *Hydrological Sciences Journal*, *48*(5), 763–780. doi: 10.1623/hysj.48.5.763.51457

Roques, C., Rupp, D. E., & Selker, J. S. (2017). Improved streamflow recession parameter estimation with attention to calculation of -dQ/dt. *Advances in Water Resources*, *108*. doi: 10.1016/j.advwatres.2017.07.013

Ruelle, D., & Takens, F. (1971). On the nature of turbulence. *Communications in Mathematical Physics*, *20*(3), 167–192. doi: 10.1007/BF01646553

Rupp, D. E., & Selker, J. S. (2006). Information, artifacts, and noise in dQ/dt - Q recession analysis. *Advances in Water Resources*, *29*(2), 154–160. doi: 10 .1016/j.advwatres.2005.03.019

Rupp, D. E., & Woods, R. A. (2008). Increased flexibility in base flow modelling using a power law transmissivity profile. *Hydrological Processes*, *22*(14), 2667–2671. doi: 10.1002/hyp.6863

Sanford, W. E., & Selnick, D. L. (2013). Estimation of evapotranspiration across the conterminous united states using a regression with climate and land-cover data1. *JAWRA Journal of the American Water Resources Association*, *49*(1), 217-230. doi: 10.1111/jawr.12010

Scott, D. W. (1992). *Multivariate Density Estimation: Theory, Practice, and Visualization.* New York: John Wiley & Sons.

Shaw, S. B., & Riha, S. J. (2012). Examining individual recession events instead of a data cloud: Using a modified interpretation of dQ/dt-Q streamflow recession in glaciated watersheds to better inform models of low flow. *Journal of*

Hydrology, *434-435*, 46–54. doi: 10.1016/j.jhydrol.2012.02.034

Shen, C., Laloy, E., Elshorbagy, A., Albert, A., Bales, J., Chang, F.-J., ... Tsai, W.-P. (2018). Hess opinions: Incubating deep-learning-powered hydrologic science advances as a community. *Hydrology and Earth System Sciences*, *22*(11), 5639–5656. doi: 10.5194/hess-22-5639-2018

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis.* London: Chapman & Hall.

Staudinger, M., Stoelzle, M., Seeger, S., Seibert, J., Weiler, M., & Stahl, K. (2017). Catchment water storage variation with elevation. *Hydrological Processes*, *31*(11), 2000-2015. doi: 10.1002/hyp.11158

Szilagyi, J., Gribovszki, Z., & Kalicz, P. (2007). Estimation of catchment-scale evapotranspiration from baseflow recession data: Numerical model and practical application results. *Journal of Hydrology*, *336*(1), 206-217. doi: 10.1016/j.jhydrol.2007.01.004

Takens, F. (1981). Detecting strange attractors in turbulence. In D. Rand & L.-S. Young (Eds.), *Dynamical systems and turbulence, warwick 1980* (pp. 366–381). Berlin, Heidelberg: Springer Berlin Heidelberg.

Tashie, A., Pavelsky, T., & Band, L. E. (2020). An Empirical Reevaluation of Streamflow Recession Analysis at the Continental Scale. *Water Resources Research*, *56*(1), 1–18. doi: 10.1029/2019WR025448

Tawfik, M., Ibrahim, A., & Fahmy, H. (1997). Hysteresis Sensitive Neural Network for Modeling Rating Curves. *Journal of Computing in Civil Engineering*, *11*(3), 206–211. doi: 10.1061/(asce)0887-3801(1997)11:3(206)

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S., & Cook, R. B. (2016). *Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3* (Tech. Rep.). Oak Ridge, Tennessee, USA: ORNL DAAC. doi: 10.3334/ORNLDAAC/1328

Troch, P. A., Berne, A., Bogaart, P., Harman, C., Hilberts, A. G. J., Lyon, S. W., ... Verhoest, N. E. C. (2013). The importance of hydraulic groundwater theory in catchment hydrology: The legacy of Wilfried Brutsaert and Jean-Yves Parlange. *Water Resources Research*, *49*(9), 5099–5116. doi: 10.1002/wrcr.20407

Vogel, R. M., & Kroll, C. N. (1992). *Regional geohydrologic-geomorphic relation-*

ships for the estimation of low-flow statistics (Vol. 28) (No. 9). doi: 10.1029/92WR01007

Wang, D., & Cai, X. (2010). Recession slope curve analysis under human interferences. *Advances in Water Resources*, *33*(9), 1053–1061. doi: 10.1016/j.advwatres.2010.06.010

Yao, W., & Li, L. (2014). A new regression model: Modal linear regression. *Scandinavian Journal of Statistics*, *41*(3), 656-671. doi: 10.1111/sjos.12054

Young, P. C. (2011). *Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner* (2nd ed.). Springer Publishing Company, Incorporated.

Young, P. C., & Beven, K. J. (1994). Data-based mechanistic modelling and the rainfall-flow non-linearity. *Environmetrics*, *5*(3), 335-363. doi: 10.1002/env.3170050311