Hysteretic behavior of flow recession dynamics: Application of machine learning and learning from the machine

Minseok Kim [1], Hannes H. Bauser [1,2], Keith Beven [3], Peter A. Troch [4]

[1]Biosphere 2, University of Arizona, Tucson, AZ, USA

[2]Institute of Environmental Physics, Heidelberg University, Heidelberg, Germany

[3]Lancaster Environment Centre, Lancaster University, Lancaster, UK

[4]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

Key Points:

- A machine learning tool can capture the hysteresis in the flow recession dynamics using the past trajectory of discharge.
- Hysteresis mainly occurs during early time recession and catchment dynamics converge to an attractor during late time recession.
- Analyzing what the machine learned and what is needed to learn is useful to characterize catchment scale flow dynamics.

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 plot of $-dQ/dt$ versus Q typically form a wide point cloud due to noise and hysteresis, and it is still unclear what information we can extract from the data points and how to understand the information. In this study, 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. While analyzing the machine learning model structure and the trained parameters is a daunting task, 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 (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). Typically, a flow recession analysis plot is constructed by plotting the rate of change in discharge $-dQ/dt$ versus discharge Q, 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 analysis 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 recession curve 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 geomorphological characteristics also can explain some patterns. While the question of which model represents reality better will probably vary 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 analysis plot (in log-log scale) 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 (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). Also, 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 analysis 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 lower envelope and the measure of central tendency. In this study, we examine if those approaches can be combined. We utilize a machine learning tool to capture dynamics represented in the recession analysis plot using the past trajectory of flow. We hypothesize 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.

2 Theoretical background, methods, and study site

2.1 Flow recession analysis

Originally, flow recession curve 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 that has been utilized predominantly instead of fully considering dS/dt.) When a power function is used to characterize the recession plot (i.e., $-dQ/dt = aQ^b$), the power function still holds for g with the exponent decreased by 1: $g(Q) = aQ^{b-1}$ (Kirchner, 2009). We will call this g vs. Q plot a recession analysis plot as well as the $-dQ/dt$ vs.

Q plot. So the name of the plot is used interchangeably. 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. The function $g(Q)$ also can be utilized to estimate a (relative) storage-discharge relationship (Kirchner, 2009). The estimated storage using the catchment sensitivity function (1) (i.e., $S(Q) = \int_{Q_0}^{Q} (1/g(Q))dQ$) is the "active" storage (relative to a certain storage at Q_0) which is the portion of the storage that drives discharge (e.g., Troch et al., 2013). (Note that the active storage is sometimes referred to as "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 exponentially 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. 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 rainfall time series (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). Another criterion used in some studies is the strict decrease in $-dQ/dt$ in raw data (Dralle et al., 2017; Tashie et al., 2020) or in 3 day moving averages of $-dQ/dt$ (Dralle et al., 2017). 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.

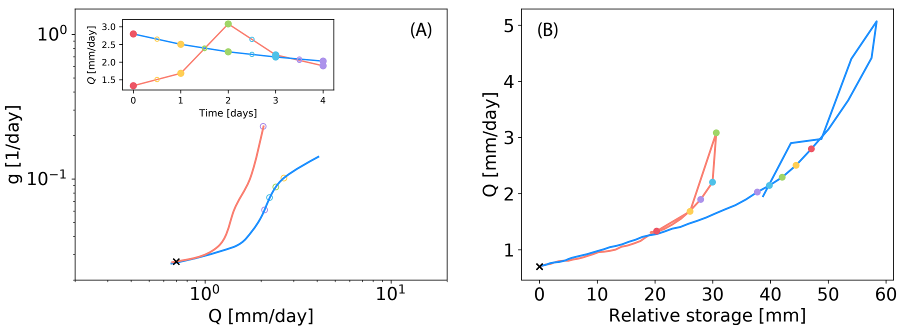

Figure 1. Illustration of the recession analysis plot and the corresponding storage-discharge relationship. (A) Two event trajectories in the recession analysis plot illustrated by different colors. The subset figure illustrates a part of the discharge time series of the two events. The empty circles in the recession analysis plot and the subset figure mark the timing of the g estimation for a few times. The empty purple circles are at a similar discharge for the two events but placed at different values of g. We hypothesize that the difference in g can be characterized by the past trajectory of discharge as shown in the subset figure. (Note that only the purple circle is illustrated for the red event because of the (not shown) rainfall event during 1.5 - 2.5 days.) (B) The corresponding storage-discharge relationship. The filled circles represent the timings corresponding to the filled circles in the subset figure in (A). 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.

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 hypothesize that g can be better characterized using the past trajectory of discharge, rather than using single discharge values. Figure 1A illustrates an example of g for two recession events and the associated discharge time series. As illustrated in the figure, the trajectory of g may vary from event to event, and the past trajectory of discharge may be used to distinguish those trajectories at similar values of Q. When the catchment sensitivity function g is hysteretic, the corresponding active storage-discharge relationship is also hysteretic, as exemplified in Figure 1B.

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 trajectory of discharge. 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: $Q(t+\Delta t) = \frac{2-g(t+\Delta t)\Delta t}{2+g(t+\Delta t)\Delta t} Q(t)$, assuming that dQ/dt is constant between the two time steps.

The functional form is similar to Beven's Holy Grail problem (Beven, 2006), 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 $-dQ/dt$).

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

We use a machine learning tool, the Long Short-Term Memory (LSTM) model (Hochreiter & Schmidhuber, 1997), to learn the function H using data. 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). 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:

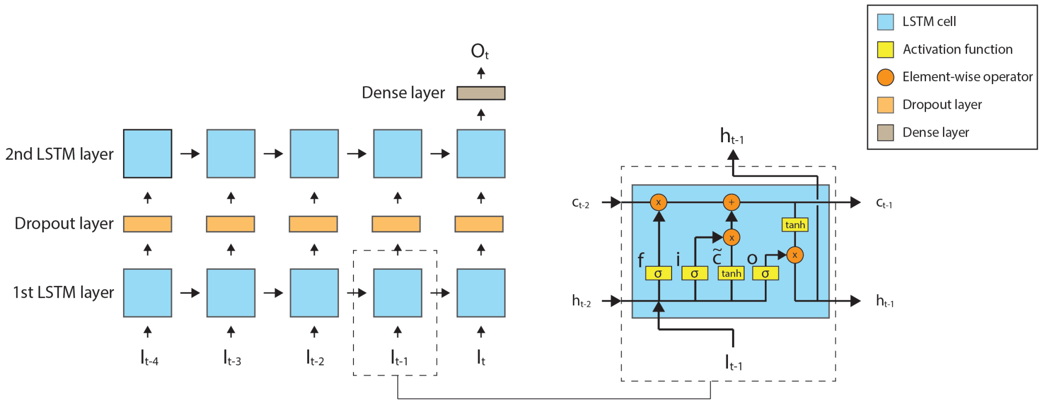

Figure 2. (Left) An example of a LSTM model structure with the dropout layer and the dense layer. The model has two layers of the recurrent LSTM layer with the dropout layer in between. Input time series I_t is fed into the first LSTM layer. The output of the second LSTM layer is fed into the dense layer, 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 avoid overfitting by setting a fraction of some variables to zero (e.g., Hochreiter & Schmidhuber, 1997). The dropout can be applied to the input sequence, to the recurrent states, or to the output of any recurrent LSTM layers. 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. If we use the model structure to estimate the function H, the illustrated model uses five 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 $q = 1$. The number of LSTM units n 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 (e.g., (3) for the LSTM layers) and the backward pass are performed to train the model using a loss function. The forward pass and the backward pass determine the gradient of the weights in those matrices and the vectors, and those weights are updated with a certain rate, the learning rate.

2.3 Study Site and Data

We use discharge data measured at the Calawah River near Fork, WA, USA (latitude 47°57'30", longitude 124°23'30", USGS gauge 12043000). The drainage area is 334 km^2, and the average topographic slope of this catchment is 0.07 (Addor et al., 2017). 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 data set also provides an estimated actual evapotranspiration rate using the Sacramento Soil Moisture Accounting (SAC-SMA) Model (Newman et al., 2015). For the period from March 1984 to December 2014, the average precipitation rate is 3,005 mm/year and the mean discharge rate is 2,819 mm/year. The actual evapotranspiration rate is 476 mm/year. The mass-balance does not close possibly due to an overestimation of the actual evaporation rate, but note that the recession plot analysis does not rely on the mass-balance and the quality of the actual evaporation time series. This catchment is wet with the aridity index of 0.25. Figure 3A shows the precipitation, the discharge, and the actual evapotranspiration rates.

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, so our analysis is a flow recession analysis rather than an analysis of the catchment sensitivity function. 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 and used all available data as we do not perform statistical analysis for each recession event separately. We applied both CTS and ETS methods to estimate the function g. The reasons for applying both methods are as follows; First, we expected the ML model to be able to find patterns in the noisy CTS method-based estimation; Second, the ETS method is a state-of-the-art method, but it relies on data smoothing.

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 was used as the loss function. The training period was from October 1980 to December 2000, and the validation period was from January 2001 to December 2014. The number of LSTM units in each cell n_u was 15 for both layers. The number of trainable pa-

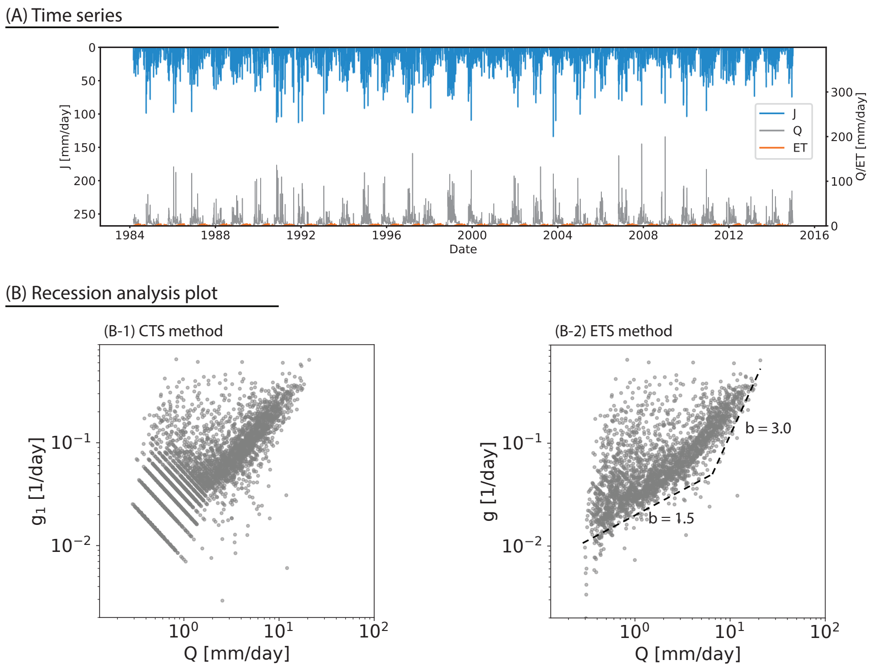

Figure 3. Time series and the flow recession analysis plots. (A) Time series of the precipitation J, the discharge Q, and the actual evapotranspiration ET. (B) The recession analysis plots that are estimated using (B-1) the CTS method and (B-2) the ETS method. Note that data points with $dQ/dt = 0$ are not shown in these log-log scale plots. The dotted lines in (B-2) are the lower envelope that was fitted to the point cloud by visual inspection.

rameters n_p is determined by the model structure and n_u as: $n_p = 12n_u^2 + 13n_u + 1 = 2896$. 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 100 iterations. The dropout rate was 0.4. 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.

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 3B illustrates the recession analysis plots. 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 be suitable for the data cloud, with $b = 3$ for high flow and $b = 1.5$ for low flow.

Figure 4A illustrates 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) for both the CTS method-based estimation and the ETS 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 for the CTS-based estimation and -0.05 for the ETS-based estimation, respectively. Figure 5A 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 are shown in Figure 4B. The model results are shown for different lengths of discharge trajectories (1 day, 3 days,

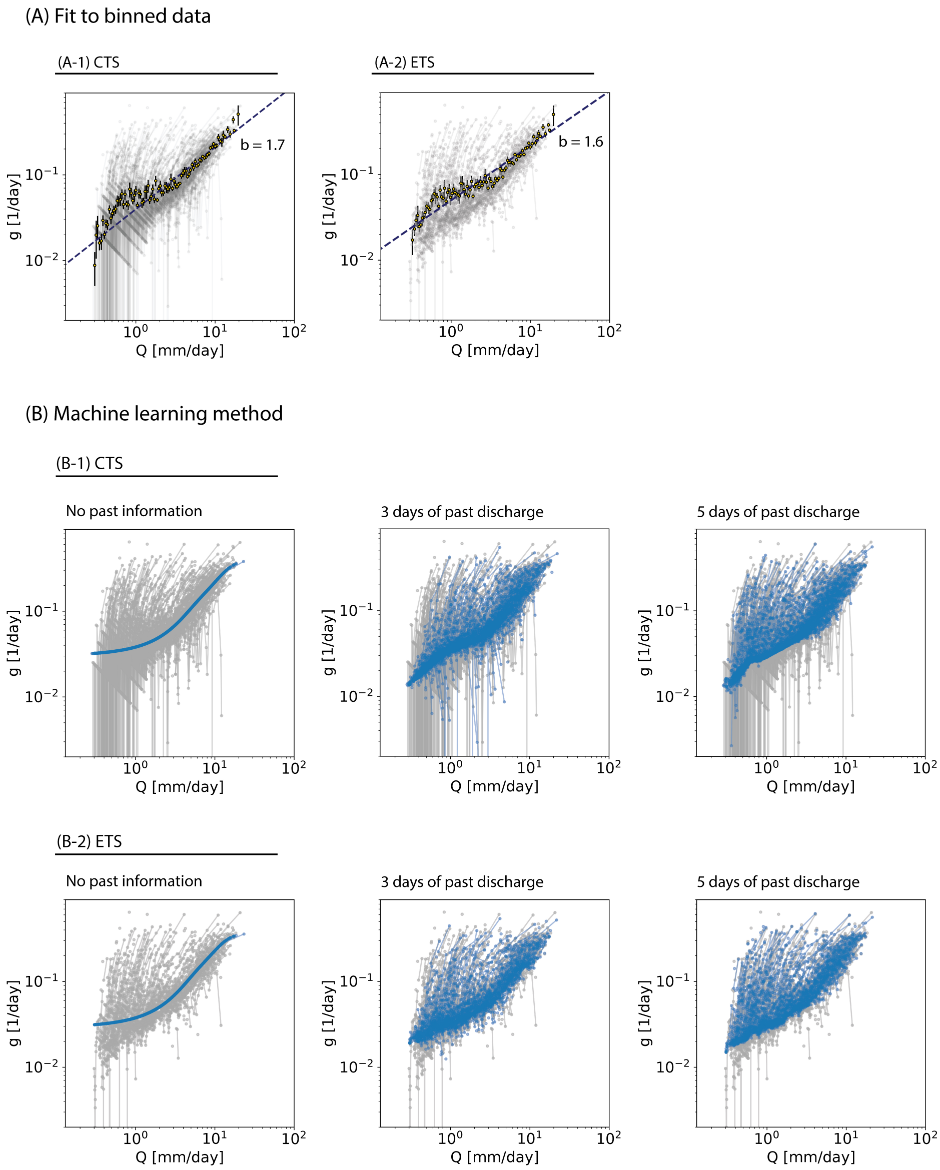

Figure 4. Estimated flow recession dynamics using (A) the central tendency and (B) the LSTM model. The CTS method-based estimation is used as observation in (A-1) and (B-1), and the ETS method-based estimation is used in (A-2) and (B-2). The yellow circles in (A) are the binned data with the error bar indicating the standard deviation of each bin. The dotted line is the power function fitted to the binned data. The grey dots are the observed data points, and the grey lines connect the points of each recession event. In (B), the blue dots are the ML model estimation and the blue lines connect the blue dots of each event. (Note that the LSTM results are shown only for the recession periods determined using the criteria that is described in the text.)

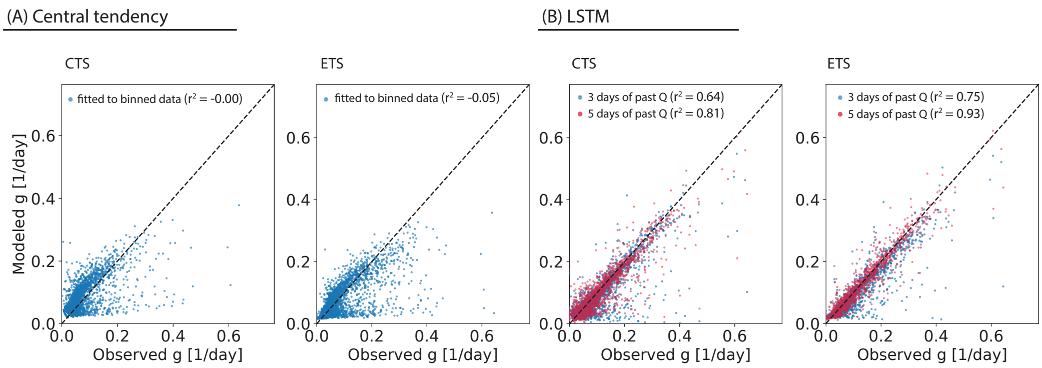

Figure 5. Comparison of the modeled g and the observed g. (A) The central tendency model, and (B) The LSTM model. The dotted black lines are 1:1 lines.

and 5 days) that were used in the function H. The LSTM model performance was similar for both training and validation periods (e.g., with the mean absolute error of 0.01 day^{-1} for both periods when 5 days of discharge was used), and the illustrated LSTM results are for both periods. 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. When the CTS method-based estimation is used as observation, the coefficient of determination r^2 is 0.64 and 0.81 for the model using 3 days and 5 days of discharge, respectively. (Note that there was no significant improvement when we increased the number of days to more than 5 days.) The LSTM model shows similar results when the ETS method-based estimation was used as observation. The coefficient of determination r^2 is 0.75 and 0.93 for the model using 3 days and 5 days of discharge, respectively. Figure 5B 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.

The LSTM model also performs decently when it is used as a forward model (updating the model input with the modeled Q as it becomes available). Figure 6 shows the simulated recession dynamics for 16 events. 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 the model was trained for the prediction of the half-step ahead g (which can be used to estimate the one-step ahead Q), the forward model performance degrades when the first few estimations are biased. Nevertheless, the model well tracks the event trajectories in the recession analysis plot which varies event-to-event. (Also, see Figure S1 that illustrates the event-to-event variation more clearly.)

4 Discussion: Learning from the machine

The results indicate that the machine has learned the nonlinear 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., 2020). 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 $r^2 \approx 0.8$). We can also investigate some machine-learned characteristics and deduce why the machine learned those features. In this study, we investigate the origin of hysteresis that appears in the plot and the origin of some areas of dense LSTM model estimation points. 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 result because the half-step ahead estimation is closer to data (see Figure 6B). Nevertheless, most of the analysis presented in this section are still valid with the forward modeling result.

We first investigate the origin of hysteresis that appears in the recession analysis plot. For example, when $Q \approx 1.0$ mm/day, the g values range from 0.03 day^{-1} to 0.4 day^{-1}. The LSTM model results indicate that the hysteresis can be explained by the past 5 days trajectory of discharge. Figures 7C and 7D show the 5 days of discharge for the points covered by each of the two areas that are indicated in the recession analysis plot. The discharge range for both areas is from 0.8 mm/day to 2.0 mm/day. The upper area is where g is greater than 0.13 day^{-1}, and the lower area is where g is less than 0.033 day^{-1}. The past trajectories of discharge are very different for the two areas. For the upper area with high g, the trajectories indicate that those recessions are from recent events. In the lower area, the trajectories of past discharge is consistently low and does not increase noticeably during the last 5 days.

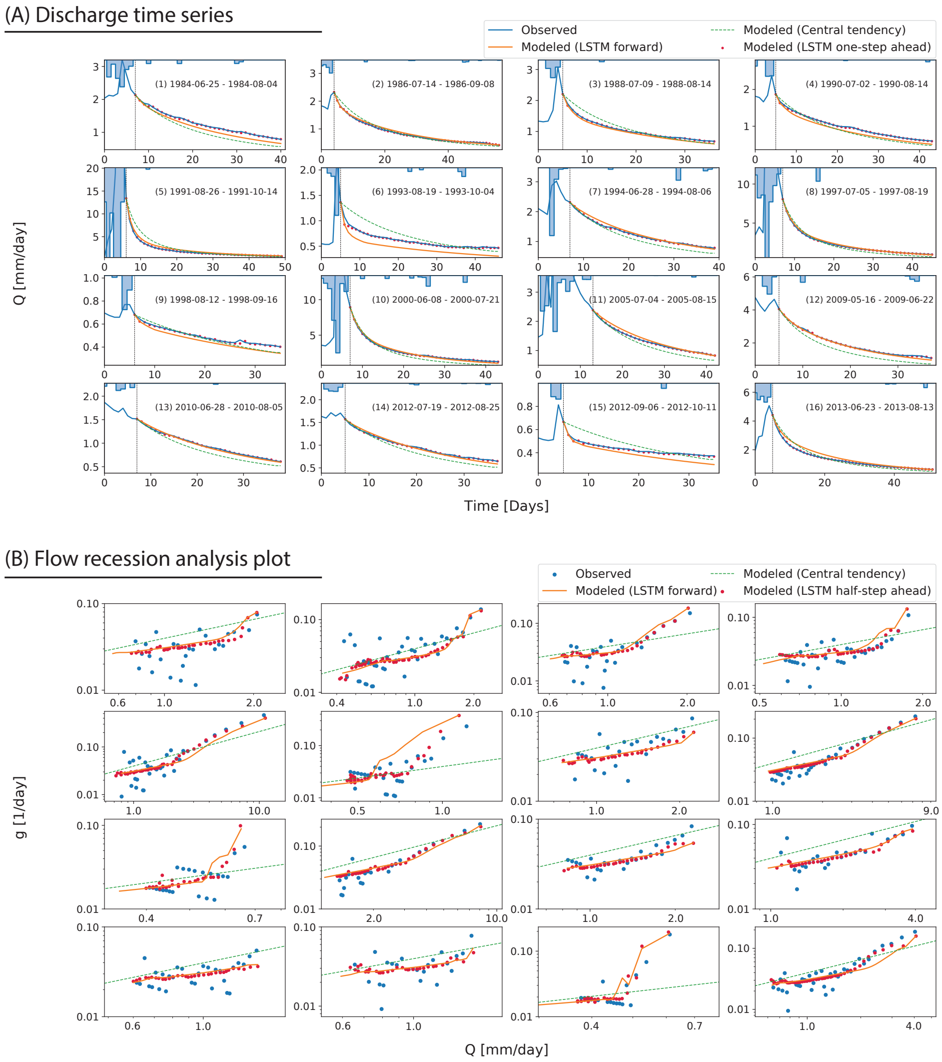

Figure 6. Forward modeling result of the LSTM model for the 16 events. (A) The simulated discharge time series, and (B) the simulated trajectory in the recession analysis plot. The forward model was run after the largest rain event (see the vertical dotted lines in (A)). The red dots represent the one-step ahead or the half-step ahead predictions, and the orange lines illustrate the forward model predictions.

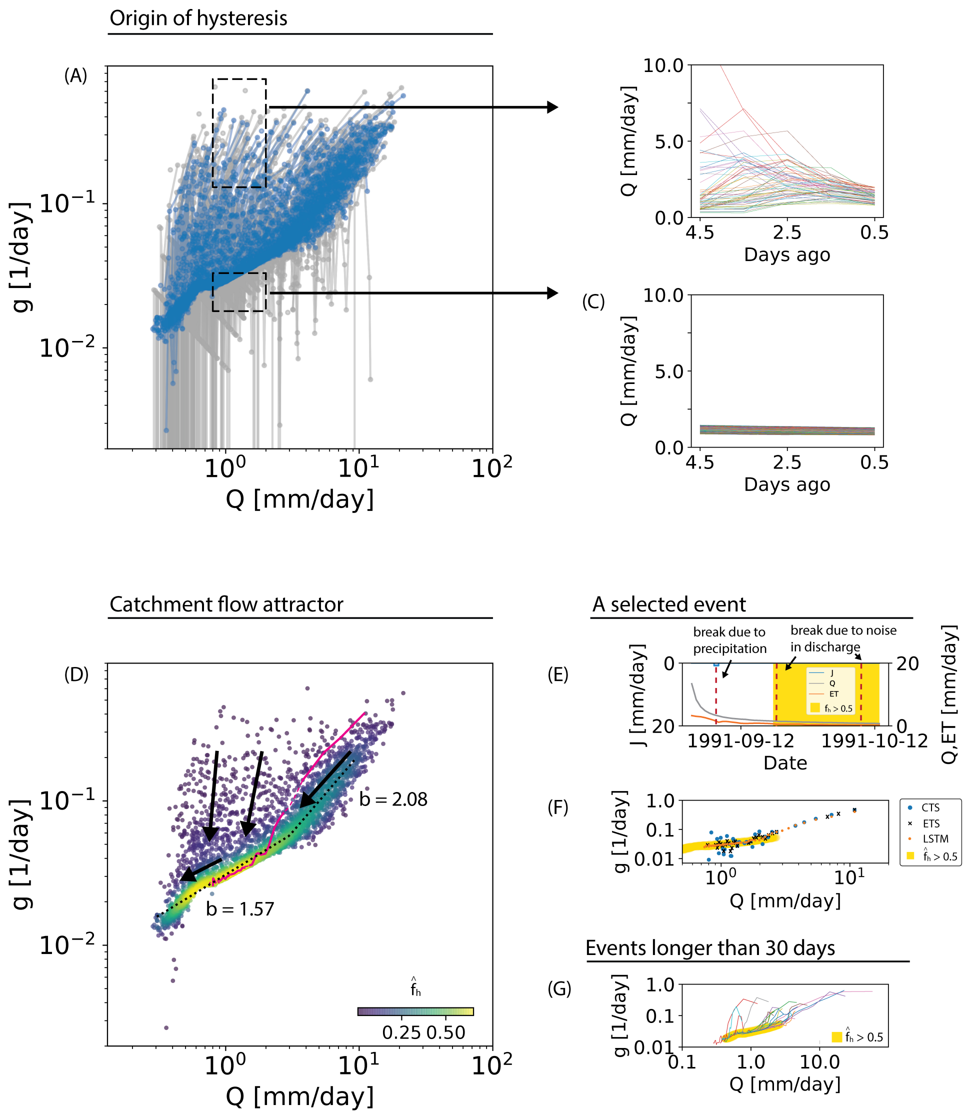

Figure 7. Learning from what machine learned. (A) LSTM model results using 5 days of discharge and the CTS method-based estimation. (B) 5 days of discharge for the events contained in the upper box in (A), that is indicated by the black dotted line, and (C) 5 days of discharge for the events contained in the lower box in (A). (D) 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 the dense area ($\hat{f}_h > 0.2$). (E) 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 ($\hat{f}_h > 0.5$) shown in (D). (F) Data points of the event that are estimated using several methods. (G) LSTM model-learned trajectories of all events longer than 30 days.

The contrasting trajectories indicate that g is high for "early" recession dynamics and is low for "late" recession dynamics. 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, and rapid contraction of variable source area. 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.

Another characteristic is that there is an area where the LSTM estimated points are densely located. Figure 7D 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. (Note that this dense area is also visible in the ETS method estimation; see Figure 3B-2.) 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 yellow area ($\hat{f}_h > 0.5$) is slow. The red line in Figure 7D 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 7E), while the line length of trajectory in the recession analysis 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 some noise (see Figure 7F).

Also, note 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 7E).

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 7G shows that all 16 recession events over 30 days, which were selective 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 S1). Figure 7G shows 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.) The early recession dynamics (that mostly appears above the yellow area) varies from event to event, depending on the spatial structure of the initial conditions (e.g. soil moisture content) for each event and the temporal and spatial patterns of external forcings (e.g. precipitation). Sometime later the dynamics of each event converge to the attractor, as the effects of those conditions and forcings vanish. This attractor will be called the "catchment flow attractor" because the attractor is a signature of catchment scale flow dynamics. The catchment flow attractor indeed is a better representation of the master recession dynamics (following the definition used in Lamb and Beven (1997)). The dynamics in the catchment flow attractor will 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 presence of the catchment flow attractor and its low slope (compared to the slope of early time dynamics) mean that the trajectory of each recession event in the recession analysis plot is, in general, concave (which means that the trajectory has a lower slope in the lower discharge range), unless the event trajectory is forced away prior to its convergence to the catchment flow attractor by external forcings. (We noticed that the slope of the catchment flow attractor is steep at the very low flow range, but the steep part still has a lower slope than most of the trajectories of the early time dynamics.) This concavity contradicts Tashie et al. (2020)'s recent study, which argues that the trajectory of each event is mostly convex (i.e., the opposite of concave) in more than 1,000 catchments in USA, with the exception of some dry and flat catchments. Nevertheless, some convex trajectories are observed in this catchment for early time recession (see the red line in Figure 7D and the blue lines in Figure 7A).

According to what we discussed so far, 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.

The upper dense area (the upper green area where $Q \gtrsim 3.0$ mm/day) indicates that many events shared similar early time recession dynamics, and the high density means that the area is a better representation of the ensemble of many early recessions than the lower envelope with the slope $b = 3$. The slope of the upper dense area is lower than the early time trajectories at low flow conditions, which is in line with the study of Jachens et al. (2020). Jachens et al. (2020) reported that recession events with lower initial discharges tend to have higher b values, while the characteristic early time dynamics of a catchment is more clearly shown at high discharge events (that constitute the upper dense area).

Overall, it seems that 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 fully considers the early time dynamics over the whole range of discharge. The full consideration results in the structure of the errors in the modeled g versus observed g plot (Figure 5), and the error in the forward simulation using the central tendency model (Figure 6). 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), but it certainly reduces the information content in data and neglect the hysteretic dynamics. The method of Brutsaert and Nieber (1977) seems to fit the data 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 $b = 3$ 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 $b = 1.57 \pm 0.00$ up to $Q = 2.99$ mm/day (see the black dotted line in Figure 7D). The value is similar to that of the late time dynamics of the Boussinesq model ($b = 1.5$). The slope of the upper line $b = 2.08 \pm 0.01$. This value is much smaller than the value of early time recession of the Boussinesq model ($b = 3$). The slope $b = 2.08$ 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). The slopes are similar to the ML model trained using the ETS method-based estimation where $b = 1.51 \pm 0.01$ for the catchment flow attractor and $b = 2.10 \pm 0.03$ for the upper dense area, indicating that the LSTM model is not very sensitive to the measurement noise and resolution. (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.)

The existence of the catchment flow attractor implies that, at some point in recession, multiple time scale dynamics reduce to simple slow dynamics. The simple 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 $b = 1.57$ in this case. When g is the power function of Q (i.e., $g = aQ^{b-1}$ and

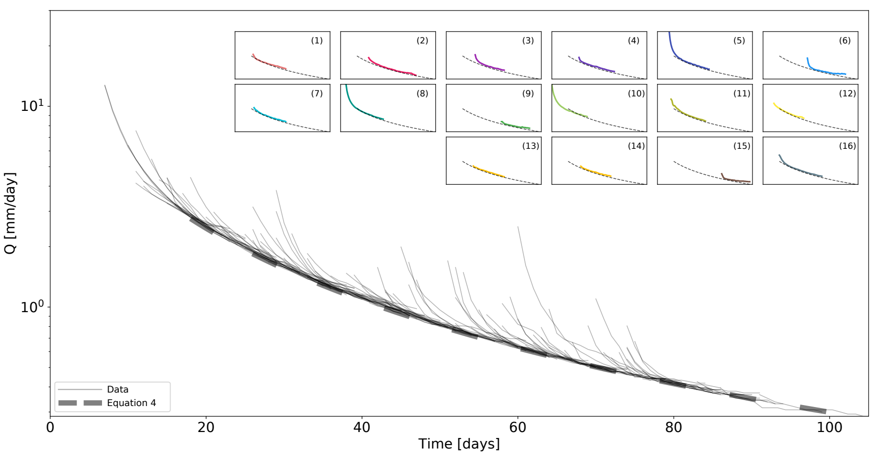

Figure 8. The attractor as the master recession curve. The thin lines illustrate the discharge time series of the all recession events longer than 8 days. The thin lines are shifted over time so that late time recessions approximately collapse to a single curved-line. The single curved-line is the master recession curve. The thick dashed line illustrates the parameterized attractor as a parameterized master recession curve. 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).

$-dQ/dt = aQ^b$), 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 8 illustrates that Equation (4) with the estimated parameters captures the late time flow recession dynamics.

Additionally, 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 inside the attractor. Using the non-hysteretic part of the relationship, the hysteretic storage-discharge relationship can be estimated if we calculate the storage using the mass-balance backward in time starting from the attractor.

Indeed, the relationship shown in Figure 1 is 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 Figure 6) estimated considering the rainfall and the discharge time series; i.e., $dS/dt = J - Q$. 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 1b 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 S2). 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 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 S2. 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. We leave a further discussion about the effect of ET on the catchment sensitivity function and the total storage-discharge relationship for future study.

5 Conclusions

The flow recession analysis has been served 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 analysis 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 analysis plot. Analyzing such hysteretic trajectories (event scale trajectories) 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 catchment flow attractor is the ensemble of many recessions during the late time flow recession dynamics. 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 hysteretic storage-discharge relationship.

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 easily customizable. 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. In this case, the model is an autoregressive (AR) model but with the past trajectory-dependent parameters. It is also generalizable to the autoregressive exogenous (ARX) model (or similar to the transfer function model) by including J as input (also ET when necessary). While analyzing the machine learning model structure and the trained parameters is a difficult task at the moment of writing, we showed that the machine learning model result provide a convenient way to extract information out of the noisy catchment scale signature, the recession analysis plot. Following the discussion in Beven (2020), 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.

Acknowledgments

The authors gratefully 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). H. H. Bauser was funded by the Deutsche Forschungsgemeinschaft (DFG) through Project BA 6635/1-1. 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.* Retrieved from `https://www.tensorflow.org/`

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

Beven, K. (2006). 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. (2020). Deep learning, hydrological processes and the uniqueness of place. *Hydrological Processes*, *34*(16), 3608-3613. doi: https://doi.org/10.1002/hyp.13805

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: https://doi.org/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

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

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

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

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

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

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.

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., ... Gupta, H. V. (2020). What role does hydrological science play in the age of machine learning? *Water Resources Research*, *n/a*(n/a), e2020WR028091. (e2020WR028091 2020WR028091) doi: https://doi.org/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

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

Rupp, D. E., & Selker, J. S. (2006). Information, artifacts, and noise in dQ/dtQ 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

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: https://doi.org/10.5194/hess-22-5639-2018

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

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

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

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