

## Hierarchical Temporal Scale Data-driven Reservoir Operation Modeling

Qianqiu Longyang<sup>1</sup>, Ruijie Zeng<sup>1\*</sup>

<sup>1</sup> School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ, 85281.

\* Corresponding author: Ruijie Zeng (ruijie.zeng.1@asu.edu)

### Contents of this file

Text S1

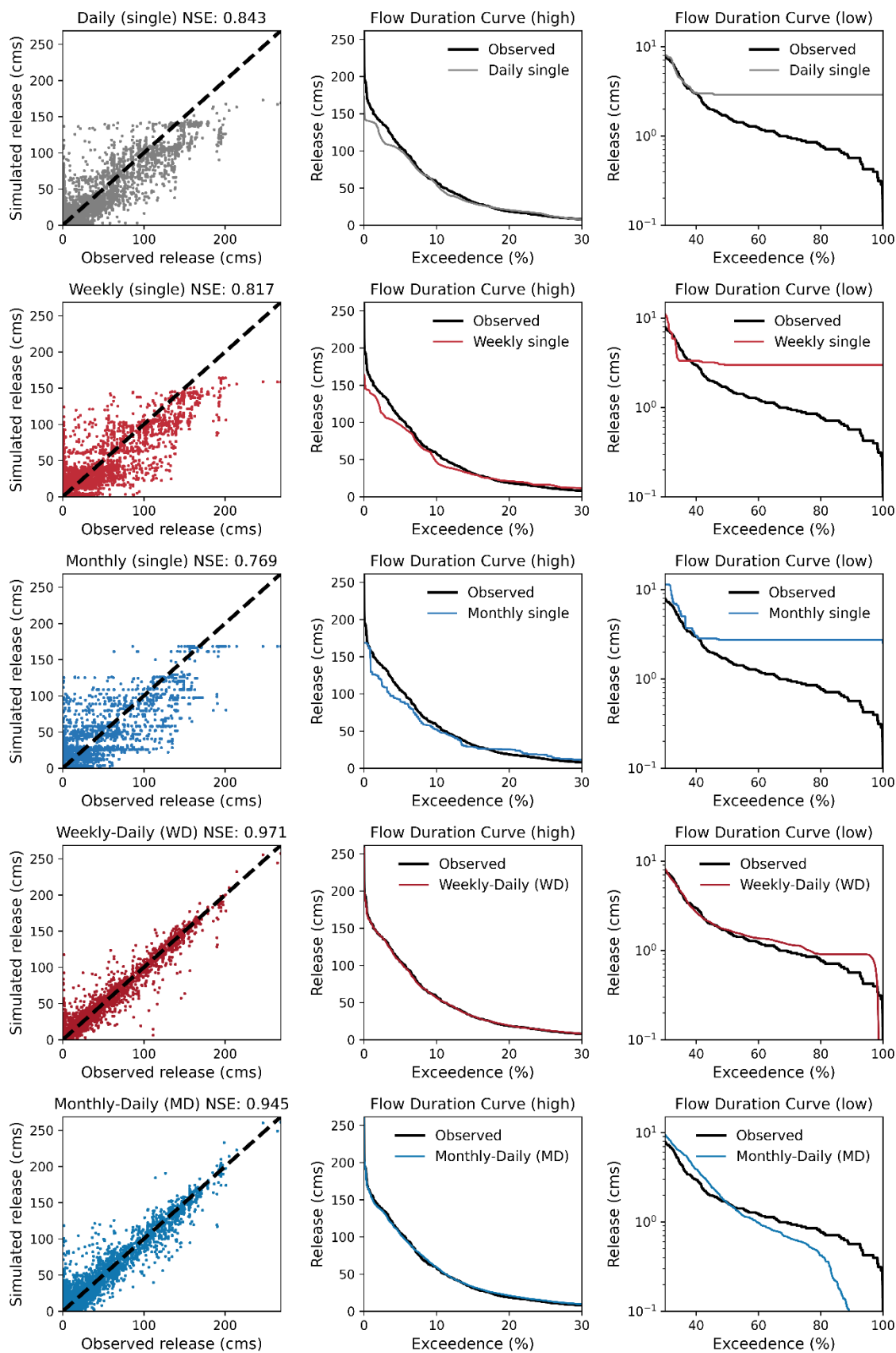
Figures S1 to S2

### Text S1.

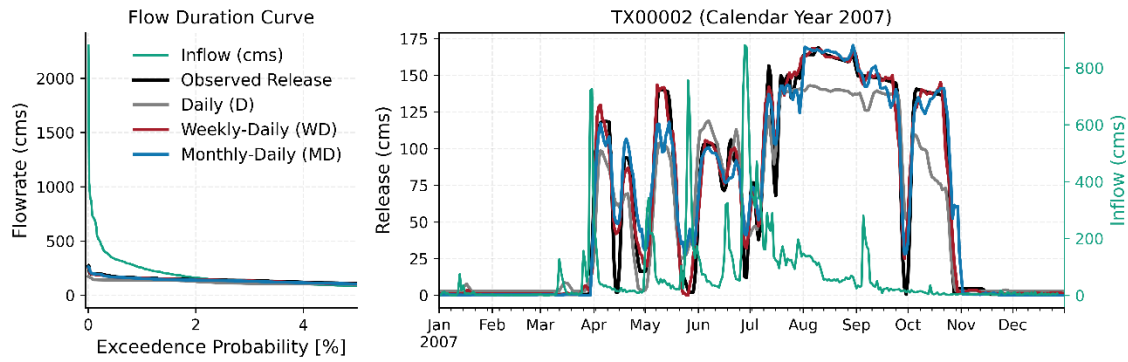
The Long-Short Term Memory (LSTM) computations are expressed as

$$\begin{aligned}i_t &= \sigma(W_{xi} \cdot x_t + W_{hi} \cdot h_{t-1} + b_i) \\f_t &= \sigma(W_{xf} \cdot x_t + W_{hf} \cdot h_{t-1} + b_f) \\g_t &= \tanh(W_{xg} \cdot x_t + W_{hg} \cdot h_{t-1} + b_g) \\o_t &= \sigma(W_{xo} \cdot x_t + W_{ho} \cdot h_{t-1} + b_o) \\c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\h_t &= o_t \odot \tanh(c_t)\end{aligned}$$

where  $W_{xi}$ ,  $W_{xf}$ ,  $W_{xg}$  and  $W_{xo}$  are learnable weights of inputs  $x_t$ ,  $W_{hi}$ ,  $W_{hf}$ ,  $W_{hg}$  and  $W_{ho}$  are learnable weights of the previous hidden states  $h_t$ , and  $b_i$ ,  $b_f$ ,  $b_o$  and  $b_g$  are biases of the four gates, respectively.  $\sigma$  means sigmoid function,  $\tanh$  is hyperbolic tangent function, and  $\odot$  represents element-wise multiplication.



**Figure S1.** Model performances of Belton Lake (TX00002). The subfigures in the first column show that Daily/Weekly/Monthly single models that are constructed on a single time scale work not so well, particularly for the high flow with the exceedance probability less than 30% (subfigures in the second column), while Weekly-Daily (WD) and Monthly-Daily (MD) hierarchical models improve the performances, both for the high flow and low flow.



**Figure S2.** Inflow, observed release, release simulated by Daily (D), Weekly-Daily (WD) and Monthly-Daily (MD) models of Belton Lake (TX00002) shown in Flow Duration Curve and hydrograph during the calendar year 2007. It illustrates that data-driven models capture the pattern how operators response to extreme flood events, and hierarchical time scale model (WD and MD) could simulate the release schemes better.