Author Response

Dear Jiawei,

Thank you very much for the careful and constructive comments regarding the implementation of the physics-constrained loss in the PCBT framework. We agree that the original manuscript did not describe the implementation of Eqs. (10) and (11), as well as the information flow during forecasting, with sufficient clarity. Your comments helped us identify several ambiguities in the mathematical formulation and its correspondence with the forecasting architecture.

We have carefully re-examined the physical constraint formulation and the forecasting procedure. In the revised manuscript, we will make the physical constraint formulation fully consistent with the actual temporal forecasting structure of PCBT, explicitly define the variables and units, and clarify the information available during forecasting. The main revisions are described below.

Comment 1: Equation (10) introduces $h(x,t)$, x, $\frac{\partial t}{\partial x}$, and $\frac{\partial Q}{\partial x}$, but the model inputs appear to consist mainly of antecedent precipitation and discharge. Please clarify how the spatial coordinates and water-depth variables are incorporated into PCBT and how $\frac{\partial Q}{\partial x}$ is calculated.

Response: Thank you for pointing out this important issue. We agree that the original formulation of Eq. (10) could give the impression that PCBT is implemented as a fully spatially resolved PDE solver, which is not the intended structure of the present study.

The final PCBT experiment is formulated as a temporal, rolling one-step-ahead, antecedent-only flood forecasting problem. The principal model inputs are the antecedent precipitation and discharge sequences, $P(t\text{-}7\text{:}t\text{-}1)$ and $Q(t\text{-}7\text{:}t\text{-}1)$, and the final forecasting task is to predict the discharge at a 24-h lead time. This input structure is explicitly given in Table 1 of the manuscript.

The revised physical constraint will therefore be formulated entirely in the temporal domain, using the observed antecedent precipitation and discharge sequences together with the predicted discharge at the target forecast time. It will not require spatial coordinates, water-depth observations, or explicit spatial derivatives. Accordingly, no spatial coordinate x or water-depth variable $h(x,t)$ is introduced into the PCBT input or computational graph.

Accordingly, we recognize that the original spatially explicit formulation involving x, $h(x,t)$, and the spatial derivative term was not described consistently with the final PCBT architecture. In the revised manuscript, we will replace this formulation with a discrete temporal hydrological consistency constraint and revise Section 2.3.1 accordingly, so that the mathematical formulation, model inputs, and forecasting architecture are fully consistent.

Comment 2: Equation (11) directly combines discharge and precipitation. Please clarify the units and preprocessing of precipitation before it is combined with discharge.

Response: Thank you for pointing out the dimensional issue in Eq. (11). We agree that the original manuscript did not clearly specify whether the rainfall-runoff constraint was evaluated in the original physical units or after data normalization.

In the actual implementation, the precipitation and discharge variables used in the neural-network training process are first normalized and then the rainfall-runoff constraint is evaluated using these normalized, dimensionless variables. Therefore, Eq. (11) is not intended to represent a direct dimensional water-balance equation in m^3/s, but rather a normalized rainfall-runoff consistency constraint used as a regularization term during model training.

In the revised manuscript, we will explicitly define the normalization procedure for both precipitation and discharge, denote the normalized variables as P_t^* and Q_t^*, and rewrite Eq. (11) accordingly: $Q_t^* = Q_t^* + P_t^* - \alpha Q_{t-1}^*$. Here, P_t^* and Q_t^* are dimensionless normalized precipitation and discharge, respectively, and α is a dimensionless runoff-loss coefficient. This clarification makes explicit that the constraint is evaluated in normalized space rather than by directly combining precipitation in mm with discharge in m^3/s. We will therefore refer to Eq. (11)

throughout the revised manuscript as a normalized rainfall-runoff consistency constraint rather than a dimensional water-balance equation.

We appreciate the reviewer for pointing out this ambiguity and will revise the corresponding formulation and explanation accordingly.

Comment 3: For D1-D7 forecasting, please clarify whether P_t represents observed future precipitation, forecast precipitation, or antecedent precipitation, and whether Q_{t-1} represents observed or predicted discharge after the first forecasting step.

Response: Thank you for raising this important point. We have further clarified the information flow and input updating mechanism used during forecasting.

In this study, we adopt a rolling one-step-ahead forecasting strategy, with 24 h as the basic forecasting horizon. At each forecasting time, the model uses only the historical precipitation and discharge observations that are available before the forecast origin to predict the discharge over the next 24 h.

At time *t*, the model input consists of the preceding seven-time steps of precipitation and discharge: $X_t = [P_{t-7;t-1}, Q_{t-7;t-1}]$, and the model produces the 24-h-ahead discharge forecast $\hat{Q}_t$.

When the next observation becomes available, the newly observed discharge Q_t^{obs} is incorporated into the input sequence, and a new 24-h-ahead forecast is generated. Therefore, the discharge input is updated using the latest observed discharge, rather than the previous model prediction. Specifically, $Q_{t-1} = Q_{t-1}^{obs}$, rather than $Q_{t-1} = \hat{Q}_{t-1}$.

Thus, the proposed forecasting procedure does not recursively feed predicted discharge values back into the model, thereby avoiding the accumulation of forecasting errors that can arise in conventional recursive multi-step forecasting.

For precipitation, no future precipitation observations or future precipitation forecasts are used during inference. Only precipitation observations that have become available before each forecast origin are included in the input sequence. As new observations become available, the precipitation and discharge input windows are updated in a rolling manner, and a new 24-h-ahead forecast is issued.

The D1-D7 analyses presented in the manuscript are therefore used to investigate the sensitivity of antecedent precipitation and discharge information at different forecast horizons and to characterize hydrological memory. They do not represent a recursive D1-D7 forecasting experiment from a single initial forecast origin. The operational forecasting experiment is formulated as a rolling 24-h-ahead flood forecasting procedure based on newly available observations.

We have clarified these points in the revised manuscript to explicitly describe the forecasting horizon, the rolling update mechanism, and the information available during each forecasting cycle.

We sincerely thank the reviewer for these thoughtful and constructive comments. The questions raised have been very helpful in clarifying the implementation of the physical constraints and the information flow in our forecasting framework. We believe that these clarifications and the corresponding revisions will substantially improve the methodological transparency, reproducibility, and overall quality of the manuscript. We greatly appreciate the reviewer's valuable suggestions and careful consideration of our work.

Best regards,

Hongyuan Fang