RC2

Prediction of basin-scale river channel migration based on landscape evolution numerical simulation

Jitian Wu, Xiankui Zeng, Qihui Wu, Dong Wang, and Jichun Wu

We sincerely thank the reviewer for their thoughtful and constructive feedback. We have carefully considered each comment, provided point-by-point responses below, and revised the manuscript accordingly. We believe these revisions have strengthened the study, and we hope the updated manuscript now meets the standards for publication in *Hydrology and Earth System Sciences* (HESS).

Note:

(1) In this response, the text in *italic type* is the original comments from the reviewers, and the text in blue, headed with "Reply", is the response from the authors.

(2) In the manuscript, the words in blue indicate the sentence is improved or revised. Some of them are mentioned in this response via the page and line number.

Comments:

1. *The manuscript should explicitly clarify the statistical/error-model assumptions behind using the average Hausdorff distance H as the calibration target and then use the same metric to provide a reach-wise (regional) difficulty diagnosis for the channel prediction/surrogate. Currently, the likelihood formulation based on H (with $Hobs = 0$) is presented, but the assumed distribution for H and the selection/estimation of the scale (variance) term are not sufficiently specified, which directly affects posterior tightness and the credibility of uncertainty bounds. Also, the paper should quantify the known spatial heterogeneity in performance by splitting the main channel into two or three reaches (e.g., upstream canyon vs downstream plain) and reporting H (and optionally mean pointwise distance) per reach for both reconstruction/validation and surrogate evaluation, since the text indicates downstream deviations are larger. This will make the Bayesian calibration statistically transparent while also giving readers a practical, spatially explicit statement of where the workflow is reliable.*

Reply: Thank you to the reviewer for pointing out these two issues.

① In this study, the average Hausdorff distance (H) is used to quantify the spatial discrepancy between the simulated and observed channel centerlines (i.e., two curves). The observed value H_{obs} represents the distance between the true channel and itself (thus $H_{obs} = 0$), whereas the simulated value H represents the distance between the modeled channel output and the true channel. In the Bayesian uncertainty analysis, we assume that the error in H follows a zero-mean, independent and identically distributed Gaussian distribution, and we construct the likelihood function accordingly. We have clarified this assumption in the revised manuscript: the mean Hausdorff distance H is treated as an average measure of the spatial deviation between the simulated and observed channel planforms, and its observation error is assumed to be normally distributed, i.e., $H \sim \mathrm{N}(0, \sigma^2)$.

② We agree with the suggestion to evaluate spatial heterogeneity in model performance by assessing metrics along different segments of the main channel. Dividing the main channel by geomorphic units—upstream canyon reach, midstream transitional reach, and downstream alluvial-plain reach—and computing the average Hausdorff distance (H) for each segment would explicitly account for spatial heterogeneity and could provide a more informative diagnosis of predictive skill. However, if segmented reaches are used within the MCMC-based parameter uncertainty analysis, the variance/covariance specification in the likelihood (Eq. 10) must be defined appropriately for each reach. Because geomorphic setting, channel stability, and migration amplitude differ substantially among reaches, the variance (or covariance-matrix elements) should be parameterized in a reach-dependent manner. An inappropriate specification would bias the likelihood evaluation and compromise the accuracy of the inferred posterior distributions and uncertainty bounds. For these reasons, we performed parameter calibration and uncertainty quantification using a basin-wide (whole-channel) computation of H. The results indicate that this whole-channel approach can effectively characterize the channel-migration process over the study period, and the resulting error level is adequate for the objectives of this study.

In future work, we will consider explicitly partitioning the main channel into three reaches (upstream canyon, midstream transition, and downstream alluvial plain) and further investigating how to specify the variance term (or covariance structure) in Eq. (10) under a segmented formulation, with the goal of enabling a finer and more robust characterization of reach-scale uncertainty and spatial heterogeneity.

2. *The LSTM surrogate section should be expanded with minimal but essential implementation details to ensure reproducibility, beyond the already provided training design (LHS sample size, train/validation split, optimizer and hyperparameters). Specifically, please add a compact description (ideally a short table plus a few sentences) of the LSTM architecture (number of layers, hidden units, dropout/regularization if any), the preprocessing applied to the 11 parameters (e.g., min–max scaling or z-score normalization), the exact output formatting of the 2,000-point planform, and the loss definition used to train the network (e.g., coordinate-wise MSE, any weighting along the channel). These additions are documentation-level and do not require new experiments, but they materially improve the scientific value of the surrogate contribution by allowing other groups to replicate and benchmark the approach.*

Reply: Thank you to the reviewer for providing this important suggestion. We have added additional implementation details for constructing the LSTM-based surrogate model in the revised manuscript. The surrogate model employs a two-layer LSTM architecture followed by a linear fully connected layer, taking normalized LE-PIHM parameters as inputs and outputting planar river-channel coordinates. The network is trained with the Adam optimizer by minimizing the RMSE between the predicted and reference channel planforms. Table 2 summarizes the specific LSTM configuration and training settings as follows:

Table 2. Architecture of the LSTM surrogate model

Training design	**Specification**
Input parameters	11 key LE-PIHM parameters
Input preprocessing	Min-max normalization applied to each parameter based on its prior range, scaled to [0, 1]
Network architecture	Two stacked LSTM layers followed by one fully connected a dense layer
LHS sample size	3000 parameter sets
Training set size	2100 samples (70% of LHS sample size)
Validation set size	900 samples (30% of LHS sample size)
LSTM layer 1	128 hidden units; return full output sequence
LSTM layer 2	256 hidden units; return last time-step output only
Model output	Planform locations of river channel represented by 2000 uniformly sampled points
Output format	(x, y) coordinates of channel points
Loss function	RMSE between surrogate predicted and reference channel coordinates
Optimizer	Adam
Learning rate	0.001
Batch size	100
Training epochs	10,000

3. *The future-scenario projection component should include a clear, concise description of how CMIP6 EC-Earth3-Veg forcings under the SSP scenarios are prepared and mapped into LE-PIHM, because scenario-to-scenario differences in projected migration—particularly the large response reported for SSP2-4.5—are sensitive to bias correction, downscaling, and temporal aggregation choices. Please state explicitly whether precipitation/temperature are used raw or bias-corrected (and name the method at a high level if applied), how spatial downscaling/interpolation to the basin model grid is performed, and what temporal resolution is used to drive the model during 2021–2100 (including whether any aggregation is performed before the landscape-evolution time step). A brief paragraph should then connect the interpretation of "threshold-like" migration behavior to these forcing-preprocessing uncertainties; this strengthens the credibility of the scenario comparison.*

Reply: Thank you for pointing out that the description of CMIP6 forcing preparation and processing for the future-scenario simulations was insufficient in the original manuscript. We agree that scenario-to-scenario differences in projected channel migration may be sensitive to forcing-preprocessing choices, including bias correction, spatial downscaling/interpolation, and temporal resolution.

Accordingly, we have added a clarifying paragraph in Section 4.2 of the revised manuscript: “In this study, future-scenario projections are driven by CMIP6 EC-Earth3-Veg precipitation and air-temperature data under SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5. The scenario forcing fields are mapped to the LE-PIHM basin computational units using bilinear interpolation. To maintain consistency with the monthly time step adopted in LE-PIHM, the forcing data are temporally aggregated to monthly resolution prior to being used as model inputs. In addition, no bias correction is applied to the CMIP6 forcing data in this study.”