



Supplement of

Improving heat transfer predictions in heterogeneous riparian zones using transfer learning techniques

Aohan Jin et al.

Correspondence to: Quanrong Wang (wangqr@cug.edu.cn)

The copyright of individual parts of the supplement might differ from the article licence.

S1. Analytical solution of the 2D heat transfer process in homogeneous streambed

In this section, the analytical model developed by Shi et al. (2023) is employed as the benchmark to generate the training dataset for the pre-training model. This model offers a balanced trade-off between analytical simplicity and accuracy, particularly in representing boundary conditions relevant to heat transport between surface water and groundwater. Moreover, this model has been validated against field data, further enhancing its credibility. The mathematical model of Shi et al. (2023) accounts for thermal convection and thermal conduction in a homogeneous streambed, which can be described as follows:

$$\frac{\partial T}{\partial t} = D_x \frac{\partial^2 T}{\partial x^2} + D_z \frac{\partial^2 T}{\partial z^2} - v_x \frac{\partial T}{\partial x} - v_z \frac{\partial T}{\partial z}, x \geq 0, z \geq 0, t \geq 0 \quad (S1)$$

$$T(x, z, t = 0) = T(x, z) \quad (S2)$$

$$T(x, z = 0, t) = \begin{cases} f(t), 0 \leq x \leq L \\ 0, otherwise \end{cases} \quad (S3)$$

$$\left. \frac{\partial T}{\partial x} \right|_{x=0} = \left. \frac{\partial T}{\partial x} \right|_{x \rightarrow \infty} = \left. \frac{\partial T}{\partial z} \right|_{z \rightarrow \infty} = 0, t \geq 0 \quad (S4)$$

where T denotes temperature [$^{\circ}\text{C}$]; t is time [T]; v_x and v_z refer to thermal front velocities [LT^{-1}] along the x - and z -axes, respectively. Specifically, $v_x = \frac{C_w}{C_s} q_x$, and $v_z = \frac{C_w}{C_s} q_z$, in which q_x and q_z denotes the components of water Darcy flux [LT^{-1}] components along the x - and z -axes, respectively; C_w and C_s are specific volumetric heat capacities [$\text{J}/(\text{m}^3 \cdot ^{\circ}\text{C})$] of water and streambed, respectively; D_x and D_z represent the effective thermal diffusivity along the x - and z - axes, respectively [L^2T^{-1}]; L denotes the half width (L) of the river.

The above mathematical model can be analytically solved using the Green's function method. The solution of the temperature field can be written as:

$$T = A(x, z, t) \left[\int_0^\infty \int_0^L E(\lambda, \eta) B(x - \lambda, z - \eta, t) d\lambda d\eta + \int_0^t g(\tau) C(x, z, t - \tau) d\tau \right] \quad (S5)$$

$$A(x, z, t) = \frac{1}{4} \exp \left[-\frac{v_z(v_z t - 2z)}{4D_z} - \frac{v_x(v_x t - 2x)}{4D_x} \right] \quad (\text{S6})$$

$$B(x - \lambda, z - \eta, t) = \frac{1}{\pi t \sqrt{D_x D_z}} \left\{ \exp \left[-\frac{(z - \eta)^2}{4D_z t} \right] - \exp \left[-\frac{(z - \eta)^2}{4D_z t} \right] \right\} \\ \left\{ \exp \left[-\frac{(x - \lambda)^2}{4D_x t} \right] + \exp \left[-\frac{(x - \lambda)^2}{4D_x t} \right] \right\} \quad (\text{S7})$$

$$C(x, z, t - \tau) = \frac{z}{\sqrt{\pi D_z (t - \tau)^3}} \exp \left[-\frac{z^2}{4D_z (t - \tau)} \right] \left\{ \operatorname{erf} \left[\frac{L - x}{\sqrt{4D_x (t - \tau)}} \right] + \operatorname{erf} \left[\frac{L + x}{\sqrt{4D_x (t - \tau)}} \right] \right\} \quad (\text{S8})$$

$$E(\lambda, \eta) = h(\lambda, \eta) \exp \left[-\frac{2\eta v_z}{4D_z} - \frac{2\lambda v_x}{4D_x} \right] \quad (\text{S9})$$

$$g(\tau) = f(\tau) \exp \left[\frac{v_z(v_z \tau)}{4D_z} + \frac{v_x(v_x \tau - 2x)}{4D_x} \right] \quad (\text{S10})$$

where λ , η , and τ represent dummy variables, which correspond to the independent variables x , z , and t , respectively. It should be noted that the groundwater flow model and heat transfer model are coupled through q_x and q_z , which are directly prescribed in this study. Therefore, we only need to consider boundary conditions for heat transfer model for analytical model. Detailed procedures of solution derivation can be found in Shi et al. (2023).

S2. Numerical solution of the 2D heat transfer process in heterogeneous streambed

A numerical model using COMSOL Multiphysics (COMSOL Inc., Burlington, MA, USA) is constructed to simulate heat transfer processes within homogeneous and heterogeneous streambeds. To improve simulation accuracy and avoid boundary effects, the semi-infinite geometry size was replaced by a finite range, and two infinite element domains were added at $x = 1 \text{ m}$ and $z = 1 \text{ m}$ to represent the infinite boundary on the x - and z -directions, respectively. The domain of interest is discretized using triangular elements, with a finer spatial discretization near the river to reduce numerical truncation errors. The mesh consists of 16,998 triangular elements in total. A constant time step of 0.01 days is employed for the numerical solution. The following values are adopted from the study of Shi et al. (2023): $C_w = 4.18 \times 10^6 \text{ J}/(\text{m}^3 \cdot ^\circ\text{C})$, $C_s = 1.82 \times 10^6 \text{ J}/(\text{m}^3 \cdot ^\circ\text{C})$, $q_x = 0.2 \text{ m/d}$, $q_z = 0.3 \text{ m/d}$, $D_x = D_z = 7.69 \times 10^{-5} \text{ m}^2/\text{s}$, $h(x, z) = 20^\circ\text{C}$. To validate the developed numerical model, Figure S1 presents a comparison of temperature-time curves between the numerical and analytical solutions at three locations. Results indicate that the numerical solution closely aligns with the analytical solution, exhibiting small mean square error (MSE) values ranging from 0.06-0.11 $^\circ\text{C}$ for selected locations, thereby confirming the accuracy of the numerical solution. In the following analysis, the nonuniform flow and temperature fields are generated utilizing COMSOL Multiphysics, incorporating various heterogeneous $\ln K$ fields generated through an autocorrelated multi-Gaussian simulator. The settings of initial and boundary conditions are shown in Figure S2.

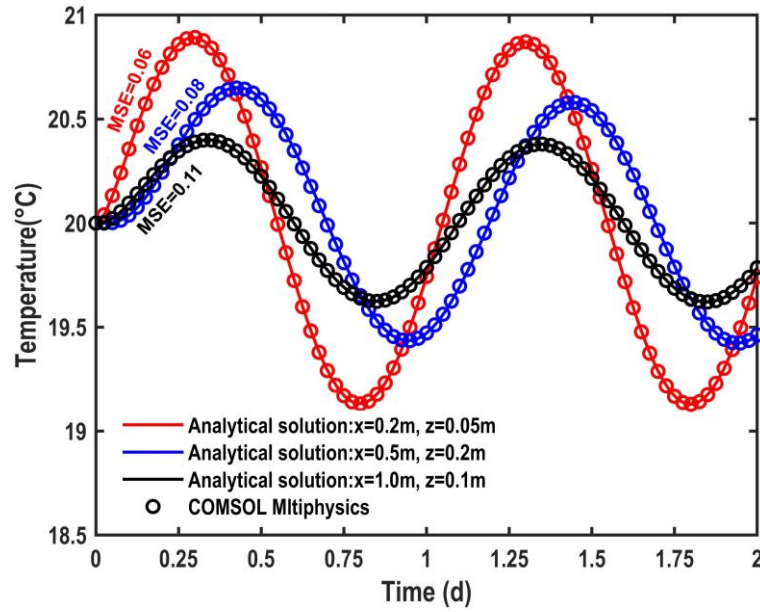


Figure S1. Comparison of temperature-time curves at three locations using the numerical solution (circle symbols) and the analytical solution of this study (solid curves).

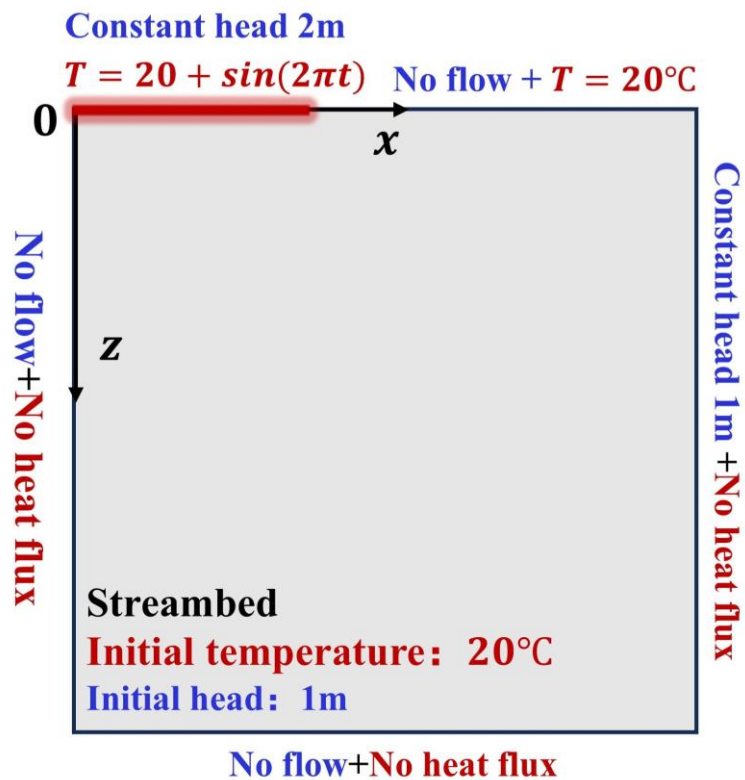


Figure S2. Conceptual model of 2D numerical model with streambed sediment, initial and boundary conditions. The red color zone represents the streambed with a half width of 0.32.

S3. Prediction results with the DNN and PDTL models using 200 observation points

In this section, 200 observation points are used to evaluate the performance of the DNN and PDTL models. Figure S3 shows the prediction results with the DNN and the PDTL models using 200 observation points for homogeneous scenario. Figure S4 - S5 present the prediction results using 200 observation points for heterogeneous scenarios with the DNN and the PDTL models, respectively. Results demonstrate that there is no significant difference in the performance of the PDTL and DNN models for both homogeneous and heterogeneous scenarios when the number of observation points increases to 200. This suggests that with enough data, DNNs can achieve similar results by learning data patterns that reduce the impact of physical information.

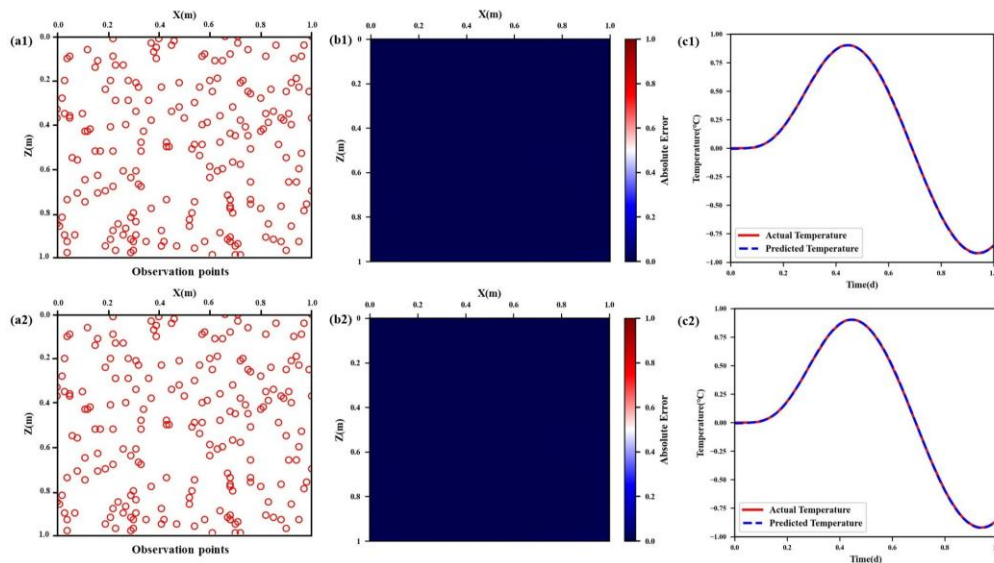


Figure S3. Prediction results with the DNN and the PDTL models using 200 observation points for homogeneous scenario. (a1) - (a2): the distribution of observation points with the DNN and the PDTL models; (b1) - (b2): the distribution of MAE predicted with the DNN and the PDTL models; (c1) - (c2): the temperature time series predicted with the PDTL and DNN models at $x = 0.5m$, $y = 0.5m$.

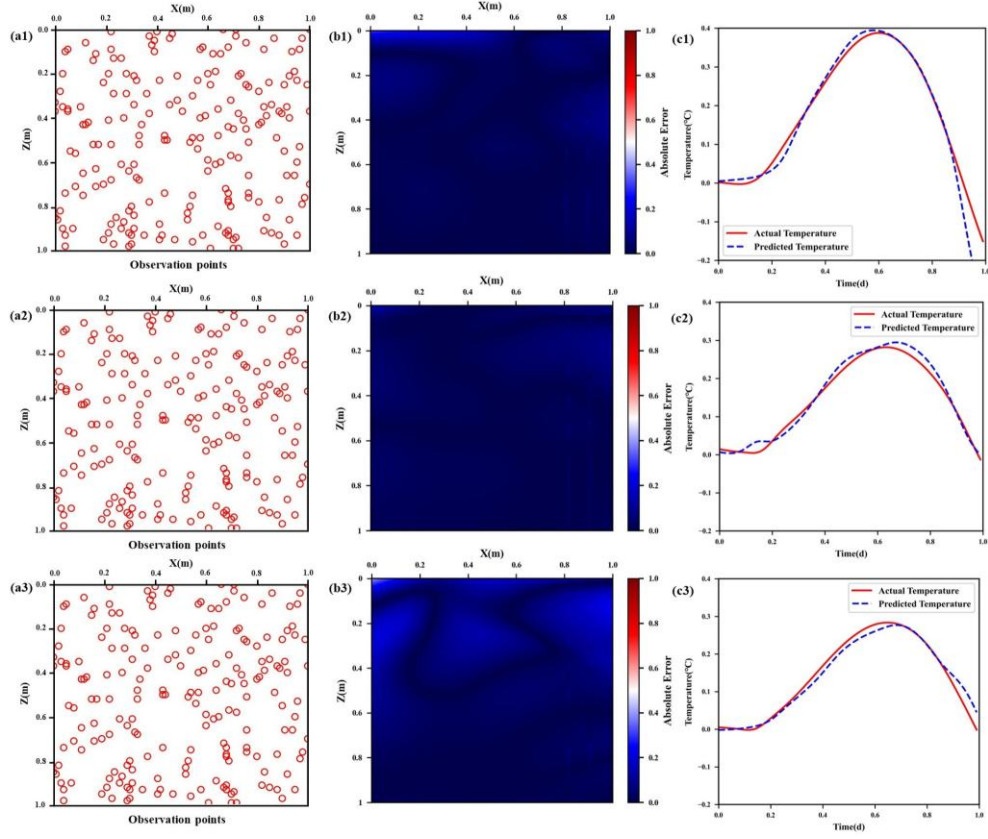


Figure S4. Prediction results with the DNN model using 200 observation points for heterogeneous scenarios. (a1) - (a3): the distribution of observation points with $\sigma_{lnk}^2 = 0.2$, 0.5, and 1.0; (b1) - (b3): the distribution of MAE predicted with $\sigma_{lnk}^2 = 0.2$, 0.5, and 1.0; (c1) - (c3): the temperature time series predicted with the DNN model at $x = 0.5m$, $y = 0.5m$ with $\sigma_{lnk}^2 = 0.2$, 0.5, and 1.0.

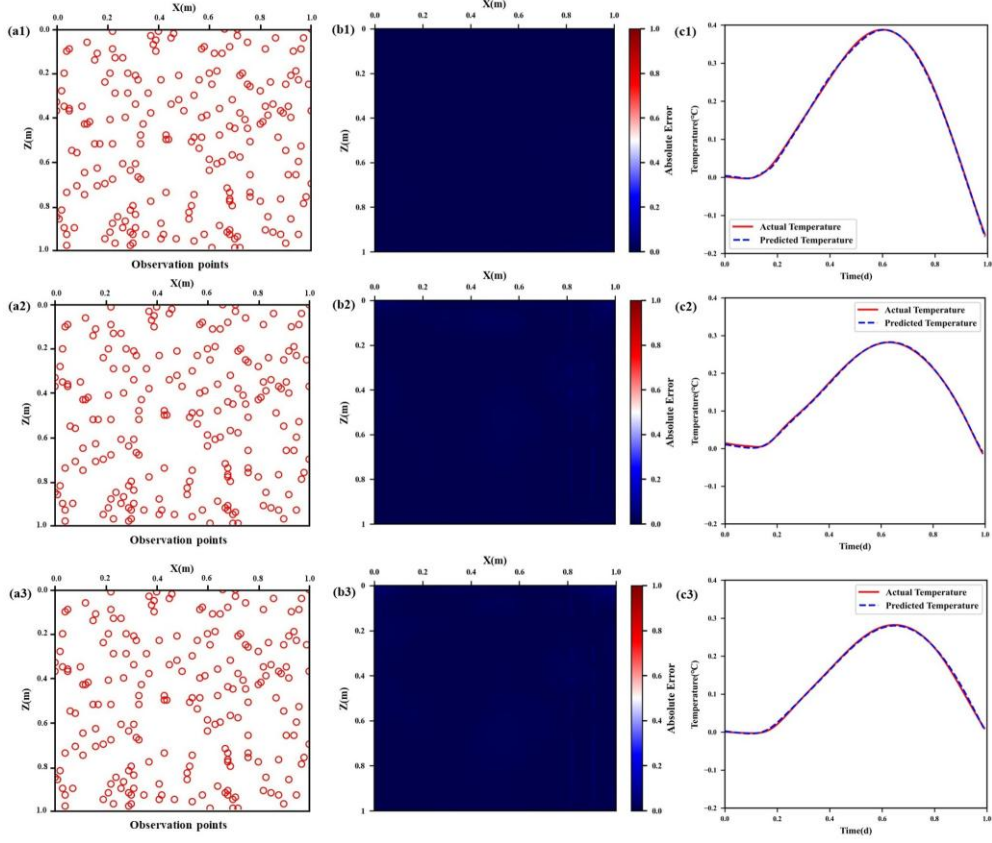


Figure S5. Prediction results with the PDTL model using 200 observation points for heterogeneous scenarios. (a1) - (a3): the distribution of observation points with $\sigma_{lnk}^2 = 0.2$, 0.5, and 1.0; (b1) - (b3): the distribution of MAE predicted with $\sigma_{lnk}^2 = 0.2$, 0.5, and 1.0; (c1) - (c3): the temperature time series predicted with the PDTL model at $x = 0.5m$, $y = 0.5m$ with $\sigma_{lnk}^2 = 0.2$, 0.5, and 1.0.

References

Shi, W., Zhan, H., Wang, Q., and Xie, X.: A two-dimensional closed-form analytical solution for heat transport with nonvertical flow in riparian zones, *Water Resources Research*, 59, e2022WR034059, <https://doi.org/10.1029/2022WR034059>, 2023.