Supplement of Biogeosciences, 23, 7939–7954, 2026
https://doi.org/10.5194/bg-23-7939-2026-supplement
© Author(s) 2026. CC BY 4.0 License.



Biogeosciences Open Access EGU

Supplement of

Riverine organic matter dynamics in the headwaters of the Yellow River under climate change

Xiaoxiao Wang et al.

Correspondence to: Xiaoxiao Wang (wxx1989@cwnu.edu.cn)

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

S1. Delta Change Method

Future climate data are often systematically biased, so the first step in using climate model data is often bias correction. This study uses delta correction for future climate following Cao et al. (2026). The delta method sums the difference between the past and present modelled climate and the present observed climate (Beyer et al., 2020). Thus, delta assumes that local (i.e., grid-cell-specific) model biases are constant over time (Maraun et al., 2017). For the temperature variable, the bias of the geographic location x is given by the difference between the current observed temperature and the modelled temperature, $T_{emp}(x,0) - T_{sim}^{raw}(x,0)$. The bias-corrected temperature at some time in the past x is estimated to be:

$$T_{sim}^{DM}(x,t) \coloneqq T_{emp}(x,0) + \left(T_{sim}^{raw}(x,t) - T_{sim}^{raw}(x,0)\right) \quad \text{(S1)}$$

$$= T_{sim}^{raw}(x,t) + \left(T_{emp}(x,0) - T_{sim}^{raw}(x,0)\right) \quad \text{(S2)}$$

The second expression ($T_{sim}^{DM}(x,t)$) illustrates an alternative approach, where the local deviation of the day is added to the local simulated temperature at time t. The deviation of the day is given by the difference between the current observed temperature and the simulated temperature.

The lower limit of precipitation is zero, and the order of magnitude of precipitation varies from region to region. Therefore, multiplicative rather than additive bias corrections are more common when applying delta to precipitation, which is equivalent to applying modelled relative changes to observed data (Maraun et al., 2017). Similar to temperature, biased precipitation is estimated as:

$$P_{sim}^{DM}(x,t) \coloneqq P_{obs}(x,0) * \frac{P_{sim}^{raw}(x,t)}{P_{sim}^{raw}(x,0)} \quad \text{(S3)}$$

$$= P_{obs}(x,t) * \frac{P_{sim}^{raw}(x,t)}{P_{sim}^{raw}(x,0)} \quad \text{(S4)}$$

We correct the future climate data for 2025-2100 using historical climate data for 2014-2024 and NEX-GDDP-CMIP6 data for the same period, and validate it with historical climate data for 2014-2024 and corrected NEX-GDDP-CMIP6 data. Future climate conditions were represented for the periods 2025-2040, 2041-2070, 2071-2100, respectively. To ensure result reliability, the results were simultaneously averaged using different nationally developed models.

Table S1 Definitions of riverine indicators sensitive to organic matter concentrations, following Yan et al. (2025).

Indicators	Equation
NDCI	$NDCI = (SR_{red} - SR_{green}) / (SR_{red} + SR_{green})$
OC_Index	$OC_Index = (SR_{red} + SR_{nir}) / (SR_{blue} + SR_{swir1})$
Hue Angel	$X = 11.053 \times SR_{ultrablue} + 6.950 \times SR_{blue} + 51.135 \times SR_{green} + 34.475 \times SR_{red}$ $Y = 1.320 \times SR_{ultrablue} + 21.053 \times SR_{blue} + 66.023 \times SR_{green} + 18.034 \times SR_{red}$ $Z = 58.083 \times SR_{ultrablue} + 34.931 \times SR_{blue} + 2.606 \times SR_{green} + 0.016 \times SR_{red}$ $x = X / (X + Y + Z)$ $y = Y / (X + Y + Z)$ $\alpha = [\arctan((y - 1/3) / (x - 1/3)) \bmod 2\pi] \times 180 / \pi$

Table S2 Test and validation scores of different models for COD_{Mn} prediction using satellite data.

No.	Model	Test score	Validation score
1	LightGBMLarge_BAG_L1	-0.22325	-0.66913
2	RandomForestMSE_BAG_L1	-0.2425	-0.661322
3	ExtraTreesMSE_BAG_L1	-0.24681	-0.672584
4	XGBoost_BAG_L1	-0.31248	-0.674913
5	CatBoost_BAG_L1	-0.34349	-0.649774
6	CatBoost_r177_BAG_L1	-0.34857	-0.652313
7	RandomForestMSE_BAG_L2	-0.36146	-0.636188
8	ExtraTreesMSE_BAG_L2	-0.36618	-0.621578
9	NeuralNetTorch_BAG_L2	-0.36816	-0.610234
10	WeightedEnsemble_L3	-0.37258	-0.602517
11	CatBoost_r177_BAG_L2	-0.38146	-0.625719
12	WeightedEnsemble_L2	-0.38381	-0.630267
13	CatBoost_BAG_L2	-0.38433	-0.623588
14	XGBoost_BAG_L2	-0.38697	-0.6598
15	LightGBM_r131_BAG_L1	-0.38935	-0.651746
16	NeuralNetFastAI_BAG_L2	-0.39335	-0.619413
17	LightGBMXT_BAG_L1	-0.39513	-0.673662
18	LightGBMXT_BAG_L2	-0.39965	-0.629565
19	LightGBMLarge_BAG_L2	-0.40136	-0.667139
20	LightGBM_BAG_L2	-0.42184	-0.637511
21	LightGBM_BAG_L1	-0.44421	-0.655706
22	NeuralNetTorch_BAG_L1	-0.51521	-0.722361
23	NeuralNetTorch_r79_BAG_L1	-0.6058	-0.68912
24	NeuralNetFastAI_BAG_L1	-0.68089	-0.773005

Table S3 Test and validation scores of different models for COD_{Mn} prediction based on geographical features.

No.	Model	Test score	Validation score	No.	Model	Test score	Validation score
1	LightGBMLarge_BAG_L1	-0.081138	-0.2644	**19**	CatBoost_r69_BAG_L1	-0.167949	-0.23341
2	XGBoost_r194_BAG_L1	-0.089043	-0.2592	**20**	LightGBM_r131_BAG_L1	-0.168802	-0.23719
3	ExtraTreesMSE_BAG_L1	-0.090097	-0.24898	**21**	LightGBM_BAG_L1	-0.176396	-0.26183
4	ExtraTrees_r42_BAG_L1	-0.092587	-0.25519	**22**	WeightedEnsemble_L2	-0.177354	-0.20715
5	RandomForest_r195_BAG_L1	-0.0953	-0.25655	**23**	NeuralNetTorch_r86_BAG_L1	-0.193151	-0.22308
6	XGBoost_r33_BAG_L1	-0.097439	-0.25877	**24**	NeuralNetTorch_r79_BAG_L1	-0.201559	-0.22065
7	RandomForestMSE_BAG_L1	-0.100664	-0.2697	**25**	NeuralNetTorch_BAG_L1	-0.20326	-0.25054
8	XGBoost_r89_BAG_L1	-0.112642	-0.2575	**26**	NeuralNetTorch_r14_BAG_L1	-0.204362	-0.2446
9	CatBoost_r9_BAG_L1	-0.116537	-0.24187	**27**	NeuralNetTorch_r30_BAG_L1	-0.212912	-0.2288
10	XGBoost_BAG_L1	-0.117449	-0.26967	**28**	NeuralNetFastAI_BAG_L1	-0.213711	-0.26241
11	CatBoost_r13_BAG_L1	-0.13718	-0.23871	**29**	NeuralNetFastAI_r191_BAG_L1	-0.214983	-0.2659
12	LightGBM_r188_BAG_L1	-0.148257	-0.24493	**30**	ExtraTrees_r172_BAG_L1	-0.217508	-0.26801
13	CatBoost_r177_BAG_L1	-0.153872	-0.23387	**31**	NeuralNetFastAI_r103_BAG_L1	-0.224872	-0.23766
14	CatBoost_r50_BAG_L1	-0.156083	-0.23575	**32**	NeuralNetFastAI_r102_BAG_L1	-0.229039	-0.26453
15	CatBoost_BAG_L1	-0.156563	-0.23661	**33**	NeuralNetTorch_r22_BAG_L1	-0.23428	-0.26398
16	LightGBMXT_BAG_L1	-0.15731	-0.24073	**34**	NeuralNetFastAI_r145_BAG_L1	-0.235515	-0.27573
17	LightGBM_r161_BAG_L1	-0.15873	-0.2379	**35**	NeuralNetFastAI_r11_BAG_L1	-0.374389	-0.45661
18	CatBoost_r137_BAG_L1	-0.160723	-0.23589	**36**	LightGBM_r96_BAG_L1	-0.428637	-0.47816

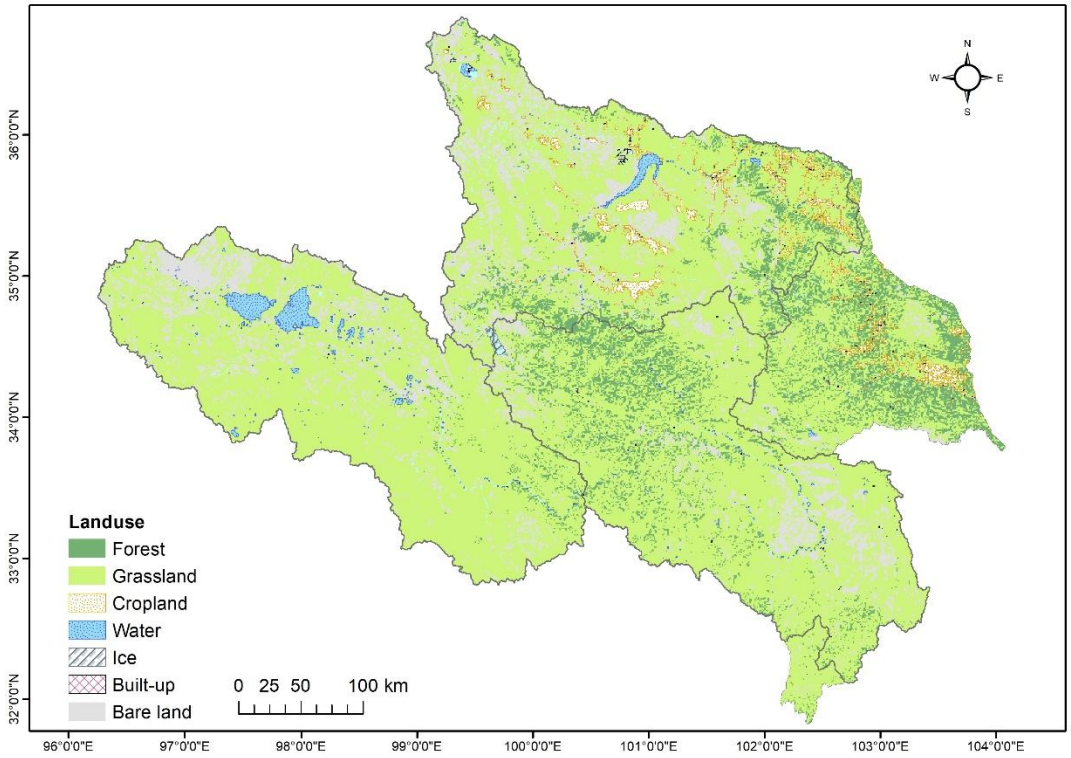


Fig.S1 Land use distribution across the three sub-basins (B1–B3) within the Yellow River source region.

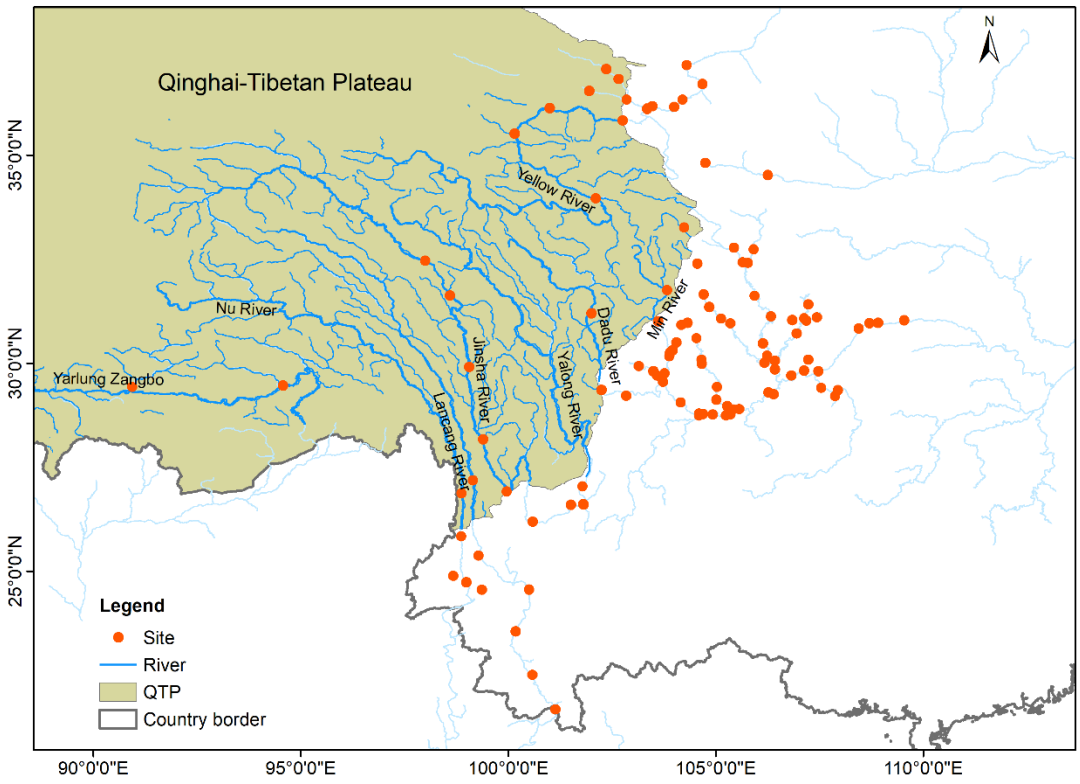


Fig.S2 Locations of data collection sites for model training across the Qinghai–Tibetan Plateau and adjacent areas.

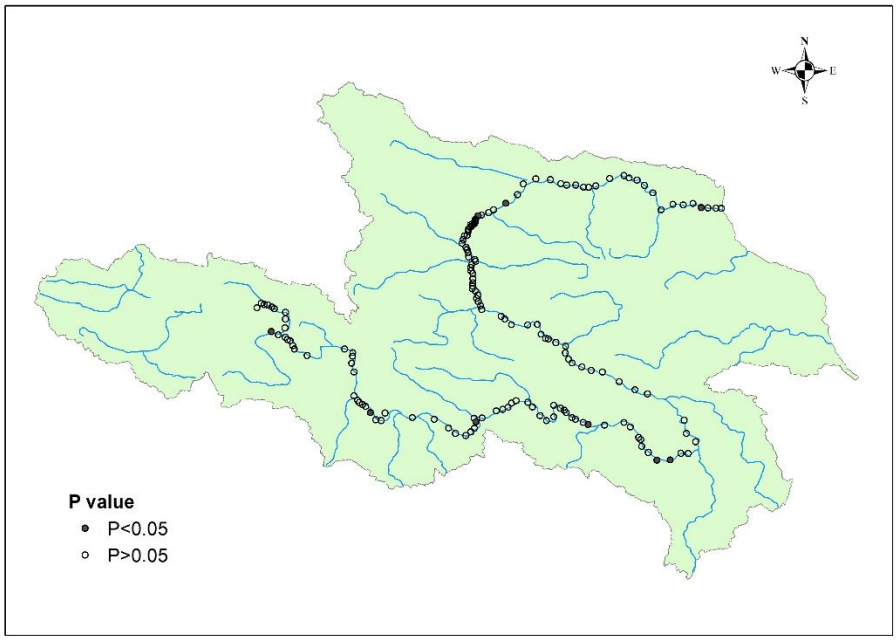


Fig.S3 Statistical significance of historical COD_{Mn} trends during 2014–2024. *P* values were obtained from the Mann–Kendall test.

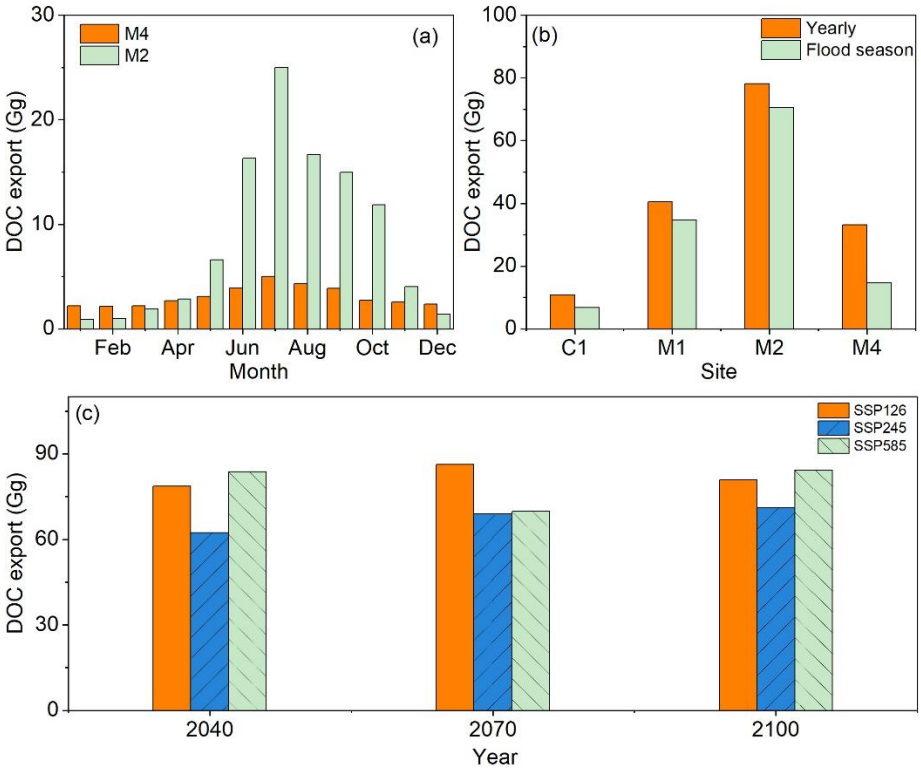


Fig.S4 Monthly average DOC export at sites M2 and M4 during 2014-2024 (a); comparison between annual and flood-season DOC exports (b); projected DOC export at M2 for 2040, 2070, and 2100 under SSP126, SSP245, and SSP585 scenarios (c).

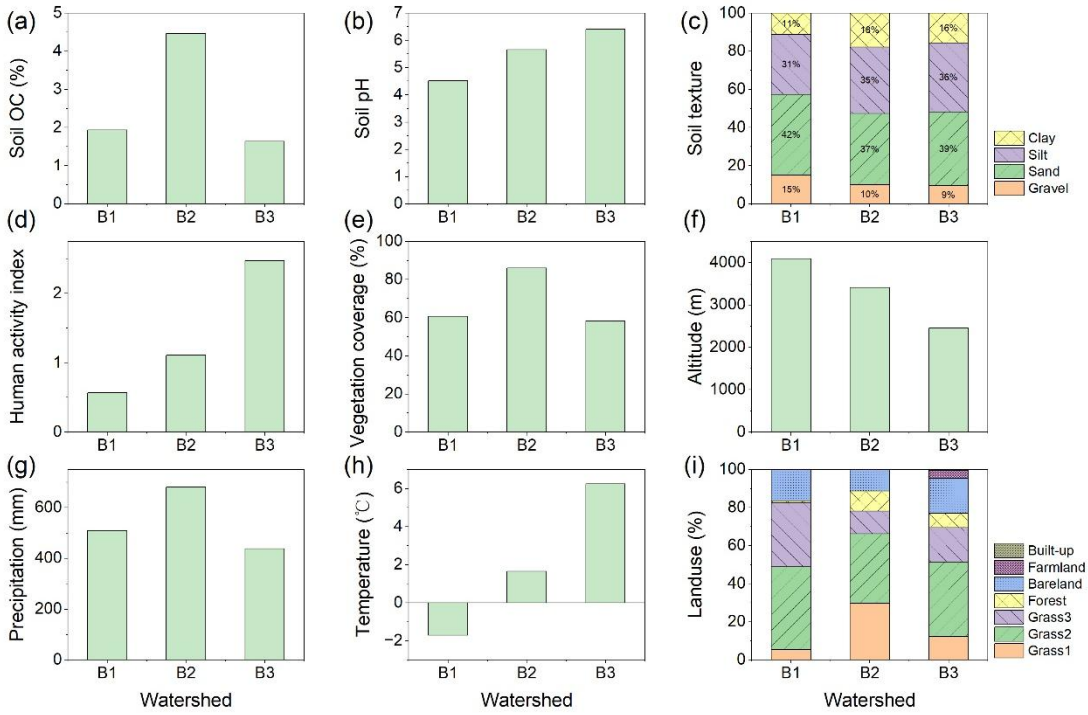


Fig.S5 Geographical characteristics of the sub-basins in the Yellow River source region. (a–c) Soil organic matter, pH, and texture; (d–f) human activity index, vegetation coverage, and altitude; (g–i) average precipitation and temperature over the last ten years, and land use in 2015.

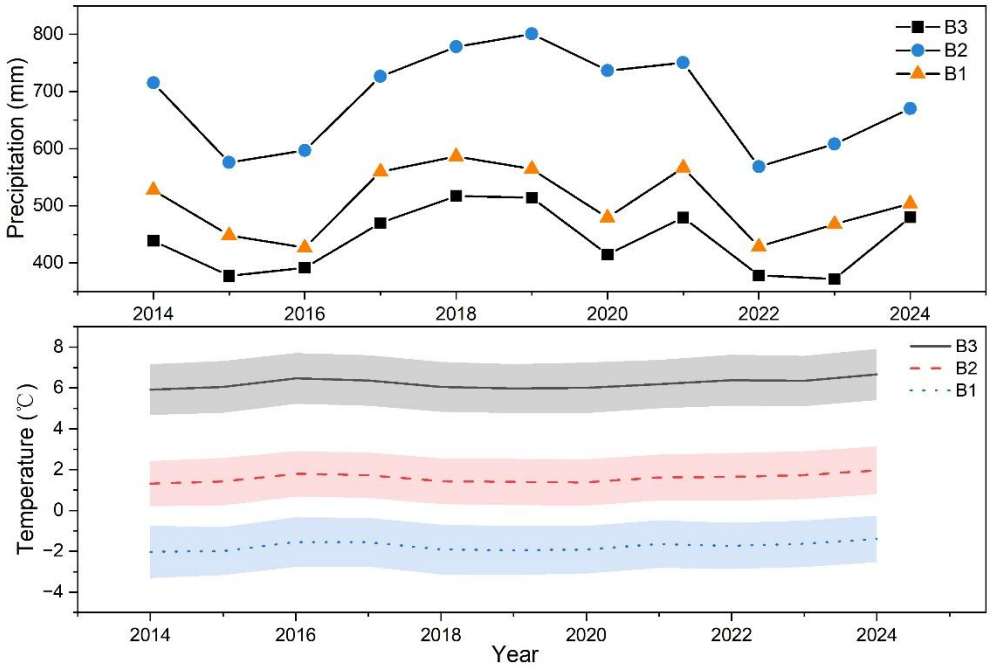


Fig.S6 Variations in precipitation and temperature averaged across different sub-basins of the Yellow River source region during 2014–2024.

References

Beyer, R., Krapp, M., Manica, A., 2020. An empirical evaluation of bias correction methods for palaeoclimate simulations. Climate of the Past 16(4), 1493–1508. https://doi.org/10.5194/cp-16-1493-2020

Cao, Y., Yao, H., Zhang, C., Gu, C., Yang, L., Hao, R., Li, X., Wang, Z., Xiao, H., Wang, J., Shi, Z., 2026. Data-driven assessment of nitrogen and phosphorus buffering capacity across 460 Chinese watersheds: Spatial patterns, drivers, and future projections. Water Research 293, 125424. https://doi.org/10.1016/j.watres.2026.125424

Maraun, D., Shepherd, T.G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J.M., Hagemann, S., Richter, I., Soares, P.M.M., Hall, A., Mearns, L.O., 2017. Towards process-informed bias correction of climate change simulations. Nature Climate Change 7(11), 764–773. https://doi.org/10.1038/nclimate3418

Yan, N., Qiu, Z., Zhang, C., Yan, Y., Liu, D., 2025. Landsat monitoring reveals the history of river organic pollution across China during 1984–2023. Water Research 275, 123210. https://doi.org/10.1016/j.watres.2025.123210