S1. Hydrological fluxes and storages

The mHM model (Samaniego et al. 2010, Kumar et al., 2013) is a spatially explicit, grid-based model that utilizes process formulations from established models such as HBV (Bergström, 1992) and VIC-3L (Gao et al., 2009). Using a grid as the basic unit, mHM represents different hydrological processes. First, canopy interception is simulated, followed by snow accumulation and melt, which are implemented using a degree-day routine. Incoming water from rainfall and snowmelt then enters a multilayer soil/root zone, where a power-law function, adapted from the HBV model, partitions it between soil moisture storage and percolation. Evapotranspiration is calculated from potential ET and reduced by soil-water stress, weighted by the root fraction in each soil layer. Runoff generation is divided into three pathways: (i) fast interflow, triggered when unsaturated storage exceeds a threshold; (ii) slow interflow, representing steady drainage from the same store; and (iii) baseflow from a lower groundwater reservoir, treated as a linear store. The combined flows are routed through the river network using the time-adaptive Muskingum–Cunge scheme (Thober et al., 2019). Full algorithmic details and mHM parameterizations can be found at www.ufz.de/mhm.

The hydrological fluxes of the lumped and landscape-explicit model structures are based on the mHM. From the mHM model, the total outflow from the soil (q_{soil} mm day^{-1}), total baseflow from groundwater ($q_{baseflow}$ mm day^{-1}), and the total groundwater recharge ($q_{recharge}$ day^{-1}) from all grid cells are calculated as follows:

$$q_{soil}(t) = \frac{\sum_{i=1}^{n}\left(q_{fast}^{i}(t)+q_{slow}^{i}(t)\right)\cdot A^{i}}{\sum_{i=1}^{n} A^{i}}, \tag{S1}$$

$$q_{baseflow}(t) = \frac{\sum_{i=1}^{n} q_{baseflow}^{i}(t)\cdot A^{i}}{\sum_{i=1}^{n} A^{i}}, \tag{S2}$$

$$q_{recharge}(t) = \frac{\sum_{i=1}^{n} I^{i}(t)\cdot A^{i}}{\sum_{i=1}^{n} A^{i}}, \tag{S3}$$

where the superscript "i" indicates the grid cell ith, n is the total number of grid cells in the catchment, t (day) is time, q_{fast}^{i}, q_{slow}^{i}, $q_{baseflow}^{i}$, and I (mm day^{-1}) are fast interflow, slow interflow, baseflow, and groundwater recharge at the grid cell level, respectively, simulated from the mHM model, and A^{i} (m^2) is the grid cell area.

The average catchment scale water storage in the soil ($S_{soil}(t)$ mm)) and groundwater ($S_{GW}(t)$ mm)) from mHM are:

$$S_{soil}(t) = \frac{\sum_{i=1}^{n}\left(x_3^{i}(t)+x_5^{i}(t)\right)\cdot A^{i}}{\sum_{i=1}^{n} A^{i}}, \tag{S4}$$

$$S_{gw}(t) = \frac{\sum_{i=1}^{n} x_6^{i}(t)\cdot A^{i}}{\sum_{i=1}^{n} A^{i}}, \tag{S5}$$

where x_3 and x_5 (mm) are water storage from different soil zones (simulated from mHM), and x_6 (mm) is groundwater storage (simulated from mHM).

Lumped model structure:

In the lumped model structure (Fig. 1a), Q_{HS}, Q_{RCH}, Q_{GW} (mm day^{-1}) are the same as q_{soil}, $q_{recharge}$, and q_{base_flow}, respectively. Total outflow (Q_{STREAM} mm day^{-1}) from the catchment is:

$$Q_{STREAM}(t) = q_{soil}(t) + q_{baseflow}(t), \quad \text{(S6)}$$

The storage of the hillslope ($S_{HS}(t)$ mm) and groundwater zone ($S_{GW}(t)$ mm) are:

$$S_{HS}(t) = S_{soil}(t) + S_{HS}^{0}, \quad \text{(S7)}$$

$$S_{GW}(t) = S_{GW}(t) + S_{GW}^{0}, \quad \text{(S8)}$$

where S_{HS}^{0} and S_{GW}^{0} (mm) are the initial water storage in the hillslope and groundwater zone, respectively.

Landscape-explicit model structure:

In the landscape-explicit model structure (Fig. 1b), Q_{HS} (mm day^{-1}) needs to be adjusted to account for the upland area fraction. As the area fraction of the upland is f_{UL} (-), therefore, flow out of the upland is:

$$Q_{UL}(t) = f_{UL} \cdot q_{soil}(t), \quad \text{(S9)}$$

Groundwater flow to the stream (Q_{GW} mm day^{-1}) and to the riparian zone (Q_{GW_RZ} mm day^{-1}), considering the upland area fraction (f_{HS}) and the fraction of groundwater flow to riparian zone ($f_{GW\rightarrow RZ}$), are:

$$Q_{GW}(t) = f_{UL} \cdot (1 - f_{GW\rightarrow RZ}) \cdot q_{baseflow}(t), \quad \text{(S10)}$$

$$Q_{GW\_RZ}(t) = f_{UL} \cdot f_{GW\rightarrow RZ} \cdot q_{baseflow}(t), \quad \text{(S11)}$$

Therefore, the amount of inflow to the riparian zone is:

$$\begin{aligned} Q_{RZ}(t) &= Q_{UL}(t) + Q_{GW\_RZ}(t) + (1 - f_{UL})\left(q_{baseflow}(t) + q_{soil}(t)\right) \\ &= f_{UL} \cdot q_{soil}(t) + f_{UL} \cdot f_{GW\rightarrow RZ} \cdot q_{baseflow}(t) + (1 - f_{UL})\left(q_{baseflow}(t) + q_{soil}(t)\right), \end{aligned} \quad \text{(S12)}$$

To ensure a consistent streamflow simulation with the mHM model, we assumed that in the riparian zone, inflow equals outflow. Therefore, the total outflow (Q_{STREAM} mm day^{-1}) from the catchment is:

$$Q_{STREAM}(t) = Q_{GW}(t) + Q_{RZ}(t) = q_{soil}(t) + q_{baseflow}(t), \quad \text{(S13)}$$

Streamflow in the lumped (Eq. S6) and landscape-explicit (Eq. S13) model structures are identical. Water storage in the upland and groundwater zone are the same as Eqs. S7-S8 while the water storage in the riparian zone ($S_{RZ}(t)$ mm) is:

$$S_{RZ}(t) = S_{soil}(t) + S_{RZ}^{0}, \quad \text{(S14)}$$

where S_{RZ}^{0} (mm) is the initial water storage in the riparian zone.

S2. Model performance metric

In addition to the KGE that was used for model calibration, we also evaluated the performance of behavioural simulations using the Nash–Sutcliffe efficiency (NSE, Nash and Sutcliffe, 1970),

$$NSE = 1 - \frac{\sum_{k=1}^{n}\left(x_k^{obs} - x_k^{sim}\right)^2}{\sum_{k=1}^{n}\left(x_k^{obs} - \bar{x}^{obs}\right)^2}, \quad \text{(S15)}$$

where x_k^{obs} and x_k^{sim} the observed and simulated streamflow (or stream DOC concentration), respectively, k is the index, n is the number of observations, and $\bar{x}^{obs}$ is the mean of observed and simulated values, respectively.

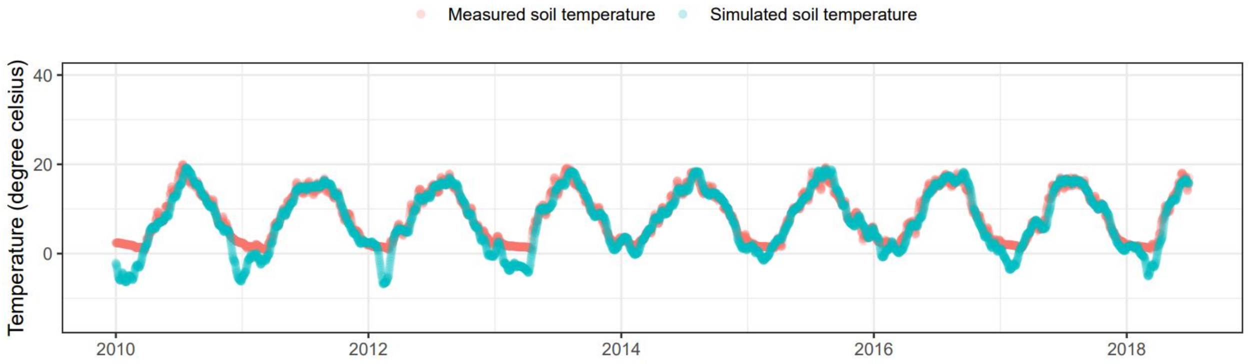

Figure S1. Measured soil temperature (at a 50 cm depth near the study area (Station ID 00656; German Weather Service) was compared with the calculated temperature using Equation (3) (with n = 35) and air temperature data from the Hassel catchment as an example. Measured soil temperature can be obtain at https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/soil_temperature/ (last accessed: 7 June 2025).

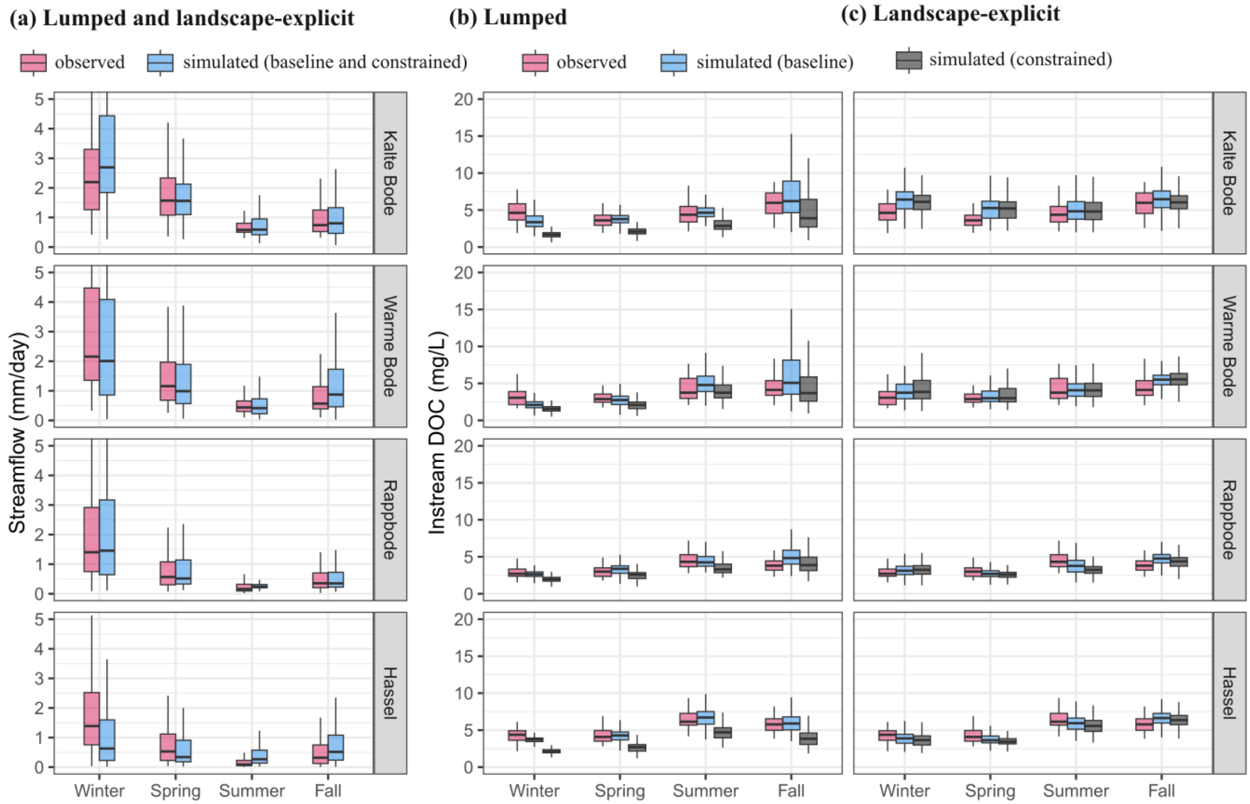

Figure S2. Boxplots of observed (obs) and simulated (sim) (a) streamflow and (b-c) stream DOC concentrations by seasons (Winter: December to February; Spring: March to May; Summer: June to August; Fall: September to November) in the four catchments and with lumped and landscape-explicit model structures under the baseline and constrained calibrations. Data were taken from both calibration and validation periods.

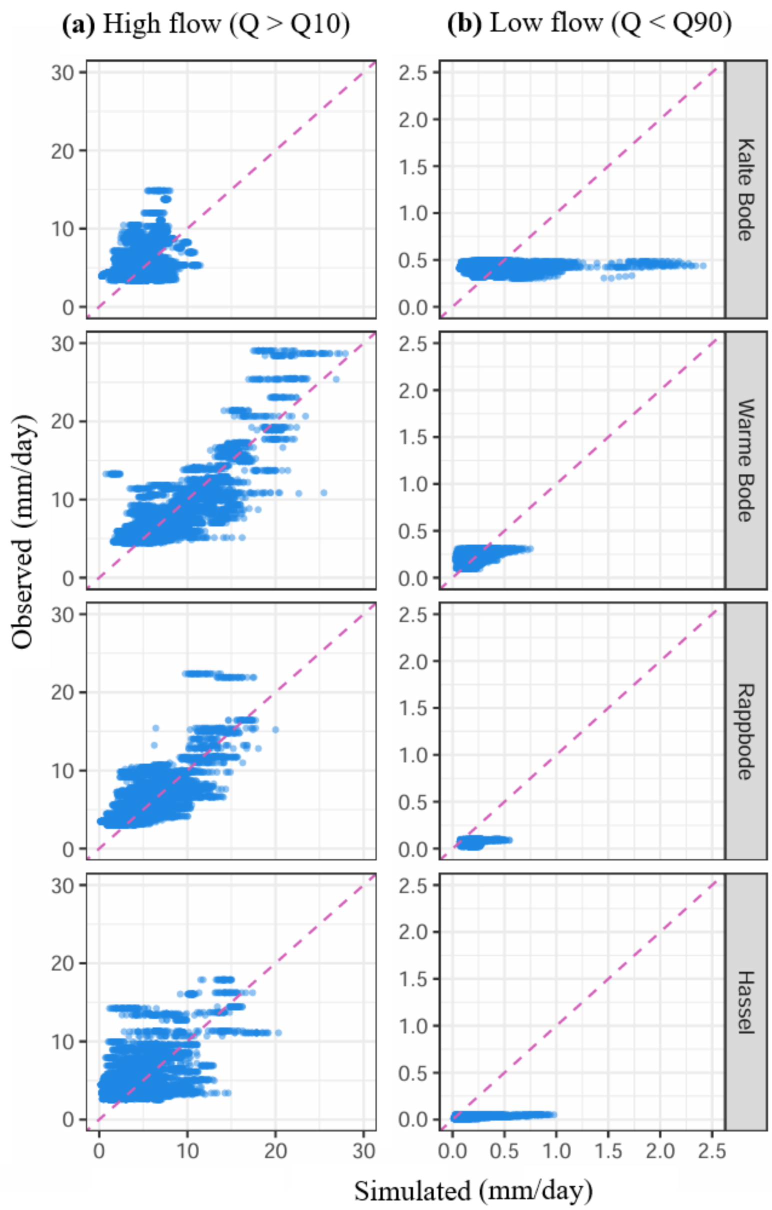

Figure S3. Simulated and observed high flows (Q > Q10: the streamflow that is equalled or exceeded 10% of the time) and low flows (Q < Q90: the streamflow that is equalled or exceeded 90% of the time). Q10 and Q90 are defined from observed data for each respective basin. Data are from both calibration and validation period.

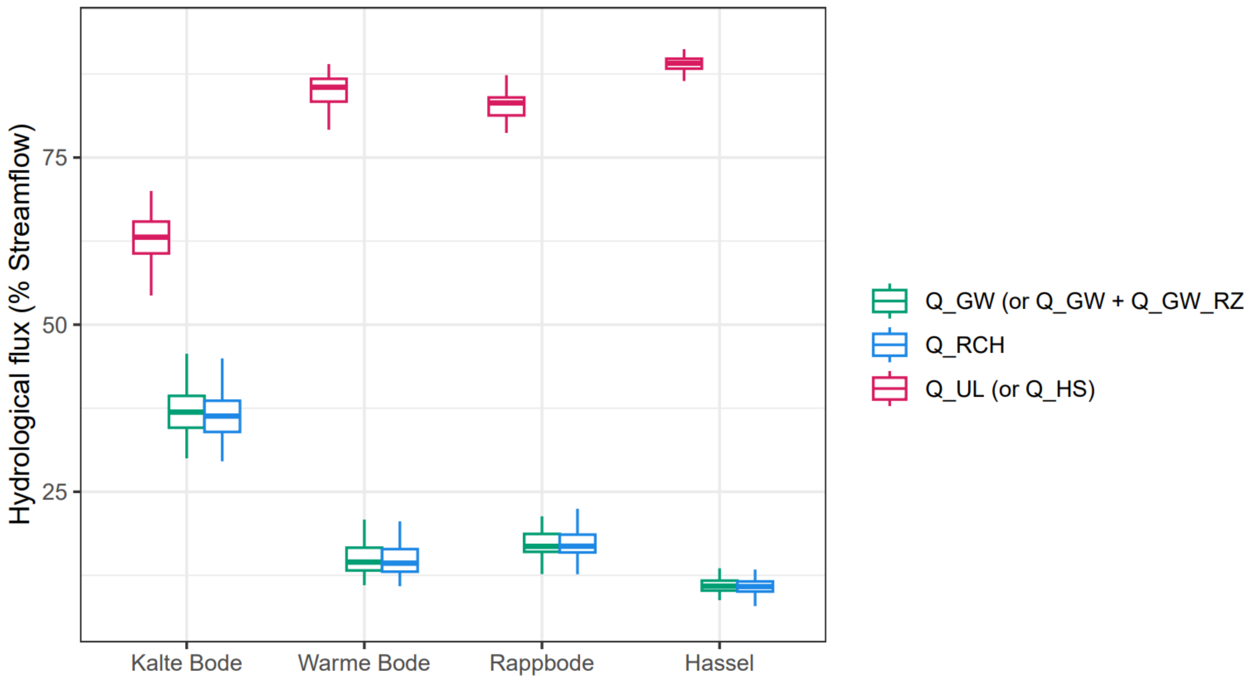

Figure S4. Boxplots of annual average hydrological fluxes (in percentage of streamflow) for both calibration and validation periods from 30 behavioural models of each catchment (baseline calibration). Q_{UL} (or Q_{HS}) and Q_{GW} (mm/year) are outflows from the upland (or hillslope when referring to the lumped model structure) and from groundwater to stream, respectively. Q_{RCH} (mm/year) is the groundwater recharge. The sum of Q_{GW} and Q_{GW_RZ} is the total outflow from groundwater to stream and groundwater to riparian zone in the landscape-explicit model.

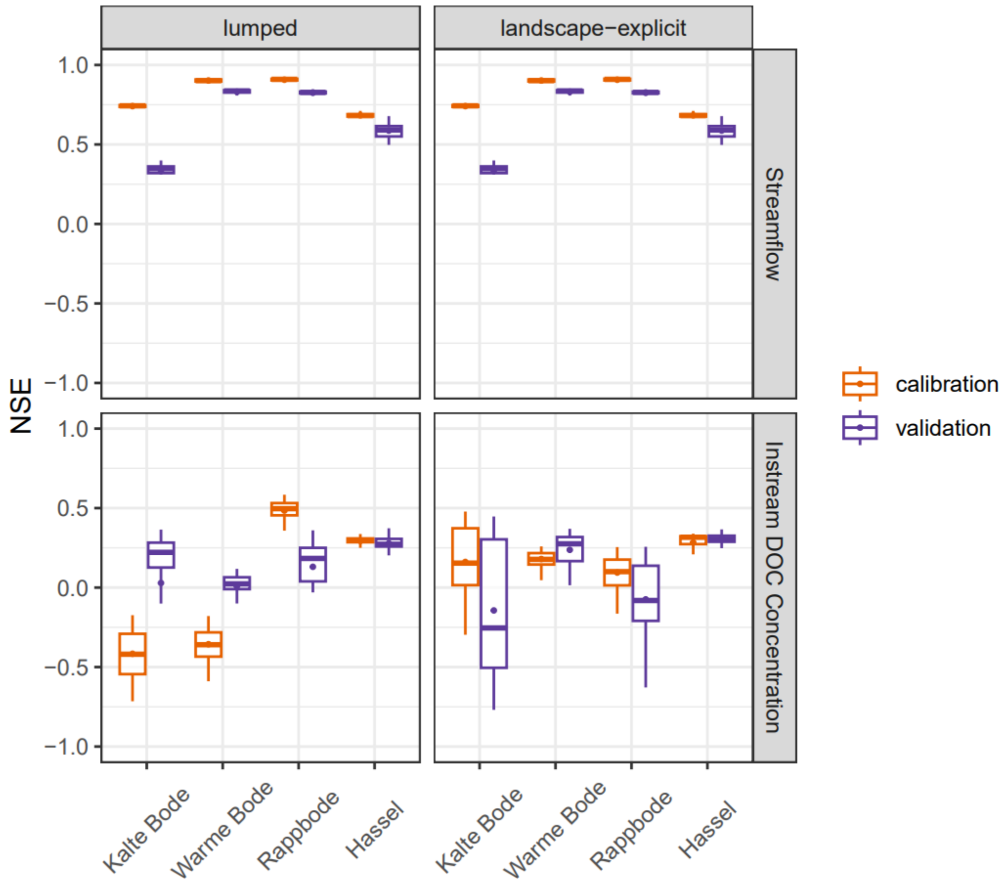

Figure S5. Boxplots of the NSE from the 30 behavioural simulations for streamflow and stream DOC under the baseline calibration. The dot represents the mean value. The orange and violet colours indicate the calibration and validation periods, respectively. Note that streamflow results are identical because hydrological modelling preceded the differentiation of model structures in DOC modelling.

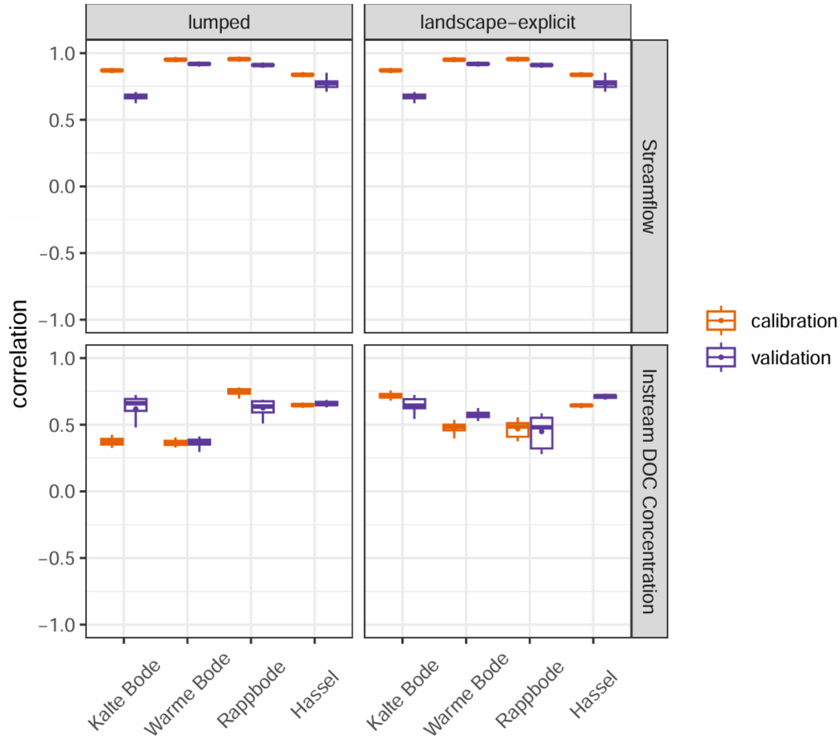

Figure S6. Boxplots of the correlation component (r) in the KGE (Eq. 14) from the 30 behavioural simulations for streamflow and stream DOC under the baseline calibration. The orange and violet colours indicate the calibration and validation periods, respectively. The dot represents the mean value. Note that streamflow results are identical because hydrological modelling preceded the differentiation of model structures in DOC modelling.

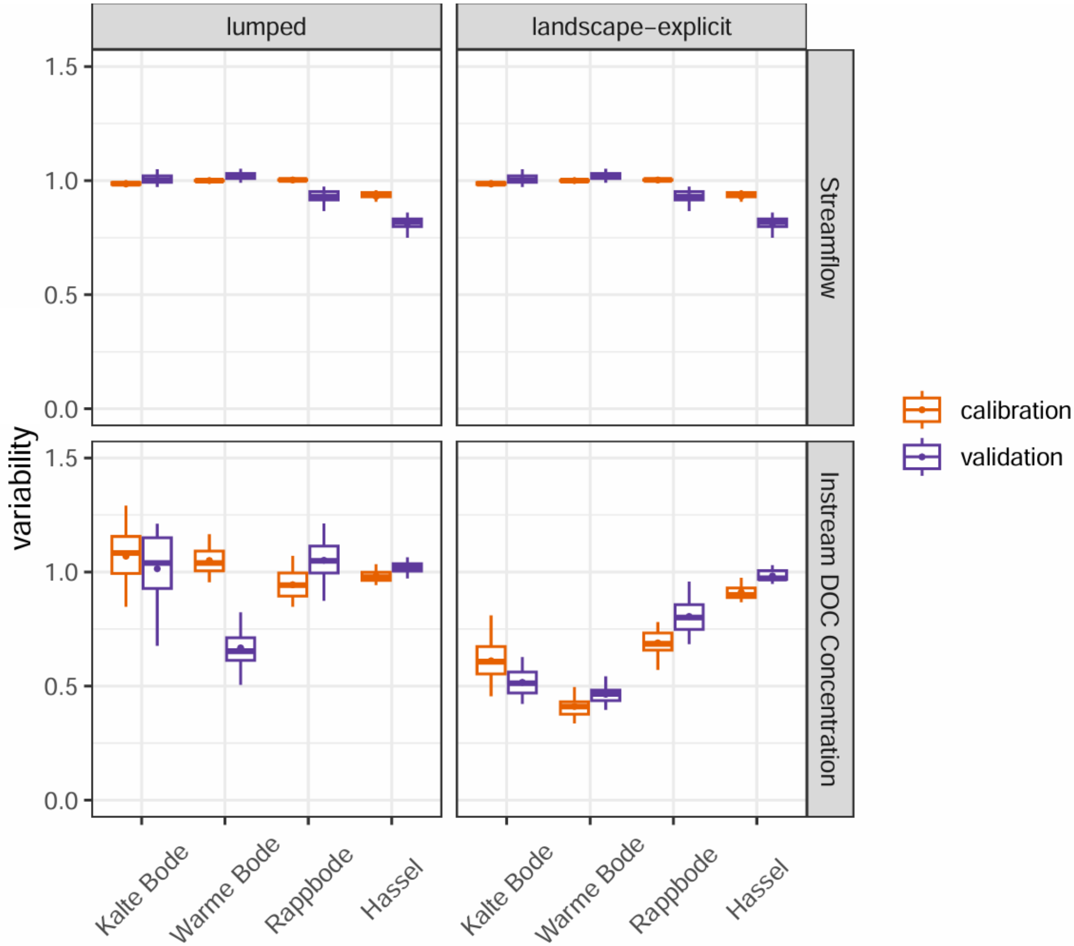

Figure S7. Boxplots of the variability component (alpha α) of the KGE (Eq. 14) from the 30 behavioural simulations for streamflow and stream DOC under the baseline calibration. The dot represents the mean value. The orange and violet colours indicate the calibration and validation periods, respectively. Note that streamflow results are identical because hydrological modelling preceded the differentiation of model structures in DOC modelling.

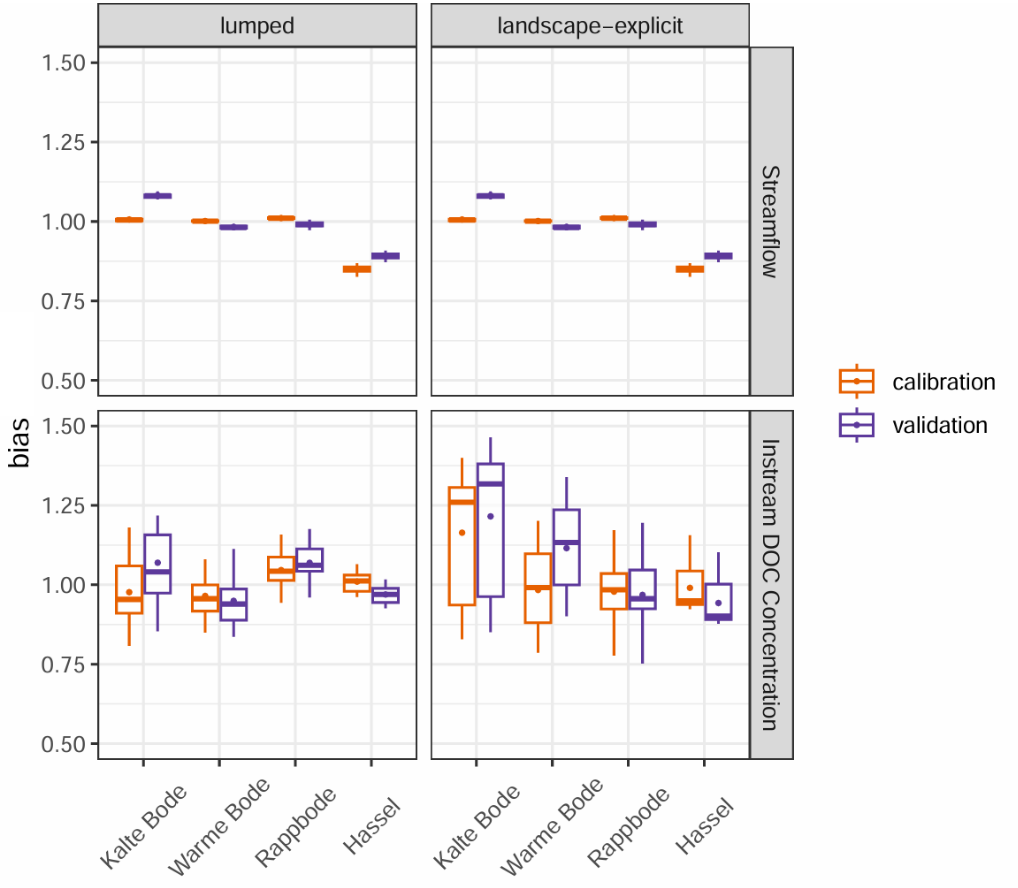

Figure S8. Boxplots of the bias component (beta β) of the KGE (Eq. 14) from the 30 behavioural simulations for streamflow and stream DOC under the baseline calibration. The dot represents the mean value. The orange and violet colours indicate the calibration and validation periods, respectively. Note that streamflow results are identical because hydrological modelling preceded the differentiation of model structures in DOC modelling.

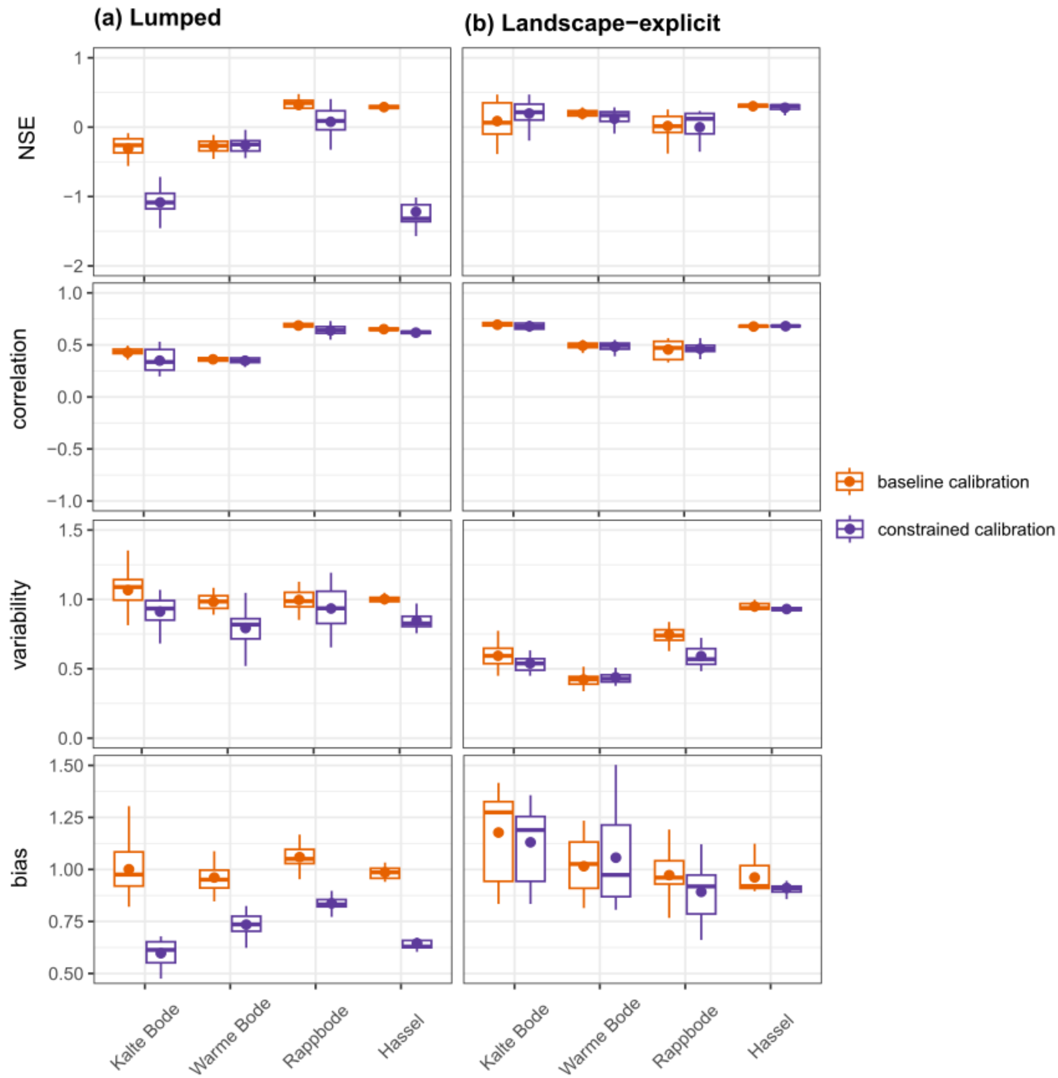

Figure S9. Boxplots of the NSE, the components of the KGE, (correlation, variability, bias) and 30 behavioural simulations for stream DOC under the baseline calibration and constrained calibration. The dot represents the mean value. The orange and violet colours indicate the baseline and constrained calibrations, respectively. The performance indices were calculated for both calibration and validation periods.

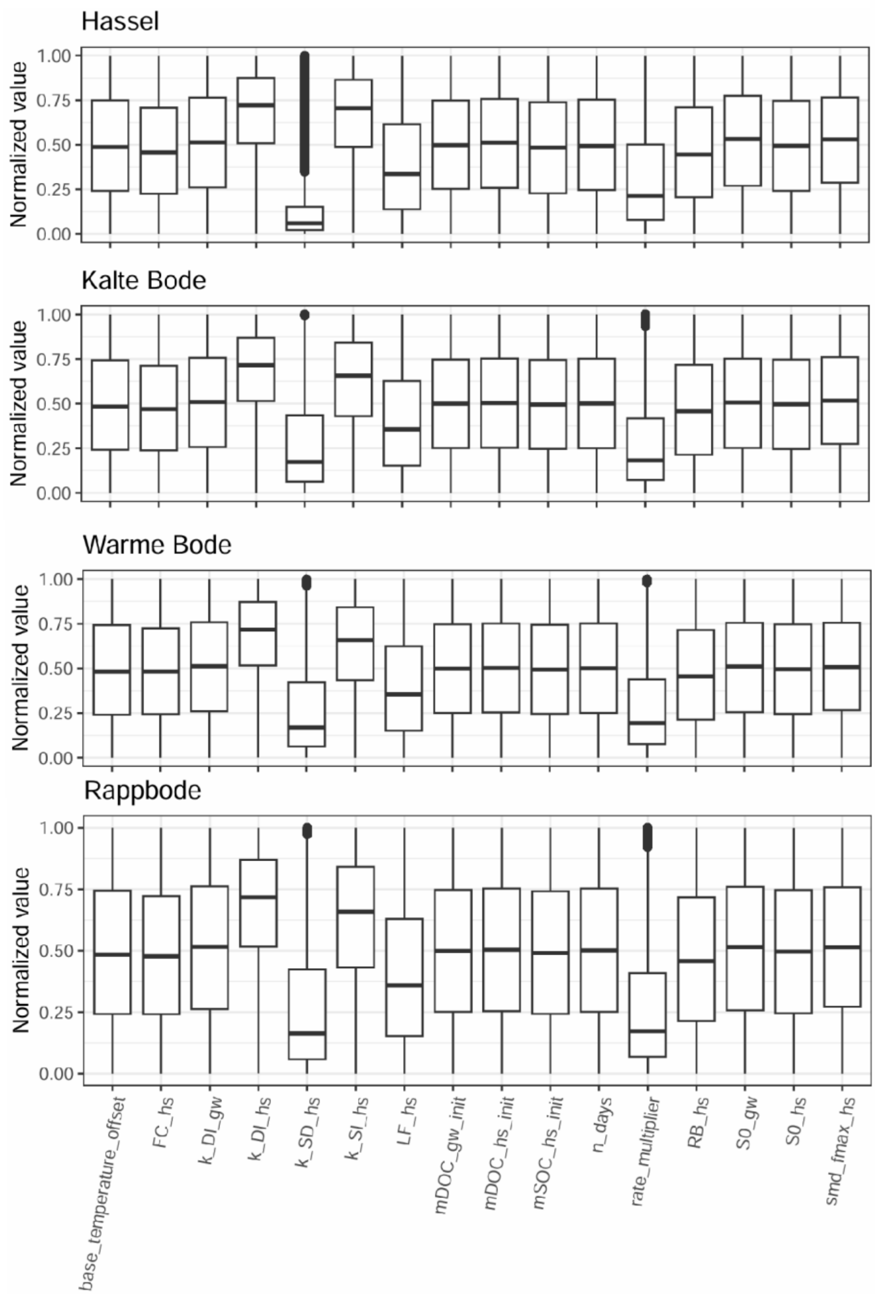

Figure S10: Boxplots of the parameters of the lumped model structure satisfying constraints 1. For visualization, each parameter range was normalized such that the minimum and maximum values correspond to 0 and 1, respectively. Therefore, the initial parameter distribution is centered at 0.5, with the first and third quartiles at 0.25 and 0.75, respectively. The parameter description, including the actual parameter ranges, is provided in the supporting document (doc_parameter.txt). These boxplots are univariate presentation the distribution of each parameter rather than a joint distribution among multiple parameters.

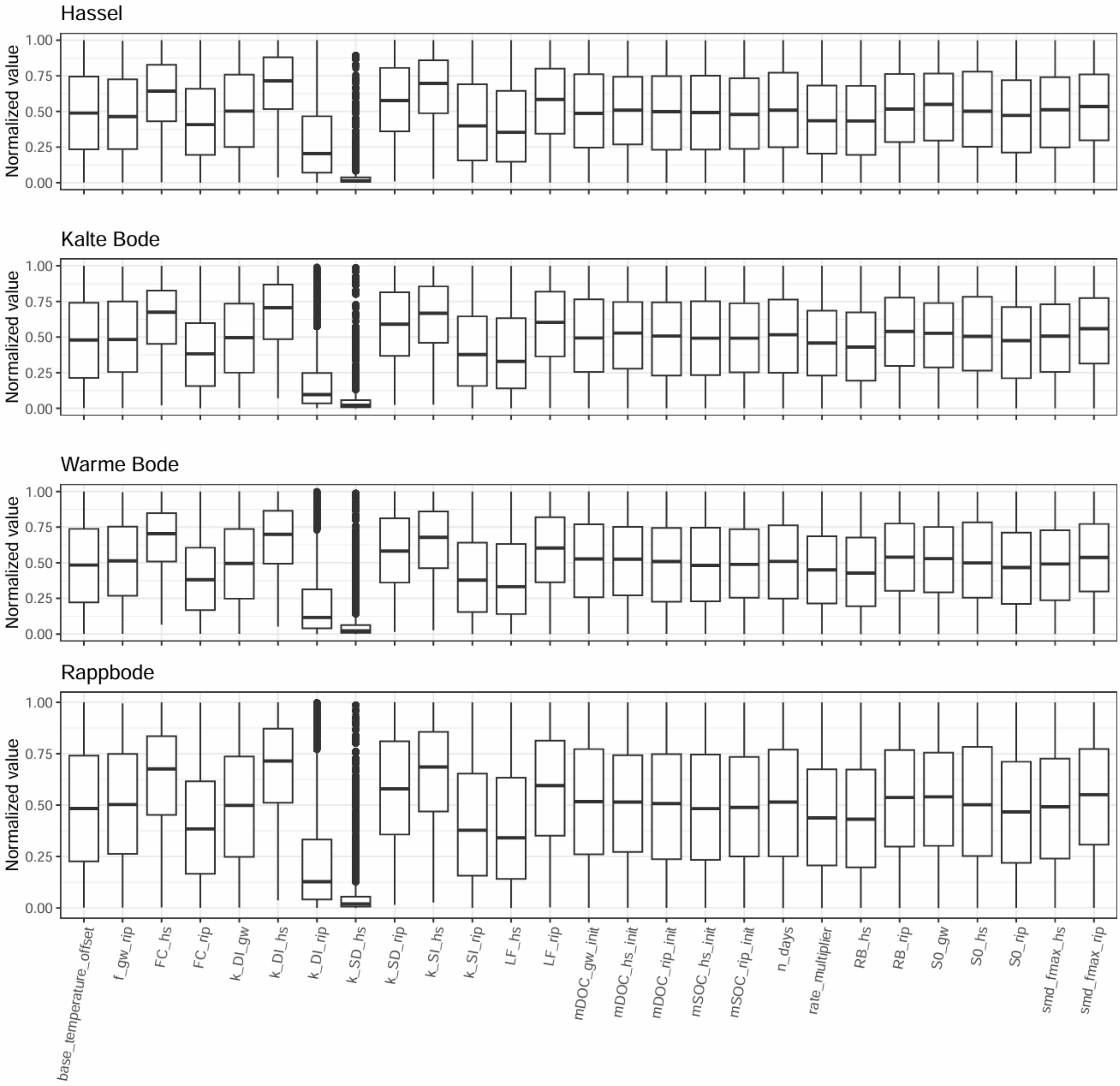

Figure S11: Boxplots of the parameters of the landscape-explicit model structure satisfying constraints 1, 2, and 3. For visualization, each parameter range was normalized such that the minimum and maximum values correspond to 0 and 1, respectively. Therefore, the initial parameter distribution is centered at 0.5, with the first and third quartiles at 0.25 and 0.75. The parameter description, including the actual parameter ranges, is provided in the supporting document (doc_parameter.txt). These boxplots are univariate presentation the distribution of each parameter rather than a joint distribution among multiple parameters

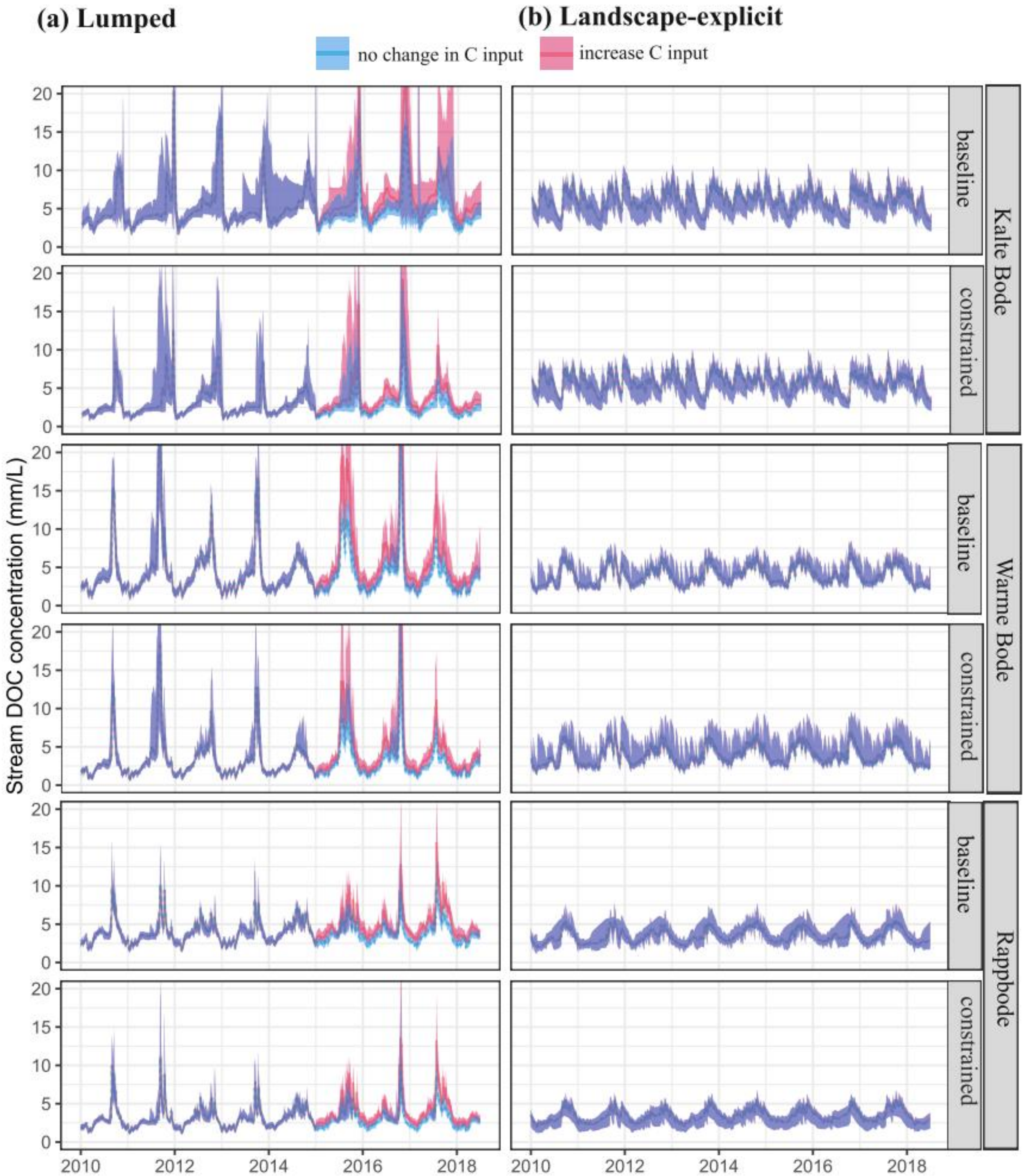

Figure S12. Simulated stream DOC concentrations in the Kalte Bode, Warme Bode, and Rappbode catchments with increasing C input by 5 kg/ha for the period 2015-2018 (a) in the hillslope of the lumped model structure and (b) in the upland of the landscape-explicit model structure and (b).

Table S1. Number of DOC observations in calibration and validation periods.

Basins	Kalte_Bode	Warme_Bode	Rappbode	Hassel
Calibration period	97	94	91	88
Validation period	34	33	119	127

S3. Effect of uncertainty in riparian zone delineation on the main findings

In the main manuscript, we acknowledge that the delineated riparian area is sensitive to the defined topographic wetness index (TWI). Here, we tested whether this uncertainty affects the main conclusions of the study. We selected the Rappbode, where the TWI threshold was defined and applied to the others. We adjusted the TWI threshold to generate riparian extents that were 25% smaller and 25% larger than the reference riparian extent. This changed the riparian area fraction of the Rappbode from 8.7% in the reference case to 6.5% and 10.9%, respectively. We then reran the landscape-explicit model structure only for the Rappbode with these new riparian zone fractions (6.5% and 10.9%) as examples, using the same baseline calibration, constrained calibration, and scenario simulation settings to test whether the main findings reported in Section 5 were maintained.

Finding 1: "Under the baseline calibration, both lumped and landscape-explicit model structures can reasonably represent stream DOC concentration dynamics, with none consistently outperforming the other across catchments and evaluation metrics. However, the simulated DOC concentrations in groundwater for both model structures were significantly higher than those from observations, indicating unrealistic internal DOC dynamics."

> Compared to the baseline simulation with a fixed value of the riparian zone (Fig. 4d), the KGE ranges are wider (lower minimum and higher maximum values; Fig. S13a). Some behavioral simulations even have higher KGE values, indicating that our conclusion that "landscape-explicit model structures can reasonably represent stream DOC concentration dynamics" is still valid. In line with previous findings, we found that while the lumped model (Fig. 4c) is worse than the landscape-explicit model structure during the calibration period, it is better during the validation period, showing that no model structure consistently outperforms the others. We also found that the performance varies across evaluation metrics. For example, the KGE shows only a slight difference between the calibration and validation periods, while the NSE (Fig. S13b) shows clearly higher performance in the calibration period than in the validation period.

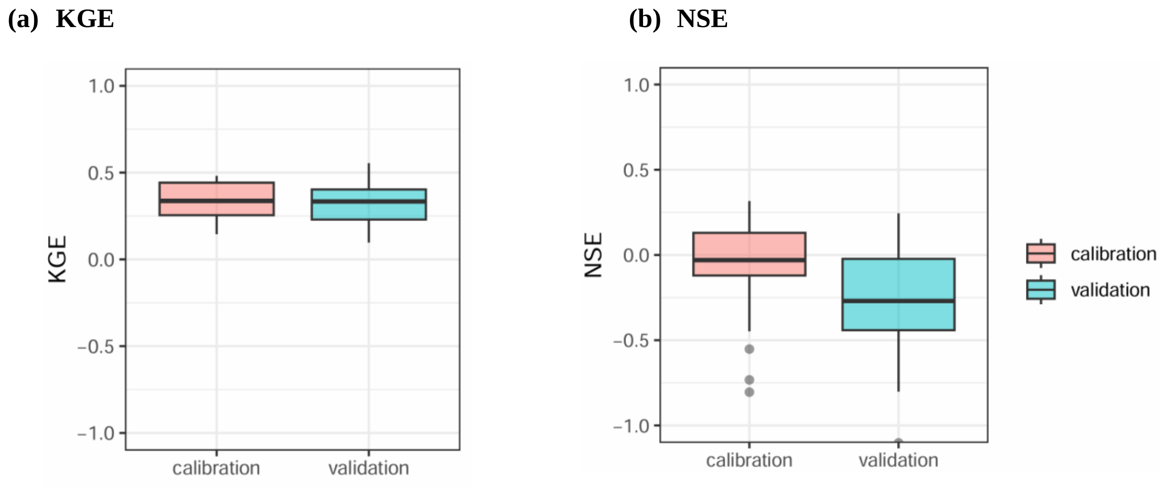

Figure S13. Boxplots of the (a) KGE and (b) NSE for stream DOC concentrations simulation for the calibration and validation periods of the Rappbode when decreasing and increasing the riparian zone area by 25%. The resulting riparian zone fractions are 6.5% and 10.9%.

Finding 2: "The constrained calibration (accounting for physically realistic DOC concentrations in the upland and groundwater compartments) slightly reduces model performance for stream DOC concentration simulation for the landscape-explicit model structure, while increasing the physical realism of the model in terms of internal DOC dynamics. In contrast, the performance of the lumped model structure is significantly reduced under the constrained calibration."

Here, we also found that, with the landscape-explicit model structure, the model performance for stream DOC concentration is slightly reduced under the constrained calibration compared to the baseline calibration. Furthermore, the model performance spreads over a wider range (wider interquartile range) when increasing and decreasing the riparian zone.

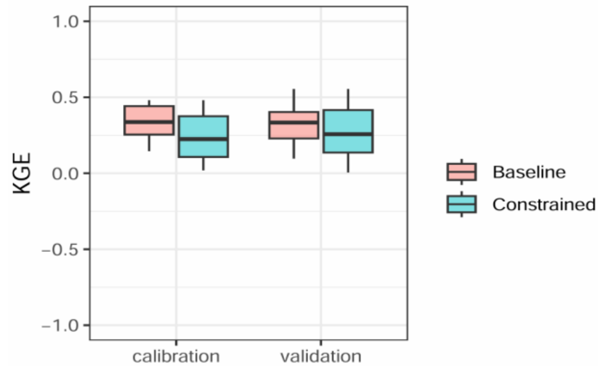

Figure S14. Boxplots of the KGE for stream DOC concentrations simulation for the baseline and constrained calibration of the Rappbode when decreasing and increasing the riparian zone area by 25%. The resulting riparian zone fractions are 6.5% and 10.9%.

Finding 3: "For evaluating the effectiveness of spatial management on stream DOC (e.g., changes that do not occur homogeneously over the entire catchment, such as increasing C input in the upland in this study), a landscape-explicit model structure provides credible results, while a lumped structure does not."

Consistent with the previous findings, after rerunning the scenario simulation with the new riparian zone fractions (6.5% and 10.9%), we found that increasing C input in the upland has no or only a very minor effect on stream DOC concentration (Fig. S15).

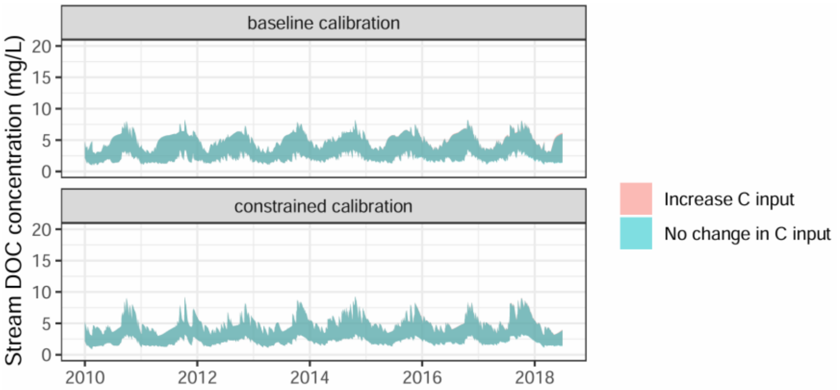

Figure S15. Effect of increasing C input (during the period 2015-2018) in the upland of the lumped model structure on simulated stream DOC concentrations in the Rappbode catchment. The results with the riparian zone fractions are 6.5% and 10.9%.

References

Bergström, S.: The HBV Model: Its Structure and Applications, Swedish Meteorological and Hydrological Institute (SMHI), Hydrology, Norrköping. https://urn.kb.se/resolve?urn=urn:nbn:se:smhi:diva-2672, 1992.

Samaniego L., R. Kumar, and Attinger, S.: Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale, Water Resour. Res., 46, W05523, https://doi.org/10.1029/2008WR007327, 2010.

Kumar, R., L. Samaniego, and S. Attinger: Implications of distributed hydrologic model parameterization on water fluxes at multiple scales and locations, Water Resour. Res., 49, https://doi.org/10.1029/2012WR012195, 2013.

Thober, S., Cuntz, M., Kelbling, M., Kumar, R., Mai, J., and Samaniego, L.: The multiscale routing model mRM v1.0: simple river routing at resolutions from 1 to 50 km, Geosci. Model Dev. 12 (6), 2501 - 2521, https://doi.org/10.5194/gmd-12-2501-2019, 2019.

Gao, H., Tang, Q., Shi, X., Zhu, C., Bohn, T. J., Su, F., Sheffieldm J., Pan, M., Lettenmaier, D., and Wood, E. F.: Water budget record from variable infiltration capacity (VIC) Model, Algorithm Theoretical Basis Document, Version 1.2., 2009.