



Supplement of

Technical note: Machine learning metamodelling for global sensitivity analysis

Patricio Yeste et al.

Correspondence to: Patricio Yeste (patricio.yeste@uni-potsdam.de)

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

S1 Model parameters

The definitions of the HBV, HyMod, and VIC parameters and their feasible ranges are provided in Tables S1–S3.

Table S1. HBV model parameters and their corresponding ranges used in the sensitivity analysis. Parameter ranges were taken from Kollat et al. (2012).

Parameter (Units)	Description	Lower bound	Upper bound
T_s (°C)	Temperature threshold for snow	−3	3
CFMAX (mm °C ^{−1} d ^{−1})	Degree-day factor	0	20
CFR (-)	Refreezing coefficient	0	1
CWH (-)	Water holding capacity of snow	0	0.8
BETA (-)	Shape coefficient	0	7
LP (-)	Soil moisture threshold for reduction of evaporation	0.3	1
FC (mm)	Maximum soil moisture	1	2000
PERC (mm d ^{−1})	Maximal flow from upper to lower groundwater box	0	100
K_0 (d ^{−1})	Recession coefficient	0.05	2
K_1 (d ^{−1})	Recession coefficient	0.01	1
K_2 (d ^{−1})	Recession coefficient	0.05	0.1
UZL (mm)	Threshold for K_0 -outflow	0	100
MAXBAS (d)	Routing, length of weighting function	1	6

Table S2. HyMod model parameters and their corresponding ranges used in the sensitivity analysis. Parameter ranges were taken from Kollat et al. (2012).

Parameter (Units)	Description	Lower bound	Upper bound
S_m (mm)	Maximum soil moisture storage	0	400
beta (-)	Contributing area curve shape parameter	0	2
alfa (-)	Fraction of effective precipitation that is fast flow	0	1
R_s (d)	Slow flow residence time	0	0.1
R_f (d)	Fast flow residence time	0.1	1

Table S3. VIC model parameters and their corresponding ranges used in the sensitivity analysis. Parameter ranges for VIC were taken from Schaperow et al. (2021), while the ranges of the gamma function parameters were taken from Yeste et al. (2023, 2024). All remaining VIC soil and vegetation parameters were obtained from the VICGlobal dataset (Schaperow et al., 2021) and held fixed during the sensitivity analysis.

Parameter (Units)	Description	Lower bound	Upper bound
b_{infiltr} (-)	Variable infiltration shape parameter	0	0.4
D_s (-)	Fraction of D_{max} where non-linear baseflow begins	0	1
D_{max} (mm d ^{−1})	Maximum baseflow	0	30
W_s (-)	Fraction of the porosity of the bottom soil layer where non-linear baseflow begins	0	1
d_1 (m)	Thickness of soil layer 2	0.01	0.5
$d_2 - d_1$ (m)	Difference of thickness between soil layer 2 and soil layer 1	0.04	0.5
d_3 (m)	Thickness of soil layer 3	0.5	2.5
γ_{shape} (-)	Shape parameter of gamma function	0	10
γ_{scale} (-)	Scale parameter of gamma function	0	2

In the VIC model, two soil parameters are subject to a structural dependency: the thickness of soil layer 1 (d_1) and the thickness of soil layer 2 (d_2). By model construction, $d_2 > d_1$. The prescribed parameter ranges for these quantities are $d_1 \in$

5 $[0.01, 0.5]$ m and $d_2 \in [0.05, 1]$ m (Schaperow et al., 2021), which implies that independent sampling of d_1 and d_2 may generate parameter combinations for which $d_2 < d_1$. Such combinations violate the physical constraints of the model and lead to model failure.

10 To preserve the independence assumption required for the validity of the approach, the Delta method described by Mai (2023) was applied. Rather than sampling d_2 directly, the difference $d_2 - d_1$ was treated as an independent model parameter and sampled from the interval $[0.04, 0.5]$ m. This range was chosen to be consistent with the original feasible ranges of d_1 and d_2 , ensuring that the resulting values of d_2 remain within the bounds commonly used in VIC applications. This reparameterization ensures that all sampled input factors are independent, while respecting the physical constraints of the VIC model.

S2 Supplementary figures

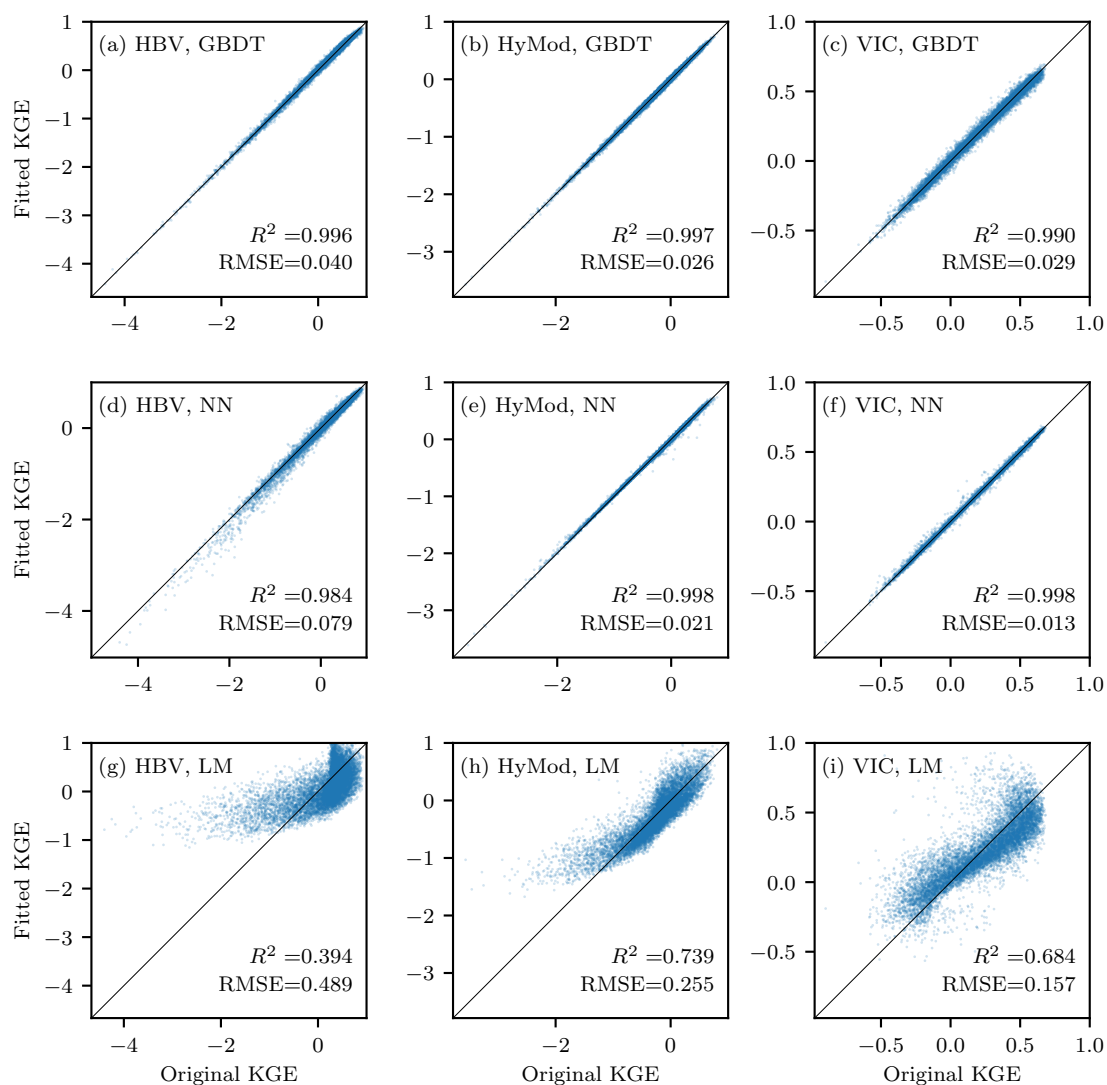


Figure S1. Predictive performance of the GBDT, NN, and LM metamodels for the three hydrologic models applied in the Schackendorf catchment. Scatter plots compare fitted and original KGE values for (a–c) GBDT, (d–f) NN, and (g–i) LM applied to HBV, HyMod, and VIC, respectively. The solid line indicates the 1:1 relationship. Coefficients of determination (R^2) and root mean square errors (RMSE) are reported in each panel.

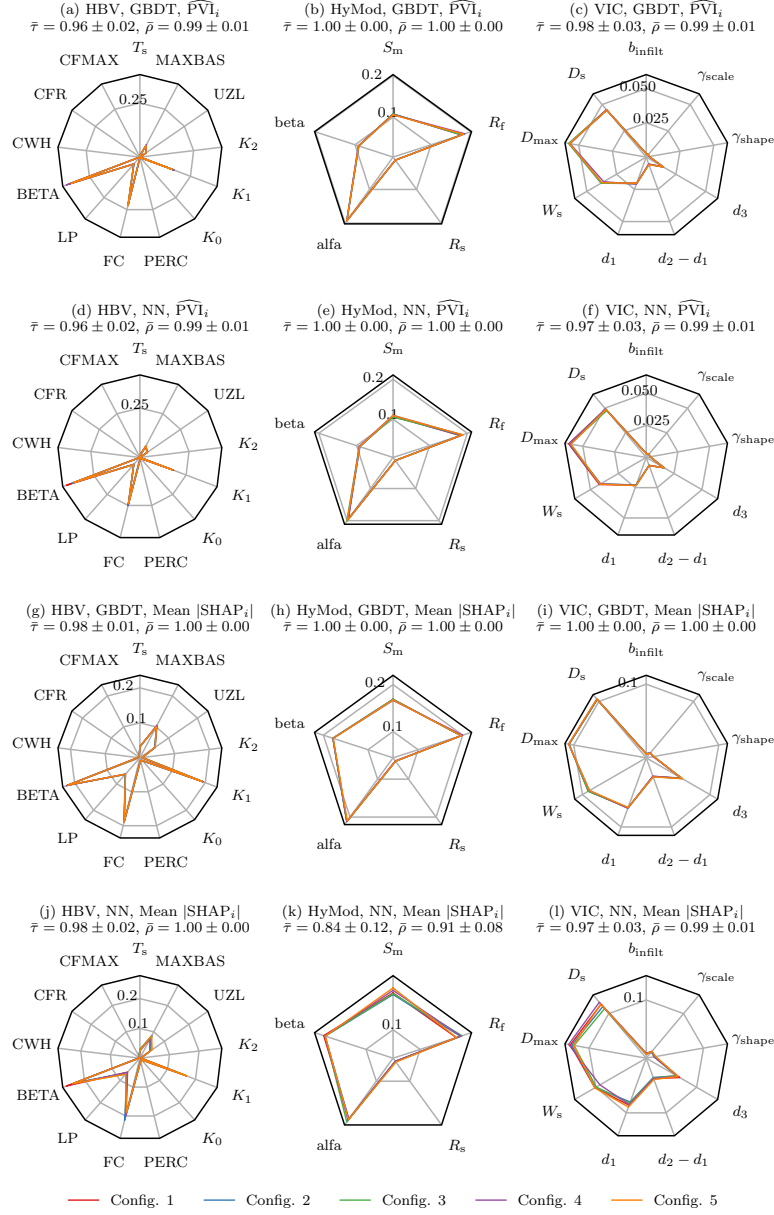


Figure S2. Sanity check of feature importance estimates for GBDT and NN metamodels across five overfitting configurations for the Schackendorf catchment. Panels (a–c) show $\widehat{PVI}_i$ and panels (g–i) show mean $|\text{SHAP}_i|$ for the GBDT metamodel applied to HBV, HyMod, and VIC, respectively, while panels (d–f) and (j–l) show the corresponding results for the NN metamodel. Each coloured line represents a different model configuration. Panel titles report the mean $\pm$ standard deviation of pairwise Kendall τ and Spearman ρ correlation between the importance vectors of the five configurations. The strong overlap of importance curves across configurations, supported by high $\bar{\tau}$ and $\bar{\rho}$ estimates, indicates that the feature importance estimates are robust to overfitting.

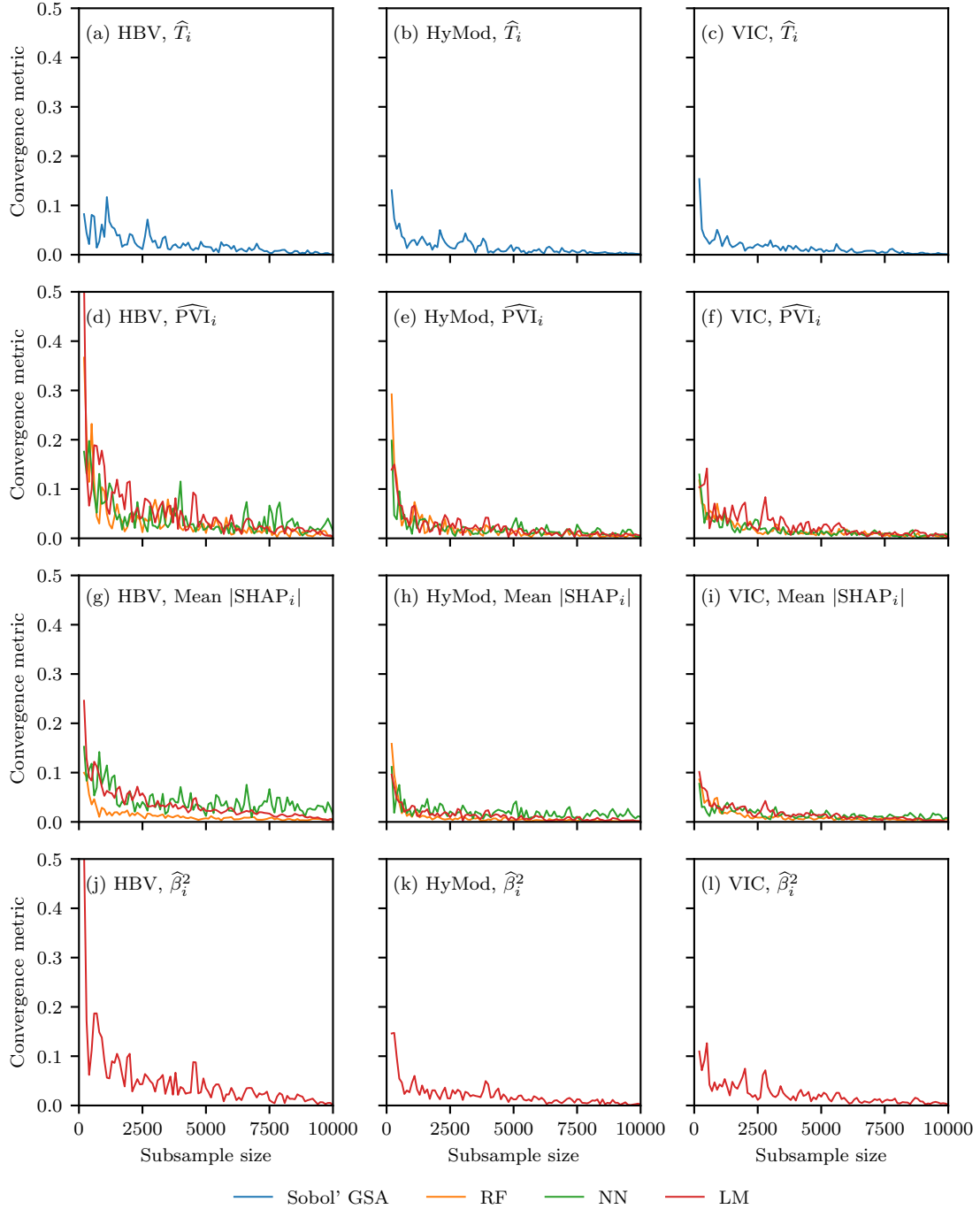


Figure S3. Convergence of sensitivity measures as a function of subsample size for the Schackendorf catchment. Panels (a–c) show the convergence of $\hat{T}_i$, panels (d–f) of $\widehat{PVI}_i$, panels (g–i) of mean $|\text{SHAP}_i|$, and panels (j–l) of $\hat{\beta}_i^2$ for HBV, HyMod, and VIC, respectively. The convergence metric is computed as the Euclidean distance between normalized sensitivity vectors obtained from consecutive subsample sizes. Normalization was performed by dividing by the sum of all sensitivities.

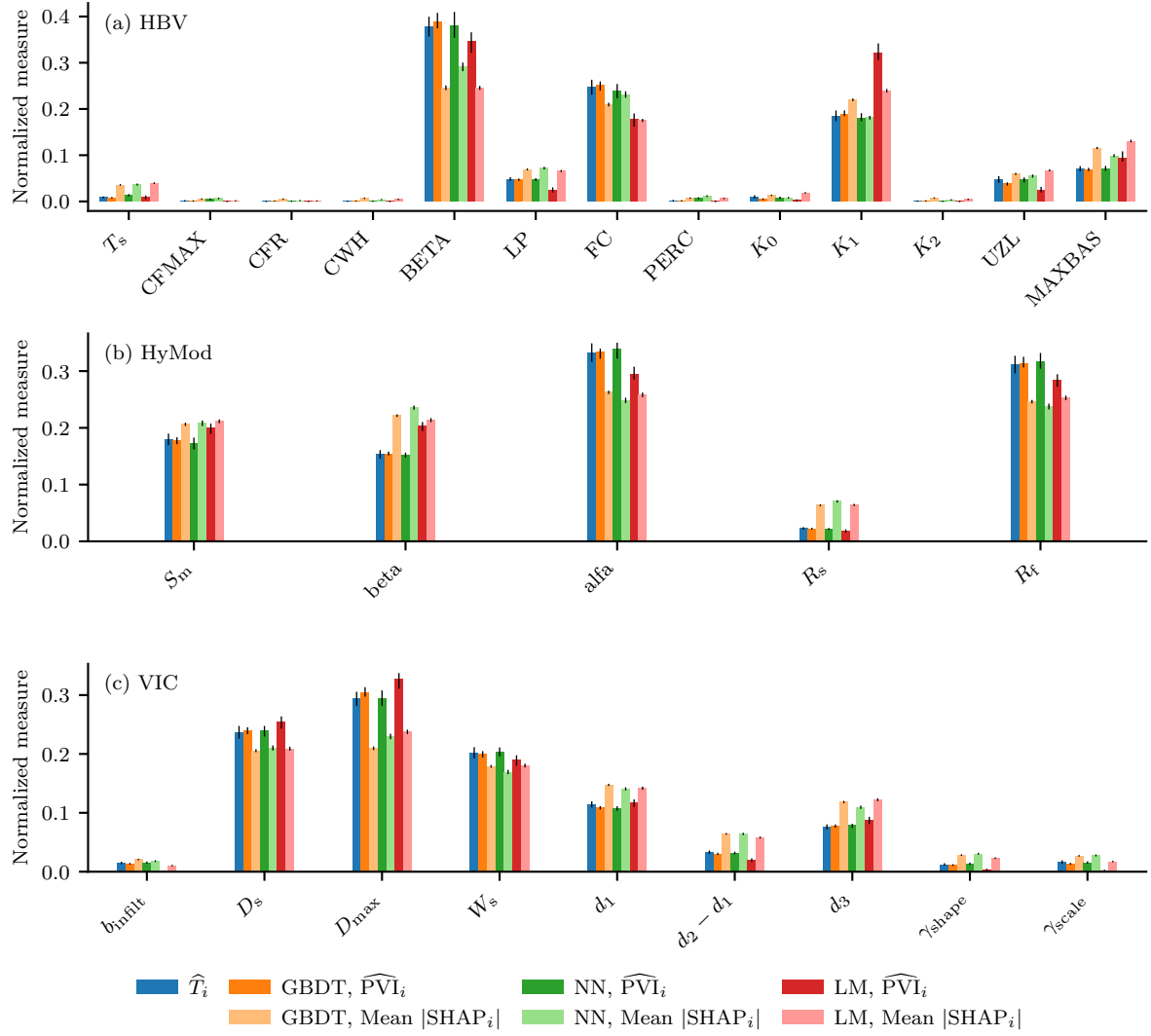


Figure S4. Parameter rankings for (a) HBV, (b) HyMod, and (c) VIC applied in the Schackendorf catchment. Bars show normalized $\hat{T}_i$ from variance-based GSA, normalized $\widehat{PVI}_i$ estimates, and normalized mean $|SHAP_i|$ with $\hat{T}_i$ obtained from GBDT, NN, and LM metamodels. Sensitivity measures were normalized by dividing each by the sum of all sensitivities, as in the convergence test. Thin black lines indicate 95% confidence intervals.

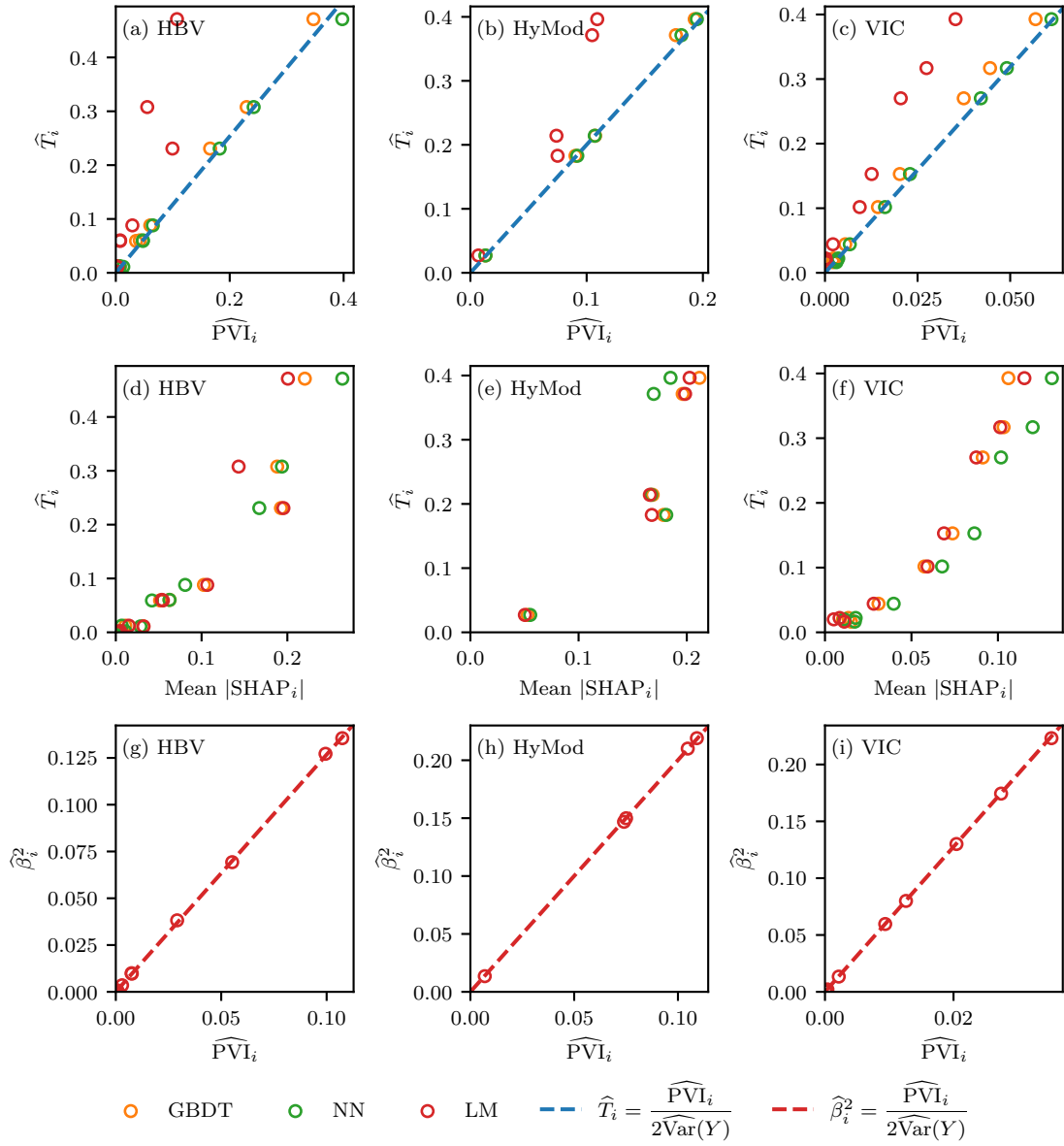


Figure S5. Comparison of different sensitivity estimates across hydrologic models for the Schackendorf catchment. Panels (a–c) compare $\widehat{PVI}_i$ with $\widehat{T}_i$ for HBV, HyMod, and VIC, respectively. Panels (d–f) compare $\widehat{PVI}_i$ with $\widehat{T}_i$, and panels (g–i) compare $\widehat{PVI}_i$ with $\widehat{\beta}_i^2$ for the LM metamodel. Dashed lines indicate the corresponding theoretical relationships. Colours denote the ML metamodel used.

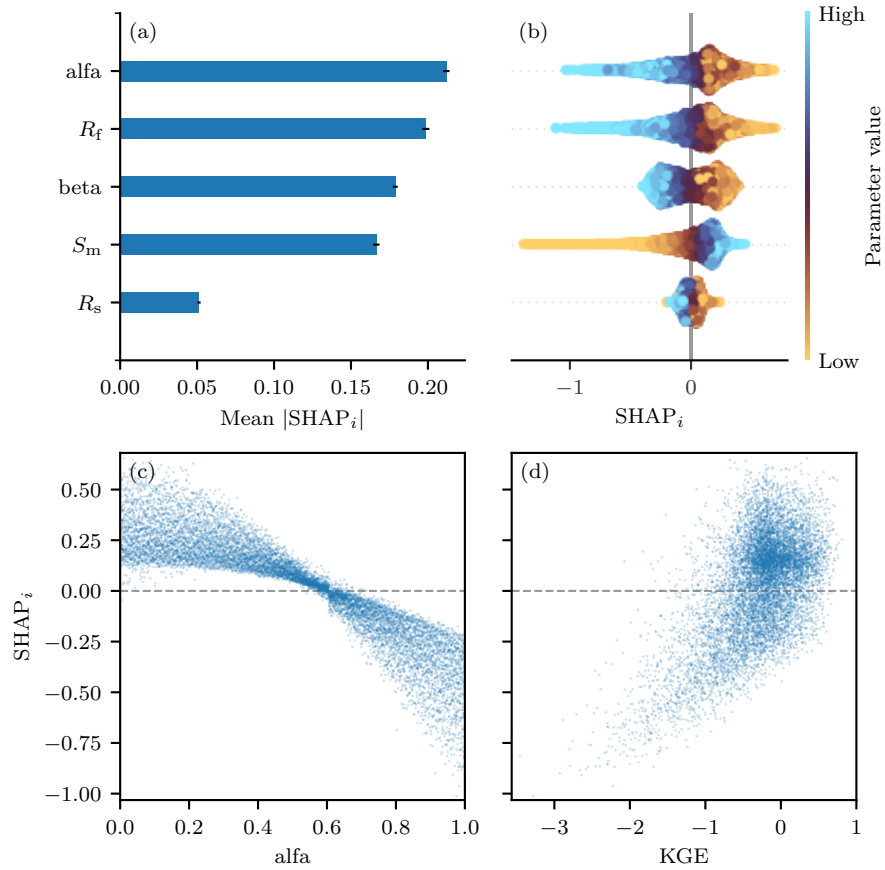


Figure S6. Detailed SHAP_i -based sensitivity analysis for HyMod applied in the Schackendorf catchment using the GBDT metamodel. Panel (a) shows the global parameter ranking based on mean $|\text{SHAP}_i|$, with thin black lines indicating 95% confidence intervals. Panel (b) displays the distribution of SHAP_i values across the parameter space, coloured by parameter value. Panel (c) illustrates the variation of SHAP_i for the most influential parameter, α , across its range. Panel (d) shows SHAP_i values for α as a function of KGE, highlighting how parameter contributions vary across model performance levels.

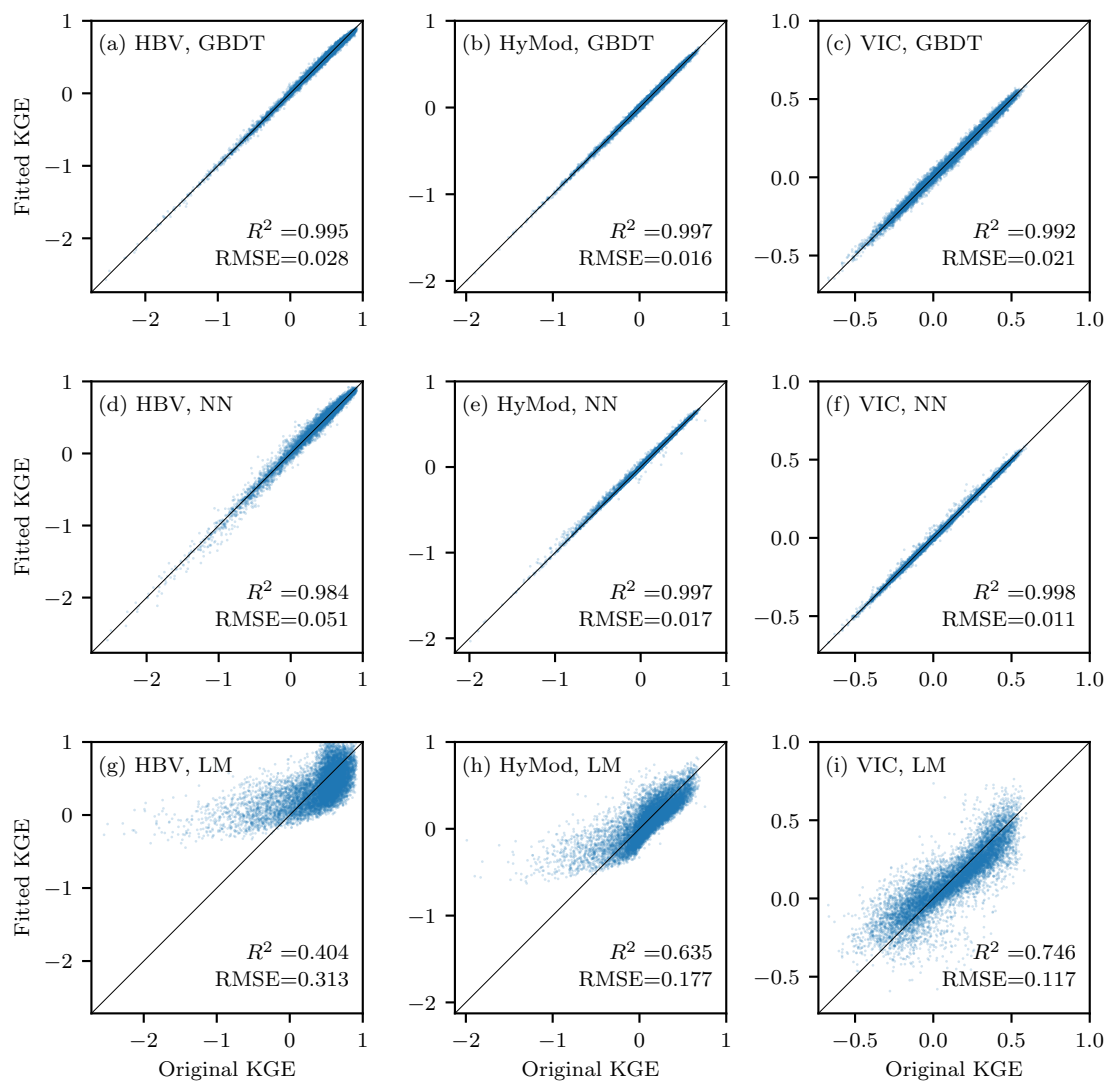


Figure S7. Same as Figure S1 but for the Augaard catchment.

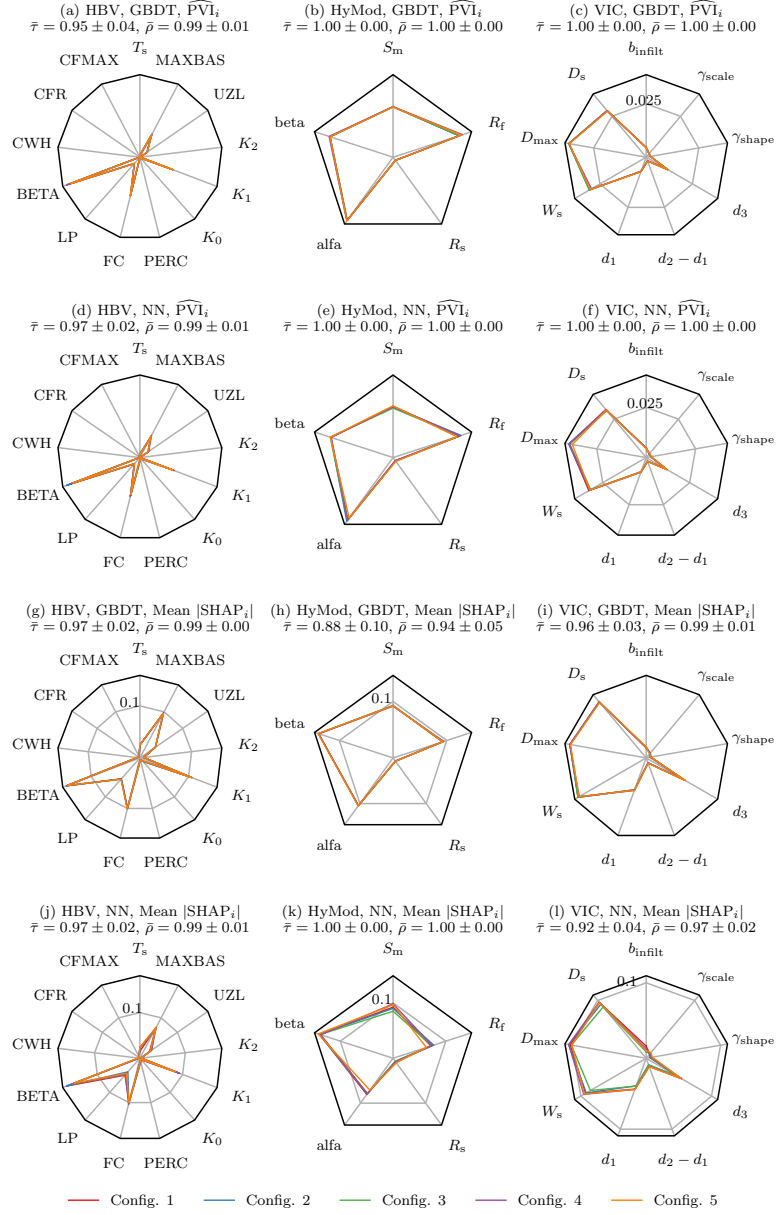


Figure S8. Same as Figure S2 but for the Augaard catchment.

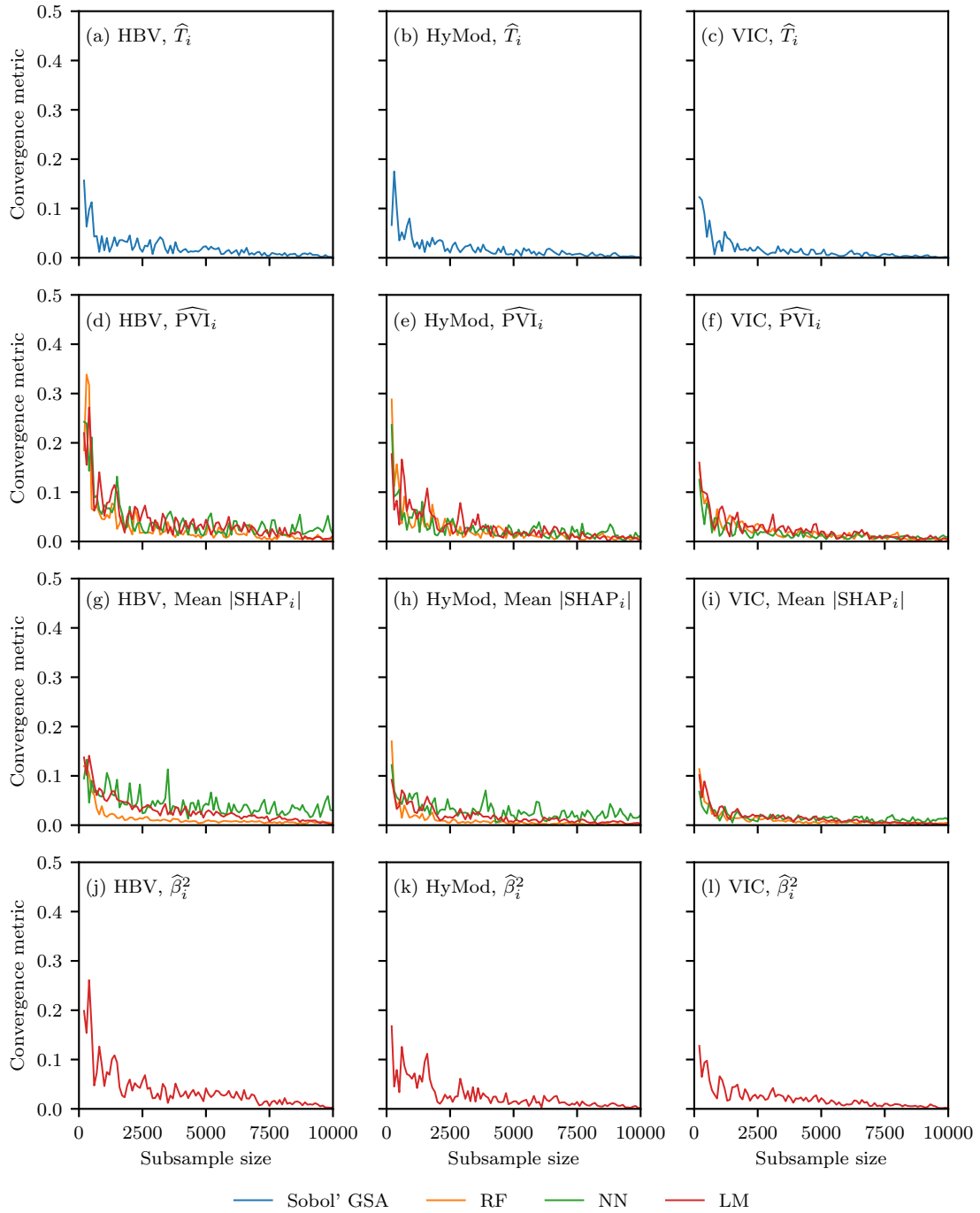


Figure S9. Same as Figure S3 but for the Augaard catchment.

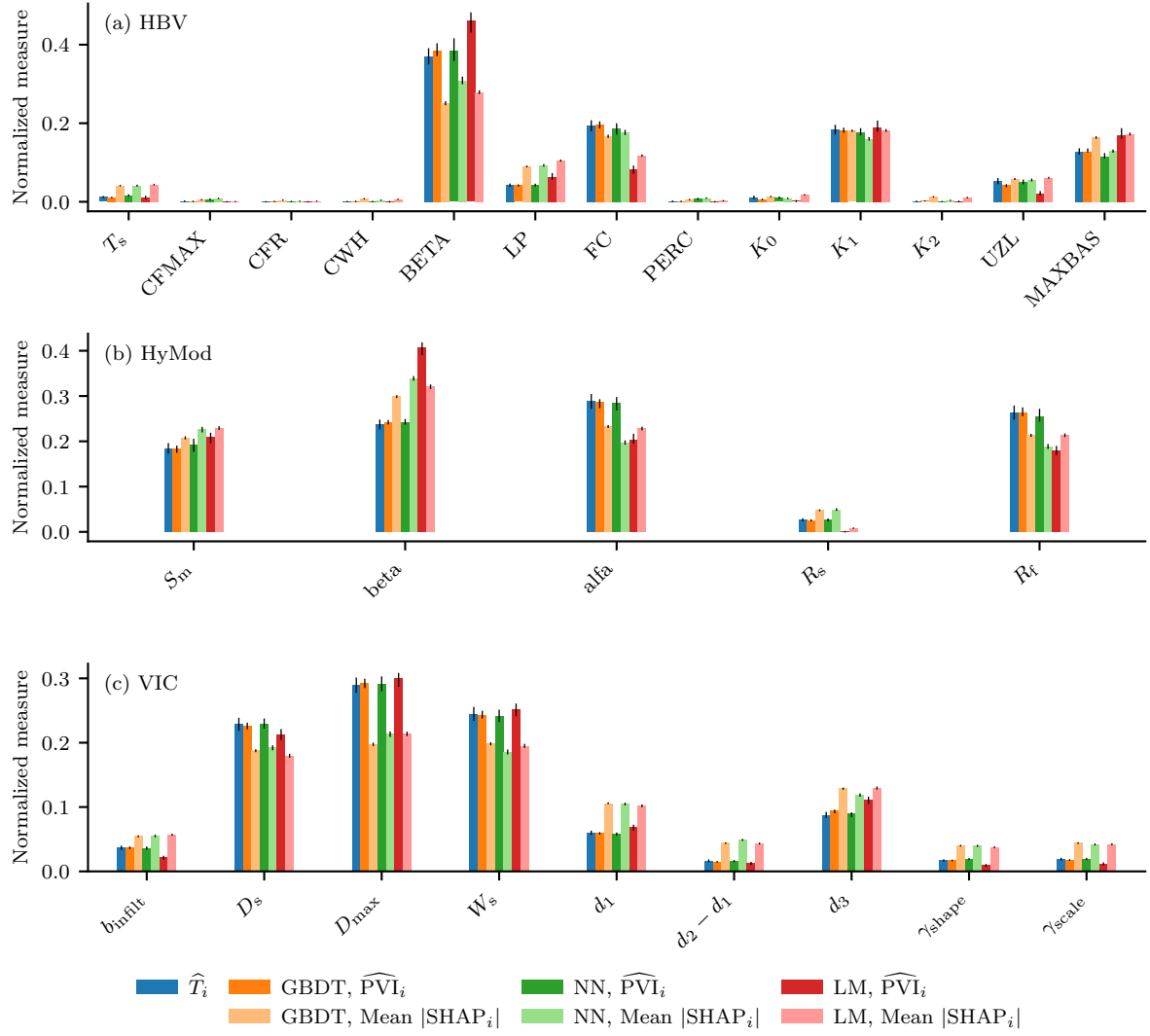


Figure S10. Same as Figure S4 but for the Augaard catchment.

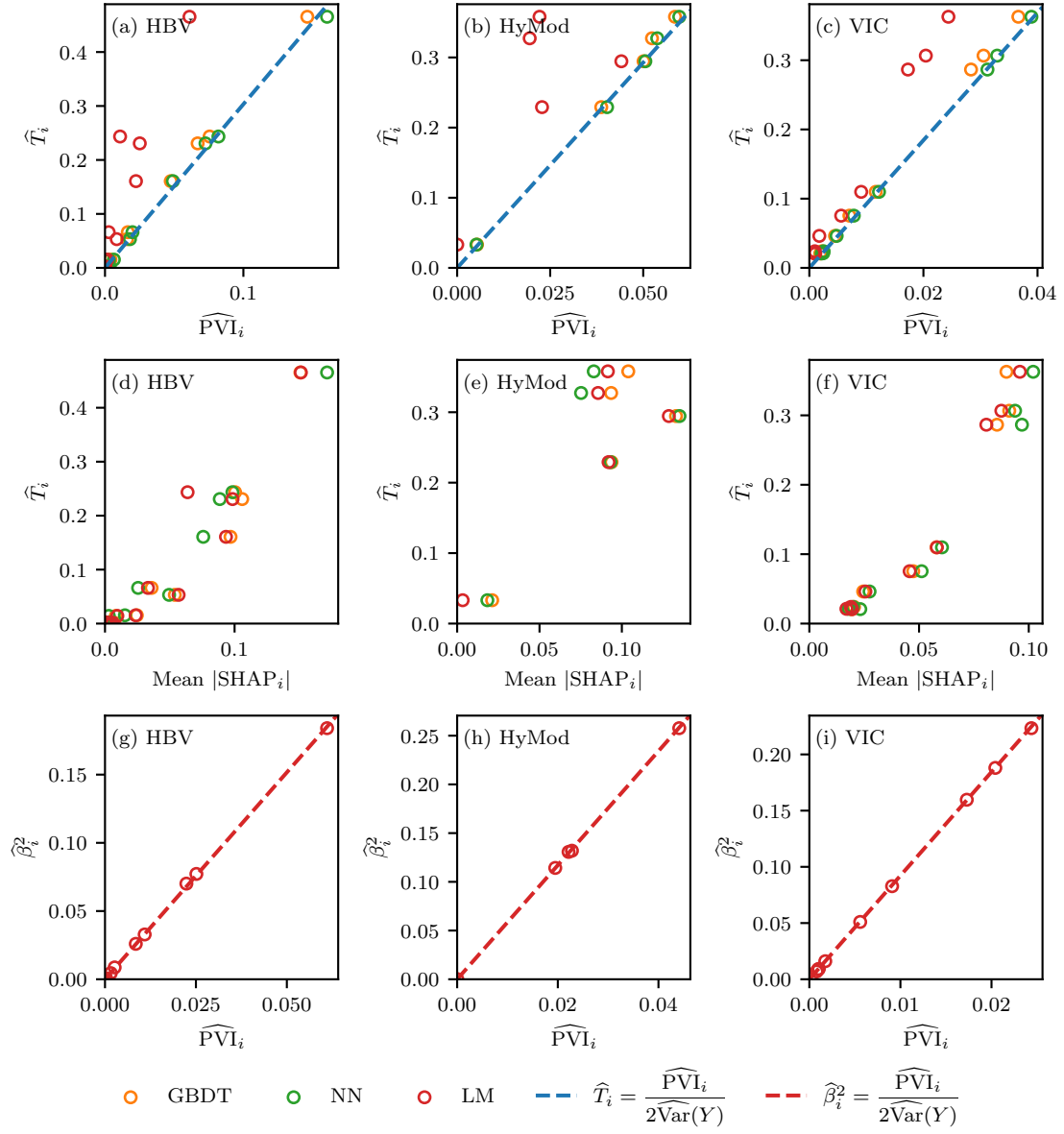


Figure S11. Same as Figure S5 but for the Augaard catchment.

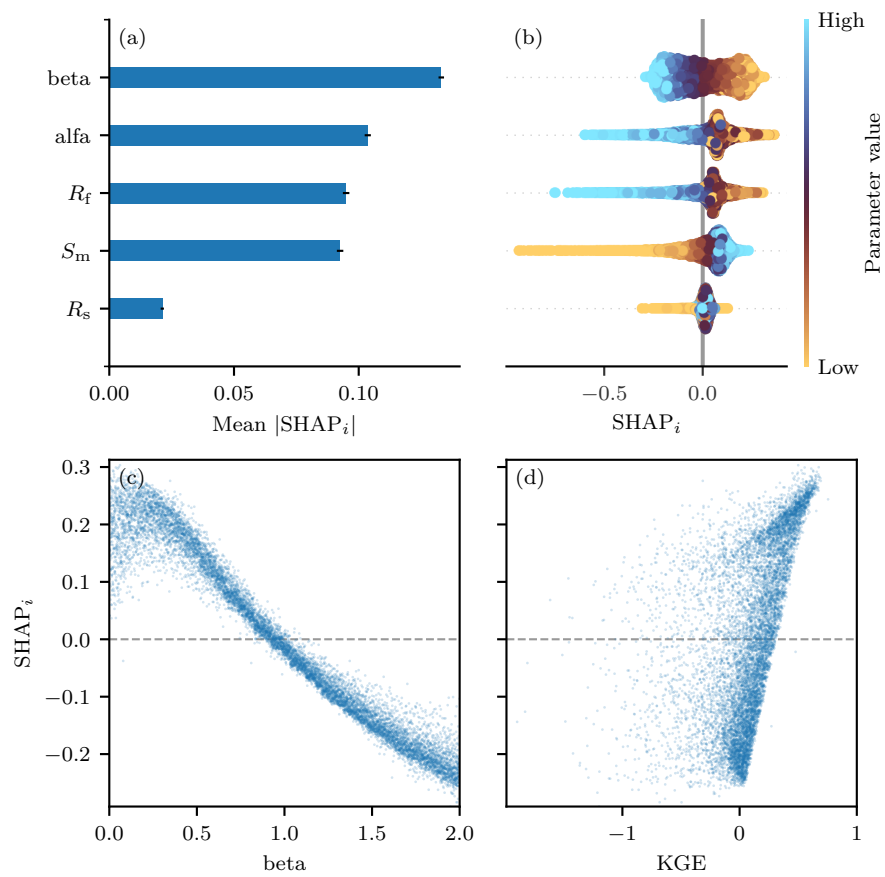


Figure S12. Same as Figure S6 but for the Augaard catchment. In this case, the most influential parameter is β .

References

- 15 Kollat, J. B., Reed, P. M., and Wagener, T.: When are multiobjective calibration trade-offs in hydrologic models meaningful?, *Water Resources Research*, 48, 1–19, <https://doi.org/10.1029/2011WR011534>, 2012.
- Mai, J.: Ten strategies towards successful calibration of environmental models, *Journal of Hydrology*, 620, 129–144, <https://doi.org/10.1016/j.jhydrol.2023.129414>, 2023.
- Schaperow, J. R., Li, D., Margulis, S. A., and Lettenmaier, D. P.: A near-global, high resolution land surface parameter dataset for the variable infiltration capacity model, *Scientific Data*, 8, 1–14, <https://doi.org/10.1038/s41597-021-00999-4>, 2021.
- 20 Yeste, P., Melsen, L. A., García-Valdecasas Ojeda, M., Gámiz-Fortis, S. R., Castro-Díez, Y., and Esteban-Parra, M. J.: A Pareto-Based Sensitivity Analysis and Multiobjective Calibration Approach for Integrating Streamflow and Evaporation Data, *Water Resources Research*, 59, e2022WR033235, <https://doi.org/10.1029/2022WR033235>, 2023.
- Yeste, P., García-Valdecasas Ojeda, M., Gámiz-Fortis, S. R., Castro-Díez, Y., Bronstert, A., and Esteban-Parra, M. J.: A large-sample modelling approach towards integrating streamflow and evaporation data for the Spanish catchments, *Hydrology and Earth System Sciences*, 28, 5331–5352, <https://doi.org/10.5194/hess-28-5331-2024>, 2024.
- 25