Supplement of Hydrol. Earth Syst. Sci., 29, 1183–1199, 2025
https://doi.org/10.5194/hess-29-1183-2025-supplement

Hydrology and Earth System Sciences
Open Access

Supplement of

Learning from a large-scale calibration effort of multiple lake temperature models

Johannes Feldbauer et al.

Correspondence to: Johannes Feldbauer (johannes.feldbauer@tu-dresden.de) and Jorrit P. Mesman (jorrit.mesman@ebc.uu.se)

S1 Additional material

S1.1 Additional Plots

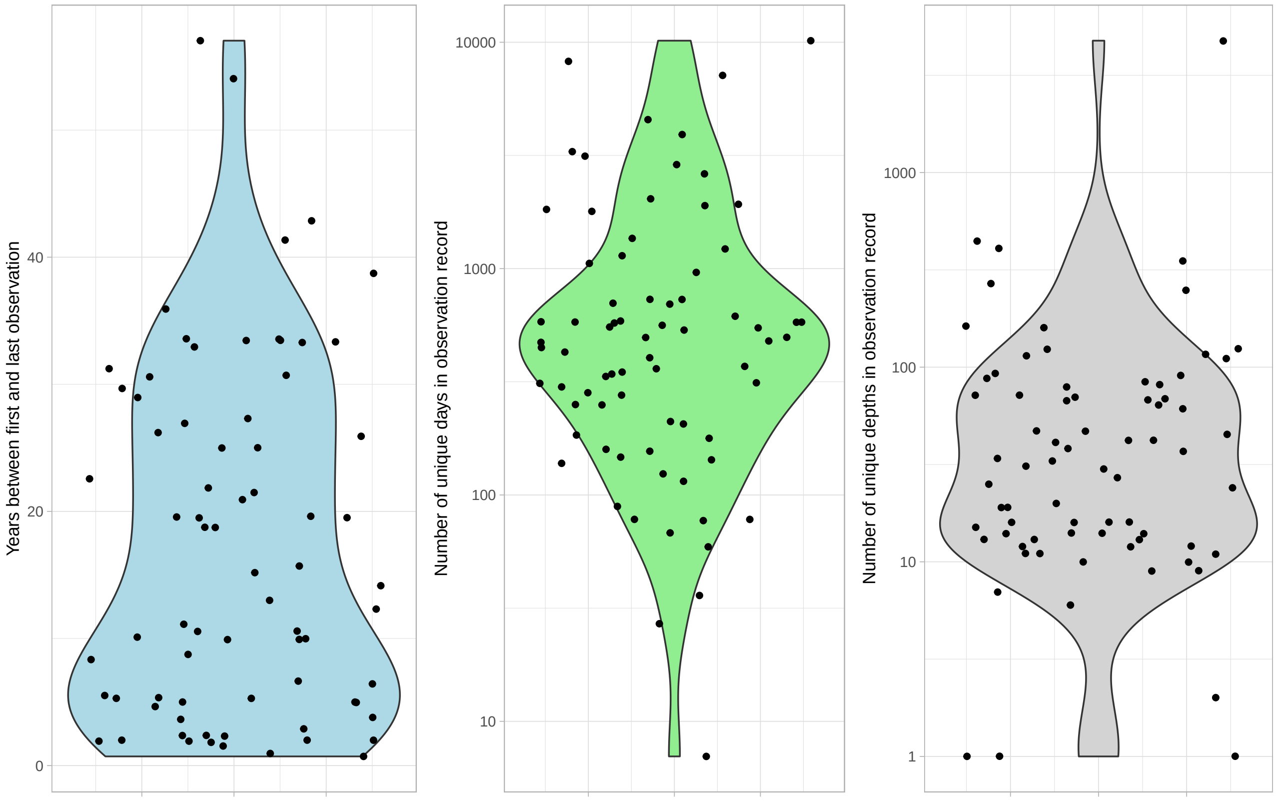

Figure S1: Violin plots showing the distribution of duration of observed data (left, calculated as years between first and last observation), Number of unique days in observation (middle), and number of unique depths in observations (right). The very high number of unique depths in observation records are caused by the fact that some of the observational data were from CTD profiles, which we aggregated to a vertical resolution of 0.1 m.

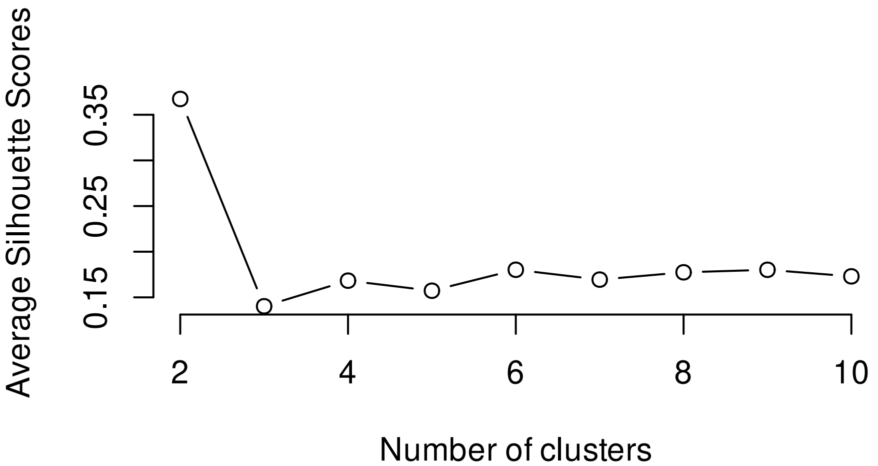

Figure S2: Silhouette plot of the cluster analysis. A larger silhouette score indicates that the clusters are well distinguishable.

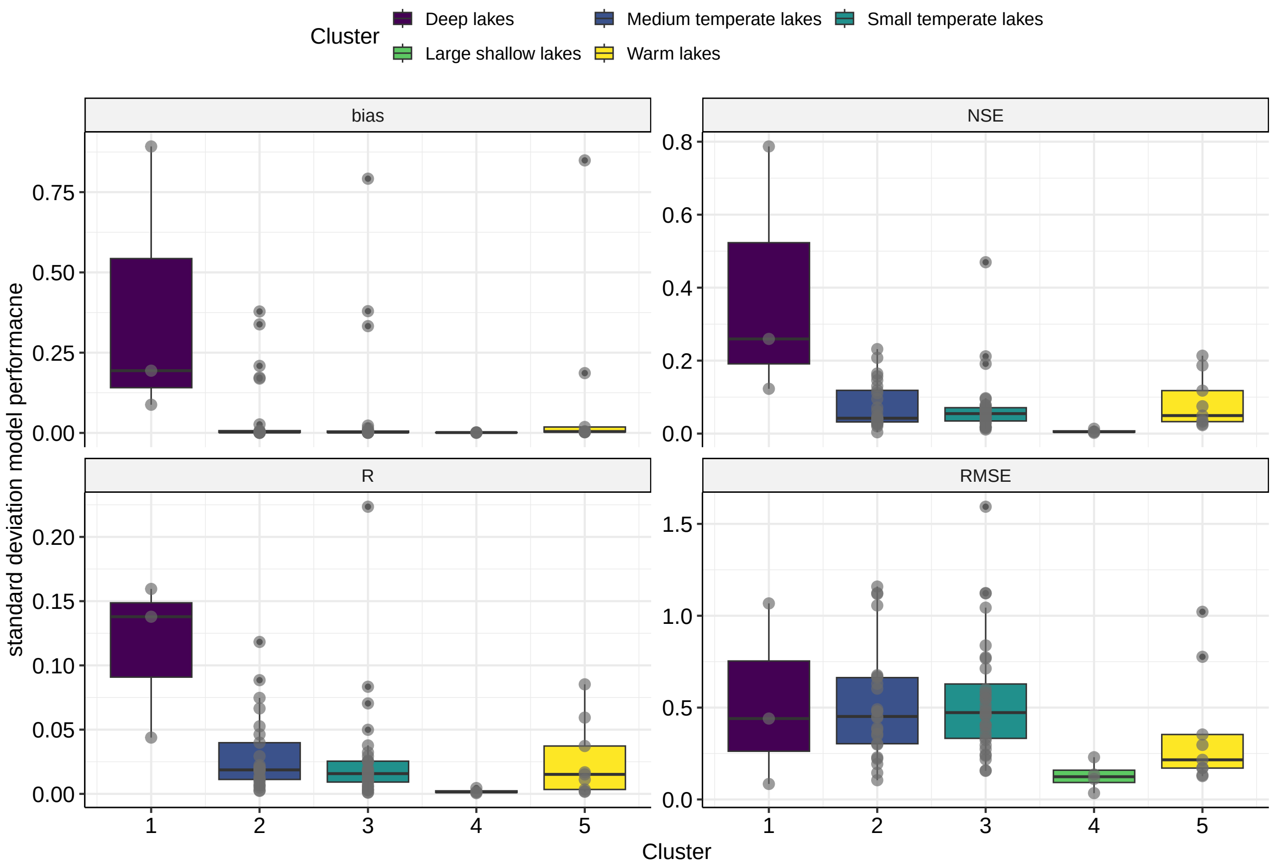

Figure S3: Boxplots of absolute difference between best and worst performing model per lake for the different metrics and lake clusters. The grey points indicate values for individual lakes.

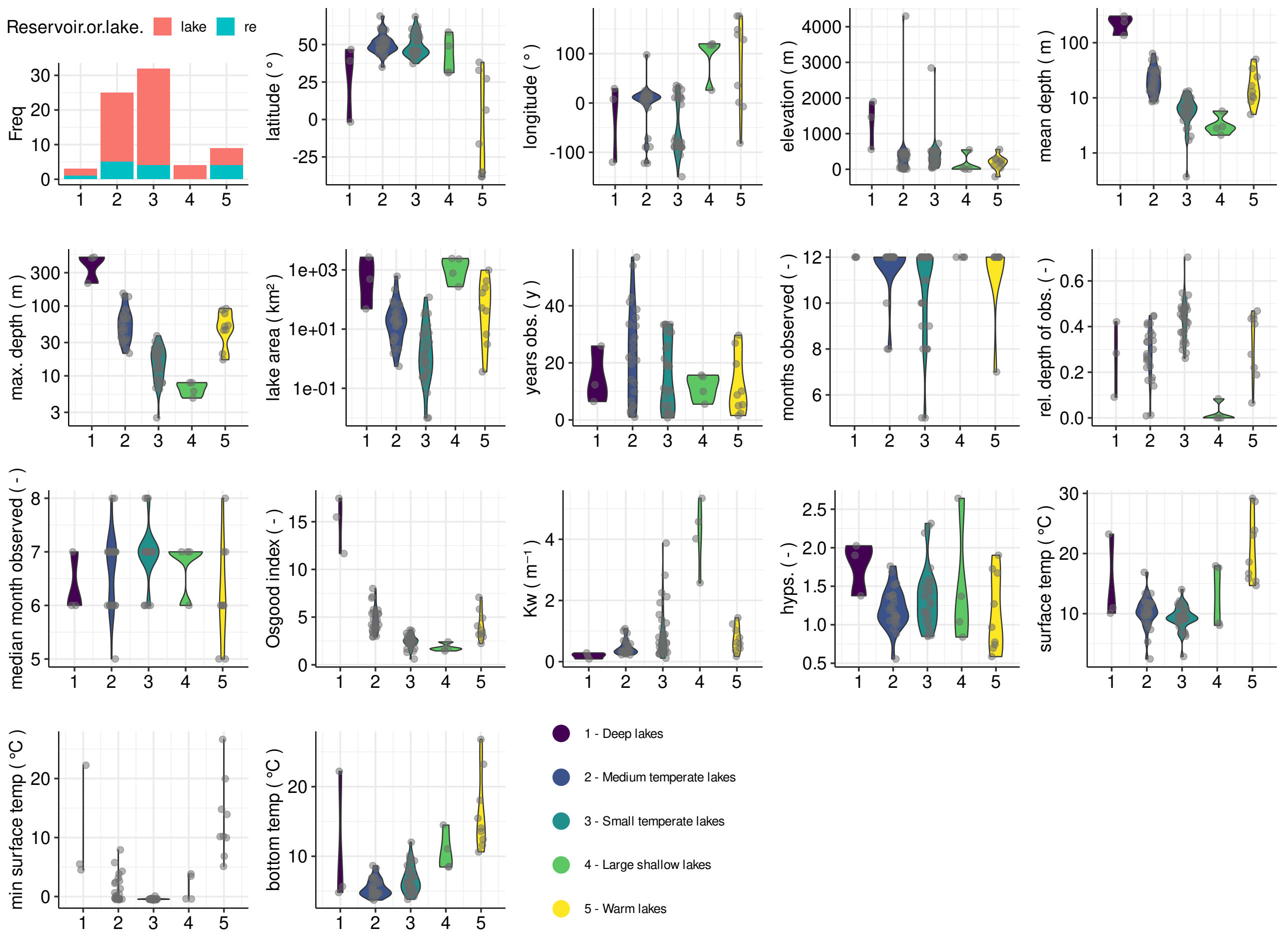

Figure S4: Distribution of lake characteristics for the clusters derived by K-means clustering. The grey points indicate values for individual lakes.

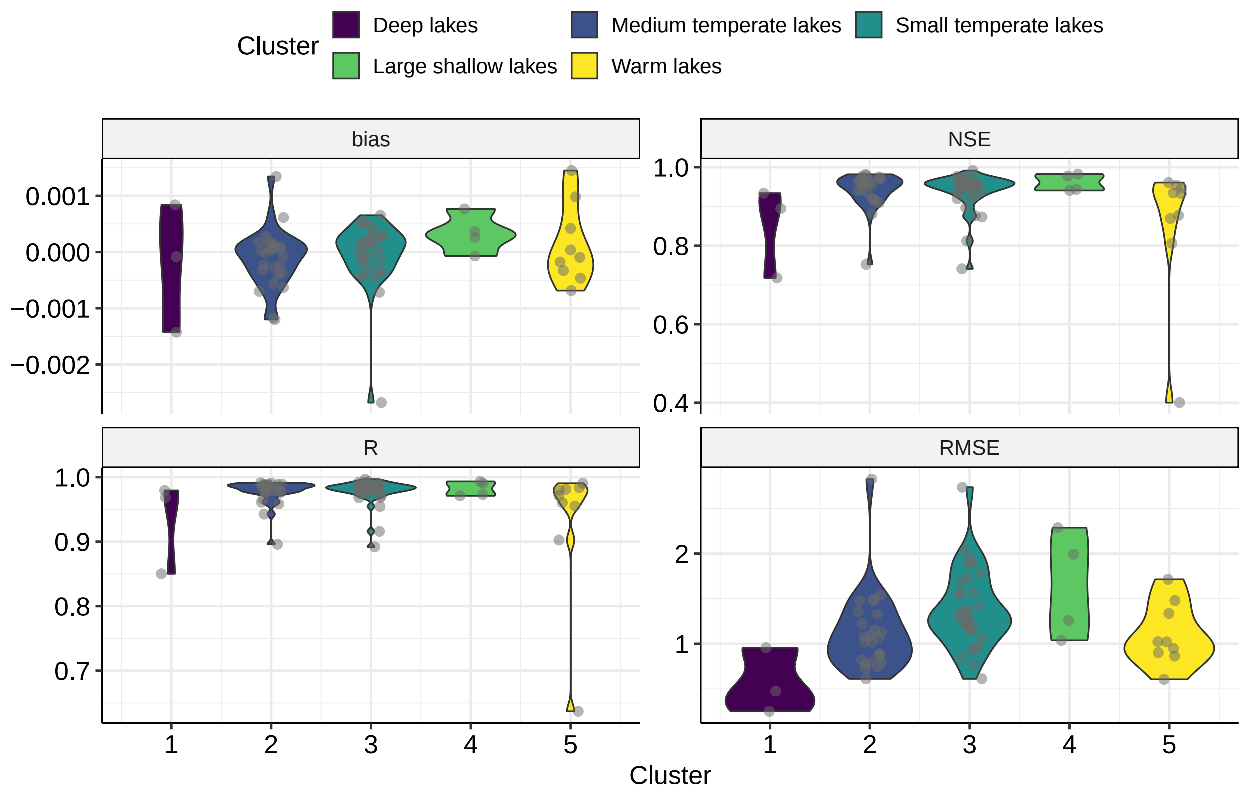

Figure S5: Violin plots of model performance of the best-performing model in each lake, split by lake clusters and performance metric. The grey points indicate values for individual lakes.

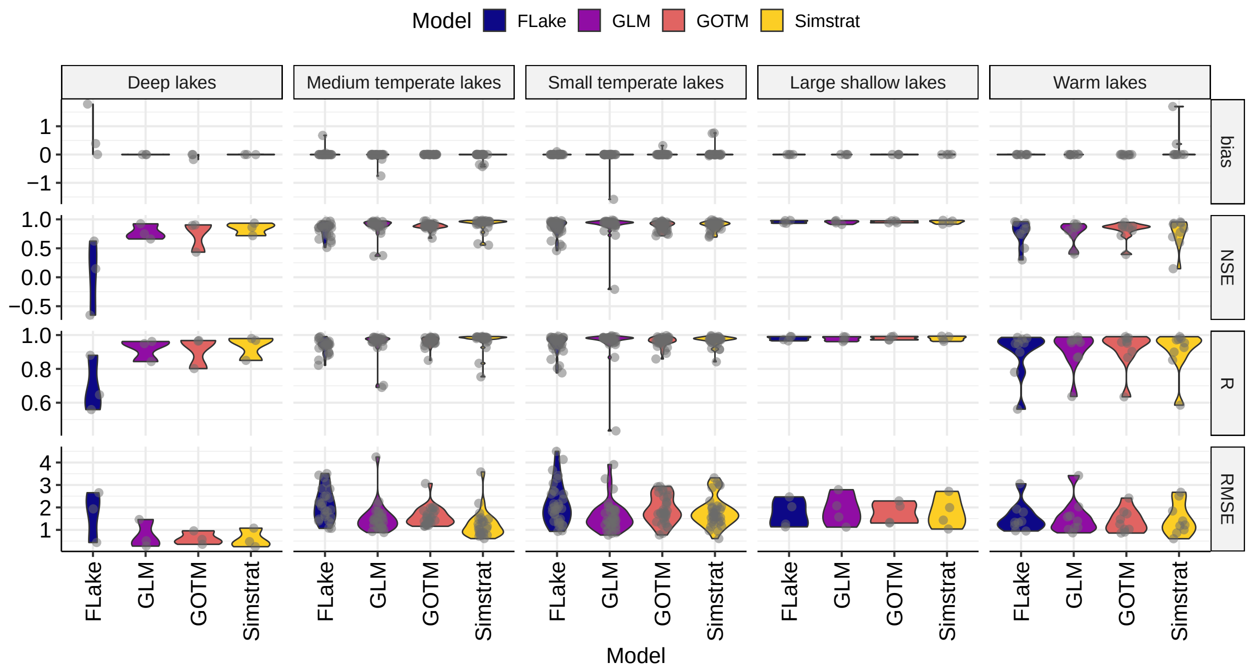

Figure S6: Violin plots of the performance metrics for best performing parameter sets per lake for each model, lake clusters, and performance metirc.

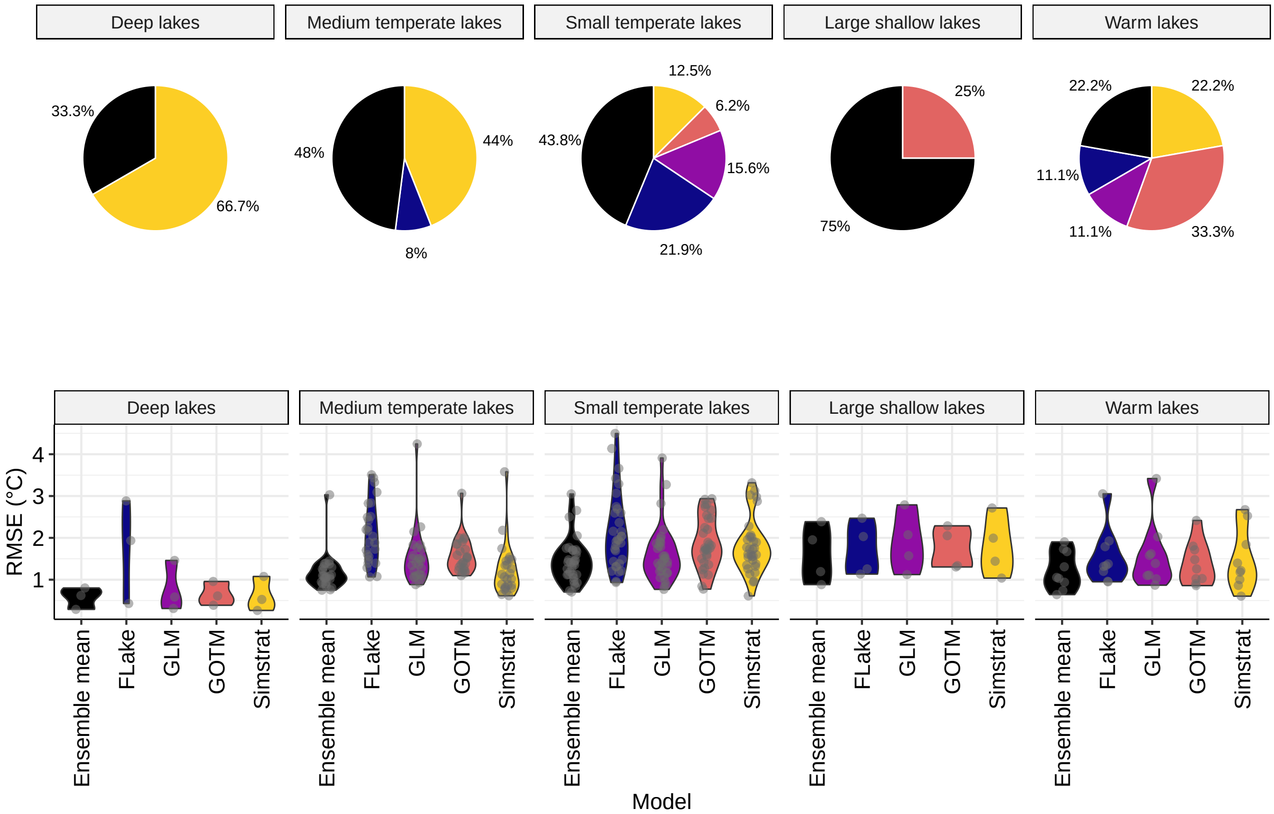

Figure S7: Percentage of lakes where each model performs best in terms of root mean squared error (RMSE), split to the lake clusters and including the ensemble mean as its own predictor (top, color refers to model or ensemble mean). Distribution of best performing parameter set (RMSE) per lake model or ensemble mean and lake clusters (below).

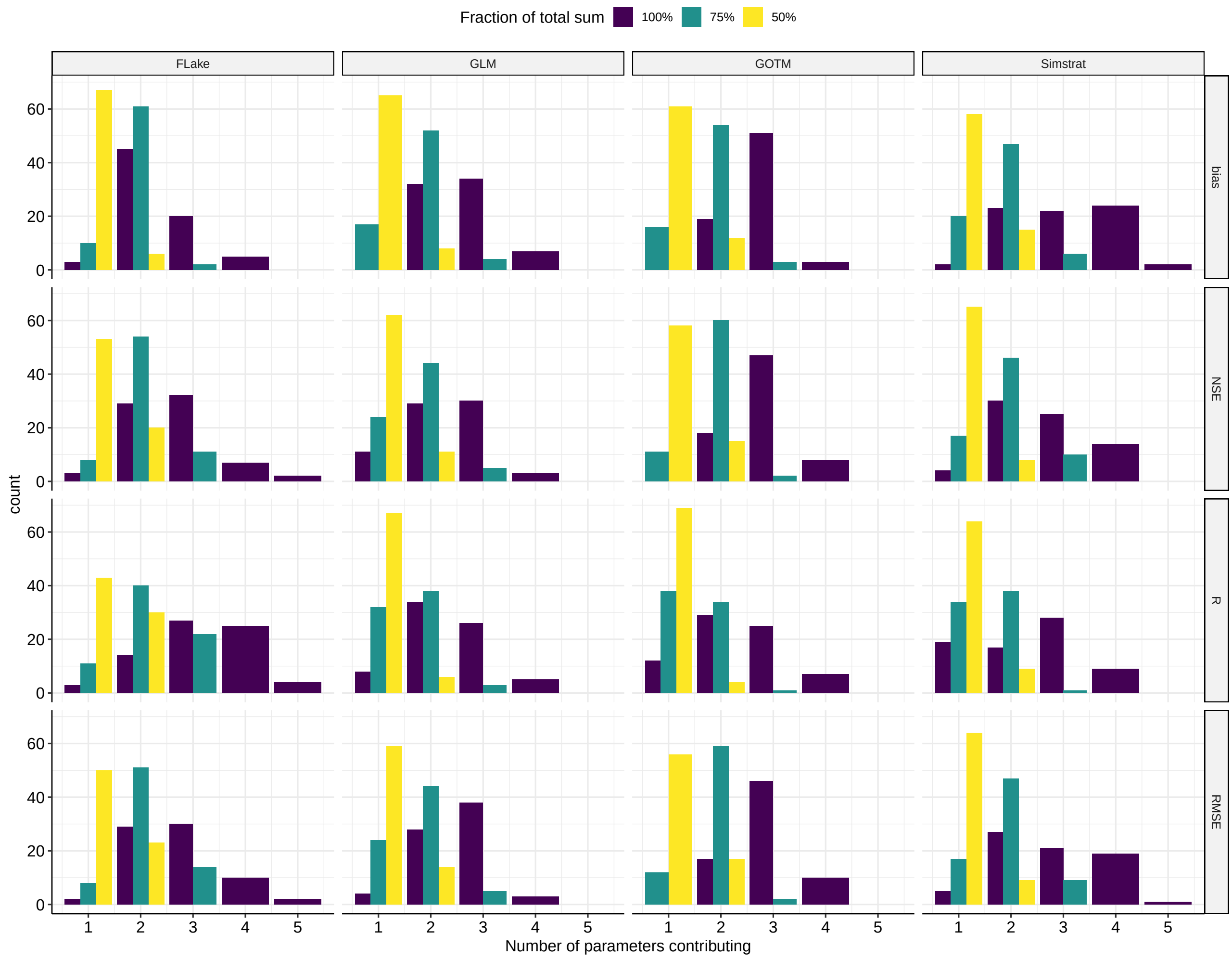

Figure S8: Distribution of the number of parameters that contribute to different fractions of the total sum of sensitivity measure value for each model and performance metric. For example, one contributing parameter for a fraction of 50% means that only one parameter contributes with at least 50% to the sum of all estimated sensitivity measures.

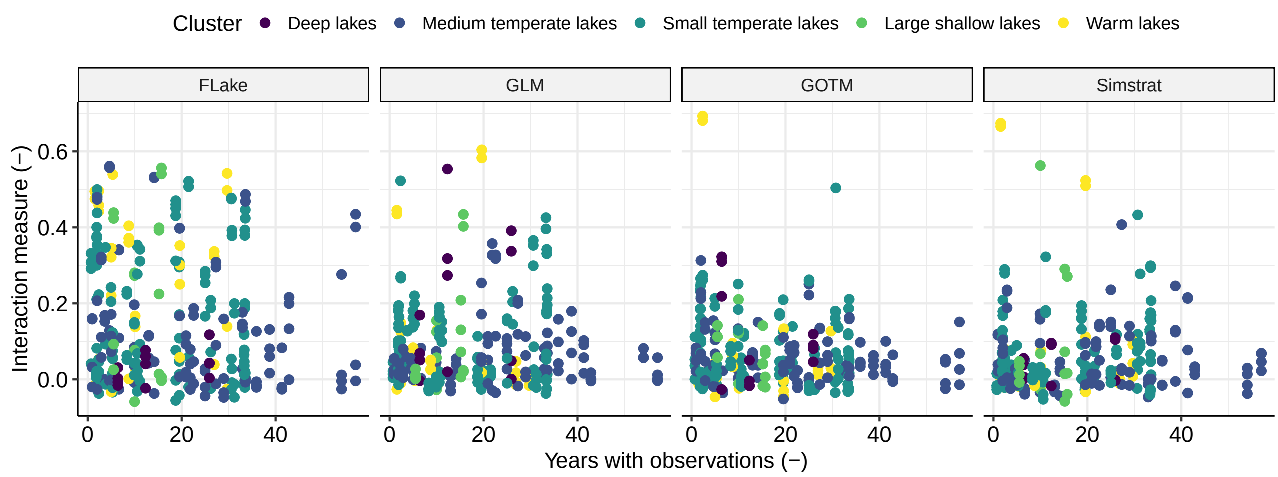

Figure S9: Interaction measure plotted against number of years with observed water temperature for the four models and lake clusters.

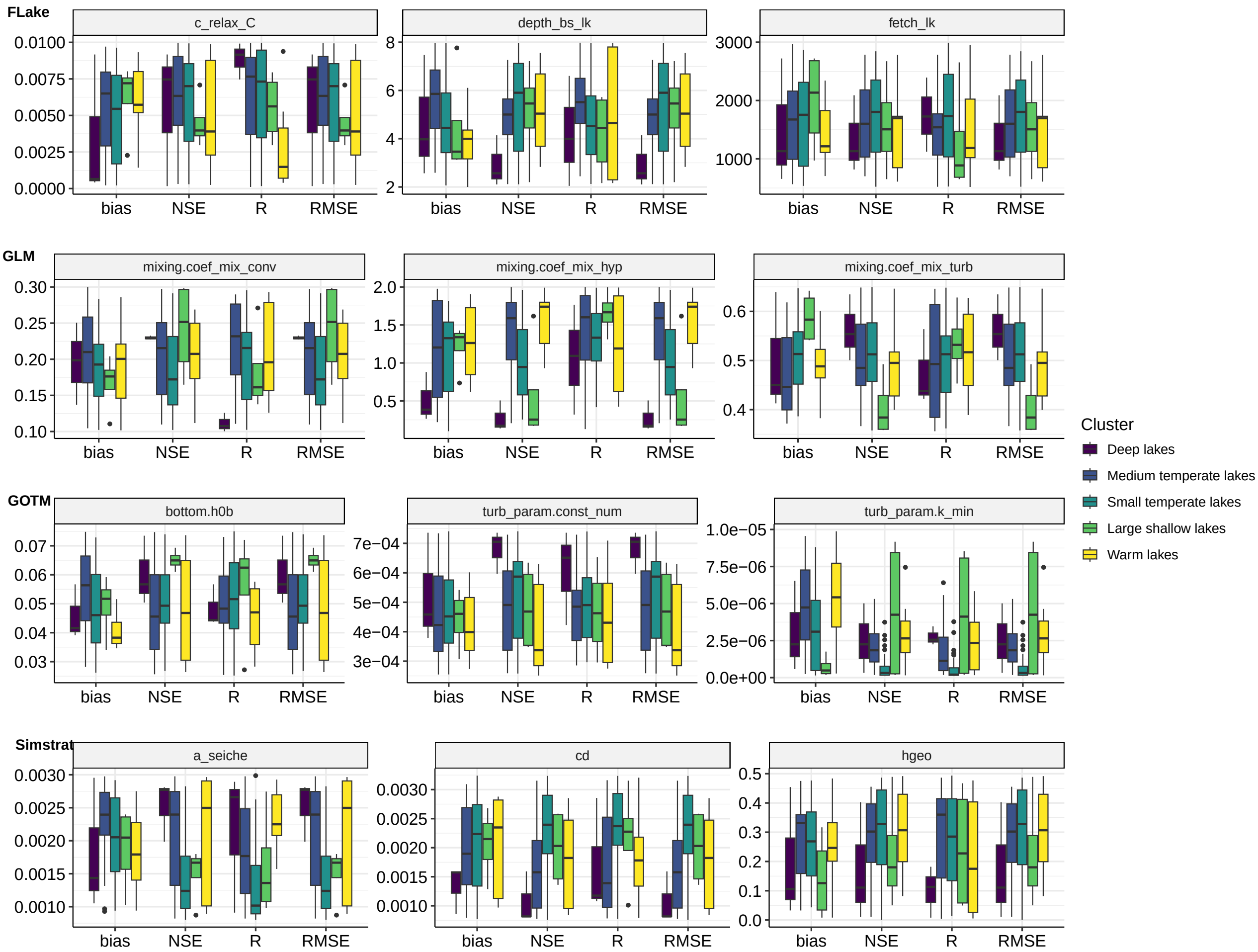

Figure S10: Boxplot of model specific parameters of the best performing parameter set for the different performance measures and lake clusters.

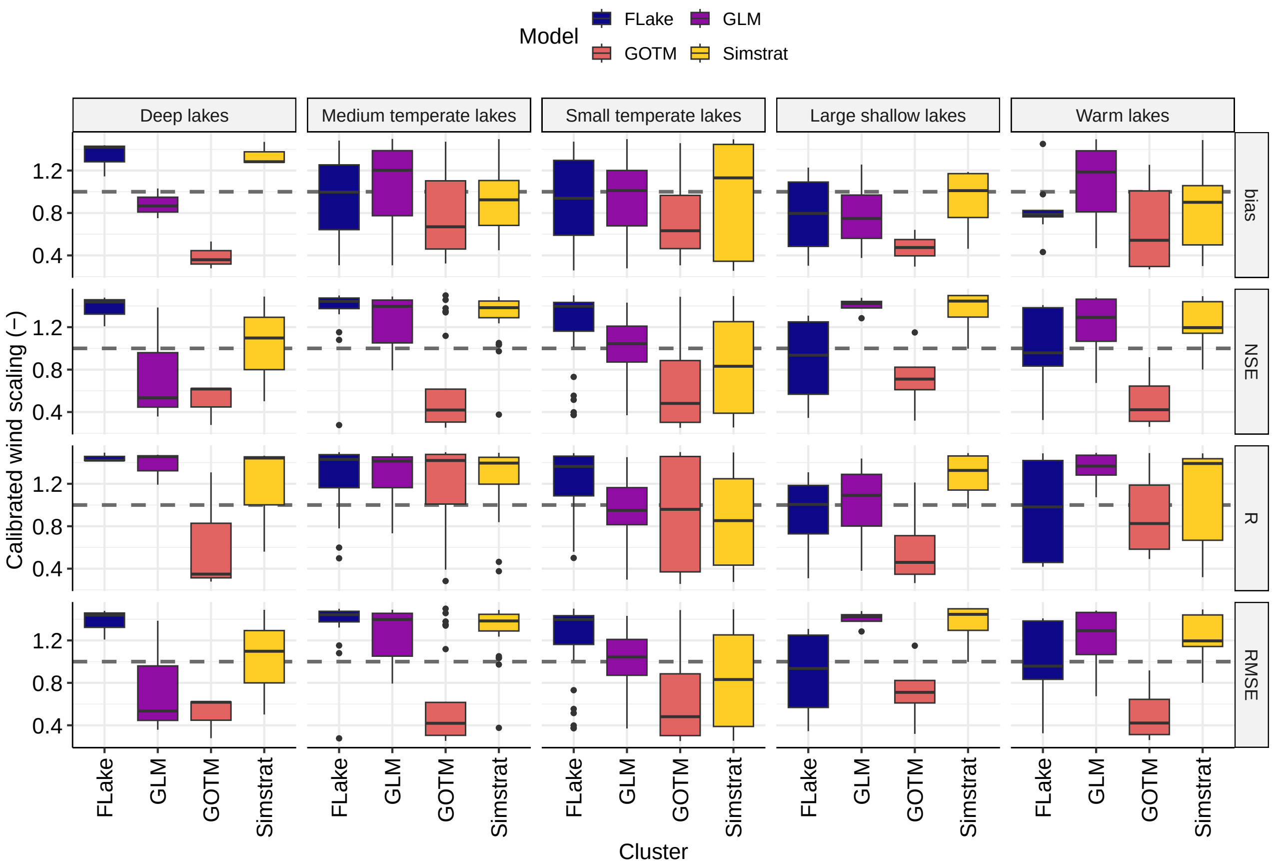

Figure S11: Boxplot of the calibrated wind speed scaling factors for the different models, clusters and based on which performance metric was choosen to optimize.

S1.2 Additional Tables

Table S1. Overview of the simulated lakes and their characteristics.

Lake Name	Country	Latitude (°)	Longitude (°)	Elevation (m)	Mean depth (m)	Max depth (m)	Lake area (km^2)	Secchi disk depth (m)	Light extinction (m^{-1})	Lake cluster
Allequash Lake	USA	46.04	-89.62	494	2.9	8	1.64	3.2		Small temperate
Alqueva Reservoir	Portugal	38.2	-7.49	152	16.6	92	250	3.15	0.69	Warm
Lake Annie	USA	27.21	-81.35	33.7	8.57	20.7	0.36			Warm
Lake Arendsee	Germany	52.89	11.47	23	28.8	49	5.11	2.89		Medium temperate
Lake Argyle	Australia	-16.31	128.68	100	10.1	51	980		0.89	Warm
Lake Biel	Switzerland	47.08	7.16	429	30	74	39.3		0.51	Medium temperate
Big Muskellunge Lake	USA	46.02	-89.61	500	7.5	21.3	3.63	6.6		Small temperate
Black Oak Lake	USA	46.16	-89.32	521.51	10.36	25.91	2.28	3.2		Small temperate
Lake Bosumtwi	Ghana	6.3	1.25	210	34	78	52	1.2		Warm
Lake Bryrup	Denmark	56.02	9.52	66	4.6	9	0.37	2.1		Small temperate
Lake Burley Griffin	Australia	-35.3	149.07	556	5	17	6.64	1.4		Warm
Lake Chao	China	31.53	117.53	4.5	3	7.98	780	0.38		Large shallow
Crystal Lake	USA	46	-89.61	501	10.4	20.4	0.38	7.5		Small temperate
Crystal Bog	USA	46.01	-89.61	501.5	1.7	2.5	0.01	1.5		Small temperate
Delavan Lake	USA	42.61	-88.6	282.55	7.61	16.46	6.96	3.53		Small temperate
Dickie Lake	Canada	45.15	-79.09	341	5	12	0.94	3.08		Small temperate
Eagle Lake	Canada	44.68	-76.7	190	10.1	31.1	6.65	4.62		Small temperate
Ekoln basin of Malaren	Sweden	59.75	17.62	0.7	11.5	50	20.18		1.09	Medium temperate
Lake Erken	Sweden	59.84	18.63	10	9	21	24	3.83		Medium temperate
Esthwaite Water	United Kingdom	54.37	-2.99	65	6.9	16	0.96		0.82	Small temperate
Falling Creek Reservoir	USA	37.31	-79.84	507	4	9.3	119		0.87	Small temperate
Lake Feeagh	Ireland	53.9	-9.5	15	14.5	44	3.9	1.74	0.98	Medium temperate
Fish Lake	USA	43.29	-89.65	265	6.6	18.9	0.8	2.4		Small temperate
Great Pond	USA	44.53	-69.89	81	6.3	21	33.83	6.7		Small temperate
Green Lake	USA	43.81	-89	243	33.55	72	29.48	4.8		Medium temperate
Harp Lake	Canada	45.38	-79.13	327	13.32	37.5	0.71	4.08		Small temperate
Hassel Predam	Germany	51.71	10.83	504	5	14	0.26	2.15		Small temperate
Lake Hulun	China	49	117.39	545.6	5.75	8	2339	0.35		Large shallow
Kilpisjarvi	Finland	69.03	20.77	473	20	57	37.3		0.3	Medium temperate
Lake Kinneret	Israel	32.82	35.58	-210	24	45	168	2.95	0.51	Warm
Lake Kivu	Rwanda/DR Congo	-1.73	29.24	1463	240	485	2700	5.21	0.27	Deep
Klicava Reservoir	Czechia	50.07	13.93	294.6	13.1	35	0.55	5.5		Medium temperate
Lake Kuivajarvi	Finland	60.47	23.51	130	6.3	13.2	0.62		0.6	Small temperate
Lake Langtjern	Norway	60.37	9.73	510	2	12	0.23	1.4	2.25	Small temperate
Laramie Lake	USA	40.62	-105.84	2843.8	0.37	6.4	0.14	0.6		Small temperate
Lower Lake Zurich	Switzerland	47.28	8.58	406	49	136	67		0.39	Medium temperate
Lake Mendota	USA	43.1	-89.41	258.5	12.8	25.3	39.61	3		Small temperate
Lake Monona	USA	43.06	-89.36	257	8.2	22.5	13.6	2.4		Small temperate
Mozhaysk reservoir	Russia	55.59	35.82	183	7	23	30.7	1		Small temperate
Mt Bold	Australia	-35.12	138.71	242.9	13	45.4	3.08	1.24	1.16	Warm
Lake Muggelsee	Germany	52.43	13.65	32.3	4.9	7.7	7.4	2	1.48	Small temperate
Lake Murten	Switzerland	46.93	7.08	429	22	45	22.8	3.49		Medium temperate
Lake Neuchatel	Switzerland	46.54	6.52	429	64	152	217		0.25	Medium temperate
Ngoring	China	34.9	97.7	4300	17.6	30.7	611		0.3	Medium temperate
Lake Nohipalo Mustjaerv	Estonia	57.93	27.34	61	3.9	8.9	0.22	0.46		Small temperate

Lake Nohipalo Valgejaerv	Estonia	57.94	27.35	62	6.2	12.5	0.07	4.52		Small temperate
Okauchee Lake	USA	43.13	-88.43	269	7.62	28.65	4.9	6.94		Small temperate
Lake Paaijarvi	Finland	61.07	25.13	102	15	85	13.44	2.2	1.15	Medium temperate
Rappbode Reservoir	Germany	51.74	10.89	415	28.6	89	3.95	4.8	0.25	Medium temperate
Rappbode Predam	Germany	51.71	10.8	533	5.3	16	0.24	2.26		Small temperate
Rimov Reservoir	Czechia	48.85	14.49	471.48	16	44	2.11	2.9		Medium temperate
Lake Rotorua	New Zealand	-38.08	176.28	280	10.8	52.9	425	2.63	0.61	Warm
Lake Sammamish	USA	47.59	-122.1	9	17.7	32	19.8	5		Medium temperate
Sau Reservoir	Spain	41.97	2.4	425	29	65	5.8	2.57	0.84	Medium temperate
Lake Scharmutzelsee	Germany	52.25	14.05	38.3	8.8	29	12.1	2		Medium temperate
Sparkling Lake	USA	46.01	-89.7	495	10.9	20	0.64	6.2		Small temperate
Lake Stechlin	Germany	53.17	13.03	59.8	23.2	69.5	2.23	8.6	0.29	Medium temperate
Lake Sunapee	USA	43.23	-72.5	333	11.4	34	16.55	8.5		Medium temperate
Lake Tahoe	USA	39.09	-120.03	1897	304.8	501	490	19.9		Deep
Lake Taihu	China	31.24	120.17	3.3	2.1	4.8	2445	0.3		Large shallow
Lake Tarawera	New Zealand	-38.21	176.43	300	50	87.5	41.3	8.34	0.18	Warm
Lake Thun	Switzerland	46.7	7.7	558	136	212	48.3	8.05		Deep
Toolik Lake	USA	68.63	-149.6	720	7	25	1.49	4.6		Small temperate
Trout Lake	USA	46.03	-89.67	491.8	14.6	35.7	15.65	4.7		Medium temperate
Trout Bog	USA	46.04	-89.69	495	5.6	7.9	0.01	1.1		Small temperate
Two Sisters Lake	USA	45.77	-89.53	481	9.14	19.2	2.91	17.75		Small temperate
Lake Vendyurskoe	Russia	62.1	33.1	131	5.3	13.4	10.4	3.5	1.5	Small temperate
Lake Vortsjarv	Estonia	58.31	26.01	33	2.8	6	270	0.86	2.76	Large shallow
Lake Washington	USA	47.64	-122.27	5	33	65.2	87.6	5.3		Medium temperate
Windermere	United Kingdom	54.31	-2.95	39	21.3	64	14.76		0.46	Medium temperate
Lake Wingra	USA	43.05	-89.43	254	2.7	6.7	1.36	0.7		Small temperate
Zlutice Reservoir	Czechia	50.09	13.11	507.95	8.5	24	1.5	2.8		Medium temperate
Lake Zurich	Switzerland	47.28	8.6	406	51	136	68.2	5.04		Medium temperate

Table S2: Characteristics of the lakes used for cluster analysis.

Reservoir.or.lake.	either lake or reservoir	-
latitude.dec.deg	coordinates of the lake latitude	°C
longitude.dec.deg	coordinates of the lake longitude	°C
elevation.m	elevation of the lake	m
mean.depth.m	mean depth of the lake	m
max.depth.m	maximum depth of the lake	m
lake.area.sqkm	lake surface area	km^2
Duration	number of years with observations	y
months_meas	number of unique months in which observations were available	-
reldepth_median	median relative depth of the observations (0 means surface, 1 means bottom)	-
months_median	median month of the observations	-
kw	Average calibrated light extinction factor	m^{-1}
vd	Volume development (Håkanson, 1981)	-
osgood	Osgood index (Osgood, 1988)	-
tsurf	average annual surface temperature for the period 1980 – 2000 from simulation	°C
min_tsurf	average annual minimum surface temperature for the period 1980 – 2000 from simulation	°C
tbot	average annual surface bottom for the period 1980 – 2000 from simulation	°C

Table S3: Median RMSE of the calibrated models and the percentage of calibrated fits lower than 2 °C

model	Median RMSE (°C)	Percentage under 2 °C
FLake	1.91	56.2
GLM	1.42	83.6
GOTM	1.58	76.7
Simstrat	1.40	80.8

S1.3 Additional simulations for GOTM and Simstrat

From the findings of the calibration runs and the discussion in the main text further questions came up and we decided to investigate some of the aspect in more detail by running an additional set of (restrained) calibrations for two of the models: GOTM and Simstrat. As these additional simulations confirmed the findings from the original calibration runs, but give further insight that might be valuable for some of the model users, we document the additional simulations here.

As we found and discuss in the main text the parameter `k_min` was the most sensitive model parameter for GOTM. We ran another round of calibrations with the same settings, except that we kept `k_min` at a constant value of 10^{-8}, which is the default value. The reasoning behind these additional simulations were to check whether or not the larger values of `k_min` are causing the different behavior of the wind speed and shortwave radiation scaling factors in GOTM (as discussed in the main text).

For Simstrat we ran an additional round of calibration with the same settings, except that we set `a_seiche` to 0. This disables the seiching module for Simstrat and we can therefore test if the better performance of Simstrat was caused by the inclusion of seiching as we speculated. In addition, disabling the seiching module allows us to better compare the impact of reducing `k_min` (which also partially covers seiching).

S1.3.1 GOTM and `k_min`

The range of `k_min` used in this study was $1.5 \cdot 10^{-7}$–$1 \cdot 10^{-5}$, which is rather large compared to the default value of $1 \cdot 10^{-8}$ in the GOTM manual. In the literature typical values of $1 \cdot 10^{-6}$ are reported for turbulent kinetic energy in the hypolimnion of lakes (e.g. Wüest and Lorke, 2003). When investigated in detail it can be seen that the calibrated `k_min` values (from the original calibration as described in the main text) are larger for deep, medium temperate, and warm lakes compared to small temperate lakes (Figure S12, subpanel A). The exceptions are the large shallow lakes that cover a wide range of `k_min` values. This could be explained by looking at the sensitivity of `k_min` for the different clusters, whereas for the large shallow lakes `k_min` has very low sensitivity values (see figure S12, subpanel B).

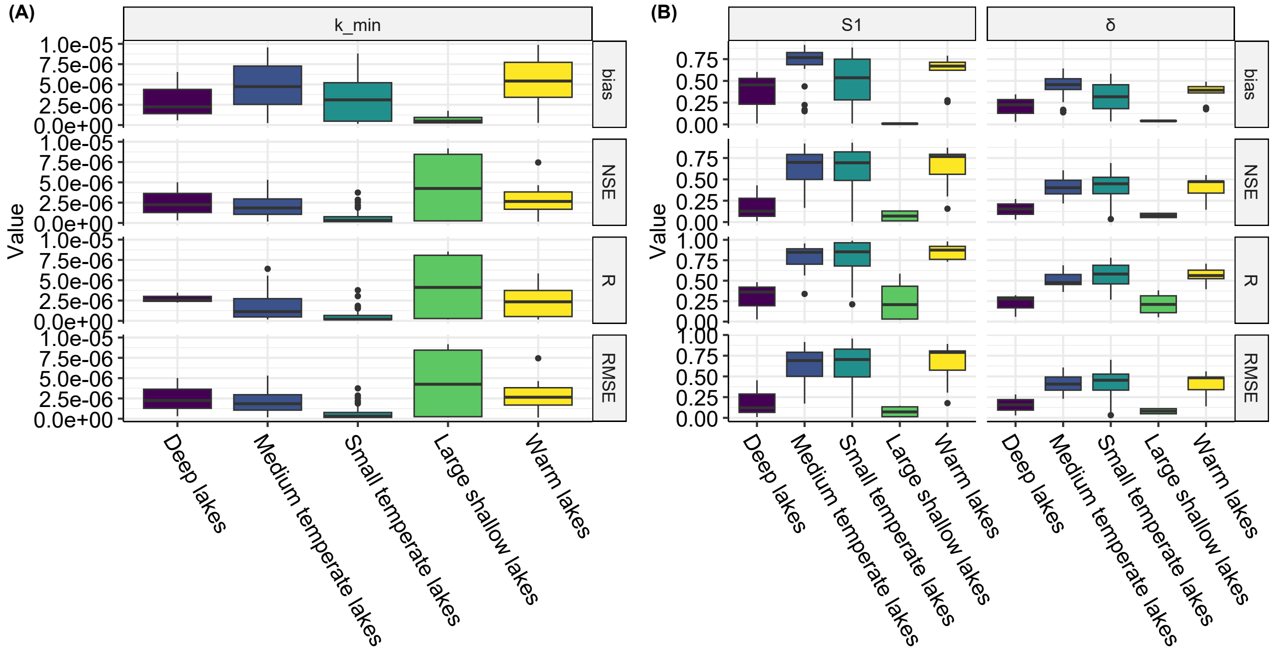

Figure S12: Boxplot of GOTMs `k_min` values for the best performing parameter set for the different performance metrics and lake clusters (A). And Boxplot of the two sensitivity metrics for GOTMs `k_min` along the different performance metrics and lake clusters (B).

S1.3.2 Additional calibration round

Under these additional calibration rounds (GOTM with `k_min` $= 1 \cdot 10^{-8}$ and Simstrat with `a_seiche` $= 0$) both of the models showed worse performance in most of the lakes, especially for the medium temperate and warm lakes. For large shallow lakes, very small differences were seen and for some of the small temperate lakes we even saw better performance for both Simstrat and GOTM (Figure S13).

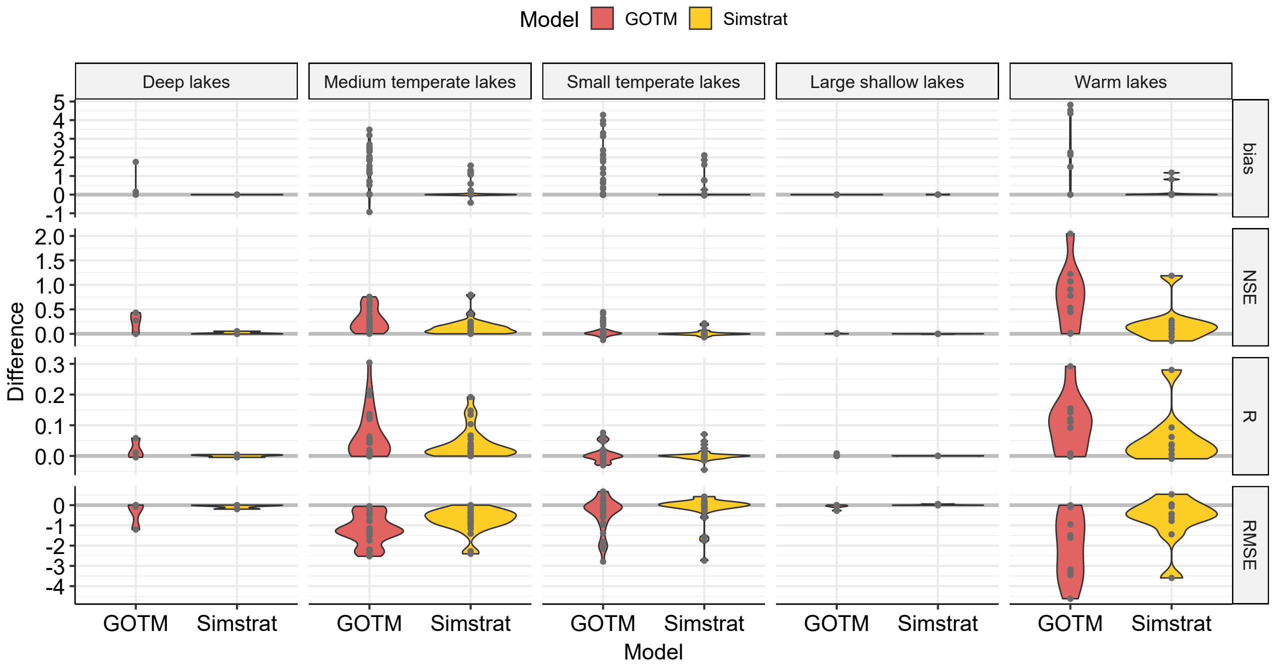

Figure S13: Difference between the performance metrics for the best performing parameter set for GOTM and Simstrat along the different performance metrics and lake clusters. The difference is calculated as original calibration minus new calibration. The new calibration rounds restricted the deep mixing by fixing GOTMs `k_min` to 10^{-8} and Simstrats `a_seiche` to 0.

However, with `k_min` kept at $1 \cdot 10^{-8}$ the calibrated wind speed scaling from GOTM (GOTM_r in figure S14) increased but, was still often smaller than in the other models.

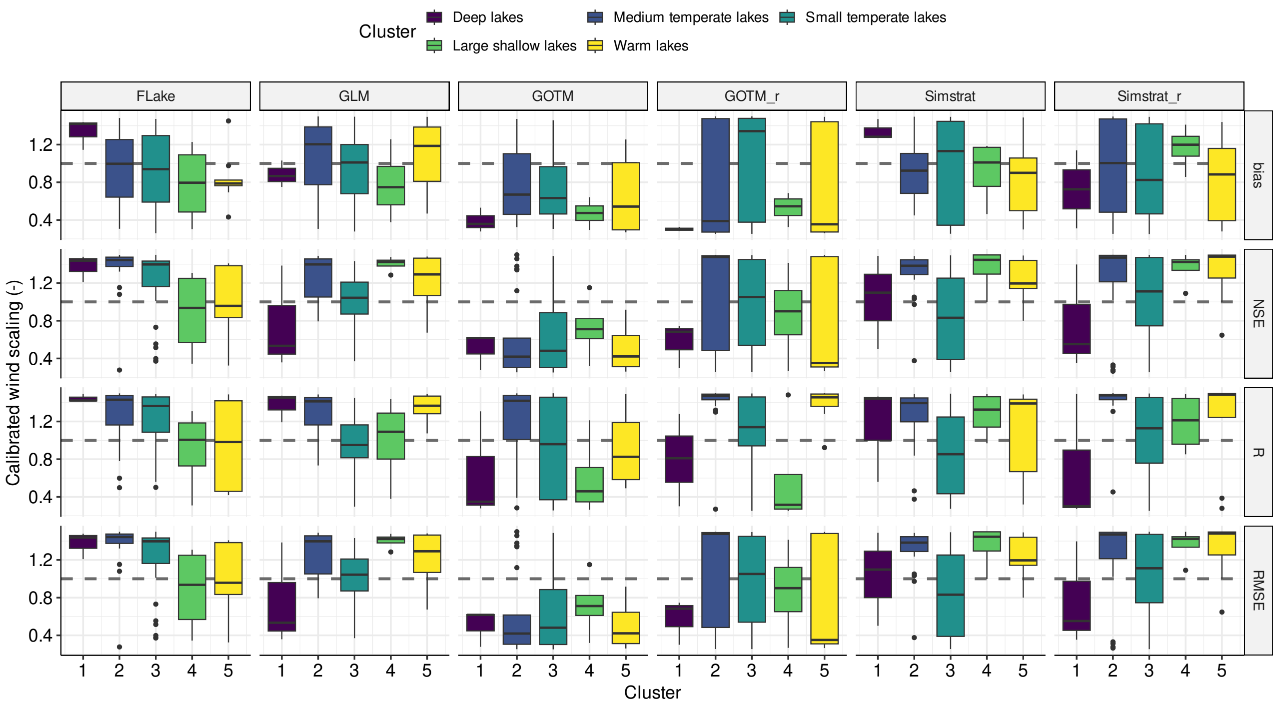

Figure S14: Boxplot of the wind speed scaling factors of the best performing parameter set for the different models, clusters and based on which performance metric was choosen to optimize. The two additional models GOTM_r and Simstrat_r denote the calibration rounds for GOTM and Simstrat with restricted the deep mixing (GOTM `k_min` $= 10^{-8}$ and Simstrats `a_seiche` $= 0$).

For the calibrated shortwave radiation scaling we did not see a reduction in the values for the best performing parameter set for the new round of calibration (GOTM_r in figure S15). The calibrated values even slightly increased when keeping `k_min` at a constant value of $1 \cdot 10^{-8}$.

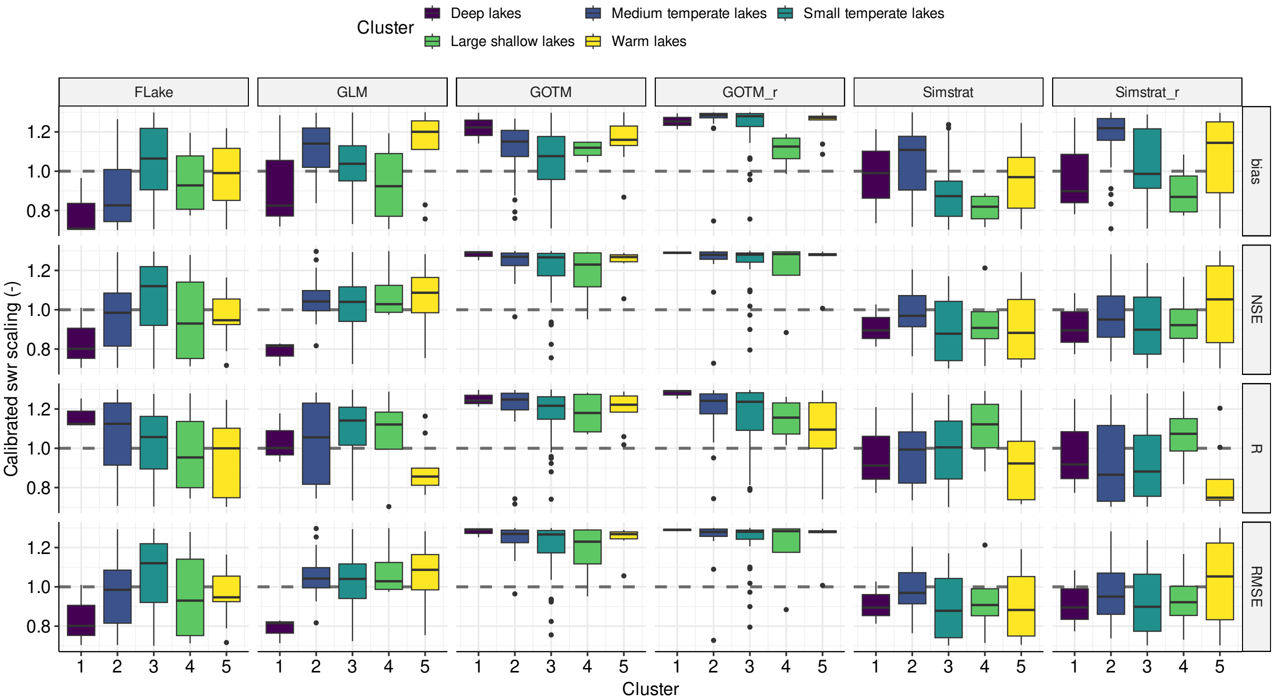

Figure S15: Boxplot of the shortwave radiation scaling factors of the best performing parameter set for the different models, clusters and based on which performance metric was choosen to optimize. The two additional models GOTM_r and Simstrat_r denote the calibration rounds for GOTM and Simstrat with restricted deep mixing (GOTM `k_min` $= 10^{-8}$ and Simstrats `a_seiche` $= 0$).

Reducing `k_min` reduced the overall model performance of GOTM for the lakes where intermediate and deep mixing (incl. internal oscillations) is of importance (as seen by the similar reduction in model performance of Simstrat when `a_seiche` $= 0$). In GOTM, increasing wind speed scaling factor can (to some degree) compensate for this, but it is not able to perform nearly as well as with larger `k_min` values. As GOTM cannot reach similar performance by increasing wind speed scaling and reducing shortwave radiation scaling, we suggest that there is potentially no strong interaction or correlation between the three parameters (as also seen by the sensitivity and interaction measure). Only for some of the small temperate lakes, a lower `k_min` value (and `a_seiche` $=$ 0 for Simstrat) is resulting in better model performance. For the practical application of the two models this could be an important hint that for smaller lakes the lower end of the calibration range of these parameters should be sufficiently low.

An open question remained as to why the values for the calibrated wind speed and shortwave radiation scaling for GOTM behave so differently compared to the other lake models (Figure S14 and S15). This is especially unexpected for Simstrat, which is the most similar to GOTM in terms of process description and even showed a similar reaction in the additional calibration round (reduced `k_min` and `a_seiche` $= 0$) where we saw decreased performance in some of the lake clusters (Figure S13).

References

Håkanson, L.: A Manual of Lake Morphometry, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-81563-8, 1981.

Osgood, R. A.: Lake mixis and internal phosphorus dynamics, Archiv für Hydrobiologie, pp. 629–638, https://doi.org/10.1127/archiv-hydrobiol/113/1988/629, publisher: Schweizerbart'sche Verlagsbuchhandlung, 1988.

Wüest, A. and Lorke, A.: SMALL-SCALE HYDRODYNAMICS IN LAKES, Annual Review of Fluid Mechanics, 35, 373–412, https://doi.org/10.1146/annurev.fluid.35.101101.161220, publisher: Annual Reviews, 2003.