Supporting Information

Prediction of river water temperature: a comparison between a new family of hybrid models and statistical approaches

Sebastiano Piccolroaz[a]*, Elisa Calamita[a], Bruno Majone[a], Aurélien Gallice[b], Annunziato Siviglia[c], and Marco Toffolon[a]

[a] Department of Civil, Environmental and Mechanical Engineering, University of Trento, via Mesiano 77, I-38123,Trento, Italy.

[b] School of Architecture, Civil and Environmental Engineering, École Polytechnique Fédérale de Lausanne, Station 2, CH-1015, Lausanne, Switzerland.

[c] Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH-Zürich, HIA C 58 Hönggerbergring 26, CH-8093, Zürich, Switzerland.

*Correspondence to: Sebastiano Piccolroaz, s.piccolroaz@unitn.it, +39 0461 282637, Department of Civil, Environmental and Mechanical Engineering, University of Trento, via Mesiano 77, I-38123,Trento, Italy

This document contains:

- five tables (S1-S5, the last two refer to external files);
- seven figures (S1-S7).

Tables

Table S1. Pairs of river and meteorological stations with elevations, distances between them, and main statistics of streamflow.

ID river station	Elevation river station [m a.s.l.]	Mean Streamflow [m^3/s]	Standard deviation of streamflow [m^3/s]	ID meteorological station	Elevation meteorological station [m a.s.l.]	Distance between river and meteorological stations [km]
2369	449	1.4	1.99	MAH	437	13
2232	1297	1.16	1.14	ABO	1320	1
2179	553	7.76	8.93	BER	553	15
2159	522	2.51	2.36	BER	553	12
2608	515	0.23	0.48	LUZ	454	11
2308	399	1.42	1.91	ARH	398	7
2327	1668	1.68	1.58	DAV	1594	5
2034	441	8	10.29	PAY	490	3
2112	769	3.02	3.56	STG	776	10
2126	466	1.73	1.74	TAE	539	4
2161	1446	15.16	20.07	ABO	1320	36
2256	1766	2.79	3.64	ROE	1895	106
2343	597	1.2	0.84	RUE	611	35
2366	1860	0.48	0.47	ROE	1895	118
2374	606	3.26	3.91	RUE	611	95
2414	682	0.1	0.14	STG	776	30
2609	840	2.27	3.18	STG	776	58
2612	536	2.8	5.79	TAE	539	136
2617	1236	2.36	1.65	ABO	1320	223
2276	767	1.86	1.87	ALT	438	5
2347	980	0.4	0.72	PLF	1042	158
2009	377	185.7	95.73	AIG	381	4
2011	484	103.58	78.13	SIO	482	2
2016	332	315.56	145.65	BAS	316	46
2019	570	36.34	23.61	INT	577	19
2029	428	242	114.9	KOP	484	25
2030	548	111.7	71.54	BER	553	28
2044	356	47.44	51.57	SHA	438	12
2056	438	42.57	34.78	ALT	438	1
2070	638	11.85	13.9	LAG	745	5
2085	437	175.21	95.36	KOP	484	33
2091	262	1036.06	431.65	BAS	316	16
2135	502	122.48	74.04	BER	553	7
2143	323	443.92	174.51	BAS	316	56
2152	432	108.99	68.66	LUZ	454	2
2372	436	32.01	20.15	LUZ	454	58
2415	336	8.07	5.21	HLL	419	14
2457	564	61.64	40.48	INT	577	2

Table S2. Mean values of NSE^* and $RMSE$ obtained in calibration and validation at daily time scale by applying all models to snow-fed rivers excluding station 2161. We note that, compared to Table 3, performances are higher especially in validation. In fact, station 2161 is a particular case (downstream a moraine, see Section 3.1 of the main text) where RWT has minimal variability during the year. Furthermore, the validation period is only two-year long, and characterized by the occurrence of a particularly warm winter. Overall, the same considerations about the relative performances of the different models considering all snow-fed stations (see Section 3.2.2 in the main text) are valid also in this case.

Model acronym	NSE^* [-] ($RMSE$ [°C])	
	Snow-fed	
	cal	val
lin	-0.96 (1.22)	-1.15 (1.21)
log	-0.53 (1.07)	-0.68 (1.08)
s-lin	-0.57 (1.07)	-0.74 (1.07)
CES	-0.02 (0.86)	-0.16 (0.89)
Teq-2	-0.80 (1.16)	-0.93 (1.15)
Teq-4	0.19 (0.76)	0.20 (0.74)
Teq-6	0.20 (0.77)	0.15 (0.76)
a2s-3	-0.31 (0.96)	-0.43 (0.97)
a2s-4	-0.26 (0.94)	-0.39 (0.96)
a2s-5	0.21 (0.76)	0.20 (0.74)
a2s-7	0.27 (0.74)	0.20 (0.74)
a2s-8	0.28 (0.73)	0.20 (0.74)

Table S3. Averaged $RMSE$ for different time scales and for the four hydrological categories.
* indicates the overall best model
† indicates the best model among those not using streamflow (excluding *air2water* models)

Model acronym	Low-land natural				Lake outlets			
	d	w	m	s	d	w	m	s
lin	1.72	1.35	1.00	0.85	2.55	2.23	1.88	1.56
log	1.46	1.27	0.96	0.91	2.24	2.06	1.78	1.57
s-lin	1.55	1.17	0.77	0.60	1.95	1.60	1.17	**0.53†**
CES	1.31	1.27	1.05	0.94	1.09	1.09	1.06	0.92
Teq-2	1.52	1.34	0.99	0.85	2.41	2.23	1.88	1.56
Teq-4	1.00	0.89	**0.66†**	**0.59†**	1.13	1.02	**0.79†**	0.58
Teq-6	0.93	0.80	**0.56***	**0.50***	1.25	1.11	0.81	**0.48***
a2s-3	1.19	0.98	0.76	0.67	1.11	1.00	0.84	0.59
a2s-4	1.18	0.96	0.74	0.66	1.08	0.97	0.82	0.59
a2s-5	**0.96†**	**0.82†**	0.72	0.63	**1.03†**	**0.95†**	0.84	0.61
a2s-7	**0.85***	**0.78***	0.67	0.57	1.07	0.98	0.83	0.55
a2s-8	**0.85***	0.79	0.65	0.56	1.04	0.96	0.81	0.54
a2w-4	-	-	-	-	0.96	0.84	0.70	0.58
a2w-6	-	-	-	-	**0.91***	**0.79***	**0.64***	0.49

Model acronym	Snow-fed				Regulated			
	d	w	m	s	d	w	m	s
lin	1.11	0.93	0.74	0.61	0.85	0.71	0.61	0.50
log	0.98	0.82	0.65	0.51	0.74	0.61	0.52	0.41
s-lin	0.98	0.79	0.58	**0.36†**	0.76	0.61	0.50	**0.30†**
CES	0.80	0.71	0.57	0.50	0.80	0.74	0.59	0.44
Teq-2	1.06	0.92	0.74	0.61	0.82	0.71	0.61	0.50
Teq-4	0.71	**0.60†**	**0.49†**	0.39	0.68	**0.56†**	**0.47†**	0.35
Teq-6	0.71	0.59	**0.44***	**0.32***	0.56	0.45	**0.35***	**0.25***
a2s-3	0.89	0.76	0.63	0.48	0.82	0.67	0.54	0.44
a2s-4	0.87	0.73	0.59	0.49	0.82	0.66	0.53	0.47
a2s-5	**0.70†**	0.60	0.49	0.39	**0.68†**	0.57	0.48	0.36
a2s-7	**0.68***	0.57	0.45	0.33	0.56	0.46	0.37	0.27
a2s-8	**0.68***	**0.56***	0.45	0.33	**0.55***	**0.45***	0.36	0.27
a2w-4	-	-	-	-	-	-	-	-
a2w-6	-	-	-	-	-	-	-	-

Table S4. Values of the model parameters (*air2stream* and *air2water* families of models) for all river stations.

See file *“SI-Piccolroaz_etal_HP2016-Parameter_values.xlsx”*.

Table S5. Values of root mean square error (RMSE) and Akaike Information Criterion (AIC) for all models and all river stations. Highlighted cells represent best models in terms of RMSE and relative AIC. The only difference concerns station 2029, where RMSE of *CES* is better than *a2s-4* for only 0.001°C, but *a2s-4* prevails in terms of AIC because of the smaller number of parameters (4 against 6). Moreover, in those cases where *a2s-7* and *a2s-8* have identical RMSEs, the former has a minor advantage having one parameter less.

See file *"SI-Piccolroaz_etal_HP2016-AIC.xlsx"*.

Note on AIC computation

AIC is an asymptotically unbiased estimator of the expected relative information distance (e.g., Posada and Buckley, 2004), which represents the amount of information lost when we use models with different number of parameters. We computed AIC using the following formula:

$$AIC = n \ln(RMSE^2) + 2k \,, \qquad \text{(S1)}$$

where n is the number of data points (observations) and k is the number of parameters characterizing the model. Absolute values of AIC are not significant because it is defined only on a relative scale. Hence, we computed AIC differences

$$\Delta AIC_i = AIC_i - AIC|_{best} \qquad \text{(S2)}$$

between for the individual case with respect to the best model (characterized by minimum AIC).

References

Posada D, Buckley TR. 2004. Model Selection and Model Averaging in Phylogenetics: Advantages of Akaike Information Criterion and Bayesian Approaches Over Likelihood Ratio Tests. *Syst. Biol.* **53**(5): 793–808, DOI: 10.1080/10635150490522304.

Figures

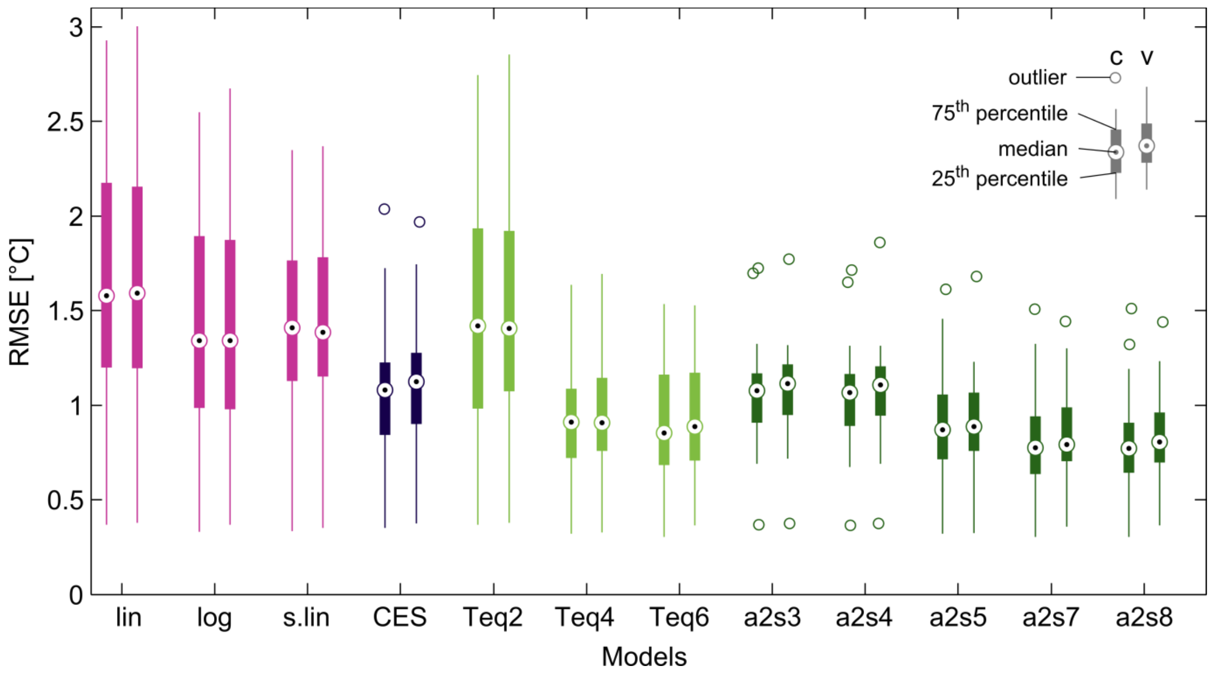

Figure S1. Box plots of $RMSE$ obtained for all statistical models and all versions of *air2stream* at daily time scale distinguishing calibration ("c") and validation ("v"), considering the entire dataset. The acronyms used to identify the models are defined in Table 2.

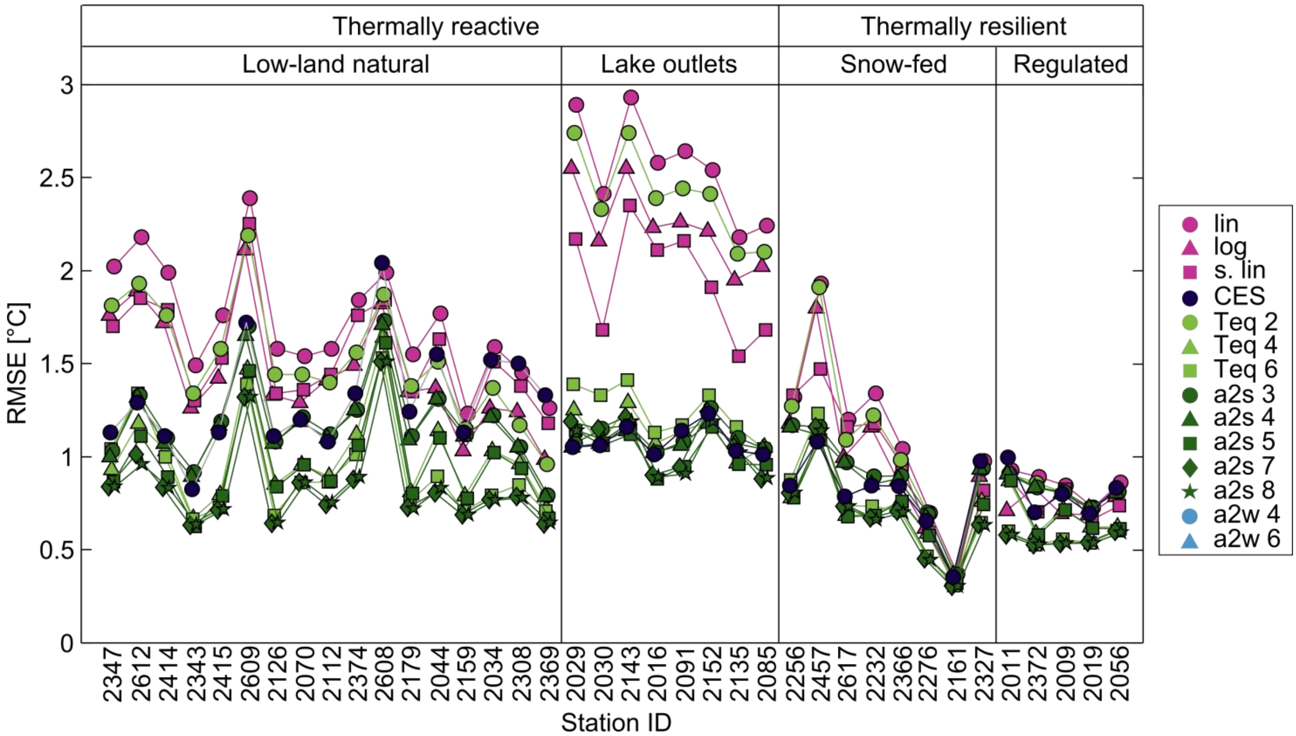

Figure S2. Comparison of $RMSE$ values obtained for each river station using different models (different symbols) at daily scale. Rivers are grouped according to their hydrological and thermal classification. $RMSE$ values refer to the calibration period.

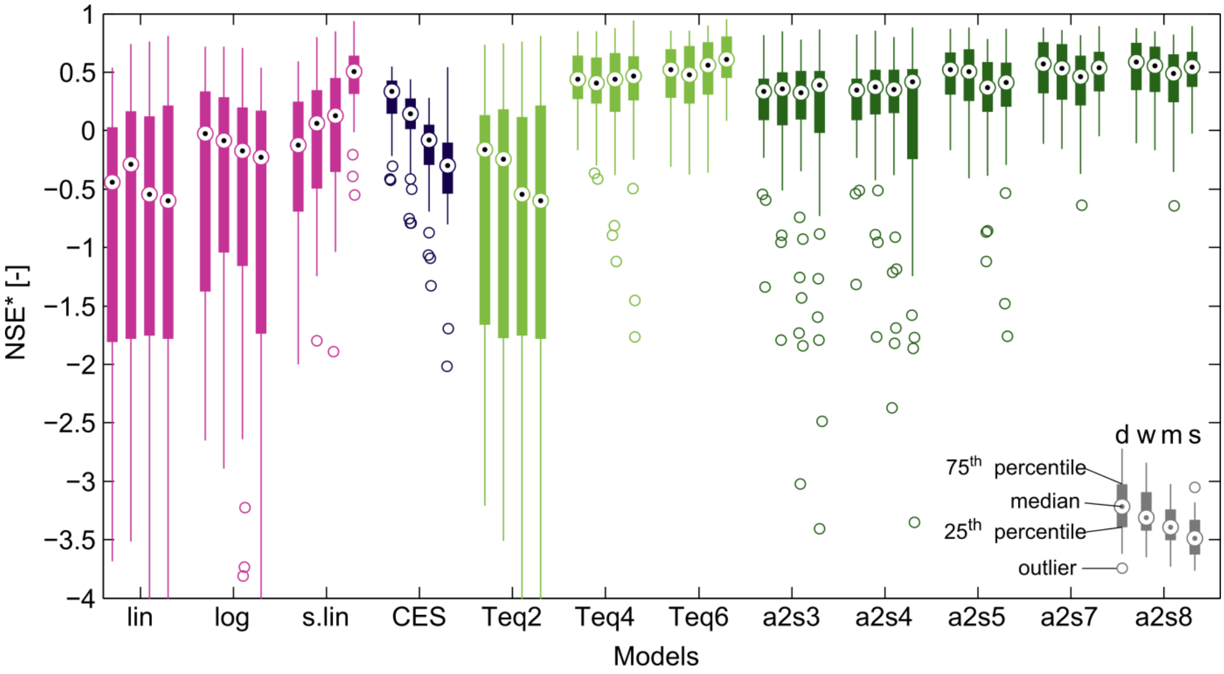

Figure S3. Box plots of NSE^* obtained for all models calibrated at daily (d), weekly (w), monthly (m), and seasonal (s) time scale, considering the entire dataset for the calibration periods.

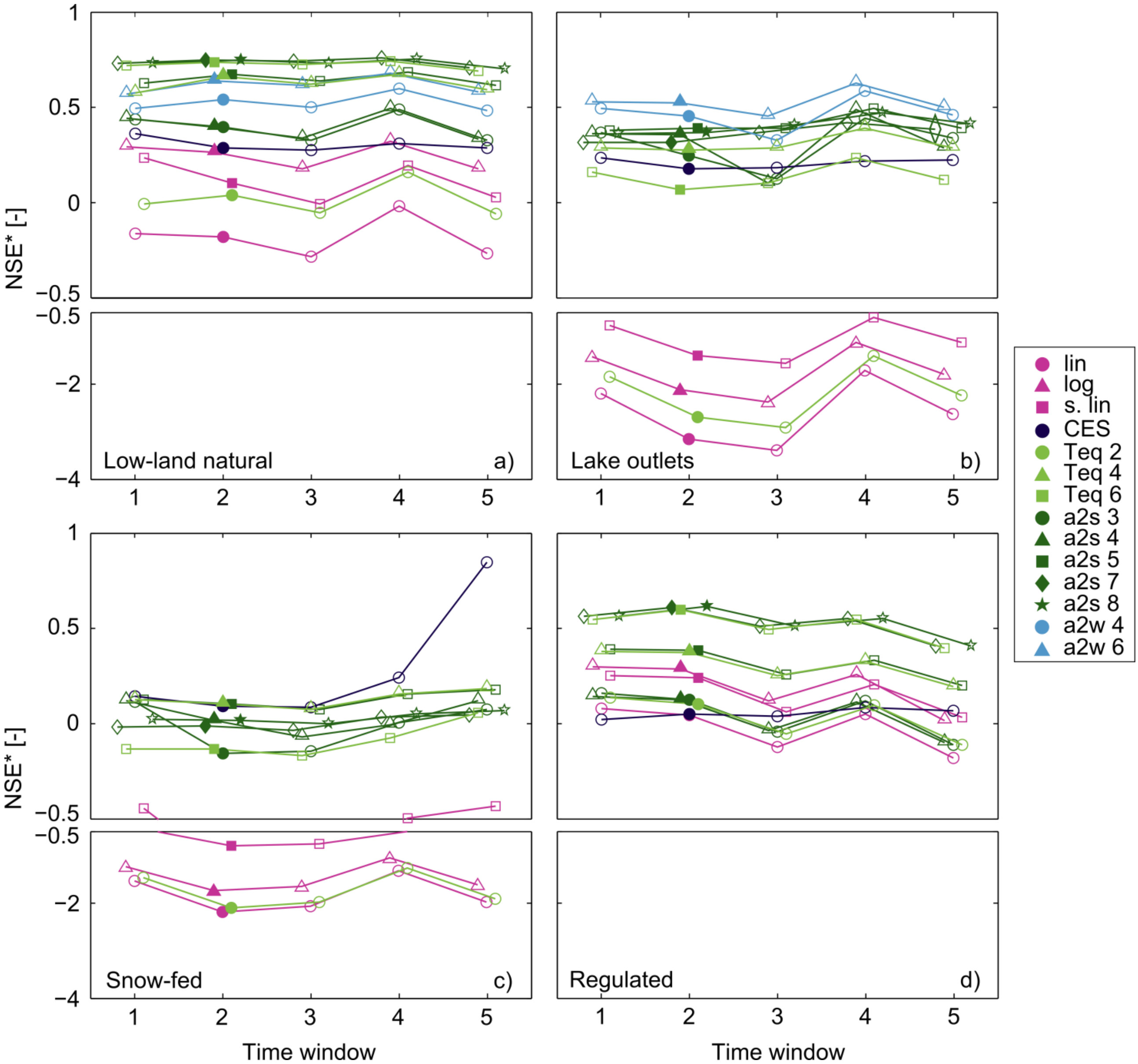

Figure S4. Cross validation of all models considering 5 time windows: NSE^{*} for all models obtained with calibration in time window 2 (1990-1995, filled symbols) and validation in the remaining period (time windows 1 and 3 to 5). Average values are presented for the four hydrological categories.

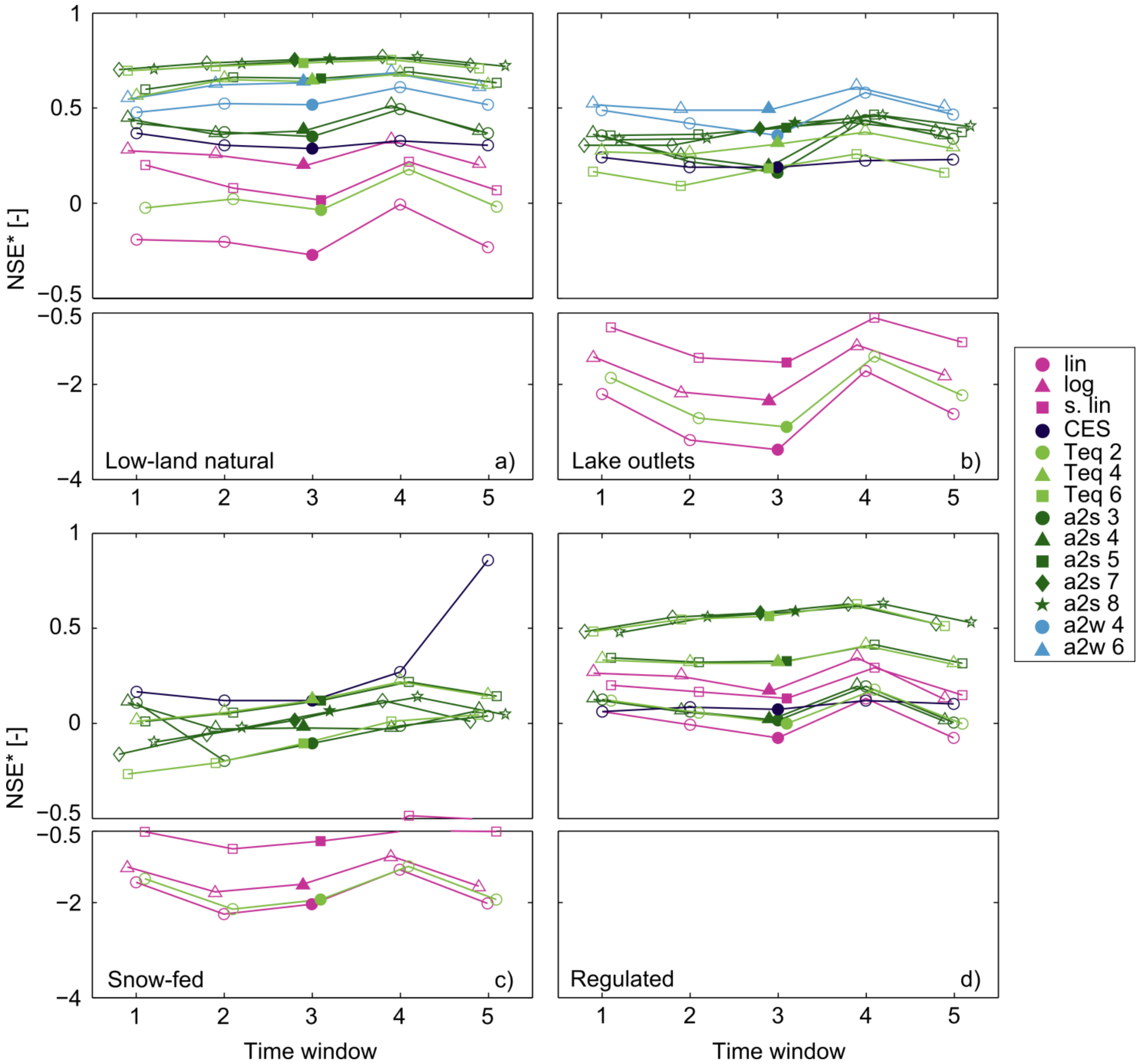

Figure S5. Cross validation of all models considering 5 time windows: NSE^{*} for all models obtained with calibration in time window 3 (1996-2001, filled symbols) and validation in the remaining period (time windows 1 to 2 and 4 to 5). Average values are presented for the four hydrological categories.

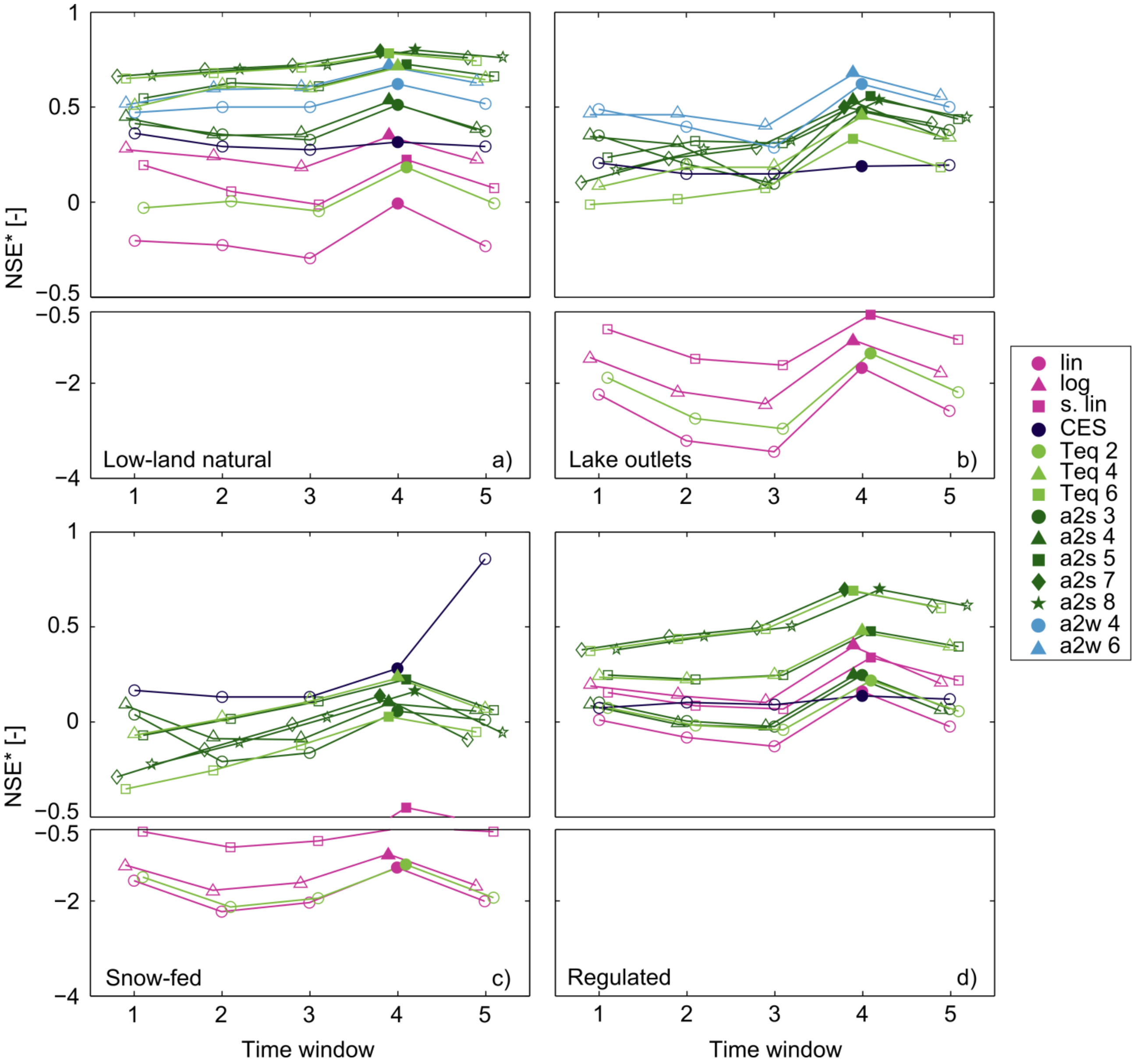

Figure S6. Cross validation of all models considering 5 time windows: NSE^* for all models obtained with calibration in time window 4 (2002-2007, filled symbols) and validation in the remaining period (time windows 1 to 3 and 5). Average values are presented for the four hydrological categories.

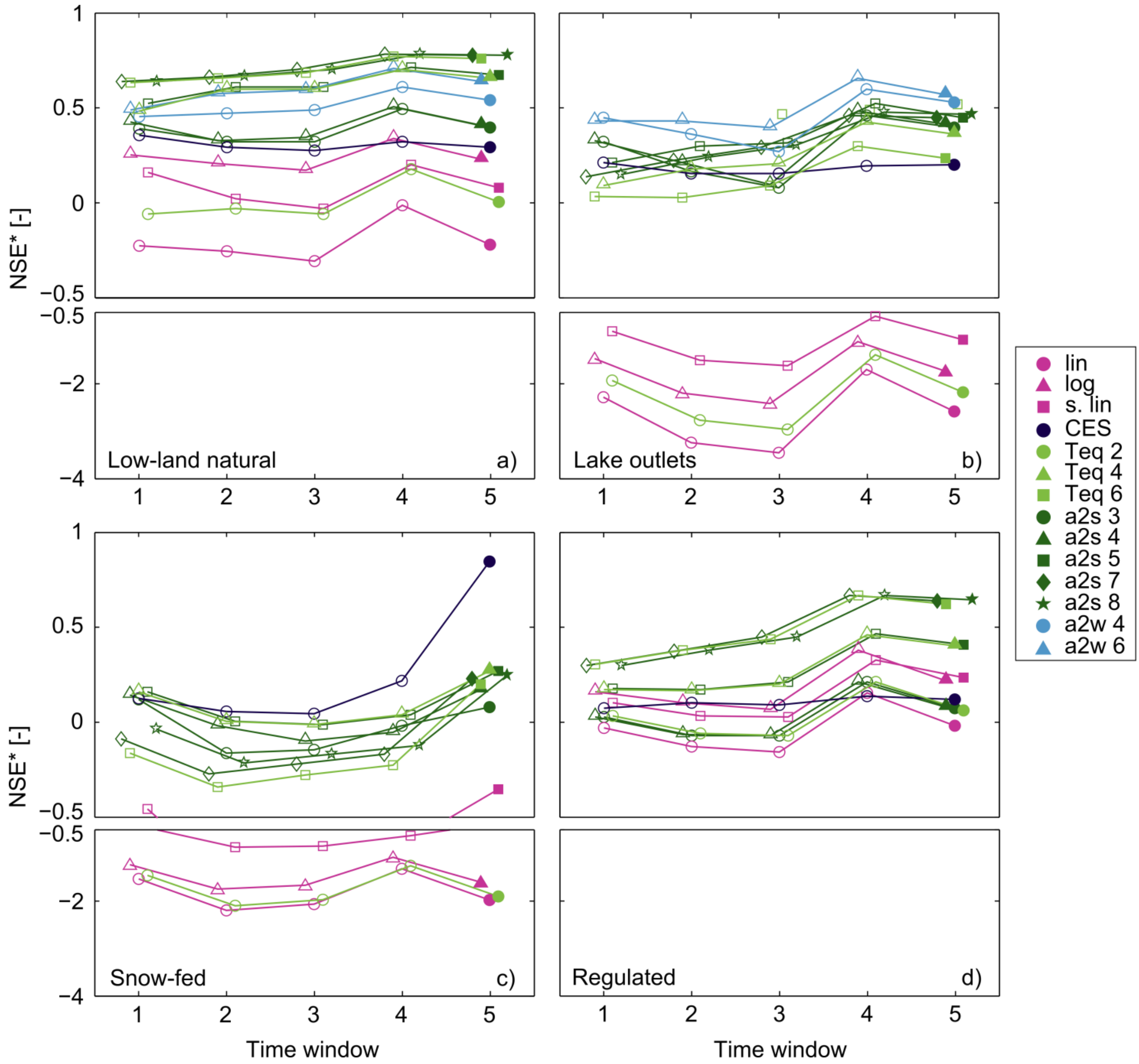

Figure S7. Cross validation of all models considering 5 time windows: NSE^* for all models obtained with calibration in time window 5 (2008-2013, filled symbols) and validation in the remaining period (time windows 1 to 4). Average values are presented for the four hydrological categories.