Supporting information: Extended range forecasting of stream water temperature with deep learning models

Ryan S. Padrón[1,2], Massimiliano Zappa[1], Luzi Bernhard[1], Konrad Bogner[1]

[1]Research Unit Mountain Hydrology and Mass Movements, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf, 8903, Switzerland
[2]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

Correspondence to: Ryan S. Padrón (ryan.padron@wsl.ch)

Station ID	Name	Area [km²]	Mean elevation [m asl]	Glacierized fraction [%]	Coord X [m]	Coord Y [m]	Station elevation [m asl]	Water temp mean [ºC]	Water temp stdev [ºC]
2009	Rhône - Porte du Scex	5238	2127	11.1	557660	133280	377	8.36	2.42
2016	Aare - Brugg	11681	1000	1.5	657000	259360	332	13.31	5.76
2018*	Reuss - Mellingen	3386	1259	1.8	662830	252580	345	13.41	6.01
2019	Aare - Brienzwiler	555	2135	15.5	649930	177380	570	7.72	2.21
2029	Aare - Brügg, Aegerten	8249	1142	2.1	588220	219020	428	13.12	5.91
2030	Aare - Thun	2459	1746	6.9	613230	179280	548	12.39	4.90
2034	Broye - Payerne, Caserne d 'aviation	416	715	0	561660	187320	441	13.01	6.87
2044	Thur - Andelfingen	1702	770	0	693510	272500	356	12.98	6.46
2056*	Reuss - Seedorf	833	2013	6.4	690085	193210	438	8.35	3.15
2068	Ticino - Riazzino	1613	1643	0.1	713670	113500	200	10.95	4.20
2070*	Emme - Emmenmatt, nur Hauptstation	443	1065	0	623610	200430	638	10.81	5.39
2084	Muota - Ingenbohl	317	1363	0	688262	206170	438	9.32	3.48
2085*	Aare - Hagneck	5112	1368	3.4	580680	211650	437	13.18	5.21
2091	Rhein - Rheinfelden, Messstation	34524	1068	0.8	627189	267845	265	13.54	5.91
2104	Linth - Weesen, Biäsche	1062	1584	1.6	725160	221380	419	12.08	5.14
2106	Birs - Münchenstein, Hofmatt	887	728	0	613570	263080	268	12.28	4.97
2109*	Lütschine - Gsteig	381	2050	13.5	633130	168200	585	7.81	2.80
2112	Sitter - Appenzell	74.4	1256	0.1	749040	244220	769	9.72	5.13
2113*	Aare - Felsenau, K.W. Klingnau (U.W.)	17687	1060	1.4	659150	271790	312	14.14	5.90
2126	Murg - Wängi	80.2	652	0	714105	261720	466	12.50	4.85
2135	Aare - Bern, Schönau	2941	1596	5.8	600710	198000	502	12.62	5.04
2150*	Landquart - Felsenbach	614	1797	0.7	765364	204910	571	8.55	4.26
2152	Reuss - Luzern, Geissmattbrücke	2254	1504	2.8	665330	211800	432	13.34	5.84
2159*	Gürbe - Belp, Mülimatt	116	846	0	604810	192680	522	12.82	5.84
2167	Tresa - Ponte Tresa, Rocchetta	609	803	0	709580	92145	268	16.13	7.10
2179	Sense - Thörishaus, Sensematt	351	1071	0	593350	193020	553	11.62	6.39
2210*	Doubs - Ocourt	1275	952	0	572530	244460	417	11.90	5.39
2243	Limmat - Baden, Limmatpromenade	2384	1131	0.7	665640	258690	351	13.83	6.51
2256*	Rosegbach - Pontresina	66.5	2704	21.7	788795	151694	1766	5.89	4.23
2269	Lonza - Blatten	77.4	2624	24.7	629130	140910	1520	5.87	2.72

2276	Grosstalbach - Isenthal	43.9	1819	6.7	685500	196050	767	8.51	2.09
2288	Rhein - Neuhausen, Flurlingerbrücke	11930	1239	0.6	689145	281975	383	13.29	6.53
2307*	Suze - Sonceboz	127	1036	0	579810	227350	642	10.03	2.79
2308	Goldach - Goldach, Bleiche, nur Hauptstation	50.4	832	0	753190	261600	399	11.89	7.01
2343	Langeten - Huttwil, Häberenbad	59.9	760	0	629560	219135	597	10.81	4.06
2351*	Vispa - Visp	786	2648	23.1	634030	125900	659	7.54	2.86
2369	Mentue - Yvonand, La Mauguettaz	105	675	0	545440	180875	449	11.41	5.98
2374*	Necker - Mogelsberg, Aachsäge	88.1	956	0	727110	247290	606	10.92	6.25
2392	Rhein (Oberwasser) - Rheinau	11950	1238	0.6	687420	277140	353	13.28	6.49
2414	Rietholzbach - Mosnang, Rietholz	3.19	794	0	718840	248440	682	9.77	4.62
2415	Glatt - Rheinsfelden	417	503	0	678040	269720	336	14.25	5.94
2432	Venoge - Ecublens, Les Bois	228	686	0	532040	154160	383	12.37	5.18
2433*	Aubonne - Allaman, Le Coulet	105	952	0	520720	147410	390	10.39	4.17
2457*	Aare - Ringgenberg, Goldswil	1138	1951	12.1	633730	171510	564	11.53	4.58
2462*	Inn - S-Chanf	616	2463	6.1	795800	165910	1645	7.09	4.26
2467	Saane - Gümmenen	1881	1131	0.1	585100	199240	473	11.61	5.35
2473*	Rhein - Diepoldsau, Rietbrücke	6299	1771	0.7	766280	250360	410	9.49	3.91
2485*	Allaine - Boncourt, Frontière	212	562	0	567830	261200	366	12.82	4.22
2493	Promenthouse - Gland, Route Suisse	120	1027	0	510080	140080	394	10.93	3.70
2604*	Biber - Biberbrugg	31.9	1003	0	697240	223280	825	9.88	6.14
2608*	Sellenbodenbach - Neuenkirch	10.4	608	0	658530	218290	515	11.75	5.51
2612	Riale di Pincascia - Lavertezzo	44.5	1705	0	708060	123950	536	8.98	5.78
2617	Rom - Müstair	128	2184	0	830800	168700	1236	7.50	3.37
2623	Rhone - Oberwald	93.3	2466	19.3	669870	154080	1368	5.25	2.42

Table S1: List of 54 stream water temperature stations from the Swiss Federal Office for the Environment monitoring network (https://www.hydrodaten.admin.ch/en/) with their corresponding characteristics used as static features when training the deep learning models. *Indicates the subset of stations not used for training the models when evaluating the predictive skill at new and ungauged stations.

15

Model	Hyperparameter	Tunning space	Tuned value
RNNED	hidden_size	$32 - 128$	116
	rnn_layers	$1 - 4$	1
	gradient_clip_val	$0.1 - 0.3$	0.144
	dropout	$0 - 0.3$	0.276
TFT	hidden_size	$32 - 128$	60
	hidden_continuous_size	$8 - 32$	12
	lstm_layers	$1 - 4$	2
	attention_head_size	$1 - 4$	1
	gradient_clip_val	$0.1 - 0.3$	0.204
	learning_rate	$0.001 - 0.1$	0.0033
	dropout	$0 - 0.3$	0.222
NHITS	hidden_size	$32 - 128$	51
	num_blocks	$1 - 3$	3
	num_stacks	$1 - 3$	3
	gradient_clip_val	$0.1 - 0.3$	0.105
	learning_rate	$0.001 - 0.1$	0.0033
	dropout	$0 - 0.3$	0.173

Table S2: Hyperparameters of the deep learning models, the explored hyperparameter space, and their tuned values. Hyperparameter names are in agreement with the PyTorch Forecasting documentation (https://pytorch-forecasting.readthedocs.io/en/stable/models.html).

Number	Forecast start date	Number	Forecast start date	Number	Forecast start date
1	2022-01-03	31	2022-04-18	61	2022-08-08
2	2022-01-06	32	2022-04-21	62	2022-08-11
3	2022-01-10	33	2022-04-25	63	2022-08-15
4	2022-01-13	34	2022-04-28	64	2022-08-18
5	2022-01-17	35	2022-05-02	65	2022-08-22
6	2022-01-20	36	2022-05-05	66	2022-08-25
7	2022-01-24	37	2022-05-09	67	2022-08-29
8	2022-01-27	38	2022-05-12	68	2022-09-01
9	2022-01-31	39	2022-05-16	69	2022-09-05
10	2022-02-03	40	2022-05-19	70	2022-09-15
11	2022-02-07	41	2022-05-23	71	2022-09-19
12	2022-02-10	42	2022-05-26	72	2022-09-22
13	2022-02-14	43	2022-05-30	73	2022-09-26
14	2022-02-17	44	2022-06-02	74	2022-09-29
15	2022-02-21	45	2022-06-06	75	2022-10-06
16	2022-02-24	46	2022-06-09	76	2022-10-10
17	2022-02-28	47	2022-06-13	77	2022-10-13
18	2022-03-03	48	2022-06-16	78	2022-10-17
19	2022-03-07	49	2022-06-20	79	2022-10-20
20	2022-03-10	50	2022-06-23	80	2022-10-24
21	2022-03-14	51	2022-06-27	81	2022-10-27
22	2022-03-17	52	2022-07-07	82	2022-10-31
23	2022-03-21	53	2022-07-11	83	2022-11-03
24	2022-03-24	54	2022-07-14	84	2022-11-07
25	2022-03-28	55	2022-07-18	85	2022-11-10
26	2022-03-31	56	2022-07-21	86	2022-11-14
27	2022-04-04	57	2022-07-25	87	2022-11-17
28	2022-04-07	58	2022-07-28	88	2022-11-21
29	2022-04-11	59	2022-08-01	89	2022-11-24
30	2022-04-14	60	2022-08-04	90	2022-11-28

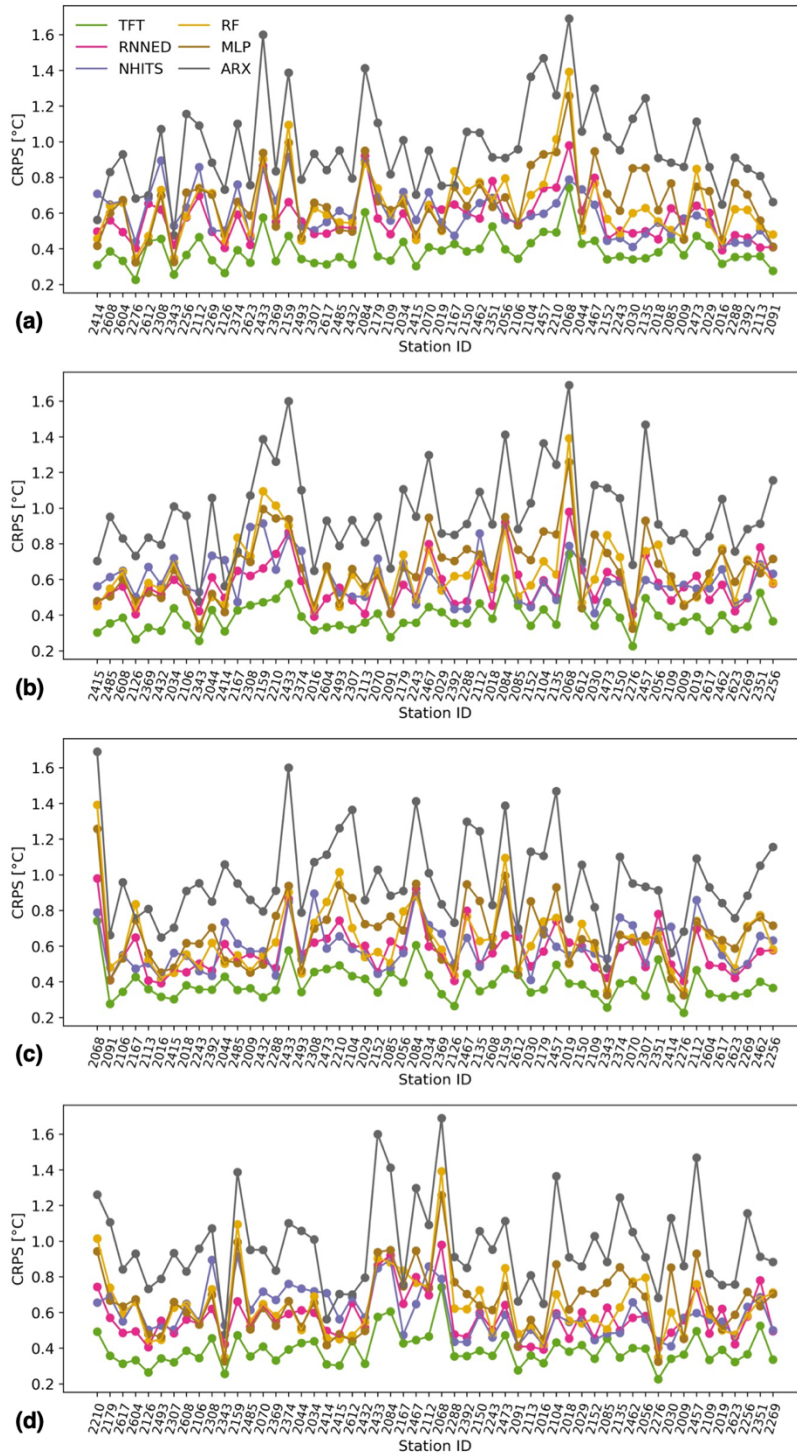
Table S3: List of forecast start dates during 2022 used for model evaluation. Forecasts are generated twice per week, with data missing for 2022-06-30, 2022-07-04, 2022-09-08, 2022-09-12, and 2022-10-03.

40

Figure S1: Model predictive skill per station when omitting the uncertainty of meteorological forecasts. Stations are sorted in ascending fashion according to **(a)** catchment area, **(b)** mean catchment elevation, **(c)** station elevation, and **(d)** glacierized fraction when omitting the uncertainty of meteorological forecasts. The CRPS is averaged over all lead times and forecasts.

45

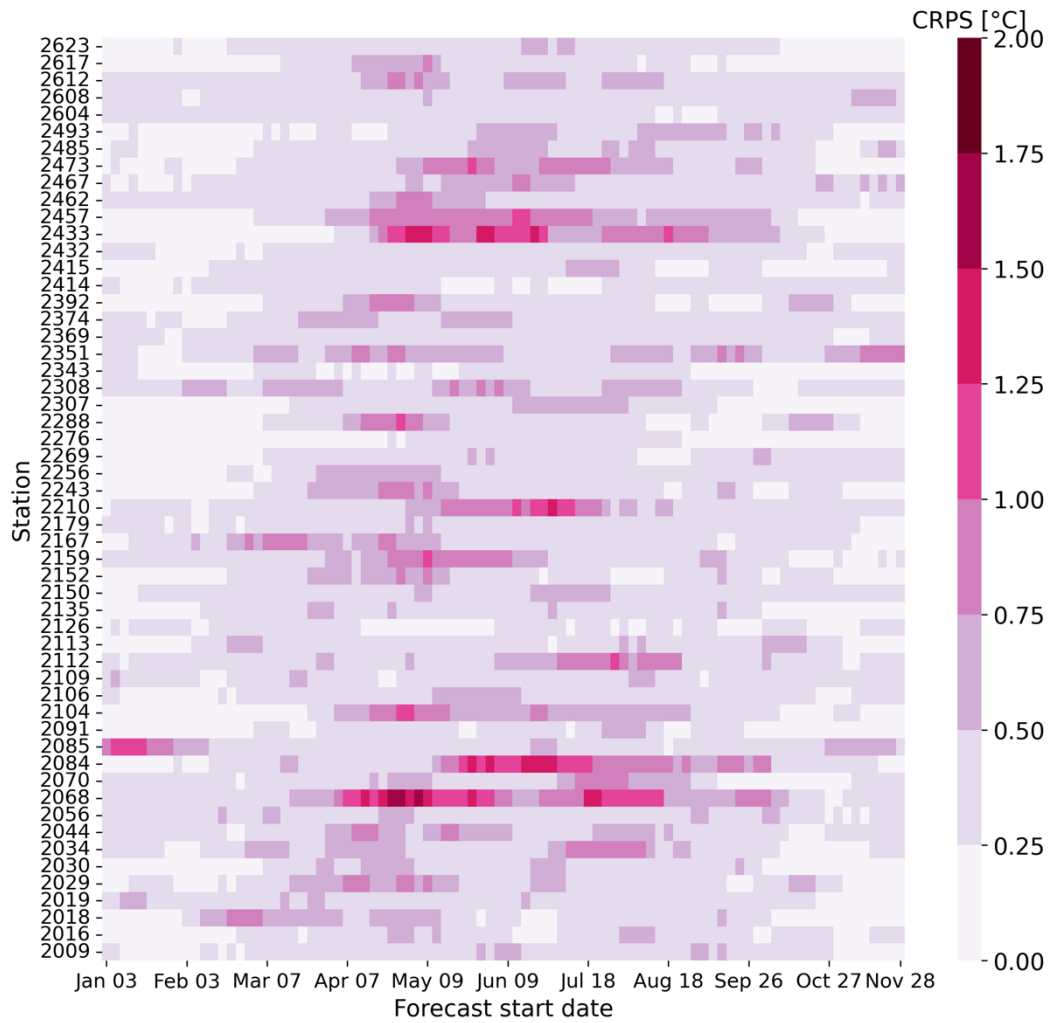



Figure S2: TFT predictive skill per forecast start date and station when omitting the uncertainty of meteorological forecasts. The CRPS is averaged over all lead times.

50

55

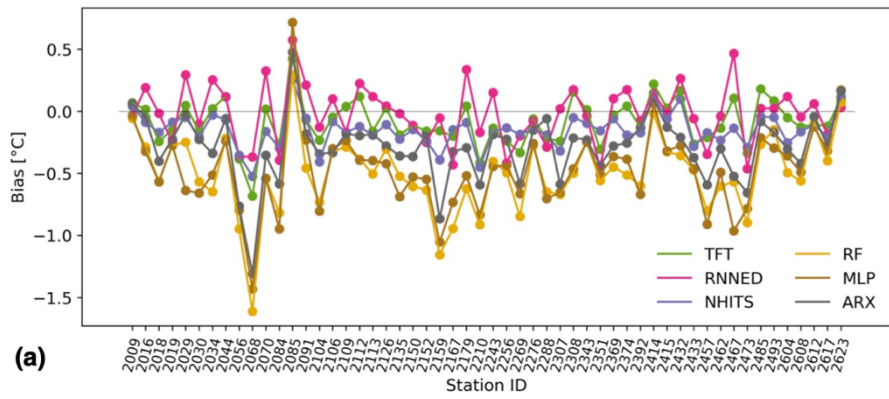
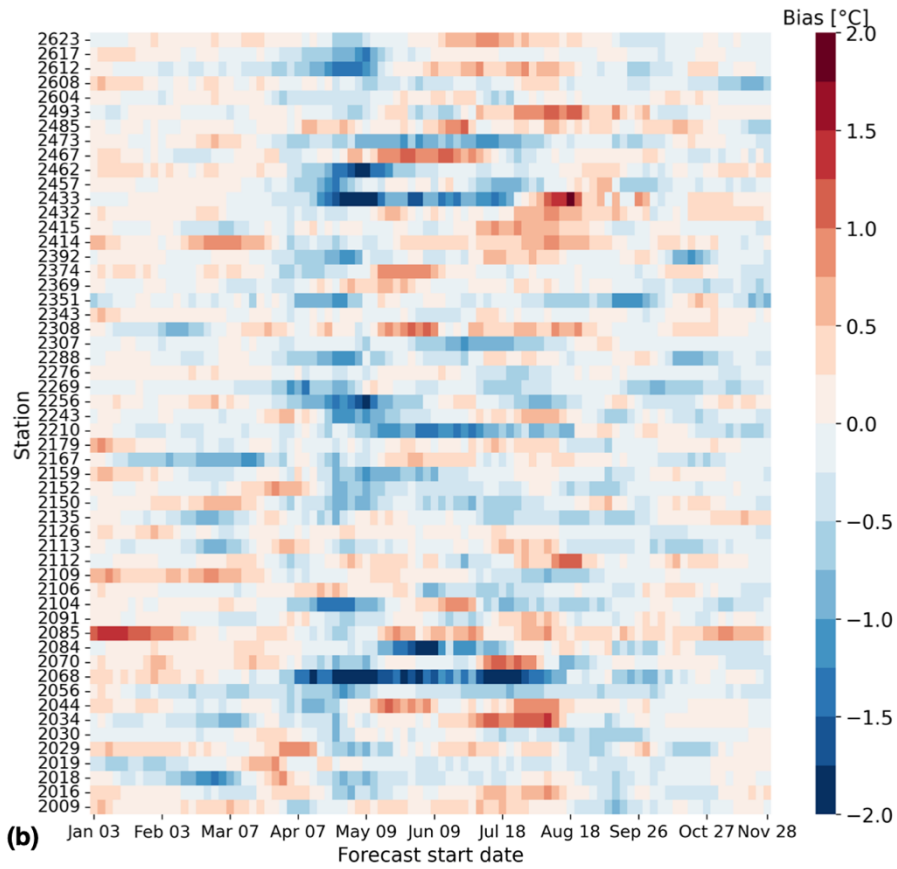



Figure S3: Water temperature forecast bias when omitting the uncertainty of meteorological forecasts. **(a)** Model comparison of forecast bias as a function of station averaged over all lead times and forecast start dates. **(b)** TFT forecast bias per forecast start date and station averaged over all lead times.

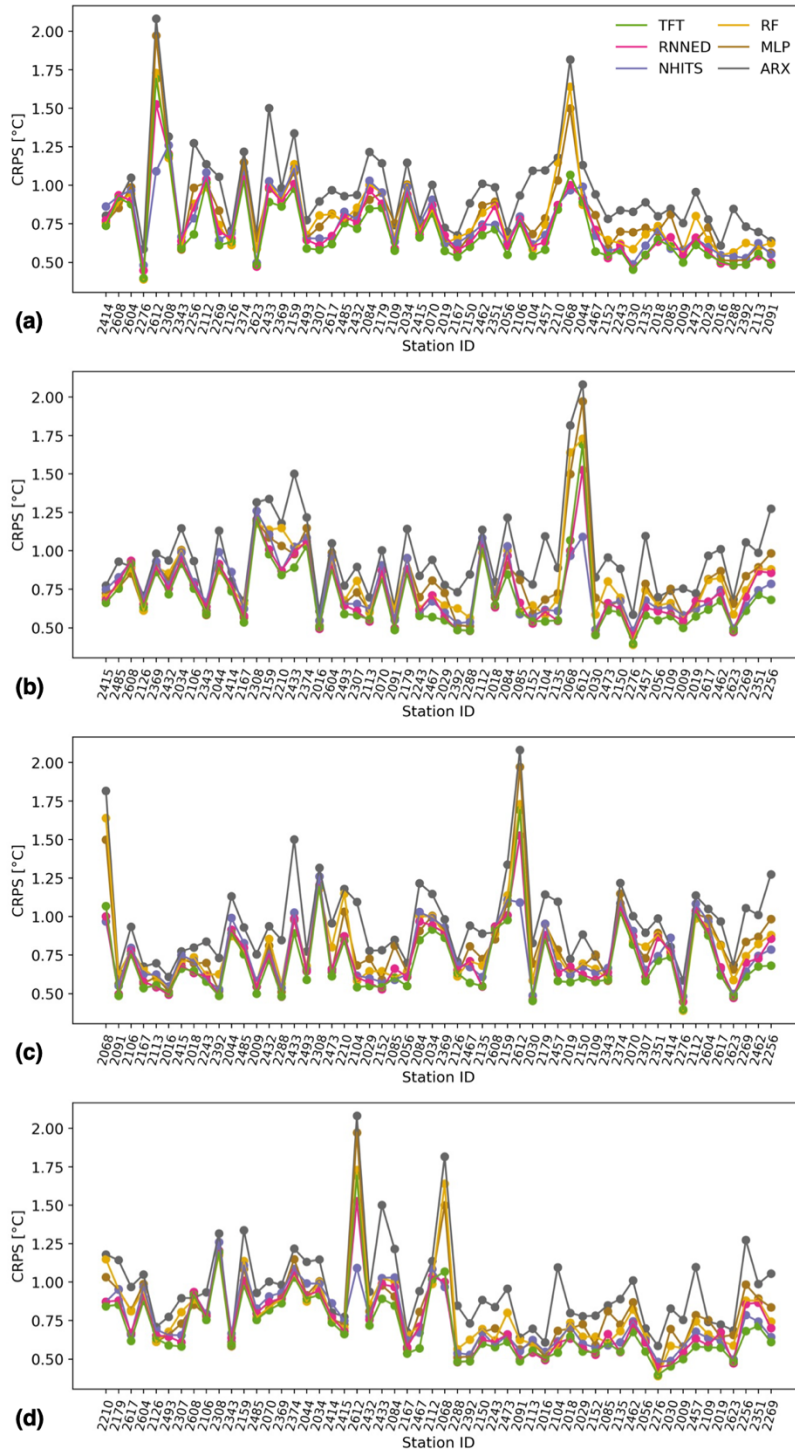



Figure S4: Model predictive skill per station when including the uncertainty of meteorological forecasts. Stations are sorted in ascending fashion according to **(a)** catchment area, **(b)** mean catchment elevation, **(c)** station elevation, and **(d)** glacierized fraction when omitting the uncertainty of meteorological forecasts. The CRPS is averaged over all lead times and forecasts.

65

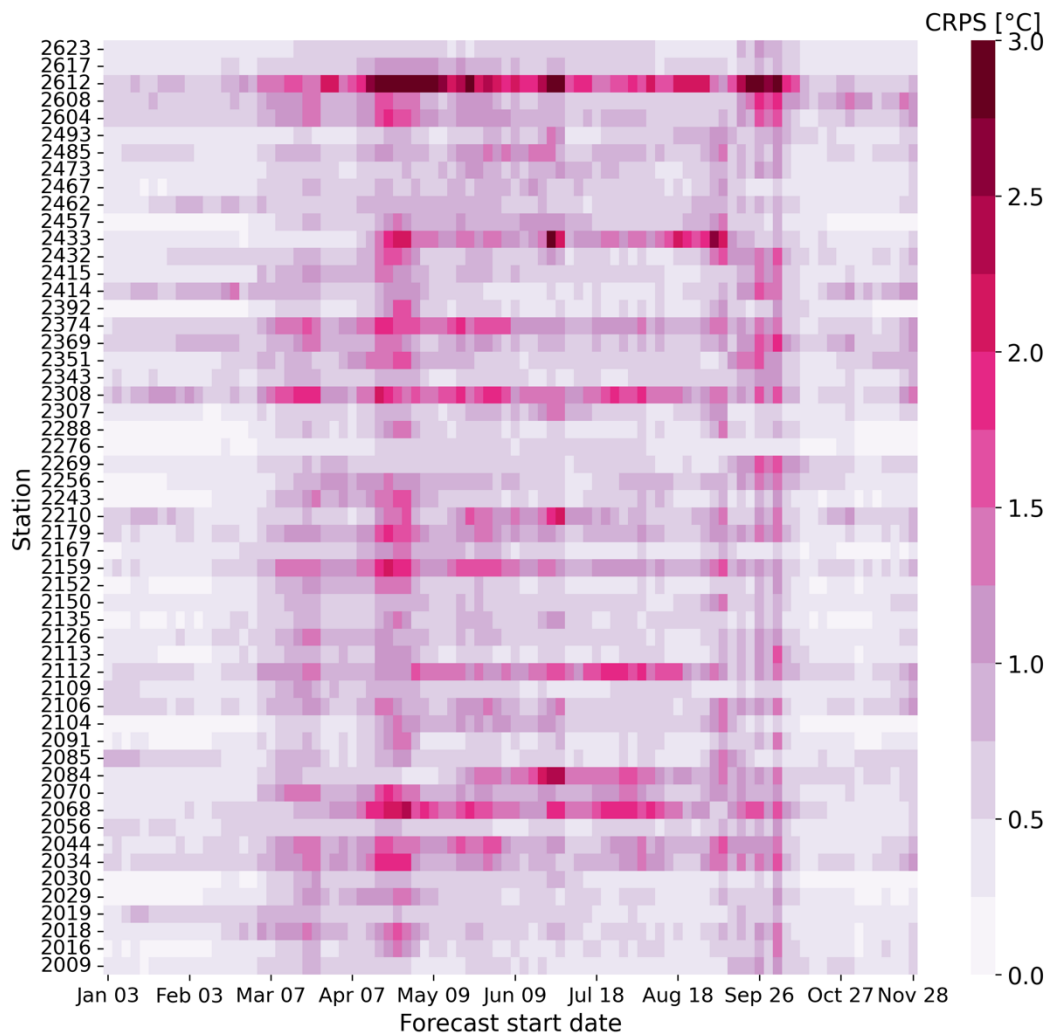



Figure S5: TFT predictive skill per forecast start date and station when including the uncertainty of meteorological forecasts. The CRPS is averaged over all lead times.

70

75

80

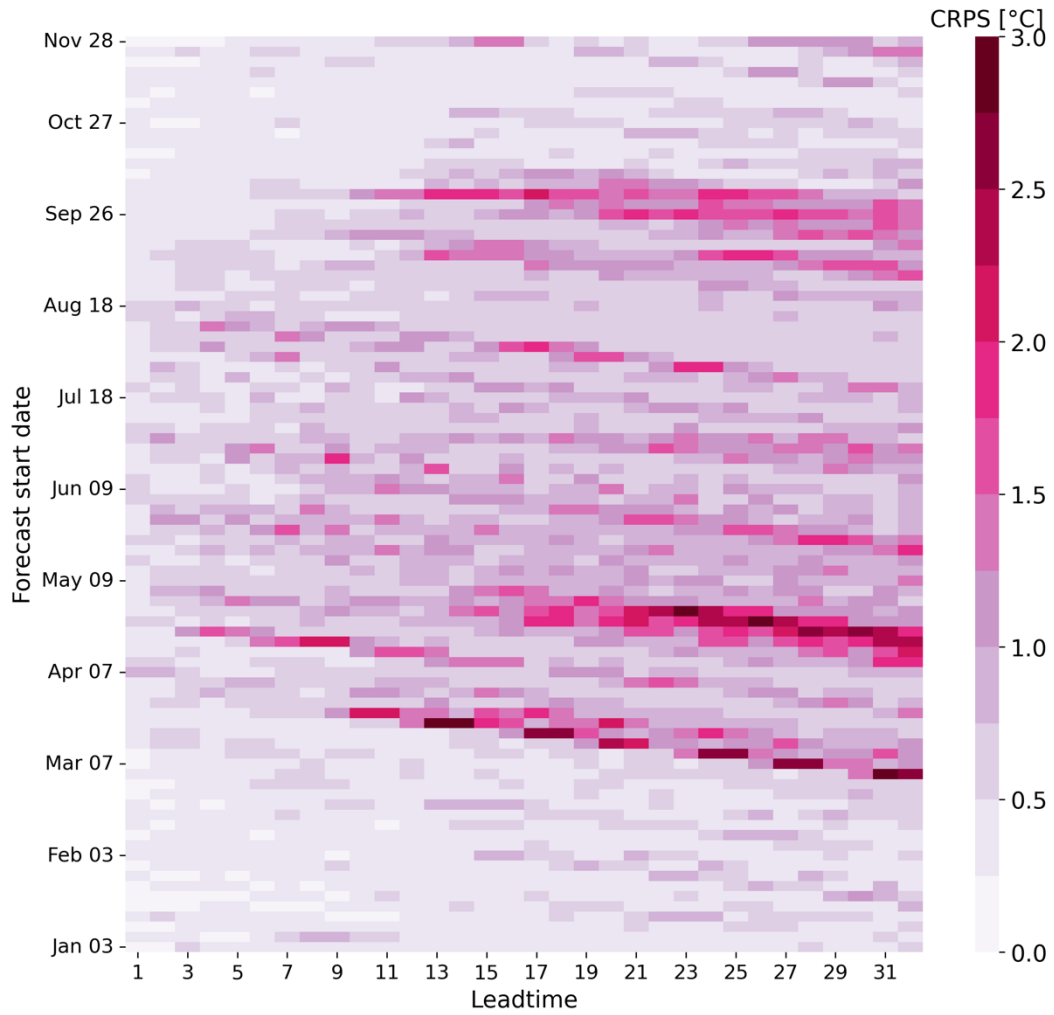



Figure S6: TFT predictive skill per lead time and forecast start date when including the uncertainty of meteorological forecasts. The CRPS is averaged over all stations.

85

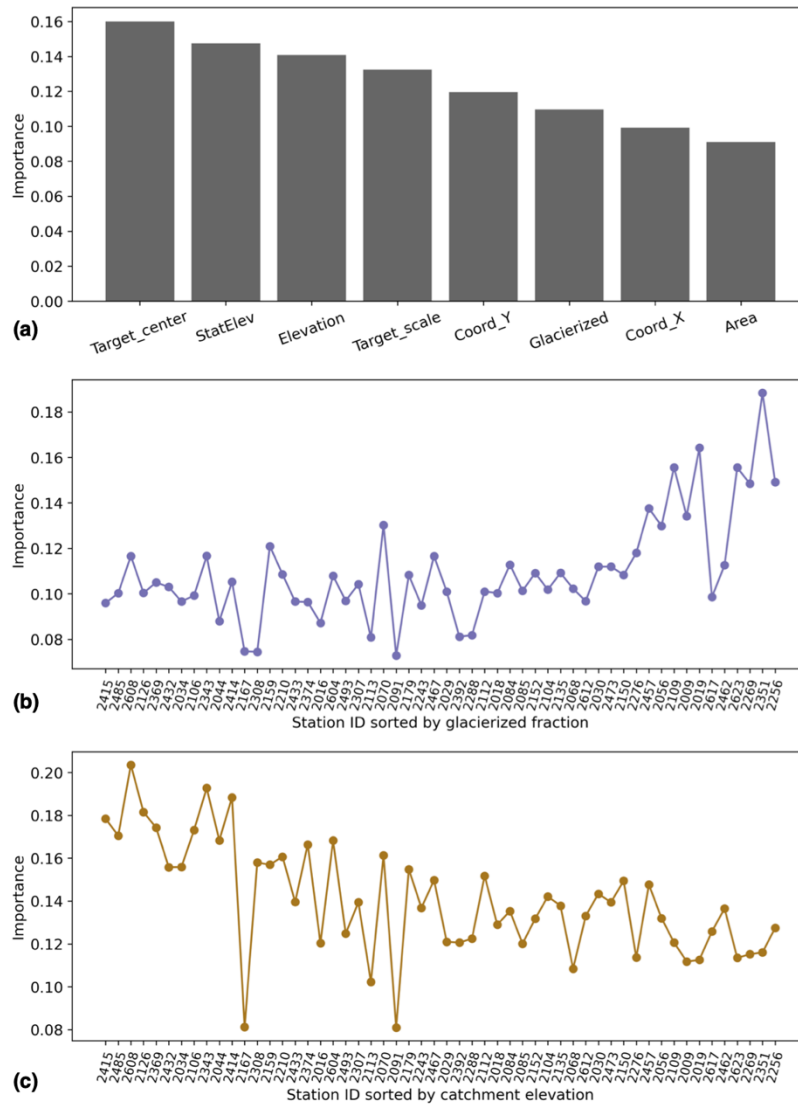



Figure S7: TFT static feature importance. **(a)** Feature importance averaged over 10 random seeds and 54 stations. **(b)** Glacierized fraction importance per station, with stations sorted according to the glacierized fraction of their catchments. **(c)** Catchment elevation importance per station, with stations sorted according to their mean catchment elevation.

100

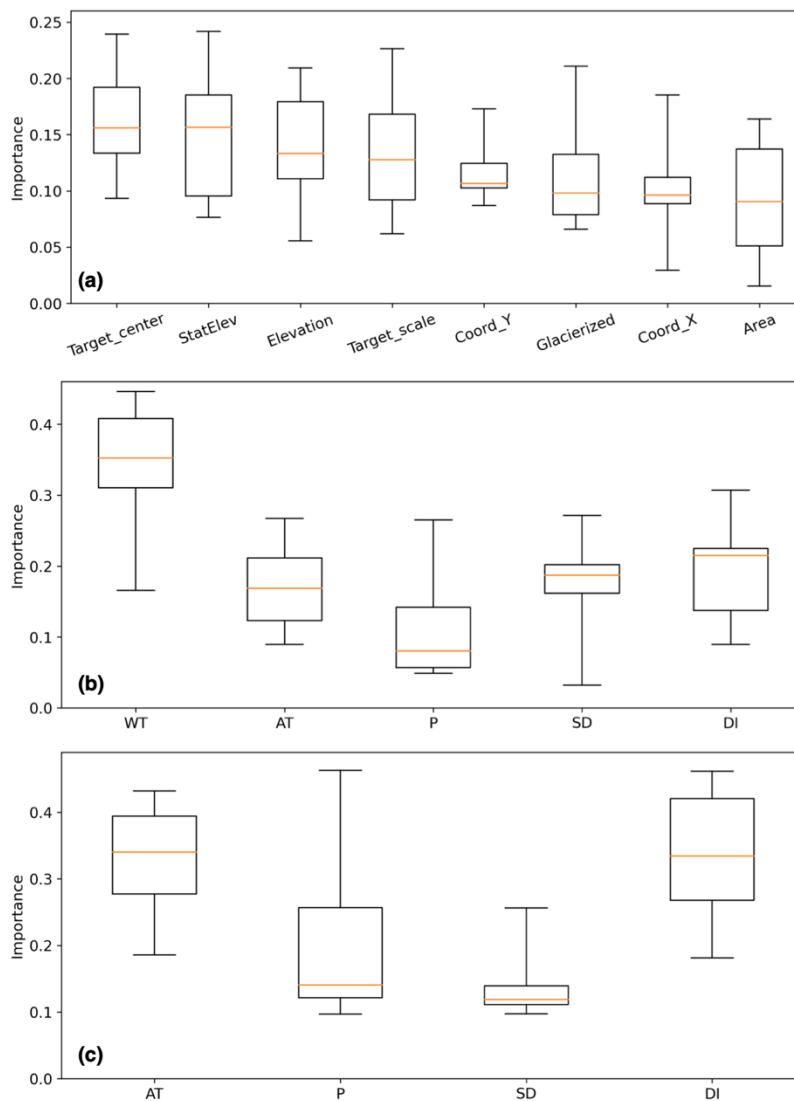
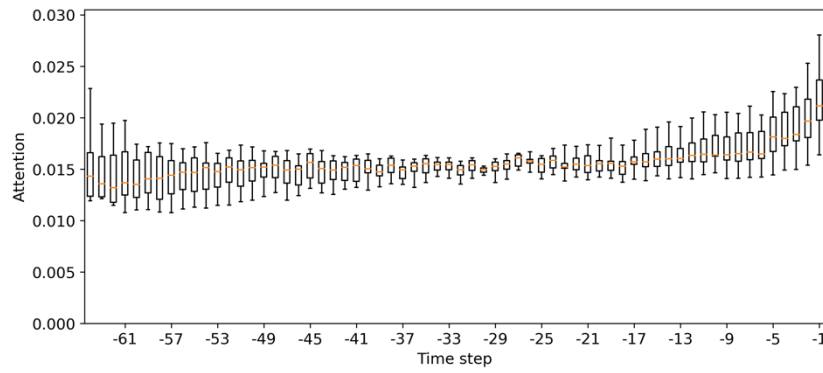



Figure S8: TFT average feature *importance* variability across 10 models trained with different random seeds. Results are shown for **(a)** static features, **(b)** encoder features, and **(c)** decoder features. Boxplots indicate the minimum, median, maximum, and interquartile range of the average *importance* across time steps, stations and forecast start dates.

105

110 **Figure S9:** TFT *attention* variability across 10 models trained with different random seeds as a function of encoder time step. Boxplots indicate the minimum, median, maximum, and interquartile range of the average *attention* across stations and forecast start dates.