Ecological Indicators 161 (2024) 111978

Contents lists available at ScienceDirect

Ecological Indicators

journal homepage: www.elsevier.com/locate/ecolind

Prediction of daily river water temperatures using an optimized model based on NARX networks

Jiang Sun [a], Fabio Di Nunno [b], Mariusz Sojka [c], Mariusz Ptak [d], You Luo [a], Renyi Xu [a], Jing Xu [a], Yi Luo [e,*], Senlin Zhu [a,*], Francesco Granata [b]

[a] College of Hydraulic Science and Engineering, Yangzhou University, Yangzhou, China
[b] Department of Civil and Mechanical Engineering (DICEM), University of Cassino and Southern Lazio, Via Di Biasio, 43, 03043, Cassino, Frosinone, Italy
[c] Department of Land Improvement, Environmental Development and Spatial Management, Poznań University of Life Sciences, Piątkowska 94E, 60-649 Poznań, Poland
[d] Department of Hydrology and Water Management, Adam Mickiewicz University, B. Krygowskiego 10, 61-680 Poznań, Poland
[e] GIS Technology Research Center of Resource and Environment in Western China, Yunnan Normal University, Kunming, China

A R T I C L E I N F O

Keywords:
River water temperature
Modeling
NARX
Air2stream
Climate change

A B S T R A C T

Water temperature is an important physical indicator of rivers because it impacts many other physical and biogeochemical processes and controls the metabolism of aquatic species in rivers. Having a good knowledge of river thermal dynamics is of great importance. In this study, an advanced machine learning based model that is fast, accurate and easy to use, namely the nonlinear autoregressive network with exogenous inputs (NARX) neural network, was coupled with Bayesian Optimization (BO) algorithm for optimizing the number of NARX hidden nodes and lagged input/target values and the Bayesian Regularization (BR) backpropagation algorithm for the NARX training, to forecast daily river water temperatures (RWT). Long-term observed data from 18 rivers of the Vistula River Basin, one of the largest rivers in Europe, were used for model testing, and model performance was compared with the air2stream model. The results showed that the NARX-based model performs significantly better than the air2stream model in the calibration and validation stages, and can better capture the seasonal pattern and peak values of RWT. Input combinations impact the performance of the NARX-based model in RWT modeling, and air temperature and the day of the year (*DOY*) are the major inputs, while streamflow and rainfall play a minor role on modeling RWT at the Vistula River Basin. Considering that future times series of air temperatures are easily accessible from climate models and *DOY* is easy to be considered in the model, the NARX-based model can serve as a promising tool to investigate the impact of climate change on river thermal dynamics.

1. Introduction

Water temperature is a pivotal indicator for rivers, exerting profound influence on both physical and biogeochemical processes, as well as the metabolic activities of aquatic species (Ouellet et al., 2020; Cai et al., 2023; Zhi et al., 2023a, b). In the context of a dynamic environment, water temperature consistently serves as sentinels of climate change (Michel et al., 2022; Rehana and Rajesh, 2023). Consequently, possessing a comprehensive understanding of the thermal dynamics within rivers becomes paramount, particularly when scrutinizing smaller time scales, such as the daily scale.

Mathematical modeling plays an important role in predicting daily river water temperature (RWT). In the past decades, tools in RWT simulation have been developed from simple statistical models (e.g., Mohseni and Stefan, 1999; Benyahya et al., 2017; Piotrowski and Napiórkowski, 2019) to complex process-based models (e.g., Dugdale et al., 2017; Gatien et al., 2023; White et al., 2023). Among these models, machine learning (ML) based models are blooming due to the fast development of computational capacity and data science, including shallow ML-based models like extreme learning machine models (e.g., Zhu et al., 2019a; Heddam et al., 2023a), artificial neural networks (e.g., Zhu et al., 2019b; Almeida and Coelho, 2023), hybrid models coupling multiple techniques (Graf et al., 2019; Heddam et al., 2023b), and deep ML-based models like deep learning models (e.g., Ikram et al., 2023;

* Corresponding authors.
E-mail addresses: mx120230642@stu.yzu.edu.cn (J. Sun), fabio.dinunno@unicas.it (F. Di Nunno), mariusz.sojka@up.poznan.pl (M. Sojka), marp114@wp.pl (M. Ptak), luoyou@yzu.edu.cn (Y. Luo), Xury@yzu.edu.cn (R. Xu), xujing7503@yzu.edu.cn (J. Xu), lysist@ynnu.edu.cn (Y. Luo), slzhu@yzu.edu.cn (S. Zhu), f.granata@unicas.it (F. Granata).

https://doi.org/10.1016/j.ecolind.2024.111978
Received 11 December 2023; Received in revised form 11 March 2024; Accepted 30 March 2024

Zwart et al., 2023). More details about these ML-based models can be found in the review of Zhu and Piotrowski (2020).

In contrast to their shallow counterparts, deep ML models, such as deep learning models, exhibit heightened complexity in both their structural design and configuration. Consequently, the ability to achieve precise daily river water temperature simulations using shallow ML-based models assumes paramount significance.

Among the advanced ML-based models, the Nonlinear Autoregressive Network with Exogenous Inputs (NARX) neural network emerges as a noteworthy choice, characterized by its high computational efficiency, accuracy, and user-friendly nature (Alsumaiei, 2020; Di Nunno et al., 2021a, 2021b, 2022; Gao et al., 2023; Zhu et al., 2023). Previous applications of this model involved coupling it with the Bayesian Optimization (BO) algorithm to optimize the number of NARX hidden nodes and lagged input/target values. Additionally, the Bayesian Regularization (BR) backpropagation algorithm was employed for NARX training, resulting in successful predictions of thermal anomalies in Polish lakes during heatwaves (Zhu et al., 2023). The efficacy of NARX-based models has been substantiated in various hydrological and environmental studies, as evidenced by works such as those by Lee and Sheridan (2018), Alsumaiei (2020), Di Nunno et al., (2021a,2022), and Gao et al. (2023).

It is noteworthy that, while the NARX model has shown widespread applications in hydrological modeling, an enhanced version of this particular neural network, enabling the optimization of both the number of NARX hidden nodes and lagged input/target values through the application of the BO algorithm, along with an optimal algorithm for the NARX training represented by the BR backpropagation, represents a notable progression in the accurate RWT prediction. Moreover, scant attention has been devoted to investigating the influence of different input combinations on model performance, particularly the role of rainfall. Notably, previous modeling endeavors predominantly employed air temperature and flow as primary input parameters, as comprehensively reviewed by Zhu and Piotrowski (2020).

To fill the above gaps, in this study we used the BO-NARX-BR model to predict daily RWT of 18 rivers (24 hydrological stations) across the Vistula River Basin, one of the largest rivers in Europe. For each hydrological station, daily RWT and flow data are available from 1991 to 2021, and daily air temperature and rainfall data from the nearby 15 meteorological stations. Model performance was compared with another widely used model for RWT modeling, namely the air2stream model (Toffolon and Piccolroaz, 2015), to evaluate its performance.

2. Materials and methods

2.1. Study area and data

The study area covers the catchment of the Vistula River, one of the largest rivers in Europe. Its length is 1022 km, and the catchment area is 193,960 km^2 (Ochrona Środowiska, 2022). In terms of catchment area, it ranks 10th in Europe, with an average population density of 114 people per km^2 (Tockner et al., 2009). Arable land occupies 48.3 %, meadows and pastures 14.0 %, forests 26.5 %, water bodies 2.5 %, and other land uses 8.7 % (Kajak, 1993). The Vistula River originates at an altitude of 1140 m a.s.l. in the Silesian Beskids, a mountain range in the Western Carpathians, flowing to the Gdańsk Bay, part of the Baltic Sea. The Vistula River and its principal tributaries play a pivotal role in influencing biodiversity, while concurrently holding significant economic importance. This significance extends to critical areas such as energy, water supply, and transportation, particularly within the context of inland waterway systems. Daily measurements for water temperature and flow were provided by the Institute of Meteorology and Water Management – National Research Institute. The former is measured at a depth of 0.4 m below the water surface, and the latter is the result of flow intensity curve analysis. Due to the extensive research area, the analysed rivers are diverse, having characteristics of both mountain rivers and lowland rivers. This has implications for the flow rate, which for the rivers analysed in the article ranges from 1.67 to 922 m^3·s^{-1}. The year-round average water temperature varied from 8.1 to 10.9 °C. Data on air temperature and precipitation also come from the Institute of Meteorology and Water Management – National Research Institute, comprising a set of standard meteorological observations in Poland. The average air temperature for the meteorological stations mentioned in the article ranged from 7.7 to 9.4 °C, while the average annual precipitation varied from 527 to 1024 mm.

Detailed information on the 18 studied rivers and analysed hydrological and meteorological stations can be found in Table 1. Fig. 1 shows the spatial locations of the hydrological and meteorological stations.

2.2. Models

2.2.1. Narx-based models

This section offers a comprehensive overview of the BO-NARX-BR model, providing information on the NARX algorithm used for RWT forecasting, along with insights related to the BO and BR algorithms applied for the optimization of the NARX model.

NARX neural networks belong to a specific category of recurrent dynamic artificial neural network (RNN), consisting of interconnected nodes inspired by the biological neural system. Each node symbolizes an artificial neuron that receives one or more inputs, processes them through a nonlinear function (referred to as the activation function), and produces an output. The primary advantages of NARX networks, compared to other ANN approaches, include faster convergence in achieving optimal weights for the connections between neurons and input parameters (Desouky and Abdelkhalik, 2019), a reduced number of the latter to calibrate, enhancing the model's effectiveness (Di Nunno et al., 2021), and the model relates the current value to current and past values of the inputs, like other RNN models.

The fundamental equation for a NARX network employed in time series prediction can be formulated as follows:

$$y(t) = f(y(t-1), y(t-2), \cdots, y(t-f_d), x(t-1), x(t-2), \cdots, x(t-p_d)) \quad (1)$$

where *x(t)* is the input layer, which includes five different input combinations (see Section 2.2.3) at a given time *t*, and *y(t)* is the target, represented by RWT at time *t*. Additionally, p_d and f_d denote the lagged values of inputs and target, respectively.

The NARX architecture comprises three layers (Fig. 2). The first layer functions as the input layer, receiving input parameters. The second layer serves as the hidden layer, responsible for computations between the input and output. Lastly, the third layer acts as the output layer, providing the predicted values. Additionally, the estimated output is looped back as input for iterative computation in the subsequent time step (as indicated by the dot-dash blue line in Fig. 2). In the hidden layer, a sigmoid activation function (f_1) was applied, given its suitability for training neural networks using back-propagation algorithms. The sigmoid function is also differentiable, facilitating the learning of neural network weights (Boussaada et al., 2018). For the output layer, a linear activation function (f_2) with a single neuron (n) was utilized.

Nevertheless, a fundamental challenge in developing any forecasting model lies in strategically selecting input variables. In this study, the determination of the optimal number of lagged values for both variables and the precise configuration of hyperparameters for the NARX model were executed using the BO algorithm (Snoek et al., 2012; Zhu et al., 2023). Specifically, the BO algorithm was employed to discover the optimal values for the number of hidden nodes ($h_1, h_2, \ldots, h_n$ in Fig. 2), p_d and f_d. The BO procedure establishes an objective function for Bayesian optimization and defines the hyperparameter search space related to the number of hidden nodes, lags, and delays. Subsequently, the NARX model can be iteratively trained and evaluated, resulting in a loss value quantified in terms of Mean Square Error (MSE) for each hyperparameter combination. Therefore, the number of hidden nodes, p_d and

Table 1
Detailed information of the studied rivers and analyzed hydrological and meteorological stations. The characters of the hydrological stations and the numbers of the meteorological stations are shown in Fig. 1.

No	River	Station	No.	Basin area (km^2)	Altitude(m asl)	Meteorological station	No.
1	Wisła	Skoczów	a	296	285.7	Bielsko-Biała	1
2	Wisła	Kępa Polska	b	168,357	57.3	Płock	13
3	Wisła	Toruń	c	180,390	32.0	Toruń	3
4	Soła	Oświęcim	c	1357	225.8	Katowice	4
5	Raba	Stróża	e	644	297.0	Kraków-Balice	5
6	Dunajec	Żabno	f	6739	172.4	Tarnów	6
7	San	Radomyśl	g	16,838	138.8	Sandomierz	7
8	Tanew	Harasiuki	h	2035	164.5	Lublin-Radawiec	15
9	Wieprz	Kośmin	i	10,293	115.0	Warszawa	2
10	Pilica	Białobrzegi	j	8665	112.0	Warszawa	2
11	Narew	Zambski Kościelne	k	27,807	79.0	Białystok	8
12	Narew	Nowogród	l	20,169	94.0	Białystok	8
13	Narew	Narew	m	1983	130.4	Lesko	14
14	Biebrza	Burzyn	n	6929	98.8	Białystok	8
15	Pisa	Ptaki	o	3576	104.9	Białystok	8
16	Omulew	Białobrzeg Bliższy	p	1788	94.4	Mława	9
17	Bug	Wyszków	q	38,395	81.5	Warszawa	2
18	Bug	Krzyczew	r	25,595	125.1	Włodawa	10
19	Bug	Strzyżów	s	8991	171.7	Włodawa	10
20	Krzna	Malowa Góra	t	3042	127.7	Włodawa	10
21	Drwęca	Brodnica	u	3540	67.4	Toruń	3
22	Wda	Czarna Woda	v	828	111.0	Chojnice	11
23	Osa	Rogóźno 2	w	1137	31.3	Toruń	3
24	Wąska	Pasłęk	x	246	6.7	Elbląg	12

f_d that minimize MSE were considered optimal for the modeling.

Moreover, other NARX parameters also require optimization. Among these, the weight (*w*) and bias (*b*) parameters are influenced by the selected training algorithm. In this context, the BR algorithm (MacKay, 1992) was chosen. This algorithm outperformed the Levenberg-Marquardt (LM) and Scaled Conjugate Gradient (SCG) algorithms, which were pre-tested. These findings align with prior studies in the literature, demonstrating that BR exhibits slower convergence but provides superior performance compared to LM and SCG (Di Nunno and Granata, 2020; Zhu et al., 2023).

Finally, the BO-NARX-BR process was halted when one of the following conditions was met (Di Nunno et al., 2023): reaching the maximum number of epochs (set at 1000), achieving the LM adjustment parameter (set at 1×10^{10}), or attaining an error gradient below a specified threshold (set at 1×10^{-7}).

2.2.2. air2stream model versions

The air2stream model is a semi-empirical, hybrid model that combines physics-based structure (e.g., energy balance) with random calibration of model parameters (Toffolon and Piccolroaz, 2015). Due to its low model input requirements and high predictive accuracy, it has been widely used to predict RWT (e.g., Toffolon and Piccolroaz, 2015; Piotrowski and Napiórkowski, 2018; Zhu et al., 2022; Almeida and Coelho, 2023).

The air2stream model is based on the lumped heat budget equation (see below, Eq. (2), and the variation of river water temperature is mainly attributed to two components: (1) the net heat flux at the river-atmosphere interface and (2) the heat flux of flow discharge.

$$\rho cpV\frac{dT_w}{dt} = AH_{net} + \rho cp\left(\sum_i QiTw,i - QTw\right) \tag{2}$$

where, t is time, ρ is water density, c_p is the specific heat capacity at constant pressure of water, A is the surface area of the river reach, H_{net} is the net heat flux at the river-atmosphere interface, T_w is water temperature, Q is flow discharge at the downstream section, Q_i and $T_{w,i}$ are flow discharge and water temperature of the i-th contributing water flux, and V is the total water volume. Toffolon and Piccolroaz (2015) employed a Taylor series expansion to rewrite the equation, introducing eight parameters into the model. This process allows the model to rely solely on the time series of air temperature and flow discharge as its exclusive inputs:

$$\frac{dTw}{dt} = \frac{1}{\theta^{a_4}}\left\{a_1 + a_2Ta - a_3Tw + \theta\left(a_5 + a_6\cos\left(2\pi\left(\frac{t}{t_y} - a_7\right)\right) - a_8T_w\right)\right\} \tag{3}$$

where T_a is air temperature, t_y is the duration of one year (365 days), θ is the dimensionless discharge, and a_1-a_8 are model parameters. The 8 parameters model version (a2s-8) is the full version of the air2stream model, which accounts for the major thermal, geometric, and hydraulic characteristics. During the model calibration phase, the 8 model parameters are determined by trial and error using optimization algorithms like particle swarm optimization.

Based on the 8 parameters model version (a2s-8), the air2stream model can be further simplified into other four versions. Disregarding the impact of fluctuating flow discharge ($\theta = 1$) and reorganizing the remaining model parameters, the model can be simplified into the 5 parameters model version (a2s-5):

$$\frac{dTw}{dt} = \left\{a_1 + a_2Ta - a_3Tw + a_6\cos(2\pi\left(\frac{t}{t_y} - a_7\right)\right\} \tag{4}$$

In some cases, the average temperature of the contributing water closely approximates to that of the river section, then the impact of flow discharge on RWT can be disregarded, and the 4 parameters model version (a2s-4) can be obtained:

$$\frac{dTw}{dt} = \frac{1}{\theta^{a_4}}\{a_1 + a_2Ta - a_3Tw\} \tag{5}$$

Assuming that flow discharge is constant ($a_4 = 0$), the a2s-8 and a2s-4 model versions can be further simplified to 7 parameters model version (a2s-7) and 3 parameters model version (a2s-3):

$$\frac{dTw}{dt} = a_1 + a_2Ta - a_3Tw + \theta\left(a_5 + a_6\cos\left(2\pi\left(\frac{t}{t_y} - a_7\right)\right) - a_8T_w\right) \tag{6}$$

$$\frac{dTw}{dt} = \{a_1 + a_2Ta - a_3Tw\} \tag{7}$$

In this study, the performance of all these five model versions will be evaluated using the observed data from the 18 rivers (24 hydrological stations).

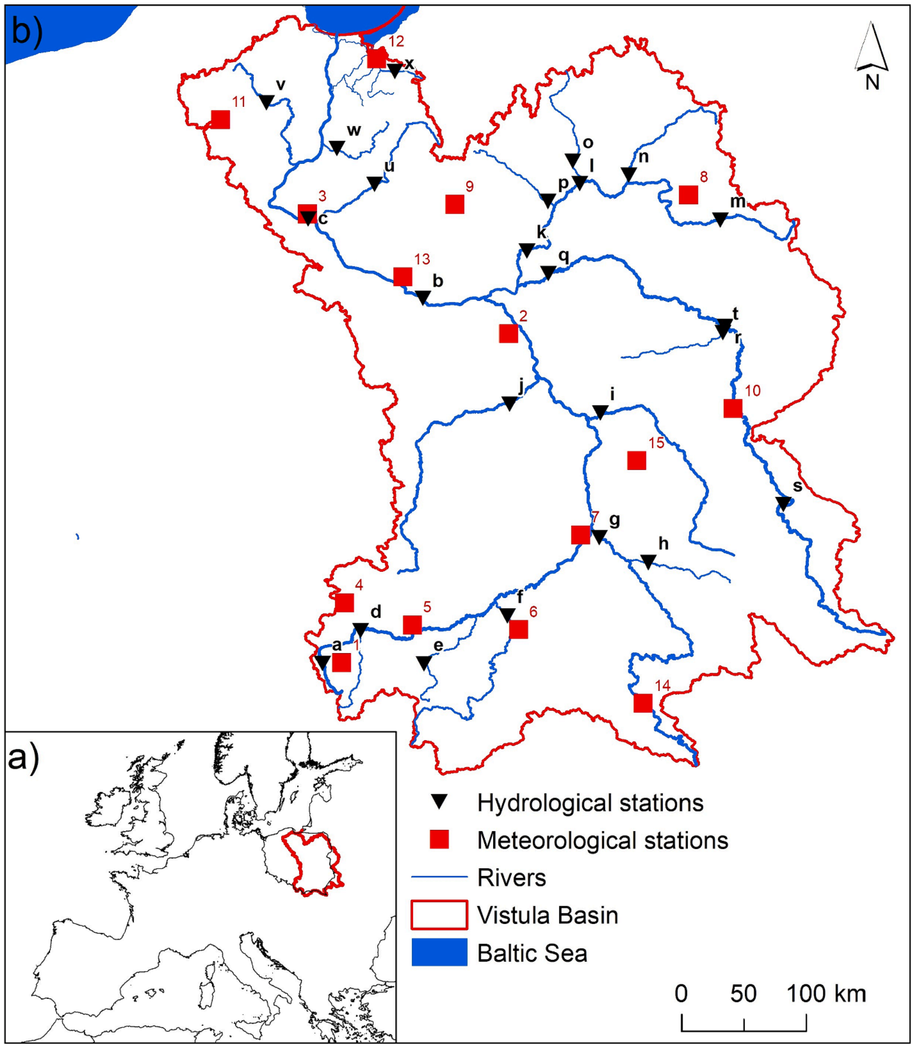

Fig. 1. Locations of the 24 studied hydrological stations and the 15 corresponding meteorological stations. The numbers and characters are corresponded with that list in Table 1.

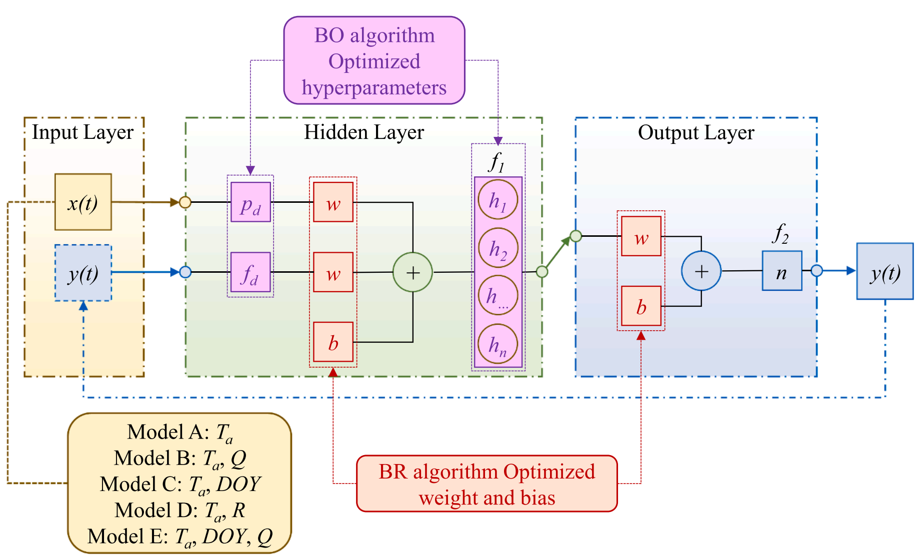

Fig. 2. Sketch of the BO-NARX-BR model: T_a – air temperature, Q – discharge, DOY – day of the year, R - rainfall.

2.2.3. Model setup and evaluation

In this study, daily dataset from 1991 to 2021 were divided into two parts: the data from 1991 to 2013 (approximately 75 % of the total dataset) were used for model calibration (training), and the remaining data from 2014 to 2021 were utilized for model validation (testing).

The BO-NARX-BR model, according to different input combinations, is divided into five model versions: (1) Model A, which used only daily air temperature (T_a) as model input, (2) Model B, which used T_a and flow discharge (Q) as model input, (3) Model C, which used T_a and the day of the year (*DOY*) as model input. Here, *DOY* is an indicator of seasonal component (see Zhu et al., 2019c; Yousefi and Toffolon, 2022). (4) Model D, which used T_a and rainfall (R) as model input, and (5) Model E, which used T_a, *DOY*, and Q as model input. As seen, Models B, C, and D are derived by adding new factors to air temperature, which serves as the most important factor for RWT modeling (Zhu et al., 2019b; Zhu and Piotrowski, 2020). Then, in the subsequent analysis, Model A is selected as the baseline. The reason for setting Model E is that this input combination is widely used in previous studies (e.g., Zhu et al., 2019c).

To assess the performance of the BO-NARX-BR and air2stream models, four widely used metrics were employed: root mean squared error (RMSE), mean absolute error (MAE), Nash-Sutcliffe efficiency coefficient (NSE), and coefficient of determination (R^2).

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(T_M^i - T_O^i\right)^2}{\sum_{i=1}^{n}\left(T_O^i - \overline{T}\right)^2} \tag{8}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}\left(T_M^i - T_O^i\right)^2}{n}} \tag{9}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|T_M^i - T_O^i\right| \tag{10}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(T_M^i - T_O^i\right)^2}{\sum_{i=1}^{n}\left(T_O^i - \overline{T}\right)^2} \tag{11}$$

where T_M^i and T_O^i are modelled and observed river water temperature for the i^{th} data, $\overline{T}$ is the average value of T_O^i, and n is the number of sample points.

Next, we will utilize these evaluation metrics to explore the impact of various input combinations on model performance.

3. Results

In this section, we compared the performance of the air2stream and BO-NARX-BR models during the calibration (training) and validation (testing) periods. The detailed results are summarized in Tables S1 to S4.

To identify the best input combination with optimal model performance in the subsequent comparisons, RMSE was used as the main indicator to analyze the BO-NARX-BR model during the training and testing phases. The detailed results are summarized in Table 2. Because all other input combinations are derived by adding new factors to air temperature, Model A is selected as the baseline, which produced a good performance (average RMSE: calibration = 0.488 °C, validation = 0.509 °C). By comparing Models B, C and D, which consider different input combinations (see Table 2, Tables S1 and S2), it is evident that Model C (average RMSE: calibration = 0.483 °C, validation = 0.502 °C) performs the best. The performance of Model D is closer to model A, suggesting that rainfall has a minor impact on the variation of RWT. As a comparison, Model E performs the worst (average RMSE: calibration = 0.530 °C, validation = 0.537 °C). Based on this observation, it is reasonable to infer that flow discharge and rainfall play a relatively minor role in regulating river thermal dynamics in the studied river basin. When introducing flow discharge, its impact on model predictions may vary based on the specific characteristics of the river. On the other hand, the introduction of *DOY* helps the model capture the seasonal variations throughout the year, thereby improving the predicted performance of the model. The detailed results are also presented using boxplots (Fig. 3), which clearly show the impact of input combinations on model performance as discussed above.

Similar to the BO-NARX-BR model, we compared the RMSE values of the five versions of the air2stream model during the calibration and validation phases (see Fig. 4, Tables S3 and S4). The version with the best performance is the a2s-5 model (average RMSE: calibration = 1.195 °C, validation = 1.249 °C). In the calibration phase, the a2s-8 and a2s-7 models have RMSE values of 1.211 °C and 1.187 °C, which are close to the a2s-5 model. However, in the validation phase, their values

Table 2

The RMSE values of the five combination inputs. The station No. corresponds with that lists in Table 1. Colorbar ranges from green (low values) to red (high values).

Station	Calibration					Validation				
	Model A	Model B	Model C	Model D	Model E	Model A	Model B	Model C	Model D	Model E
1	0.572	0.594	0.549	0.582	0.643	0.577	0.606	0.560	0.598	0.651
2	0.344	0.348	0.337	0.345	0.345	0.364	0.373	0.360	0.370	0.372
3	0.360	0.352	0.353	0.352	0.370	0.367	0.373	0.355	0.363	0.373
4	0.635	0.650	0.631	0.634	0.681	0.662	0.659	0.658	0.670	0.690
5	0.760	0.814	0.762	0.758	0.876	0.633	0.630	0.627	0.673	0.676
6	0.545	0.562	0.545	0.554	0.596	0.624	0.627	0.625	0.643	0.641
7	0.573	0.584	0.577	0.577	0.604	0.676	0.690	0.667	0.668	0.709
8	0.323	0.352	0.353	0.358	0.409	0.449	0.454	0.459	0.468	0.532
9	0.450	0.463	0.441	0.462	0.471	0.515	0.505	0.476	0.506	0.517
10	0.799	0.806	0.801	0.807	0.817	0.828	0.853	0.836	0.842	0.878
11	0.440	0.441	0.439	0.436	0.442	0.466	0.480	0.463	0.462	0.463
12	0.457	0.485	0.447	0.455	0.509	0.470	0.453	0.464	0.476	0.481
13	0.544	0.561	0.535	0.547	0.614	0.559	0.544	0.554	0.566	0.579
14	0.433	0.452	0.435	0.445	0.472	0.456	0.445	0.464	0.484	0.468
15	0.412	0.403	0.401	0.407	0.423	0.432	0.421	0.412	0.419	0.445
16	0.665	0.682	0.652	0.666	0.718	0.682	0.677	0.682	0.688	0.712
17	0.452	0.452	0.426	0.451	0.447	0.464	0.484	0.453	0.477	0.482
18	0.555	0.561	0.551	0.559	0.572	0.573	0.621	0.570	0.583	0.623
19	0.424	0.431	0.428	0.431	0.444	0.441	0.448	0.437	0.446	0.454
20	0.671	0.691	0.665	0.676	0.731	0.727	0.737	0.721	0.731	0.760
21	0.423	0.425	0.411	0.418	0.448	0.440	0.429	0.432	0.443	0.459
22	0.395	0.400	0.392	0.391	0.441	0.480	0.467	0.454	0.465	0.468
23	0.238	0.245	0.225	0.256	0.292	0.133	0.127	0.118	0.121	0.155
24	0.261	0.248	0.240	0.235	0.339	0.202	0.214	0.195	0.186	0.296

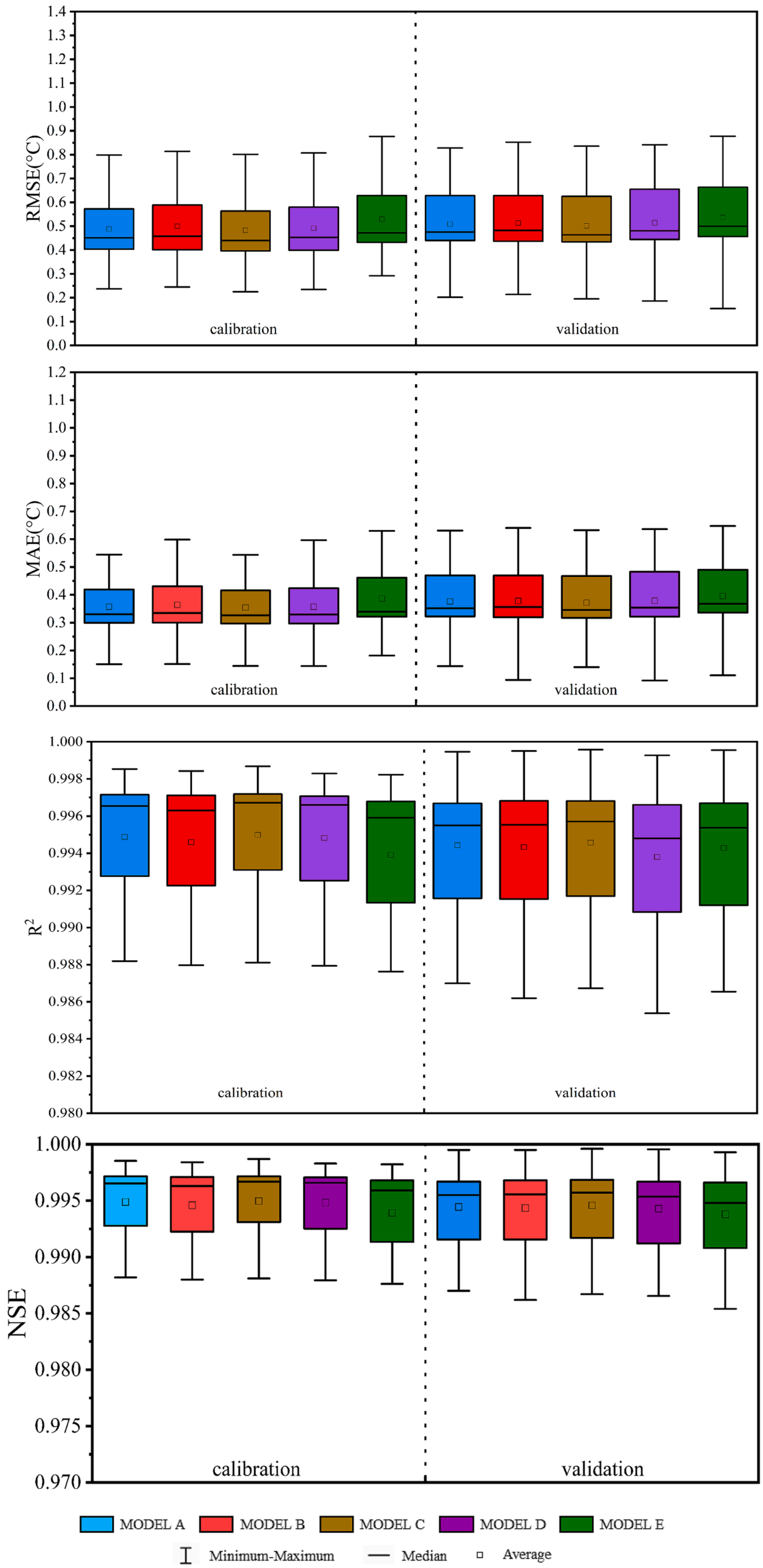

Fig. 3. Impact of input combinations on model performance of the BO-NARX-BR model.

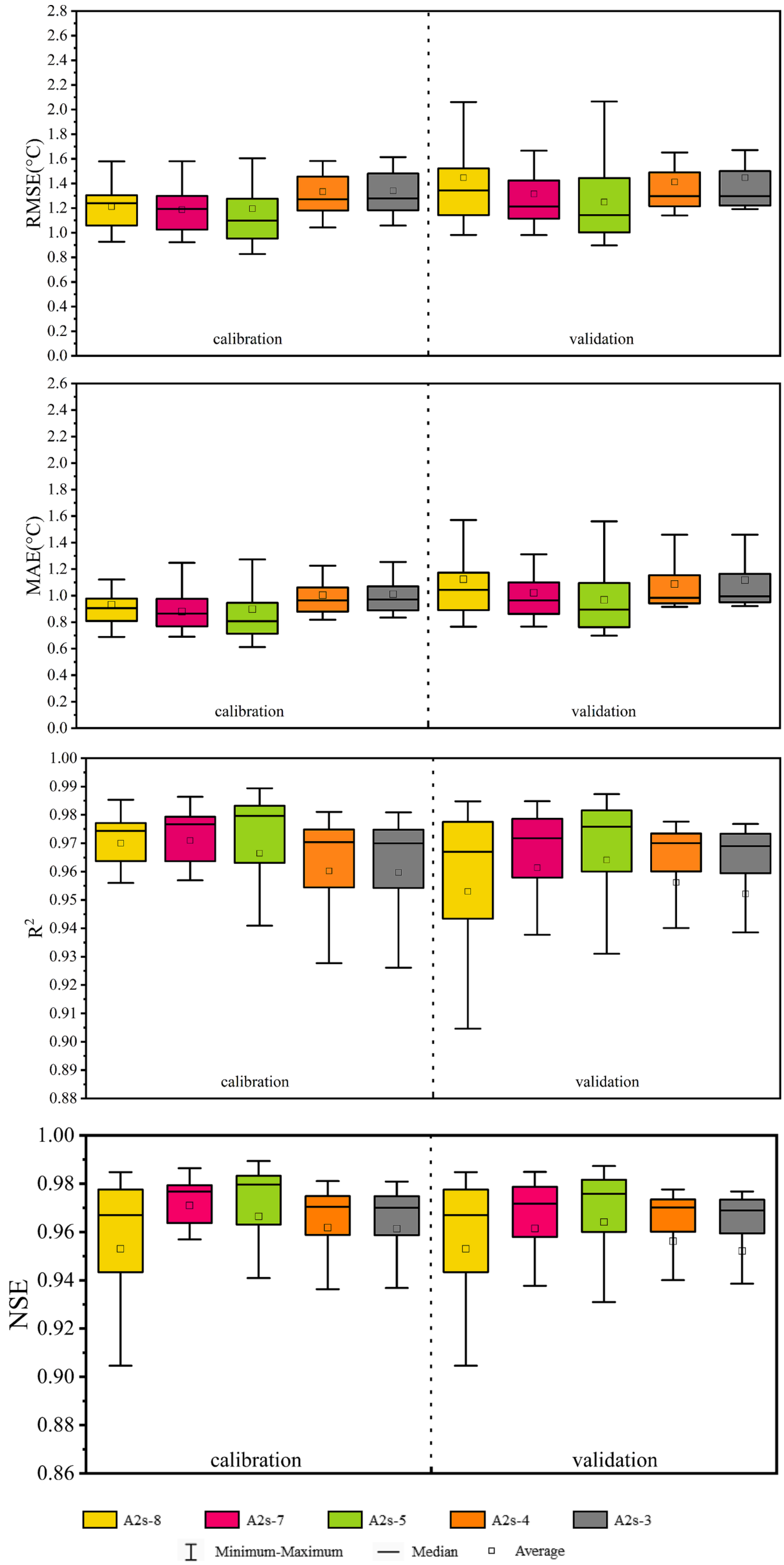

Fig. 4. Performance of different versions of the air2stream model.

are 1.448 °C and 1.320 °C respectively. This indicates that the predictive capabilities of the a2s-8 and a2s-7 models are worse than the a2s-5 model. Additionally, the performance of both the a2s-3 and a2s-4 models is poor during both the calibration phase (RMSE: 1.333 °C and 1.341 °C, respectively) and the validation phase (RMSE: 1.411 °C and 1.448 °C, respectively).

For the BO-NARX-BR model which exhibited superior performance (Model C), the training phase manifested RMSE values within the range of 0.225 °C to 0.801 °C (average: 0.483 °C), MAE values varying from 0.145 °C to 0.600 °C (average: 0.354 °C), R^2 values spanning 0.985 to 0.999 (average: 0.995), and all NSE values surpassing 0.985 (range: 0.985–0.999, average: 0.995). The model's robust performance extends to the testing phase, with RMSE ranging from 0.118 °C to 0.836 °C (average: 0.502 °C), MAE fluctuating between 0.089 °C and 0.633 °C (average: 0.372 °C), R^2 values ranging from 0.987 to 1.0 (average: 0.995), and NSE varying between 0.987 and 1.0 (average: 0.995).

In stark contrast, the widely employed air2stream model (specifically the a2s-5 model version) exhibited notably inferior performance. During the calibration period, the air2stream model registered RMSE values ranging from 0.827 °C to 2.830 °C (average: 1.195 °C), MAE values varying between 0.612 °C and 2.348 °C (average: 0.898 °C), and NSE values spanning 0.820 to 0.989 (average: 0.966). As the analysis extended to the validation period, the performance disparity between the air2stream model and the BO-NARX-BR model became more evident (air2stream model: RMSE range 0.897–2.065 °C, average: 1.249 °C; MAE range 0.698–1.634 °C, average: 0.969 °C; NSE range 0.886–0.987, average: 0.964).

These findings unequivocally indicate that the BO-NARX-BR model significantly outperforms the air2stream model in the prediction of daily river water temperature. This can be better seen from Fig. 5, which presents the predicted and observed daily values of RWT for three rivers, including Vistula River at Kępa Polska (a large-sized river, character "b" at Fig. 1), Narew River at Zambski Kościelne (a medium-sized river, character "k" at Fig. 1), and Wąska River at Pasłęk (a small-sized river, character "x" at Fig. 1). As shown in the figure, the NARX-based model significantly outperforms the air2stream model in capturing the seasonal pattern and the peak values of RWT, and the predicted values of the air2stream model are more scattered, while the predicted RWT of the NARX-based model is closer to the 1:1 line.

Fig. 6 presents the predicted and observed annual average values of RWT for three rivers, including Vistula River at Kępa Polska (a large-sized river, character "b" at Fig. 1), Narew River at Zambski Kościelne (a medium-sized river, character "k" at Fig. 1), and Wąska River at Pasłęk (a small-sized river, character "x" at Fig. 1). As shown in the figure, the NARX-based model can well reproduce the warming trend of RWT, with the nearly overlapped trend lines and annual average RWT values.

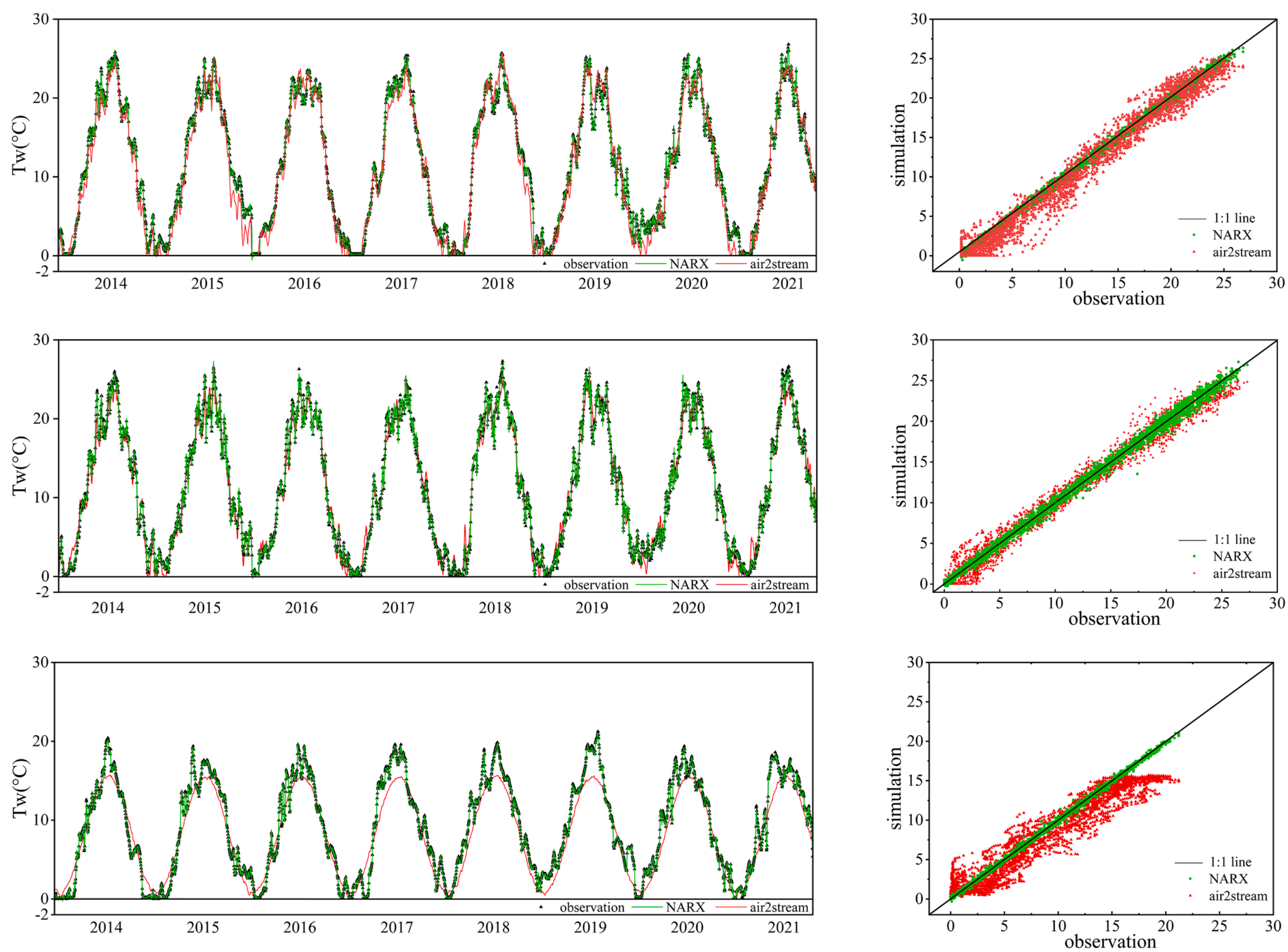

Fig. 5. The predicted and observed daily values of water temperatures (Tw) for three rivers: Vistula River at Kępa Polska (upper), Narew River at Zambski Kościelne (middle), and Wąska River at Pasłęk (bottom).

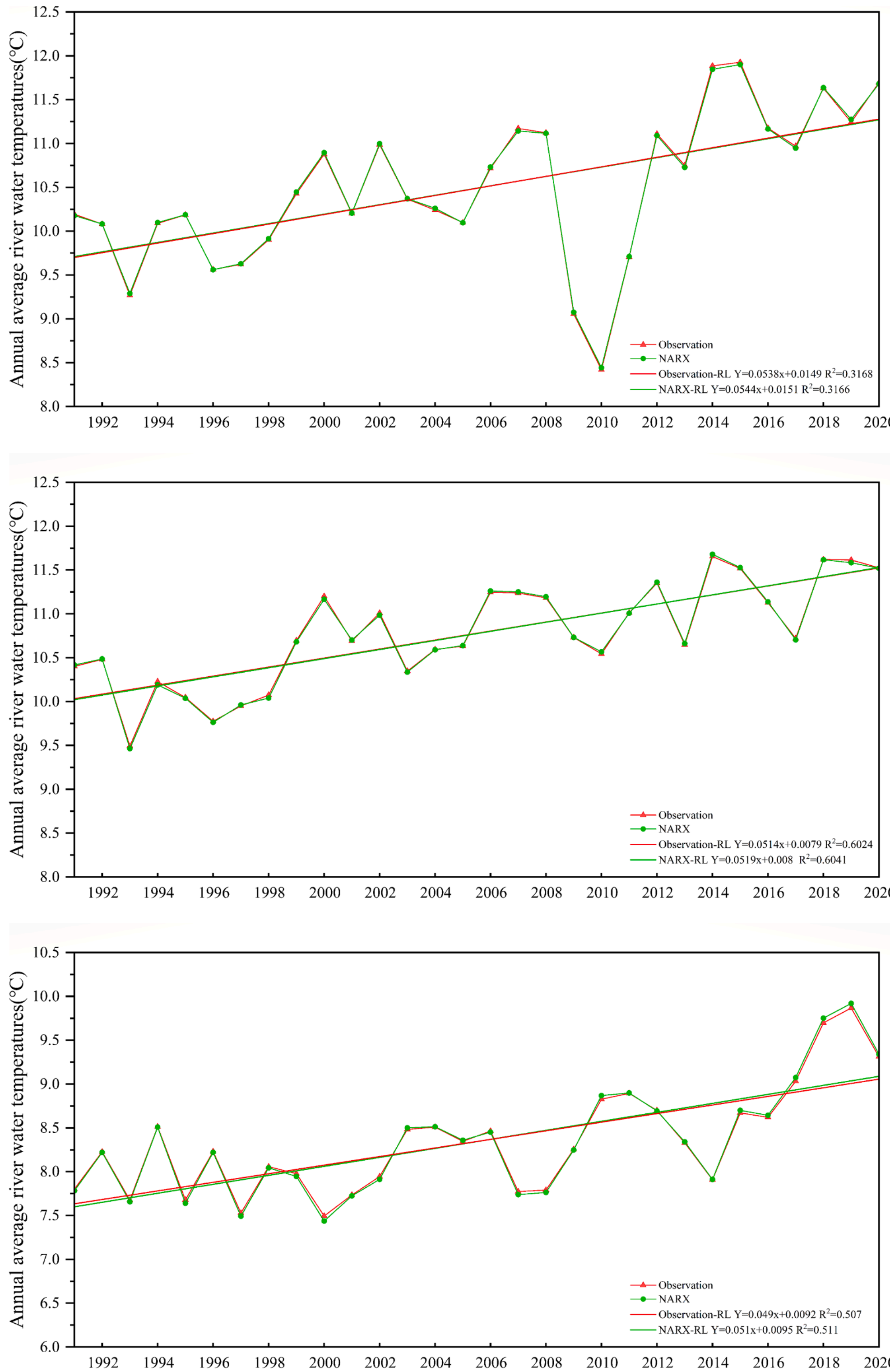

Fig. 6. The predicted and observed annual average values of river water temperatures for three rivers: Vistula River at Kępa Polska (upper), Narew River at Zambski Kościelne (middle), and Wąska River at Pasłęk (bottom).

4. Discussion

4.1. Selection of air2stream model versions

As shown in the modeling results, overall, the version with the best performance is the a2s-5 model version (average RMSE: calibration = 1.195 °C, validation = 1.249 °C), slightly outperforming the a2s-8 model version. The results are consistent with that in Piccolroaz et al. (2016), assessing the model performance in 38 Swiss rivers. The results indicated that the full model version (a2s-8) considering more impact factors (e.g., flow) might not the best option, and the simplified 5 parameters model version (a2s-5) by disregarding the impact of fluctuating flow discharge might be a better choice. For individual rivers, the a2s-8 version might outperform other versions, including a2s-5, for example, the station No. 8, as shown in Tables S3 and S4, which indicates that the selection of the appropriate model version of the air2stream model might depend on the studied rivers of interest.

4.2. Impact of input combinations on model performance

Based on the results of model validation and verification, it can be observed that all five input combinations perform generally well for the 24 hydrological stations (RMSE < 0.9 °C, MAE < 0.7 °C, NSE > 0.98), especially when comparing with the air2stream model. Furthermore, regardless of the input combination, the BO-NARX-BR model exhibits better performance during the training phase compared to the testing period. Overall, the BO-NARX-BR model demonstrates outstanding performance in predicting daily river water temperatures.

However, when analyzing the performance of different input combinations, differences among the five combinations can be identified. The performance of Model D and Model A are nearly identical, suggesting that rainfall plays a secondary role in regulating river thermodynamics. In comparison, Models A and B show minor differences in performance during both the training and testing periods. This shows that adding streamflow as model input didn't help to improve model performance, indicating that streamflow plays a minor role in controlling river thermal dynamics in the Vistula River Basin. Previous studies showed that streamflow played a relevant role mainly in snow-fed and regulated rivers with hydropower reservoirs, however, for lowland rivers, it improved to a lower extent in model performance (see Zhu et al., 2019b). The rivers in the Vistula River Basin, evaluated in this study, are typical lowland rivers, and the modeling results further indicated that streamflow played a minor role in these rivers.

Additionally, compared to Model A, Model C, with the introduction of *DOY*, though yielding insignificant improvement in average RMSE, outperformed other input combinations in most cases among the studied rivers, which is consistent with the modeling results in Heddam et al., (2023a) and Zhu et al., (2019c) for RWT modeling and Yousefi and Toffolon (2022) for the forecasting of lake surface water temperatures. Note that Model E with three inputs (T_a, Q and *DOY*), which has shown the best performance in previous studies (e.g., Zhu et al., 2019c), performs the worst in this study. This result might indicate that choice of the best input combination depends on the studied rivers of interest.

4.3. Potential applications of the proposed model on climate change studies

In the context of climate change studies, procuring future time series data for air temperatures poses minimal challenges, given the ready accessibility of such information through climate models like EURO-CORDEX (e.g., Piccolroaz et al., 2021). Additionally, the incorporation of *DOY* into analyses is a straightforward consideration. However, when it comes to other critical factors, such as streamflow, ensuring the reliability of future streamflow values becomes a formidable challenge. This study delves deeper into unravelling the repercussions of climate change on river thermal dynamics within the Vistula River Basin, which has shown significant warming trend in the past decades (Ptak et al., 2022). The modeling outcomes presented herein underscore the robustness and reliability of the proposed BO-NARX-BR model. Notably, T_a and DOY emerge as the two pivotal inputs crucial for modeling RWT within the Vistula River Basin.

The results unveiled through our study not only affirm the efficacy of the BO-NARX-BR model but also position it as a promising tool for meticulously investigating the ramifications of climate change on river thermal dynamics. This nuanced approach allows for a more comprehensive understanding of the interplay between various environmental factors, offering valuable insights for both scientific research and practical applications in the context of climate change impact assessment within river ecosystems.

In extending the application of data-driven modeling to long-term predictions, considerations arise regarding the inherent limitations associated with models trained on historical data. The BO-NARX-BR model, showcased for its effectiveness in short to medium-term projections, may face challenges in accurately forecasting scenarios significantly deviating from observed conditions. This limitation is particularly relevant when confronted with the dynamic nature of climate change.

Considering climate models like EURO-CORDEX as a strategy to enhance the reliability of long-term predictions could be a key factor. These models project future scenarios, providing critical input data for variables such as air temperatures under different Representative Concentration Pathway (e.g., RCP 4.5 and RCP 8.5, Di Nunno and Granata, 2023). Therefore, a combined approach, based on data-driven models and climate model projections, can bridge the gap between historical context and future expectations. It could mitigate potential errors associated with relying solely on historical data, offering a more comprehensive understanding of the evolving environmental conditions.

5. Conclusions

This study employed an optimized model based on NARX networks for predicting daily river water temperatures for 18 rivers of the Vistula River Basin, one of the largest rivers in Europe. Model performance was compared with another widely used model air2stream. The modeling results lead to the following conclusions:

(1) The BO-NARX-BR model significantly outperforms the air2stream model in the calibration and validation phases, and the model can better capture the seasonal pattern and the peak values of RWT. The coupling of BO and BR with the NARX model in the BO-NARX-BR framework significantly enhances the modeling process. BO strategically selects input variables and fine-tunes hyperparameters, such as hidden nodes and lagged values, ensuring the effectiveness of the NARX model. Simultaneously, BR optimizes weight and bias parameters, crucial for training algorithms, thereby ensuring the robustness of the overall model.
(2) The full model version considering more factors might not be the best option, and the selection of the appropriate model version of the air2stream model might depend on the studied rivers of interest.
(3) Input combinations impact the performance of the BO-NARX-BR model in daily RWT modeling. Air temperature and *DOY* can work as the major model inputs, and streamflow and rainfall play a minor role in modeling river water temperatures at the Vistula River Basin.
(4) The BO-NARX-BR model is a promising tool to investigate the impact of climate change on river thermal dynamics.

6. Data and code availability

The codes of the BO-NARX-BR model are available at: https://github.com/Fabiodinunno1989/LSWT-prediction/tree/main. The observed data for one river station (River WISŁA at SKOCZÓW) are available at: https://github.com/slzhu1989/Lake-heatwaves.

CRediT authorship contribution statement

Jiang Sun: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation. **Fabio Di Nunno:** Writing – review & editing, Writing – original draft, Methodology, Investigation. **Mariusz Sojka:** Investigation, Methodology, Writing – original draft, Writing – review & editing. **Mariusz Ptak:** Investigation, Methodology, Writing – original draft, Writing – review & editing. **You Luo:** Writing – review & editing, Writing – original draft. **Renyi Xu:** Writing – review & editing, Writing – original draft. **Jing Xu:** Writing – review & editing, Writing – original draft. **Yi Luo:** Writing – review & editing, Writing – original draft, Funding acquisition. **Senlin Zhu:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Data curation, Conceptualization. **Francesco Granata:** Writing – review & editing, Writing – original draft, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors acknowledge the Institute of Meteorology and Water Management–National Research Institute in Poland for providing the data used in this study. This study is funded by the Yunnan Province Innovation Team Project (202305AS350003), and Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJB170023).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ecolind.2024.111978.

References

Almeida, M.C., Coelho, P.S., 2023. Modeling river water temperature with limiting forcing data: Air2stream v1.0.0, machine learning and multiple regression. Geosci. Model Dev. 16 (14), 4083–4112.
Alsumaiei, A.A., 2020. A nonlinear autoregressive modeling approach for forecasting groundwater level fluctuation in urban aquifers. Water 12 (3), 820.
Benyahya, L., Caissie, D., St-Hilaire, A., Ouarda, T.B.M.J., Bobée, B., 2017. A review of statistical water temperature models. Canadian Water Resources Journal 32 (3), 179–192.
Boussaada, Z., Curea, O., Remaci, A., Camblong, H., Mrabet Bellaaj, N., 2018. A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation. Energies 11 (3), 620.
Cai, X., Yao, L., Hu, Y., Wang, S., Wang, Z., Dahlgren, R.A., 2023. Water temperature and organic carbon control spatio-temporal dynamics of particle-attached and free-living bacterial communities in a hypereutrophic urban river network. Freshw. Biol. 68 (9), 1627–1645.
Desouky, M.A.A., Abdelkhalik, O., 2019. Wave prediction using wave rider position measurements and NARX network in wave energy conversion. Appl. Ocean Res. 82, 10–21.
Di Nunno, F., Granata, F., 2020. Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network. Environ. Res. 190.
Di Nunno, F., Granata, F., 2023. Future trends of reference evapotranspiration in Sicily based on CORDEX data and Machine Learning algorithms. Agric Water Manag 280 (1–2).
Di Nunno, F., Granata, F., Gargano, R., de Marinis, G., 2021a. Forecasting of extreme storm tide events using NARX neural network-based models. Atmos. 12 (4), 512.
Di Nunno, F., Granata, F., Gargano, R., de Marinis, G., 2021b. Prediction of spring flows using nonlinear autoregressive exogenous (NARX) neural network models. Environ. Monit. Assess. 193 (6), 350.
Di Nunno, F., Race, M., Granata, F., 2022. A nonlinear autoregressive exogenous (NARX) model to predict nitrate concentration in rivers. Environ. Sci. Pollut. Res. 29, 40623–40642.
Di Nunno, F., Zhu, S., Ptak, M., Sojka, M., Granata, F., 2023. A stacked machine learning model for multi-step ahead prediction of lake surface water temperature. Sci. Total Environ. 890, 164323.
Dugdale, S.J., Hannah, D.M., Malcolm, I., 2017. River temperature modelling: a review of process-based approaches and future directions. Earth Sci. Rev. 175, 97–113.
Gao, M., Wu, Q., Li, J., et al., 2023. Temperature prediction of solar greenhouse based on NARX regression neural network. Sci. Rep. 13 (1), 1563.
Gatien, P., Arsenault, R., Martel, J., St-Hilaire, A., 2023. Using the ERA5 and ERA5-Land reanalysis datasets for river water temperature modelling in a data-scarce region. Canadian Water Resources Journal 48 (2), 93–110.
Graf, R., Zhu, S., Sivakumar, B., 2019. Forecasting river water temperature time series using a wavelet–neural network hybrid modelling approach. J. Hydrol. 578, 124115.
Heddam, S., Kim, S., Mehr, A.D., Zounemat-Kermani, M., Ptak, M., Elbeltagi, A., Malik, A., Tikhamarine, Y., 2023a. Bat algorithm optimised extreme learning machine (Bat-ELM): A novel approach for daily river water temperature modelling. Geogr. J. 189 (1), 78–89.
Heddam, S., Merabet, K., Difi, S., Kim, S., Ptak, M., Sojka, M., Zounemat-Kermani, M., Kisi, O., 2023b. River water temperature prediction using hybrid machine learning coupled signal decomposition: EWT versus MODWT. Eco. Inform. 78, 102376.
Ikram, R.M.A., Mostafa, R.R., Chen, Z., Parmar, K.S., Kisi, O., Zounemat-Kermani, M., 2023. Water temperature prediction using improved deep learning methods through reptile search algorithm and weighted mean of vectors optimizer. Journal of Marine Science and Engineering 11, 259.
Kajak, Z., 1993. The Vistula River and its riparian zones. Hydrobiologia 251, 149–157.
Lee, C.C., Sheridan, S.C., 2018. A new approach to modeling temperature-related mortality: non-linear autoregressive models with exogenous input. Environ. Res. 164, 53–64.
Michel, A., Schaefli, B., Wever, N., Zekollari, H., Lehning, M., Huwald, H., 2022. Future water temperature of rivers in Switzerland under climate change investigated with physics-based models. Hydrol. Earth Syst. Sci. 26 (4), 1063–1087.
Mohseni, O., Stefan, H.Z., 1999. Stream temperature/air temperature relationship: a physical interpretation. J. Hydrol. 218 (3), 128–141.
Ouellet, V., St-Hilaire, A., Dugdale, S.J., Hannah, D.M., Krause, S., Ouellet-Proulx, S., 2020. River temperature research and practice: Recent challenges and emerging opportunities for managing thermal habitat conditions in stream ecosystems. Sci. Total Environ. 736, 139679.
Piccolroaz, S., Calamita, E., Majone, B., Gallice, A., Siviglia, A., Toffolon, M., 2016. Prediction of river water temperature: a comparison between a new family of hybrid models and statistical approaches. Hydrol. Process. 30 (21), 3901–3917.
Piccolroaz, S., Zhu, S., Ptak, M., Sojka, S., Du, X., 2021. Warming of lowland Polish lakes under future climate change scenarios and consequences for ice cover and mixing dynamics. J. Hydrol.: Reg. Stud. 34, 100780.
Piotrowski, A.P., Napiórkowski, J.J., 2018. Performance of the air2stream model that relates air and stream water temperatures depends on the calibration method. J. Hydrol. 561, 395–412.
Piotrowski, A.P., Napiórkowski, J.J., 2019. Simple modifications of the nonlinear regression stream temperature model for daily data. J. Hydrol. 572, 308–328.
Ptak, M., Sojka, M., Graf, R., Choiński, A., Zhu, S., Nowak, B., 2022. Warming of the Vistula River – Effect of Climate and Local Conditions on the Scale of the Process in one of the Largest Rivers in Europe. Journal of Hydrology and Hydromechanics 70, 1–11.
Rehana, S., Rajesh, M., 2023. Assessment of impacts of climate change on Indian riverine thermal regimes using hybrid deep learning methods. Water Resources Research, 59 (2), e2021WR031347.
Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian optimization of machine learning algorithms. In Advances in Neural Information Processing Systems 2951–2959.
Ochrona Środowiska. Główny Urząd Statystyczny, 2022, Warszawa.
Tockner, K., Tonolla, D., Uehlinger, U., Siber, R., Robinson, C.T., Peter, F.D., 2009. Introduction to European Rivers. Rivers of Europe, 1-21. Academic Press, San Diego.
Toffolon, M., Piccolroaz, S., 2015. A hybrid model for river water temperature as a function of air temperature and discharge. Environ. Res. Lett. 10 (11), 114011.
White, J.C., Khamis, K., Dugdale, S., Jackson, F.L., Malcolm, I.A., Krause, S., Hannah, D. M., 2023. Drought impacts on river water temperature: A process-based understanding from temperate climates. Hydrol. Process. 37 (10), e14958.
Yousefi, A., Toffolon, M., 2022. Critical factors for the use of machine learning to predict lake surface water temperature. J. Hydrol. 606 (5), 127418.
Zhi, W., Ouyang, W., Shen, C., Li, L., 2023a. Temperature outweighs light and flow as the predominant driver of dissolved oxygen in US rivers. Nature Water 1, 249–260.
Zhi, W., Klingler, C., Liu, J., Li, L., 2023b. Widespread deoxygenation in warming rivers. Nat. Clim. Chang. 13, 1105–1113.
Zhu, S., Bonacci, O., Oskoruš, D., Hadzima-Nyarko, M., Wu, S., 2019a. Long term variations of river temperature and the influence of air temperature and river discharge: case study of Kupa River watershed in Croatia. Journal of Hydrology and Hydromechanics 67 (4), 305–313.
Zhu, S., Heddam, S., Nyarko, E.K., Hadzima-Nyarko, M., Piccolroaz, S., Wu, S., 2019b. Modeling daily water temperature for rivers: comparison between adaptive neuro-fuzzy inference systems and artificial neural networks models. Environ. Sci. Pollut. Res. 26, 402–420.
Zhu, S., Heddam, S., Wu, S., Dai, J., Jia, B., 2019c. Extreme learning machine-based prediction of daily water temperature for rivers. Environ. Earth Sci. 78, 202.
Zhu, S., Luo, Y., Graf, R., et al., 2022. Reconstruction of long-term water temperature indicates significant warming in Polish rivers during 1966–2020. J. Hydrol.: Reg. Stud. 44, 101281.
Zhu, S., Di Nunno, F., Ptak, M., Sojka, M., Granata, F., 2023. A novel optimized model based on NARX networks for predicting thermal anomalies in Polish lakes during heatwaves, with special reference to the 2018 heatwave. Sci. Total Environ. 905, 167121.
Zhu, S., Piotrowski, A.P., 2020. River/stream water temperature forecasting using artificial intelligence models: a systematic review. Acta Geophys. 68, 1433–1442.
Zwart, J.A., Oliver, S.K., Watkins, W.D., et al., 2023. Near-term forecasts of stream temperature using deep learning and data assimilation in support of management decisions. J. Am. Water Resour. Assoc. 59 (2), 317–337.