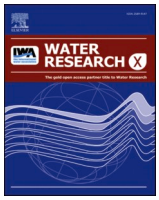
Water Research X 23 (2024) 100228

Contents lists available at ScienceDirect

Water Research X

journal homepage: www.sciencedirect.com/journal/water-research-x



Spatiotemporal estimation of groundwater and surface water conditions by integrating deep learning and physics-based watershed models

Soobin Kim [a], Eunhee Lee [b], Hyoun-Tae Hwang [c,d], JongCheol Pyo [e], Daeun Yun [f], Sang-Soo Baek [g], Kyung Hwa Cho [h,*]

[a] Disposal Safety Evaluation R&D Division, Korea Atomic Energy Research Institute (KAERI), 111, Daedeok-daero 989 beon-gil, Yuseong-gu, Daejeon 34057, Republic of Korea
[b] Korea Institute of Geoscience and Mineral Resources, 124 Gwahak-ro, Yuseong-gu, Daejeon 34132, Republic of Korea
[c] Aquanty, Inc., 600 Weber St. N., Unit B, Waterloo, ON N2V 1K4, Canada
[d] Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
[e] Department of Environmental Engineering, Pusan National University, Busandaehak-ro 63 beon-gil 2, Geumjeong-gu, Busan 46241, South Korea
[f] School of Civil, Urban, Earth, and Environmental Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Ulju-gun, Ulsan 44919, Republic of Korea
[g] Department of Environmental Engineering, Yeungnam University, 280 Daehak-Ro, Gyeongsan-Si, Gyeongbuk 38541, South Korea
[h] School of Civil, Environmental, and Architectural Engineering, Korea University, Seoul 02841, South Korea

ARTICLE INFO

Keywords:
Climate change impact
Convolutional neural network
Fully distributed hydrological model
HydroGeoSphere

ABSTRACT

The impacts of climate change on hydrology underscore the urgency of understanding watershed hydrological patterns for sustainable water resource management. The conventional physics-based fully distributed hydrological models are limited due to computational demands, particularly in the case of large-scale watersheds. Deep learning (DL) offers a promising solution for handling large datasets and extracting intricate data relationships. Here, we propose a DL modeling framework, incorporating convolutional neural networks (CNNs) to efficiently replicate physics-based model outputs at high spatial resolution. The goal was to estimate groundwater head and surface water depth in the Sabgyo Stream Watershed, South Korea. The model datasets consisted of input variables, including elevation, land cover, soil type, evapotranspiration, rainfall, and initial hydrological conditions. The initial conditions and target data were obtained from the fully distributed hydrological model HydroGeoSphere (HGS), whereas the other inputs were actual measurements in the field. By optimizing the training sample size, input design, CNN structure, and hyperparameters, we found that CNNs with residual architectures (ResNets) yielded superior performance. The optimal DL model reduces computation time by 45 times compared to the HGS model for monthly hydrological estimations over five years (RMSE 2.35 and 0.29 m for groundwater and surface water, respectively). In addition, our DL framework explored the predictive capabilities of hydrological responses to future climate scenarios. Although the proposed model is cost-effective for hydrological simulations, further enhancements are needed to improve the accuracy of long-term predictions. Ultimately, the proposed DL framework has the potential to facilitate decision-making, particularly in large-scale and complex watersheds.

1. Introduction

In recent decades, urbanization and climate change have exacerbated water stress in groundwater and surface water systems (Kaandorp et al., 2018; Mehran et al., 2017; Taylor et al., 2013). Climate change has led to changes in precipitation patterns and an increased frequency of extreme weather events (Papalexiou and Montanari, 2019; Schwartz and Randall, 2003). To address these issues, hydrological variations in watersheds must be estimated (Apurv and Cai, 2020; Delpla et al., 2009; Inyinbor Adejumoke et al., 2018). Physics-based, fully distributed hydrological models (such as, ParFLow and HydroGeoSphere) are valuable for simulating intricate watershed hydrological conditions (Hwang et al., 2014; Loague et al., 2006; Maxwell et al., 2015; VanderKwaak and Loague, 2001) and excel in capturing spatiotemporal characteristics and

* Corresponding author.
E-mail address: khcho80@korea.ac.kr (K.H. Cho).

https://doi.org/10.1016/j.wroa.2024.100228
Received 26 October 2023; Received in revised form 7 May 2024; Accepted 8 May 2024
Available online 16 May 2024

key hydrological processes, such as evapotranspiration and groundwater-surface water interactions (Guevara Ochoa et al., 2020; Sudicky et al., 2008; Vieux et al., 2004).

However, the application of fully distributed integrated models can be problematic due to their data intensity and computational expense, particularly for large-scale watersheds. These challenges arise from the complexity of surface water and groundwater models, specifically the inherent non-linearities (for example, Saint-Venant and Richards' equations) that are present in hydrogeological and hydrological systems (Buitink et al., 2020; Ekmekcioğlu et al., 2022; Ocio et al., 2019; Sinha et al., 2022). Therefore, the development of computationally efficient and simplified approaches is crucial for effective water resource management. Strategies such as parallel computing and spatial/temporal resolution adjustments have been proposed to alleviate the computational burden (Vivoni et al., 2011; Wang et al., 2018). Parallel computing distributes computations across multiple processors or cores, notably reducing processing time (Kollet et al., 2010; Wu et al., 2002). Adjusting the spatial and temporal resolutions can optimize the balance between model accuracy and computational demands (Wang et al., 2018).

Although these methods offer some solutions, physics-based modeling approaches may be somewhat limited when dealing with highly complex systems that require extensive input data and domain expertise (Alvi et al., 2023). Consequently, data-driven deep learning (DL) techniques are now viewed as a promising alternative compared with conventional hydrologic model simulations (Sabzipour et al., 2023; Shen and Lawson, 2021). DL algorithms can learn intricate relationships and features autonomously (Szegedy et al., 2015; Zhou et al., 2017), simplify complex dynamics, such as hydrogeological processes, handle vast datasets efficiently (L'heureux et al., 2017; Rasp et al., 2018; Talukdar et al., 2023), and provide rapid and accurate predictions of hydrological conditions. (Cheng et al., 2007; L'heureux et al., 2017).

In the field of hydrology, DL algorithms, such as long short-term memory networks (LSTMs) (Kratzert et al., 2018; Sabzipour et al., 2023), convolutional neural networks (CNNs) (Leonarduzzi et al., 2022; Liao et al., 2023), and graph neural networks (GNNs) (Gai et al., 2023; Liu et al., 2022b) have been widely applied in previous research. Maqsood et al. (2022) utilized LSTM and CNN to estimate the variation in evapotranspiration under climate change scenarios, and Gai et al. (2023) employed GNNs to simulate spring discharge and capture the spatial dependence of groundwater propagation and precipitation infiltration. LSTMs effectively process sequential and temporal data, addressing the vanishing gradient problem (Abbas et al., 2023; Kratzert et al., 2018; Pyo et al., 2023). CNNs capture spatiotemporal patterns from multi-dimensional data and the nonlinearity inherent within hydrological processes (Chen et al., 2020; Maqsood et al., 2022; Prasad et al., 2022). Additionally, GNNs demonstrate proficiency in handling graph-structured data, such as river networks (Bentivoglio et al., 2023; Sun et al., 2022).

However, hydrological research using DL modeling has typically focused on conducting time-series simulations at specific monitoring stations (Lange and Sippel, 2020; Maqsood et al., 2022; Shen et al., 2021). In addition, enhancing the grid resolution of DL modeling for spatial analysis is required (Iqbal et al., 2022; Lim and Wang, 2022; Liu et al., 2022a; Wang et al., 2022). Sun et al. (2019) integrated a physics-based model with deep CNNs to predict spatiotemporal groundwater storage anomalies; the authors emphasized the necessity for further investigations employing higher-resolution grids to achieve outputs with heightened precision, as the grid resolution used in their study was relatively coarse. Therefore, we propose a DL framework utilizing the CNN algorithm to effectively simulate high-resolution spatiotemporal groundwater and surface water distributions on a watershed scale. Few studies have combined DL with fully distributed watershed models to enhance the efficiency and resolution of hydrological simulations.

Our focus was on estimating the groundwater heads and surface water depths in the Sabgyo Stream Watershed, South Korea (Fig. S1). We employed HydroGeoSphere (HGS) (Aquanty Inc., 2023), a physics-based, fully distributed hydrological model (Fig. 1A), and generated DL model datasets using topographical, geographical, and hydrometeorological information from field observations and HGS simulations (Fig. 1B and Table 1). We then examined the effect of spatial data quantities (n) on DL model performance (Fig. 1C) and optimized the input data design, CNN architecture, and hyperparameters, which are crucial factors affecting DL model performance and computational cost (Bilali et al., 2021; Justus et al., 2018). We compared the optimal CNN model estimates with those generated by the HGS model (Fig. 1F) and assessed the accuracy of our model in predicting hydrological responses under future climate scenarios (Fig. 1G). We propose that this novel DL approach offers a computationally efficient solution in high-spatial-resolution hydrological simulations. These findings can aid in decision-making regarding complex and long-term hydrological phenomena in large-scale watersheds.

2. Results and discussion

2.1. DL performance with respect to training and validation

Our DL modeling framework was developed to efficiently reproduce the spatiotemporal flow conditions derived from a fully distributed hydrological model (HGS). We estimated the regular patterns of groundwater heads and surface water depths in the Sabgyo Stream Watershed. We assessed the computational efficiency of the DL model by training it with a smaller dataset compared with that required for constructing the HGS model. Specifically, we varied the spatial data quantity (n) of the datasets (Fig. S2), which were generated by extracting data from 300, 500, and 1000 grid cells within the watershed (Fig. 1C). When estimating groundwater heads, the DL model achieved the root mean square error (RMSE) of 1.70–3.39 m and coefficient of determination (R^2) of 0.999–1.000 for the training dataset. In the validation dataset, the RMSE ranged from 2.98 to 4.93 m with an R^2 of 0.999 (Fig. S2A). When estimating surface water depths, the DL model achieved an RMSE of 0.04–0.08 m and R^2 of 0.997–1.000 for the training set. In the validation dataset, the RMSE was 0.10–0.26 m with an R^2 of 0.970–0.995 (Fig. S2B). The RMSE values provide insight into the level of deviation between the predicted values generated by our DL models and the physics-based modeling results. In practical terms, higher RMSE values indicate larger prediction errors, suggesting that the predictions generated by the DL model may deviate more significantly from the ground truth.

In general, the performance of the DL model improved as n increased and converged at a certain point. Dataset size is a crucial factor affecting DL model performance and computational cost (Heinen et al., 2020; Xu and Goodacre, 2018). Here, the groundwater model converged when n was 1000, whereas the surface water model converged when n exceeded 500 (Fig. S2). The results suggest that the DL input resolutions were 1.65 km^2 for simulating the groundwater heads and 0.20 km^2 for the surface water depths in the Sabgyo Stream Watershed. These values were calculated by dividing the watershed drainage area (1650 km^2) and the stream network area (102 km^2) by their respective optimal n. This information provides insights into the impact of the dataset size on the spatial distribution estimation of hydrological components. Previous studies have demonstrated that the accuracy of neural networks is influenced by the training dataset size and the amount of spatial information (Bilali et al., 2021; Lim and Wang, 2022; Pyo et al., 2020). The authors reported that their models exhibited improved performance with increasing training size and spatial information density up to a certain extent.

2.2. Optimization of the input data and DL model

The training data size, input window size (wd), look-back size (lk),

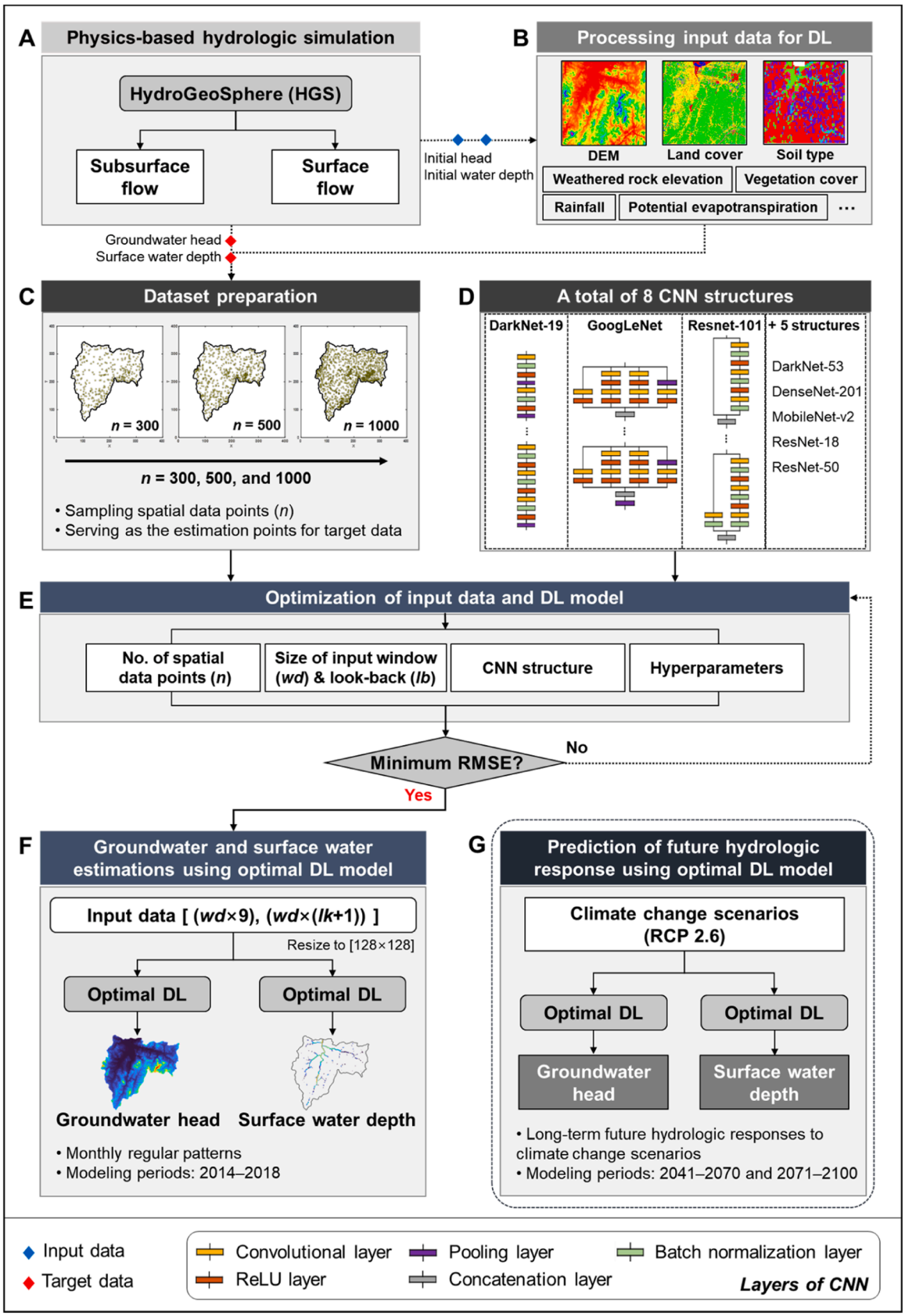


Fig. 1. Steps in estimating groundwater head and surface water depth: (A) hydrological simulation using the physics-based model HGS, (B–C) preparation of the DL dataset, (D) configuration of the DL model, (E) optimization of input data and the DL model, (F) estimation of groundwater and surface water conditions using the optimal DL model, and (G) prediction of future hydrological responses under climate change scenarios. In (D), the colored rectangular blocks represent multiple layers comprising a CNN structure (for example, convolutional, pooling, ReLU, and normalization layers). Detailed information on these layers can be found in the Supplementary Information (Appendix B.2).

Table 1
Input data in the DL model: topographical, geographical, hydrometeorological, and hydrological information.

	Input variable	Data type	Min	Max	Source
Topographical data	Digital elevation model	Numerical	2.31	605.21	National Geographic Information Institute[1)]
	Weathered rock elevation	Numerical	−19.82	564.10	Geotechnical Information DB System[2)] National Groundwater Information Center[3)]
Geographical data	Land cover	Categorical	–	–	Ministry of Environment[4)]
	Vegetation cover	Categorical	–	–	National Institute of Agricultural Sciences[5)]
	Soil type	Categorical	–	–	
Hydrometeorological data	Rainfall	Numerical	2.22×10^{-9}	1.65×10^{-7}	Korea Meteorological Administration[6)]
	Potential evapotranspiration	Numerical	6.40×10^{-9}	5.92×10^{-8}	Calculated using the simplified FAO Penman-Monteith equation (Valiantzas, 2006)
Hydrological data	Initial hydraulic head	Numerical	1.54	561.36	Simulated using the HGS model[¶]
	Initial water depth	Numerical	2.88×10^{-10}	5.48	

[1)] https://www.ngii.go.kr/kor/main.do;
[2)] https://www.geoinfo.or.kr/;
[3)] http://www.gims.go.kr/;.
[4)] http://me.go.kr/home/web/main.do;.
[5)] https://www.naas.go.kr/naas_index.do;.
[6)] https://www.weather.go.kr/w/index.do;.
[¶] indicates the Sabgyo Stream Watershed model developed by Lee et al. (2023).

CNN structure, and hyperparameters were tuned to enhance the DL model performance (Fig. 1C–E). The optimal DL model was determined by obtaining the minimum RMSE value for mapping. Among the eight CNN structures evaluated (DarkNet-19, DarkNet-53, DenseNet-201, GoogLeNet, MobileNet-v2, ResNet-18, ResNet-50, and ResNet-101), the ResNet structures demonstrated the best performance (Table 2, Fig. 1D). The optimal model for groundwater head estimation was ResNet-18, comprising 18 deep layers when the dataset comprises $n = 1000$, $wd = 3$, and $lk = 12$. This model was trained with a mini-batch size of 128 and learning rate of 0.001 using the RMSProp optimizer (Table 2A). Similarly, for surface water depth estimation, the optimal DL model was ResNet-18 with $n = 500$, $wd = 3$, and $lk = 12$, trained with a mini-batch size of 128 and a learning rate of 0.005 using the RMSProp optimizer (Table 2B). ResNet is renowned for its incorporation of residual and skip connections, which facilitate the direct transfer of output from layers to both subsequent and earlier layers (Gao et al., 2019; He et al., 2016). This mechanism has significantly enhanced model performances by effectively addressing challenges such as vanishing gradient and overfitting (Huang et al., 2017; Targ et al., 2016). The superior performance of ResNet has been recognized in previous studies; Ardakani et al. (2020) found that ResNet-101 and Xception outperformed nine other CNN structures, whereas Canziani et al. (2016) observed comparable model accuracies for GoogLeNet, ResNet-18, and ResNet-101.

Our DL input design consisted of $wd \times wd$ spatial grids and temporal data spanning the previous lk months to the present time. These findings highlight the significance of utilizing spatiotemporal input data measuring 600 $m \times 600$ m (3×3 grids) and encompassing the preceding 12 months to estimate the hydrological conditions in our study. In a previous study using a CNN, spatial information of less than 5×5 grids and temporal data spanning over 30 d was observed to be suitable for simulating harmful algal blooms within the study site (Baek et al., 2021).

2.3. Estimating watershed hydrological conditions using optimal DL

2.3.1. Spatiotemporal groundwater heads

Our DL modeling approach reliably reproduced the spatiotemporal distribution of groundwater heads across the entire watershed (Fig. 2). The mapping results of the monthly groundwater heads are provided in supplementary video clips (Vid. 1–3). These clips enabled a comparison of the estimates between the optimal DL and HGS models from 2014 to 2018. During this period, the optimal model achieved an RMSE of 2.35

Table 2
Optimal settings and DL model performance for estimating (A) groundwater head and (B) surface water depth. The following parameters were optimized: number of spatial data points (n), input window size (wd), look-back size (lk), CNN structure, and hyperparameters (i.e., mini-batch size, optimizer, and learning rate). The optimal DL model was determined based on mapping accuracy.

	Spatial	Input design		CNN type	Hyperparameter			Model performance					
								RMSE (m)			R^2		
	n	wd	lk	Structure	Mini-batch size	Optimizer	Learning rate	Tr[¶]	Val[¶¶]	Map[¶¶¶]	Tr[¶]	Val[¶¶]	Map[¶¶¶]
A	300	5	12	ResNet-18	64	RMSProp	0.003	3.10	4.93	3.97	0.999	0.998	0.997
	500	7	12	ResNet-101	128	RMSProp	0.003	3.39	4.35	4.24	0.999	0.999	0.997
	1000	3	12	ResNet-18	128	RMSProp	0.001	1.70	2.98	2.35	1.000	0.999	0.999
B	300	7	8	ResNet-101	128	Adam	0.008	0.04	0.26	0.36	0.999	0.970	0.905
	500	3	12	ResNet-18	128	RMSProp	0.005	0.08	0.17	0.29	0.997	0.989	0.941
	1000	3	12	ResNet-50	192	Adam	0.003	0.04	0.10	0.32	1.000	0.995	0.924

[¶] denotes training accuracy.
[¶¶] denotes validation accuracy.
[¶¶¶] denotes mapping accuracy.
Colored highlighting indicate the input design, CNN structure, hyperparameters, and model performance of optimal DL.

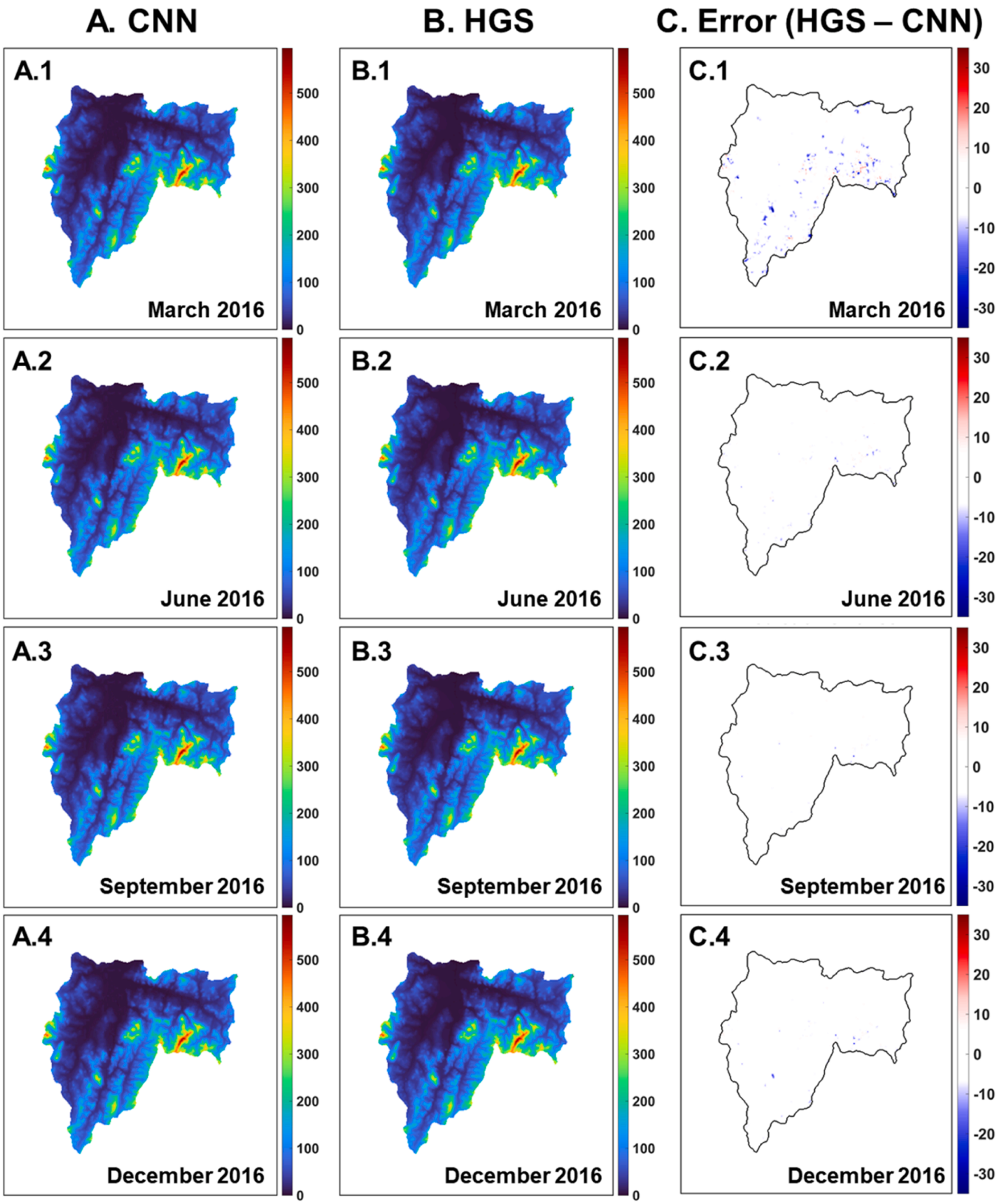


Fig. 2. Spatiotemporal maps of estimated groundwater heads using the optimal DL model (A), HGS model (B), and their prediction discrepancy (C). Subplots 1–4 present mapping results for March, June, September, and December 2016, respectively. (Mapping results for other periods can be found in a supplementary video clip). In panels (A) and (B), the color bar indicates groundwater head (m), with red representing higher values and blue representing lower values. In panel (C), the color bar displays the estimation discrepancy between the HGS and optimal DL models (i.e., ground-truth - DL prediction), with red indicating underestimations and blue indicating overestimations of the DL model compared to the HGS model.

m and R^2 of 0.999 (Table 2A). Taccari et al. (2022) also demonstrated a high DL performance, accurately predicting the groundwater head and its patterns compared to a physics-based model (MODFLOW), with an R^2 value of 0.99.

Furthermore, our DL model provides substantial computational advantages for watershed-scale simulation. The optimal DL model reproduced the results at a resolution of 200 m in 0.115 h on a single graphics processing unit (GPU). Conversely, the computation time of the HGS was 5.137 h, involving four central processing units (CPUs), to simulate a mean resolution of 420 m. DL modeling performs complex hydrological simulations approximately 45 times faster than would the conventional HGS model. Additional details regarding the computer hardware used in our study can be found in Section 4.2.3. Our research findings suggest that our approach can effectively alleviate the computational burden and reduce the time needed to simulate complex, large-scale hydrological processes, such as climate change impacts. This enables efficient execution of high-spatial-resolution hydrological modeling and predictions. Previous studies have emphasized the superior computational efficiency of DL simulations over conventional, physics-based models (Kakuda et al., 2021; Tran et al., 2021). In Leonarduzzi et al. (2022), their 2D-CNN model for soil moisture estimation achieved a computational speed of 500 times that of a physics-based hydrological model (ParFlow-CLM) at a resolution of 1000 m.

In addition, we conducted a detailed temporal analysis of model estimates at specific locations over the period 2014 to 2018 (Fig. 4A). The estimated groundwater heads were compared at three distinct points (GW1–3) located at various elevations (Fig. 4C). At GW3, the estimation discrepancy was significant, with an RMSE of 8.09 m (Fig. 4A.3), compared to 0.43 m at GW1 and 0.34 m at GW2 (Fig. 4A.1, 2). Our DL model tended to underestimate groundwater heads in mountainous terrains, such as GW3, compared to lower regions (e.g., GW1–2). This underestimation was observed at elevations above 400 m (Fig. 2C), indicating inherent challenges in accurately predicting peak values using data-driven DL models. As discussed in previous studies (Sudheer et al., 2003; Wu et al., 2009; Yang et al., 2019), these challenges can be attributed to the scarcity of extreme values, which impedes the training of DL models for peak values.

2.3.2. Spatiotemporal surface water depths

The proposed DL modeling framework demonstrated robust performance in capturing spatiotemporal variations in surface water depth across the entire water body (Fig. 3). In mapping the surface water depths from 2014 to 2018, the optimal DL model achieved an RMSE of 0.29 m and R^2 of 0.941 (Table 2B). The mapping results are provided as supplementary video clips (Vid. 4–6). To investigate temporal variations, this study examined the model outputs at specific spatial points (Fig. 4B). Monthly results were collected from 2014 to 2018 at three spatial points, denoted as SW1–3 (Fig. 4C). At SW1, the estimation discrepancy was an RMSE of 0.07 m, whereas it was 0.08 m at SW2–3. The optimal DL model underestimated the surface water depths at SW1–2 (Fig. 4B.1–2), suggesting a potential underestimation near the watershed outlet and midstream area. These estimation discrepancies could be caused by dataset size and imbalanced data distribution (Batista et al., 2004; Bilali et al., 2021). Previous research by Hussain et al. (2020) found that their 1D-CNN model underestimated the peak values for daily and monthly streamflow forecasts. Additionally, the data processing step in our study, such as the conversion of HGS data from a triangular mesh format to gridded data, may affect the DL model performance. Data quality has been identified as a potential factor influencing the DL model performance (Budach et al., 2022; Gudivada et al., 2017; Jain et al., 2020).

Overall, our DL approach serves as a valuable tool for efficient analysis and rapid decision making. Nonetheless, it is crucial to evaluate areas with significant estimation differences when considering the practical implications. Specifically, higher RMSE values in our DL estimations suggest less reliable forecasts. This could lead to suboptimal decisions, particularly in areas such as resource allocation and climate change mitigation strategies. Conversely, lower RMSE values indicate higher estimation accuracy, comparable to that of the physics-based HGS model, enabling more informed decision-making.

2.4. Exploratory analysis of the optimal DL model

2.4.1. Effect of input variables on the DL model estimations

A parametric sensitivity analysis was conducted to examine the effect of time-varying input variables (rainfall and potential evapotranspiration) on the estimations generated by the optimal DL model. This study employed the Elementary Effect Test (EET) method (Appendix D.1) with Latin hypercube-one-factor-at-a-time (LH–OAT) sampling (Nossent and Bauwens, 2012; Saltelli et al., 2008; Xu et al., 2016). The results of this analysis are presented in Fig. S3 showing the mean and standard deviation of elementary effect (EE) values. These metrics provide insights into the direct effect of the input variable and the interaction effect between the variables on the DL model outputs. A detailed analysis of the results is provided in Appendix D.2 of the Supplementary Information.

2.4.2. Predictions of future hydrology using the optimal DL model

The predictive performance of the optimal DL model was evaluated under future climate change scenarios (RCP 2.6) (Fig. 1G). The assessment was conducted for two periods: 2041–2070 (the 2050s) and 2071–2100 (the 2080s). In this analysis, the monthly average prediction results of the optimal DL and HGS models were compared at specific sites (GW1–3 and SW1–3) (Fig. S4). In terms of predictions of long-term groundwater heads, the optimal DL model yielded RMSE and R^2 values of 0.19 m and 0.88 for 2041–2070 (Fig. S4A.1) and 0.17 m and 0.90 for 2071–2100 at GW1 (Fig. S4B.1). However, the performance of our model deteriorated in high-elevation regions (e.g., GW3). At GW3, the estimation discrepancy resulted in an RMSE of 8.09 m for the period 2014–2018 (Fig. 4A.3), whereas it increased to RMSE values of 12.45 m for 2041–2070 (Fig. S4A.3) and 14.33 m for 2071–2100 (Fig. S4B.3). This diminishing DL performance is consistent with previous research findings regarding the prediction of future events and extreme values. Gumiere et al. (2020) reported that despite outperforming physics-based models, the DL model prediction accuracy notably decreased towards the end of a 3-d lead time, in contrast to the relatively small errors observed in physics-based models over time.

For long-term surface water prediction, the optimal DL model showed reliable performance in the watershed outlet region (e.g., SW1). At SW1, the RMSE and R^2 values were 0.09 and 0.84 m for 2041–2070 (Fig. S4C.1) and 0.10 and 0.83 m for 2071–2100 (Fig. S4D.1). However, the optimal DL model performance deteriorated when predicting the surface water depths in the upstream region (for example, SW3). At SW3, the estimation discrepancy was an RMSE of 0.08 m for the period 2014–2018 (Fig. 4B.3), increasing to 0.34 m for 2041–2070 (Fig. S4A.3) and 2071–2100 (Fig. S4B.3). In addition, the optimal DL model exhibited a larger variation in groundwater prediction over 30 years than the HGS model, whereas it demonstrated less variation in surface water prediction. These variations in long-term predictions may be attributed to out-of-distribution uncertainties in the DL model, in which the input data for future climate conditions differ significantly from the training data (Liu et al., 2020; Saria and Subbaswamy, 2019). According to Wan et al. (2022), traditional DL methods have limitations in predicting extreme pollutant values under extreme meteorological conditions primarily because they rely on historical data. Therefore, further improvements to our optimal DL model are recommended to enhance its predictive performance, particularly for long-term hydrological estimations.

3. Conclusions

With this study we proposed a novel DL modeling framework for

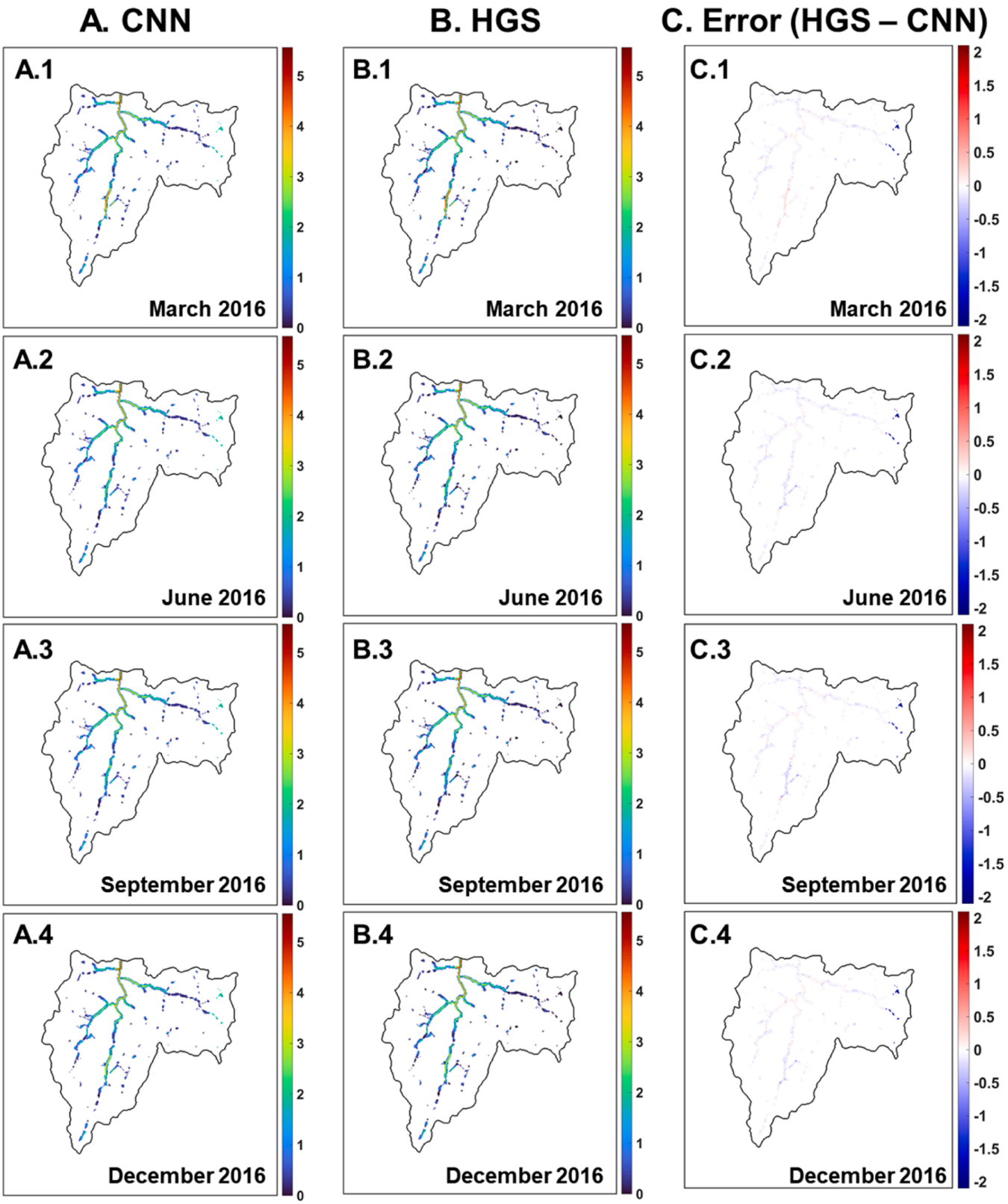


Fig. 3. Spatiotemporal maps of estimated surface water depths using the optimal DL model (A), HGS model (B), and their prediction discrepancy (C). Subplots 1–4 present mapping results for March, June, September, and December 2016, respectively, while mapping results for other periods can be found in a supplementary video clip. In panels (A) and (B), the color bar indicates surface water depth (m), with red representing higher values and blue representing lower values. In panel (C), the color bar displays the estimation discrepancy between the HGS and optimal DL models (i.e., ground-truth - DL prediction), with red indicating underestimations and blue indicating overestimations of the DL model compared to the HGS model.

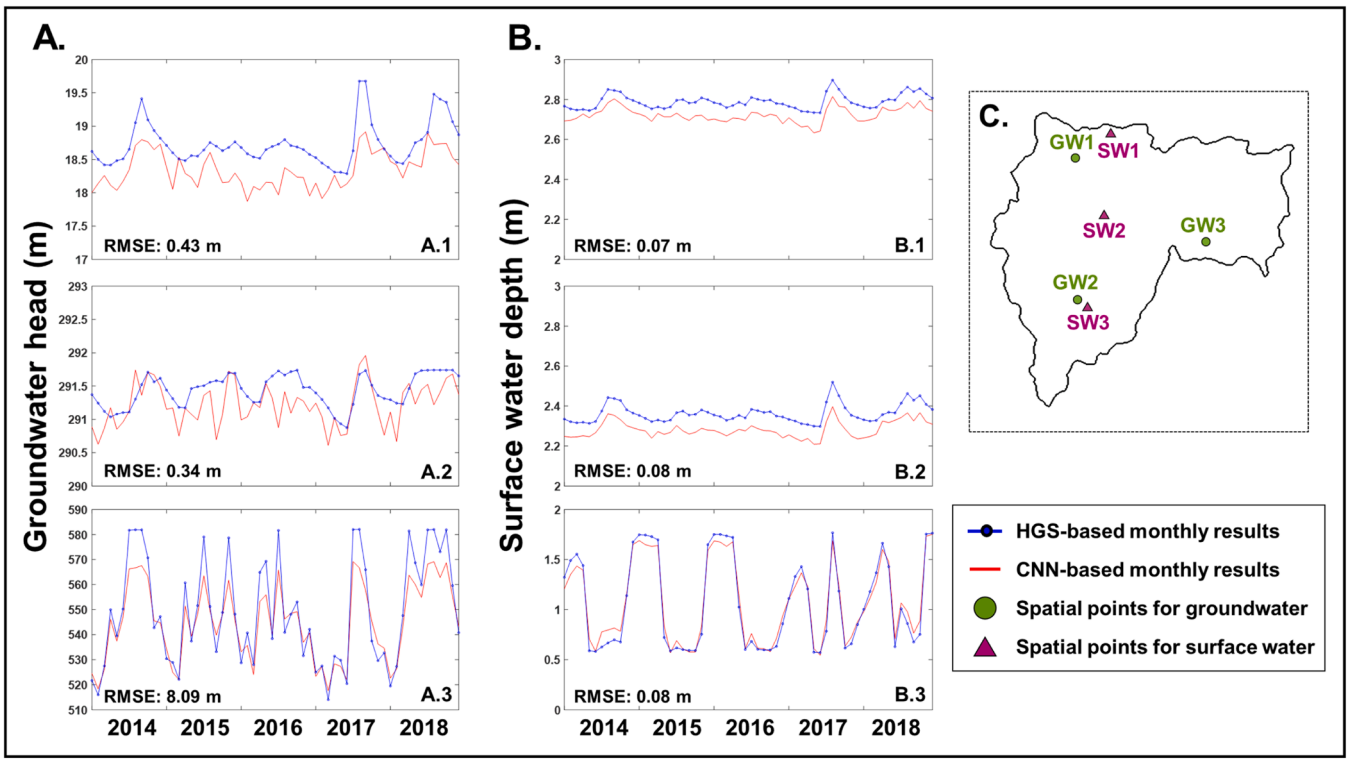


Fig. 4. Time-series plots of estimated groundwater head (A) and surface water depth (B). Hydrological conditions from 2014 to 2018 were collected at spatial points (C). The blue dotted line indicates the HGS model, while the red line represents the optimal CNN model. Light green circles indicate spatial points for groundwater (GW1–3), while fuchsia triangles indicate spatial points for surface water (SW1–3). Subplots A.1–3 correspond to the results at GW1–3, respectively, and subplots B.1–3 represent the results at SW1–3, respectively.

estimating groundwater and surface water conditions using topographical, geographical, hydrometeorological, and hydrological data. The DL model estimates were validated against the HGS model outputs and the optimal DL model was determined by optimizing the training samples, input design, CNN structure, and hyperparameters.

- The optimal DL model, based on ResNet-18 with optimal input designs, achieved RMSE values of 2.35 and 0.29 m for mapping groundwater heads and surface water depths, respectively.
- The optimal DL model achieved a computational speed that is 45 times faster than the HGS model when estimating the hydrological conditions in the watershed (Sabgyo Stream Watershed, South Korea).
- However, our DL model performance decreased when predicting future hydrological responses to climate change scenarios.

Overall, our DL model offers rapid hydrological predictions, reducing computational costs while maintaining a high spatial resolution compared with the physics-based, fully distributed HGS model. This capability is valuable for water resource management, particularly when making urgent decisions regarding large-scale and complex watersheds. However, further enhancements are required in our DL model for long-range analyses, such as evaluating the long-term effects of climate change on hydrology.

4. Materials and methods

4.1. Study area

The study site was the Sabgyo Stream Watershed (36.395796°–36.911621° N, 126.596445°–127.213928° E) located in the midwestern part of South Korea (Fig. S1). This watershed has a drainage area of approximately 1650 km^2 and a stream length of 65 km. The major soil texture types are sand (50.9 %) and sandy clay loam (26.3 %), followed by sandy loam (12.8 %) and silt loam (4.8 %) (National Institute of Agricultural Sciences, 2020). The study site is covered by forests (44.5 %), croplands (42.9 %), and urban areas (4.7 %) (Ministry of Environment, 2020). The Sabgyo Stream Watershed has a monsoon period with a dry season from October to May and a wet season from June to September. From 2000 to 2019, the average annual temperature at the study site was 12.1 ± 9.8 °C and rainfall was 1203.1 ± 126.3 mm/year (Korea Meteorological Administration, 2020). This watershed has 19- and 15-gauge stations for monitoring groundwater and discharge, respectively (Korea Rural Community Corporation, 2020; National Groundwater Information Center (2020); Water Resources Management Information System, 2020).

4.2. DL modeling for simulating groundwater and surface water

A DL modeling framework was proposed to efficiently reproduce spatiotemporal groundwater–surface water flow conditions. Among DL techniques, the CNN algorithm was adopted to capture the spatiotemporal patterns and nonlinearity inherent within hydrologic processes. Our DL modeling focused on estimating the regular patterns of groundwater heads and surface water depths within the study site. In addition, we assessed the predictive capabilities of future climate change impacts. The implementation of DL modeling comprises six steps: (1) acquisition of hydrological data from the HGS model (Fig. 1A), (2) preparation of the DL datasets (Fig. 1B, C), (3) configuration of DL model (Fig. 1D), (4) optimization of input data and DL model (Fig. 1E), (5) estimation of hydrological conditions using the optimal DL model (Fig. 1F), and (6) prediction of future hydrologic responses under climate change scenarios (Fig. 1G). The HGS model simulates hydrological conditions, specifically groundwater heads and surface water depths (Fig. 1A). These data were considered as the initial conditions and estimation targets during the DL modeling. Additionally, observed

topographical, geographic, and hydrometeorological data were utilized as DL model inputs (Fig. 1B and Table 1). This study investigated model performance by examining the effect of training data volume and varying the number of spatial data points (Fig. 1C). We optimized the CNN structure, input design (input window size (*wd*) and lookback size (*lk*)), and hyperparameters (Fig. 1D and E). These factors are crucial, because they affect the model and its computational costs (Heinen et al., 2020; Justus et al., 2018; Loussaief and Abdelkrim, 2018; Xu and Goodacre, 2018; Yu and Zhu, 2020). Subsequently, the optimal DL model was determined by comparing its estimates with those of the HGS model (Fig. 1F). Additionally, we assessed the predictive capabilities of the model by examining hydrological responses under a future climate scenario (RCP 2.6) (Fig. 1G).

4.2.1. Fully distributed hydrological model

HGS is a 3D-integrated surface-subsurface hydrological model that is widely used for water resource management (Persaud et al., 2020; Xu et al., 2021). It employs the 2D Saint-Venant equation for surface flow simulation and Richards' equation for variably saturated flow in subsurface simulations. In this study, the HGS model developed by Lee et al. (2023) was used to conduct hydrological simulations of the Sabgyo Stream Watershed. This HGS model consisted of 89530 nodes and 156420 elements in the 3D domain, while the 2D domain comprised 17380 nodes with an average length of 420 m (Fig. S1B). The model parameters and their values were adopted from previous studies (Chow, 1959; Freeze and Cherry, 1979; Hwang, 2021; Myneni et al., 2015; Panday and Huyakorn, 2004) (Table S1). To calibrate the HGS model, simulated groundwater levels were compared with monthly averaged observations from 19 monitoring wells and stream discharges were compared with data from 15 gauge stations (Fig. S1A). The calibration process used a 20-year observational dataset from 2000 to 2019. The simulation period was from 2012 to 2018, and monthly climate data were used as inputs. Comparisons between the simulated and observed groundwater levels and surface water flow rates are shown in Fig. S5, demonstrating a reasonable agreement with the observed data. Additional details regarding the HGS model and calibration are provided in Appendix A of the Supplementary Information.

4.2.2. Data processing for DL model

The DL model was developed using topological, geographical, hydrometeorological, and simulated hydrological data (Table 1). Topological data consist of a digital elevation model (DEM) and weathered rock elevation, and geographical data provide spatial information on soil types and land cover. The hydrometeorological data included rainfall and potential evapotranspiration (Fig. 1B). These data have been recognized as important factors in the hydrological modeling of watershed systems (Archfield et al., 2015; Jiang et al., 2022; Nguyen et al., 2023; Thapa et al., 2018). Additionally, the simulated groundwater heads and surface water depths were used as the initial conditions and ground truth for the DL model. Monthly data were synthesized using HGS model outputs from 2013 to 2018. Data from 2014 to 2018 were used for the training and validation periods, 2013 served as the lookback period, and December 2012 was used as the initial condition. The DL datasets were processed in three main steps: (1) conversion of the HGS model outputs into a gridded format; (2) extraction of spatiotemporal information from the study site; and (3) transformation of the input data from 3D to 2D. A detailed description of the processing dataset is provided in Appendix B.1 of the Supplementary Information. The DL datasets were divided into training (80 %) and validation (20 %) datasets. To address potential data imbalances during dataset generation, a random sampling approach with uniform distribution was utilized, following the methodology employed in previous studies (Brion et al., 2002; Kim et al., 2023; Tan and Beklioglu, 2006).

4.2.3. Convolutional neural networks

A CNN is a type of DL algorithm specialized in capturing spatiotemporal features from multi-dimensional data (Ajuria Illarramendi et al., 2022; LeCun et al., 2015). The convolutional filters and local connectivity in CNNs facilitate a better capture of the influence of the local features of the input data on the results (Liao et al., 2023). A typical CNN structure comprises of multiple layers, including convolutional, pooling, dropout, and fully connected layers (Basha et al., 2020; Wu and Gu, 2015). The convolutional and pooling layers extract the features, whereas the ReLU and normalization layers perform linear and normalization calculations, respectively (Ide and Kurita, 2017). Model performance can be enhanced by combining these layers (Khan et al., 2019; Szegedy et al., 2016). Detailed information on the layers is provided in Appendix B.2 of the Supplementary Information. Eight types of CNN architectures were used in this study: DarkNet-19, DarkNet-53 (Redmon, 2013), DenseNet-201 (Huang et al., 2017), GoogLeNet (Szegedy et al., 2015), MobileNet-v2 (Sandler et al., 2018), ResNet-18, ResNet-50, and ResNet-101 (He et al., 2016) (Fig. 1D). These architectures, which were recognized in the ImageNet Large-Scale Visual Recognition Challenge, are among the leading algorithms used for image recognition and classification (Ardakani et al., 2020; Tang et al., 2022).

Our computational setup featured a high-performance computing workstation comprising an Intel ® Core i9–10,900 with a 2.80 GHz CPU with 128 GB of RAM. The system was equipped with an NVIDIA GeForce RTX 3090 Ti GPU. Model development and training were conducted using MATLAB with DL toolboxes (R2022b) operating on the Microsoft Windows platform. The loss function during the DL modeling was the mean squared error (MSE) (Eq. S1) to quantify the difference between DL estimations and HGS outputs during DL training. The DL models were evaluated using RMSE (Eq. S2), and R^2 (Eq. S3). Detailed equations for these metrics can be found in Appendix C of the Supplementary Information.

4.2.4. Optimization of input data and DL model

Model optimization was conducted to enhance the performance of the DL model for estimating groundwater and surface water flow conditions. The optimization process employs a pattern-search algorithm known for its efficiency in global optimization and direct search methods (Fatemifar et al., 2021; Findler et al., 1987). This algorithm minimizes the inequalities by determining the minimum value of an objective function using straightforward numerical operations (Palacio-Morales et al., 2021; Park et al., 2014). We optimized the amount of training data, input window size (*wd*), lookback size (*lk*), CNN structure, and hyperparameters (Fig. 1E). The DL datasets were sampled with spatial data quantities (*n*) of 300, 500, and 1000 grid cells within the watershed (Fig. 1C). The input design, represented by *wd* and *lk*, determines whether the DL model considers spatial information with *wd* × *wd* grids and temporal data from the previous *lk* months to the current time. The DL model architecture was selected from the following eight CNN structures: DarkNet-19, DarkNet-53, DenseNet-201, GoogLeNet, MobileNet-v2, ResNet-18, ResNet-50, and ResNet-101 (Fig. 1D). The optimized hyperparameters included the mini-batch size, optimizer, and learning rate (Kandel and Castelli, 2020; Smith, 2018). Three optimizer types were compared: Stochastic Gradient Descent with Momentum (Ruder, 2016), Root Mean Square Propagation (RMSProp) (Hinton et al., 2012), and Adam (Kingma and Ba, 2014). Other optimizing parameters were varied within predefined ranges: 3–11 (odd number) for the *wd* size, 1–12 (months) for the *lk* size, 64–256 for the mini-batch size, and 0.0001–0.01 for the learning rate. The objective function for optimization was the RMSE, which quantified the discrepancies between the HGS and DL outputs to minimize the RMSE value during model validation. The DL model was trained for 150 epochs (Fig. 1E, F).

4.3. Exploratory model analysis

An exploratory analysis was conducted to investigate the parametric sensitivity and prediction ability of the optimal DL model. EET with LH–OAT sampling has been used as a sensitivity analysis method

(Morris, 1991; Saltelli et al., 2008; Xu et al., 2016). This is because of its simplicity and computational efficiency, which require a smaller sample size than other methods (Campolongo et al., 2007; Pianosi et al., 2016). A brief description of the EET method is provided in Appendix D.1 of the Supplementary Information. In addition, we examined the capability of our DL model to predict long-term hydrological responses. Predicting future hydrological patterns in watersheds is crucial for effective management of water resource sustainability. The predictive performance was measured by comparing our model's prediction results with those of the HGS model under a future climate change scenario (Fig. 1G).

4.3.1. *Sensitivity analysis of the optimal DL model*

The EE value (Eq. S4) in the EET method represent the change in the model output induced by a change in the input variable. We focused on analyzing the influence of time-dependent input variables as they vary over predictions of the present and future values. Rainfall and evapotranspiration exhibited temporal variability across the predictions of present and future values, whereas variables such as elevation and land cover remained static throughout simulations of the present and future conditions. In the EET analysis, we explicitly computed the EEs of rainfall and evapotranspiration from the current time (t) to the preceding 12 months (t-12) without considering static variables. Consequently, we assessed the changes in groundwater head and surface water depth affected by temporal variations in rainfall and evapotranspiration. The range of the input variables is listed in Table 1, and the EET analysis was performed using the MATLAB toolbox for global sensitivity analysis developed by Pianosi et al. (2015). In the EET method, the sensitivities of the input variables are measured based on the mean and standard deviation of the EE values. The mean EEs represent the direct effect of the input variable on the hydrological condition estimations from our DL model, whereas the standard deviation of the EEs reflects the interaction effect between variables. In this study, the sensitivity rankings were determined according to the mean EE values in descending order. The variable with the highest mean EE represents the most sensitive input condition.

4.3.2. *Predicting hydrological responses to future climate scenarios*

Using the optimal DL model, we estimated the long-term hydrological responses under future climate change impacts. This analysis investigated the ability of our model to predict regular patterns of future hydrological conditions within a watershed. For this analysis, input datasets were generated using simulated future hydrometeorological and hydrological conditions as well as present-day topographical and geographical information. Future hydrometeorological and hydrological data will include rainfall and potential evapotranspiration under a climate change scenario (i.e., representative concentration pathway (RCP) 2.6 (IPCC, 2013; Van Vuuren et al., 2011). These future data were derived from the long-term HGS model simulations (from 2011 to 2100) conducted by Lee et al. (2023). The projected climatic conditions are summarized in Table S2. Details of the long-term simulation, including the procedures and analysis, are thoroughly described by Lee et al. (2023). The topographical and geographical information and initial hydrological conditions were the same as those listed in Table 1. In this study, the predictive ability was assessed for two periods: 2041–2070 (the 2050s) and 2071–2100 (the 2080s).

CRediT authorship contribution statement

Soobin Kim: Writing – original draft, Visualization, Software, Methodology, Formal analysis. **Eunhee Lee:** Writing – review & editing, Software, Resources. **Hyoun-Tae Hwang:** Writing – review & editing, Formal analysis. **JongCheol Pyo:** Resources, Data curation. **Daeun Yun:** Resources, Data curation. **Sang-Soo Baek:** Writing – review & editing, Supervision, Data curation. **Kyung Hwa Cho:** Writing – review & editing, Supervision, Funding acquisition.

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

Acknowledgment

This work was supported by the Korea Environment Industry & Technology Institute (KEITI) through the Advanced Technology Development Project for Predicting and Preventing Chemical Accidents Project, funded by the Korea Ministry of Environment (MOE) (2022003620001), the Basic Research Project of the Korea Institute of Geoscience and Mineral Resources (23-3411), and MSIT through Sejong Science Fellowship, funded by National Research Foundation of Korea (NRF) (No. 2022R1C1C2003649).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.wroa.2024.100228.

References

Abbas, A., Park, M., Baek, S.-S., Cho, K.H., 2023. Deep learning-based algorithms for long-term prediction of chlorophyll-a in catchment streams. J. Hydrol. (Amst.) 626, 130240.

Ajuria Illarramendi, E., Bauerheim, M., Cuenot, B, 2022. Performance and accuracy assessments of an incompressible fluid solver coupled with a deep convolutional neural network. Data-Centric Eng. 3, e2.

Alvi, M., Batstone, D., Mbamba, C.K., Keymer, P., French, T., Ward, A., Dwyer, J., Cardell-Oliver, R., 2023. Deep learning in wastewater treatment: a critical review. Water Res. 245, 120518.

Apurv, T., Cai, X., 2020. Drought propagation in contiguous US watersheds: a process-based understanding of the role of climate and watershed properties. Water Resour. Res. 56 (9), e2020WR027755.

Aquanty Inc., 2023. HydroGeoSphere. A three-dimensional numerical model describing fully-integrated subsurface and surface flow and solute. transport, Waterloo, ON, Canada.

Archfield, S.A., Clark, M., Arheimer, B., Hay, L.E., McMillan, H., Kiang, J.E., Seibert, J., Hakala, K., Bock, A., Wagener, T., 2015. Accelerating advances in continental domain hydrologic modeling. Water Resour. Res. 51 (12), 10078–10091.

Ardakani, A.A., Kanafi, A.R., Acharya, U.R., Khadem, N., Mohammadi, A., 2020. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks. Comput. Biol. Med. 121, 103795.

Baek, S.-S., Pyo, J., Kwon, Y.S., Chun, S.-J., Baek, S.H., Ahn, C.-Y., Oh, H.-M., Kim, Y.O., Cho, K.H., 2021. Deep learning for simulating harmful algal blooms using ocean numerical model. Front. Mar. Sci. 8, 729954.

Basha, S.S., Dubey, S.R., Pulabaigari, V., Mukherjee, S., 2020. Impact of fully connected layers on performance of convolutional neural networks for image classification. Neurocomputing 378, 112–119.

Batista, G.E., Prati, R.C., Monard, M.C., 2004. A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD Explor. Newslett. 6 (1), 20–29.

Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R., 2023. Rapid spatio-temporal flood modelling via hydraulics-based graph neural networks. EGUsphere 2023, 1–24.

Bilali, A.E., Taleb, A., Bahlaoui, M.A., Brouziyne, Y., 2021. An integrated approach based on Gaussian noises-based data augmentation method and AdaBoost model to predict faecal coliforms in rivers with small dataset. J. Hydrol. (Amst) 599, 126510.

Brion, G.M., Neelakantan, T., Lingireddy, S., 2002. A neural-network-based classification scheme for sorting sources and ages of fecal contamination in water. Water Res. 36 (15), 3765–3774.

Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., Naumann, F. and Harmouch, H. 2022. The effects of data quality on machine learning performance. arXiv preprint. arXiv:2207.14529.

Buitink, J., Melsen, L.A., Kirchner, J.W., Teuling, A.J., 2020. A distributed simple dynamical systems approach (dS2 v1.0) for computationally efficient hydrological modelling at high spatio-temporal resolution. Geosci. Model. Dev. 13 (12), 6093–6110.

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Modell. Softw. 22 (10), 1509–1518.

Canziani, A., Paszke, A. and Culurciello, E. 2016. An analysis of deep neural network models for practical applications. arXiv preprint. arXiv:1605.07678.
Chen, H., Chen, A., Xu, L., Xie, H., Qiao, H., Lin, Q., Cai, K., 2020. A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. Agric. Water. Manage. 240, 106303.
Cheng, H., Chen, H., Jiang, G., Yoshihira, K., 2007. Nonlinear Feature Selection By Relevance Feature Vector Machine. Springer, pp. 144–159.
Chow, V., 1959. T. 1959 Open-Channel Hydraulics. MCGraw Hiu.
Delpla, I., Jung, A.-V., Baures, E., Clement, M., Thomas, O., 2009. Impacts of climate change on surface water quality in relation to drinking water production. Environ. Int. 35 (8), 1225–1233.
Ekmekcioğlu, Ö., Demirel, M.C., Booij, M.J., 2022. Effect of data length, spin-up period and spatial model resolution on fully distributed hydrological model calibration in the Moselle basin. Hydrol. Sci. J. 67 (5), 759–772.
Fatemifar, S., Awais, M., Akbari, A., Kittler, J., 2021. Particle Swarm and Pattern Search Optimisation of an Ensemble of Face Anomaly Detectors. IEEE, pp. 3622–3626.
Findler, N.V., Lo, C., Lo, R., 1987. Pattern search for optimization. Math Comput Simul 29 (1), 41–50.
Freeze, R.A., Cherry, J.A., 1979. Groundwater Prentice-Hall, Eaglewood Cliffs, New Jersey.
Gai, Y., Wang, M., Wu, Y., Wang, E., Deng, X., Liu, Y., Jim Yeh, T.-C., Hao, Y, 2023. Simulation of spring discharge using graph neural networks at Niangziguan Springs. Chin. J. Hydrol. 625, 130079.
Gao, H., Tao, X., Shen, X., Jia, J., 2019. Dynamic Scene Deblurring With Parameter Selective Sharing and Nested Skip Connections, pp. 3848–3856.
Gudivada, V., Apon, A., Ding, J., 2017. Data quality considerations for big data and machine learning: going beyond data cleaning and transformations. Int. J. Ad. Softw. 10 (1), 1–20.
Guevara Ochoa, C., Medina Sierra, A., Vives, L., Zimmermann, E., Bailey, R., 2020. Spatio-temporal patterns of the interaction between groundwater and surface water in plains. Hydrol. Process. 34 (6), 1371–1392.
Gumiere, S.J., Camporese, M., Botto, A., Lafond, J.A., Paniconi, C., Gallichand, J., Rousseau, A.N., 2020. Machine learning vs. physics-based modeling for real-time irrigation management. Front. Water. 2.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep Residual Learning For Image Recognition, pp. 770–778.
Heinen, S., Schwilk, M., von Rudorff, G.F., von Lilienfeld, O.A., 2020. Machine learning the computational cost of quantum chemistry. Mach. Learn. Sci. Technol. 1 (2), 025002.
Hinton, G., Srivastava, N., Swersky, K., 2012. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. Cited 14 (8), 2.
Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely Connected Convolutional Networks, pp. 4700–4708.
Hussain, D., Hussain, T., Khan, A.A., Naqvi, S.A.A., Jamil, A., 2020. A deep learning approach for hydrological time-series prediction: a case study of Gilgit river basin. Earth Sci. Inform. 13 (3), 915–927.
Hwang, H.-T., et al., 2021. Estimating anthropogenic effects on a highly-controlled basin with an integrated surface-subsurface model. Journal of Hydrology 603, 126963.
Hwang, H.-T., Park, Y.-J., Sudicky, E., Forsyth, P.A., 2014. A parallel computational framework to solve flow and transport in integrated surface–subsurface hydrologic systems. Environmental modelling & software 61, 39–58.
Ide, H., Kurita, T., 2017. Improvement of Learning For CNN With ReLU Activation By Sparse Regularization. IEEE, pp. 2684–2691.
Inyinbor Adejumoke, A., Adebesin Babatunde, O., Oluyori Abimbola, P., Adelani Akande Tabitha, A., Dada Adewumi, O., Oreofe Toyin, A., 2018. Water pollution: effects, prevention, and climatic impact. Water Challeng. Urban. World 33, 33–47.
IPCC, 2013. Climate Change 2013: The physical Science Basis. Cambridge University Press, Cambridge, UKNew York.
Iqbal, Z., Shahid, S., Ismail, T., Sa'adi, Z., Farooque, A., Yaseen, Z.M., 2022. Distributed Hydrological Model Based on Machine Learning Algorithm. Assess. Clim. Change Impact Floods Sustain. 14 (11), 6620.
Jain, A., Patel, H., Nagalapatti, L., Gupta, N., Mehta, S., Guttula, S., Mujumdar, S., Afzal, S., Sharma Mittal, R., Munigala, V., 2020. Overview and Importance of Data Quality For Machine Learning Tasks, pp. 3561–3562.
Jiang, P., Son, K., Mudunuru, M.K., Chen, X., 2022. Using mutual information for global sensitivity analysis on watershed modeling. Water Resour. Res. 58 (10), e2022WR032932.
Justus, D., Brennan, J., Bonner, S., McGough, A.S., 2018. Predicting the Computational Cost of Deep Learning Models. IEEE, pp. 3873–3882.
Kaandorp, V.P., Molina-Navarro, E., Andersen, H.E., Bloomfield, J.P., Kuijper, M.J.M., de Louw, P.G.B., 2018. A conceptual model for the analysis of multi-stressors in linked groundwater–surface water systems. Sci. Total Environ. 627, 880–895.
Kakuda, K., Morimasa, Y., Enomoto, T., Okaniwa, W., Miura, S., 2021. Data-Driven Fluid Flow Simulations by Using Convolutional Neural Network. Springer, pp. 14–19.
Kandel, I., Castelli, M., 2020. The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. ICT Exp. 6 (4), 312–315.
Khan, R.U., Zhang, X., Kumar, R., 2019. Analysis of ResNet and GoogleNet models for malware detection. J. Comput. Virol. Hack. Tech. 15, 29–37.
Kim, S., Abbas, A., Pyo, J., Kim, H., Hong, S.M., Baek, S.-S., Cho, K.H., 2023. Developing a data-driven modeling framework for simulating a chemical accident in freshwater. J. Clean. Prod. 425, 138842.
Kingma, D.P. and Ba, J. 2014. Adam: a method for stochastic optimization. arXiv preprint. arXiv:1412.6980.
Kollet, S.J., Maxwell, R.M., Woodward, C.S., Smith, S., Vanderborght, J., Vereecken, H., Simmer, C., 2010. Proof of concept of regional scale hydrologic simulations at hydrologic resolution utilizing massively parallel computer resources. Water Resour. Res. 46 (4).
Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018. Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrol. Earth. Syst. Sci. 22 (11), 6005–6022.
L'heureux, A., Grolinger, K., Elyamany, H.F., Capretz, M.A., 2017. Machine learning with big data: challenges and approaches. IEEE Access. 5, 7776–7797.
Lange, H., Sippel, S., 2020. Machine learning applications in hydrology. Forest Water Interact. 233–257.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Lee, E., Lee, H., Park, D., Hwang, H.-T., Park, C., 2023. Application of different weighting schemes and stochastic simulations to parameterization processes considering observation error: implications for climate change impact analysis of integrated watershed models. Water. (Basel) 15 (10), 1880.
Leonarduzzi, E., Tran, H., Bansal, V., Hull, R.B., De la Fuente, L., Bearup, L.A., Melchior, P., Condon, L.E., Maxwell, R.M., 2022. Training machine learning with physics-based simulations to predict 2D soil moisture fields in a changing climate. Front. Water. 4.
Liao, Y., Wang, Z., Chen, X., Lai, C., 2023. Fast simulation and prediction of urban pluvial floods using a deep convolutional neural network model. J. Hydrol. (Amst.) 624, 129945.
Lim, T., Wang, K., 2022. Comparison of machine learning algorithms for emulation of a gridded hydrological model given spatially explicit inputs. Comput. Geosci. 159, 105025.
Liu, Q., Gui, D., Zhang, L., Niu, J., Dai, H., Wei, G., Hu, B.X., 2022a. Simulation of regional groundwater levels in arid regions using interpretable machine learning models. Sci. Total Environ. 831, 154902.
Liu, W., Wang, X., Owens, J., Li, Y., 2020. Energy-based out-of-distribution detection. Adv. Neural Inf. Process. Syst. 33, 21464–21475.
Liu, Y., Hou, G., Huang, F., Qin, H., Wang, B., Yi, L., 2022b. Directed graph deep neural network for multi-step daily streamflow forecasting. J. Hydrol. (Amst.) 607, 127515.
Loague, K., Heppner, C.S., Mirus, B.B., Ebel, B.A., Ran, Q., Carr, A.E., BeVille, S.H., VanderKwaak, J.E., 2006. Physics-based hydrologic-response simulation: foundation for hydroecology and hydrogeomorphology. Hydrol. Process. Int. J. 20 (5), 1231–1237.
Loussaief, S., Abdelkrim, A., 2018. Convolutional neural network hyper-parameters optimization based on genetic algorithms. Int. J. Adv. Comput. Sci. Appl. 9 (10).
Maqsood, J., Farooque, A.A., Abbas, F., Esau, T., Wang, X., Acharya, B., Afzaal, H., 2022. Application of artificial neural networks to project reference evapotranspiration under climate change scenarios. Water Resour. Manage. 36 (3), 835–851.
Maxwell, R., Condon, L., Kollet, S., 2015. A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. Geosci. Model. Dev. 8 (3), 923–937.
Mehran, A., AghaKouchak, A., Nakhjiri, N., Stewardson, M.J., Peel, M.C., Phillips, T.J., Wada, Y., Ravalico, J.K., 2017. Compounding impacts of human-induced water stress and climate change on water availability. Sci. Rep. 7 (1), 6282.
Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics. 33 (2), 161–174.
Myneni, R., Knyazikhin, Y., Park, T., 2015. MOD15A2H MODIS/terra leaf area index. FPAR 8-day L4 Global 500 m SIN Grid V006 Data Set.
Nguyen, B.Q., Vo, N.D., Le, M.-H., Nguyen, Q.-D., Lakshmi, V., Bolten, J.D., 2023. Quantification of global digital elevation model (DEM)–A case study of the newly released NASADEM for a river basin in Central Vietnam. J. Hydrol. Region. Stud. 45, 101282.
Nossent, J., Bauwens, W., 2012. Multi-variable sensitivity and identifiability analysis for a complex environmental model in view of integrated water quantity and water quality modeling. Water Sci. Technol. 65 (3), 539–549.
Ocio, D., Beskeen, T., Smart, K., 2019. Fully distributed hydrological modelling for catchment-wide hydrological data verification. Hydrol. Res. 50 (6), 1520–1534.
Palacio-Morales, J., Tobón, A., Herrera, J., 2021. Optimization based on pattern search algorithm applied to ph non-linear control: application to alkalinization process of sugar juice. Processes 9 (12), 2283.
Panday, S., Huyakorn, P.S., 2004. A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. Advances in water Resources 27 (4), 361–382.
Papalexiou, S.M., Montanari, A., 2019. Global and regional increase of precipitation extremes under global warming. Water Resour. Res. 55 (6), 4901–4914.
Park, Y., Cho, K.H., Kang, J.-H., Lee, S.W., Kim, J.H., 2014. Developing a flow control strategy to reduce nutrient load in a reclaimed multi-reservoir system using a 2D hydrodynamic and water quality model. Sci. Total Environ. 466, 871–880.
Persaud, E., Levison, J., MacRitchie, S., Berg, S.J., Erler, A.R., Parker, B., Sudicky, E., 2020. Integrated modelling to assess climate change impacts on groundwater and surface water in the Great Lakes Basin using diverse climate forcing. J. Hydrol. (Amst.) 584, 124682.
Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: a systematic review with practical workflow. Environ. Modell. Softw. 79, 214–232.
Pianosi, F., Sarrazin, F., Wagener, T., 2015. A matlab toolbox for global sensitivity analysis. Environmental Modell. Softw. 70, 80–85.
Prasad, D.V.V., Venkataramana, L.Y., Kumar, P.S., Prasannamedha, G., Harshana, S., Srividya, S.J., Harrinei, K., Indraganti, S., 2022. Analysis and prediction of water quality using deep learning and auto deep learning techniques. Sci. Total Environ. 821, 153311.
Pyo, J., Pachepsky, Y., Kim, S., Abbas, A., Kim, M., Kwon, Y.S., Ligaray, M., Cho, K.H., 2023. Long short-term memory models of water quality in inland water environments. Water. Res. X., 100207

Pyo, J., Park, L.J., Pachepsky, Y., Baek, S.-S., Kim, K., Cho, K.H., 2020. Using convolutional neural network for predicting cyanobacteria concentrations in river water. Water Res. 186, 116349.
Rasp, S., Pritchard, M.S., Gentine, P., 2018. Deep learning to represent subgrid processes in climate models. Proc. Natl. Acad. Sci. 115 (39), 9684–9689.
Redmon, J., 2013. Darknet: Open source Neural Networks in c.
Ruder, S. 2016. An overview of gradient descent optimization algorithms. arXiv preprint. arXiv:1609.04747.
Sabzipour, B., Arsenault, R., Troin, M., Martel, J.-L., Brissette, F., Brunet, F., Mai, J., 2023. Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. J. Hydrol. (Amst.) 627, 130380.
Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity analysis: the Primer. John Wiley & Sons.
Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv2: Inverted residuals and Linear Bottlenecks, pp. 4510–4520.
Saria, S. and Subbaswamy, A. 2019. Tutorial: safe and reliable machine learning. arXiv preprint. arXiv:1904.07204.
Schwartz, P., Randall, D., 2003. An Abrupt Climate Change Scenario and Its Implications For United States National security, California inst of Tech Pasadena Jet Propulsion Lab.
Shen, C., Chen, X., Laloy, E., 2021. Editorial: broadening the use of machine learning in hydrology. Front. Water. 3.
Shen, C., Lawson, K., 2021. Applications of deep learning in hydrology. Deep Learn. Earth Sci. 283–297.
Sinha, S., Hammond, A., Smith, H., 2022. A comprehensive intercomparison study between a lumped and a fully distributed hydrological model across a set of 50 catchments in the United Kingdom. Hydrol. Process. 36 (3), e14544.
Smith, L.N. 2018. A disciplined approach to neural network hyper-parameters: part 1–learning rate, batch size, momentum, and weight decay. arXiv preprint. arXiv:1803.09820.
Sudheer, K., Nayak, P., Ramasastri, K., 2003. Improving peak flow estimates in artificial neural network river flow models. Hydrol. Process. 17 (3), 677–686.
Sudicky, E.A., Jones, J.P., Park, Y.-J., Brookfield, A.E., Colautti, D., 2008. Simulating complex flow and transport dynamics in an integrated surface-subsurface modeling framework. Geosci. J. 12 (2), 107–122.
Sun, A.Y., Jiang, P., Yang, Z.L., Xie, Y., Chen, X., 2022. A graph neural network (GNN) approach to basin-scale river network learning: the role of physics-based connectivity and data fusion. Hydrol. Earth Syst. Sci. 26 (19), 5163–5184.
Sun, A.Y., Scanlon, B.R., Zhang, Z., Walling, D., Bhanja, S.N., Mukherjee, A., Zhong, Z., 2019. Combining physically based modeling and deep learning for fusing GRACE satellite data: can we learn from mismatch? Water Resour. Res. 55 (2), 1179–1195.
Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going Deeper With Convolutions, pp. 1–9.
Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016. Rethinking the Inception Architecture For Computer Vision, pp. 2818–2826.
Taccari, M.L., Nuttall, J., Chen, X., Wang, H., Minnema, B., Jimack, P.K., 2022. Attention U-Net as a surrogate model for groundwater prediction. Adv. Water Resour. 163, 104169.
Talukdar, S., Ahmed, S., Naikoo, M.W., Rahman, A., Mallik, S., Ningthoujam, S., Bera, S., Ramana, G., 2023. Predicting lake water quality index with sensitivity-uncertainty analysis using deep learning algorithms. J. Clean. Prod. 406, 136885.
Tan, C.O., Beklioglu, M., 2006. Modeling complex nonlinear responses of shallow lakes to fish and hydrology using artificial neural networks. Ecol. Modell. 196 (1), 183–194.
Tang, P., Zhang, D., Li, H., 2022. Predicting permeability from 3D rock images based on CNN with physical information. J. Hydrol. (Amst.) 606, 127473.
Targ, S., Almeida, D., Lyman, K., 2016. arXiv preprint.
Taylor, R.G., Scanlon, B., Döll, P., Rodell, M., Van Beek, R., Wada, Y., Longuevergne, L., Leblanc, M., Famiglietti, J.S., Edmunds, M., 2013. Ground water and climate change. Nat. Clim. Change 3 (4), 322–329.
Thapa, R., Gupta, S., Gupta, A., Reddy, D., Kaur, H., 2018. Use of geospatial technology for delineating groundwater potential zones with an emphasis on water-table analysis in Dwarka River basin, Birbhum, India. Hydrogeol. J. 26 (3), 899–922.
Tran, H., Leonarduzzi, E., De la Fuente, L., Hull, R.B., Bansal, V., Chennault, C., Gentine, P., Melchior, P., Condon, L.E., Maxwell, R.M., 2021. Development of a deep learning emulator for a distributed groundwater-surface water model: parflow-ML. Water. (Basel) 13 (23), 3393.
Van Vuuren, D.P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G.C., Kram, T., Krey, V., Lamarque, J.-F., 2011. The representative concentration pathways: an overview. Clim. Change 109 (1), 5–31.
VanderKwaak, J.E., Loague, K., 2001. Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model. Water Resour. Res. 37 (4), 999–1013.
Vieux, B.E., Cui, Z., Gaur, A., 2004. Evaluation of a physics-based distributed hydrologic model for flood forecasting. J. Hydrol. (Amst.) 298 (1), 155–177.
Vivoni, E.R., Mascaro, G., Mniszewski, S., Fasel, P., Springer, E.P., Ivanov, V.Y., Bras, R. L., 2011. Real-world hydrologic assessment of a fully-distributed hydrological model in a parallel computing environment. J. Hydrol. (Amst.) 409 (1–2), 483–496.
Wan, H., Xu, R., Zhang, M., Cai, Y., Li, J., Shen, X., 2022. A novel model for water quality prediction caused by non-point sources pollution based on deep learning and feature extraction methods. J. Hydrol. (Amst.) 612, 128081.
Wang, D., Liu, Y., Kumar, M., 2018. Using nested discretization for a detailed yet computationally efficient simulation of local hydrology in a distributed hydrologic model. Sci. Rep. 8 (1), 5785.
Wang, Z., Wang, J., Han, J., 2022. Spatial prediction of groundwater potential and driving factor analysis based on deep learning and geographical detector in an arid endorheic basin. Ecol. Indic. 142, 109256.
Wu, C., Chau, K.W., Li, Y.S., 2009. Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques. Water Resour. Res. 45 (8).
Wu, H., Gu, X., 2015. Towards dropout training for convolutional neural networks. Neural Netw. 71, 1–10.
Wu, Y.-S., Zhang, K., Ding, C., Pruess, K., Elmroth, E., Bodvarsson, G., 2002. An efficient parallel-computing method for modeling nonisothermal multiphase flow and multicomponent transport in porous and fractured media. Adv. Water Resour. 25 (3), 243–261.
Xu, S., Frey, S.K., Erler, A.R., Khader, O., Berg, S.J., Hwang, H.T., Callaghan, M.V., Davison, J.H., Sudicky, E.A., 2021. Investigating groundwater-lake interactions in the Laurentian Great Lakes with a fully-integrated surface water-groundwater model. J. Hydrol. (Amst.) 594, 125911.
Xu, X., Sun, C., Huang, G., Mohanty, B.P., 2016. Global sensitivity analysis and calibration of parameters for a physically-based agro-hydrological model. Environ. Modell. Software 83, 88–102.
Xu, Y., Goodacre, R., 2018. On splitting training and validation set: a comparative study of cross-validation, bootstrap and systematic sampling for estimating the generalization performance of supervised learning. J. Anal. Test. 2 (3), 249–262.
Yang, S., Yang, D., Chen, J., Zhao, B., 2019. Real-time reservoir operation using recurrent neural networks and inflow forecast from a distributed hydrological model. J. Hydrol. (Amst) 579, 124229.
Yu, T. and Zhu, H. 2020. Hyper-parameter optimization: a review of algorithms and applications. arXiv preprint. arXiv:2003.05689.
Zhou, L., Pan, S., Wang, J., Vasilakos, A.V., 2017. Machine learning on big data: opportunities and challenges. Neurocomputing 237, 350–361.