IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 December 2023, accepted 24 January 2024, date of publication 30 January 2024, date of current version 23 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3360337

RESEARCH ARTICLE

A New Approach Based on Deep Neural Networks and Multisource Geospatial Data for Spatial Prediction of Groundwater Spring Potential

VIET-HA NHU[1], DUONG CAO PHAN[2,3], PHAM VIET HOA[4], PHAM THE VINH[5], AND DIEU TIEN BUI[6]

[1]Department of Geological-Geotechnical Engineering, Hanoi University of Mining and Geology, Bac Tu Liem, Hanoi 100000, Vietnam
[2]Centre for AI and Applied Data Analytics, School of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland
[3]Hydraulic Construction Institute, Vietnam Academy for Water Resources, Dong Da, Hanoi 116765, Vietnam
[4]Ho Chi Minh City Institute of Resources Geography, Vietnam Academy of Science and Technology, Ho Chi Minh City 700000, Vietnam
[5]Center of Environmental Technology Science and Ecology, Southern Institute of Water Resources Research, Ho Chi Minh City 700000, Vietnam
[6]GIS Group, Department of Business and IT, University of South-Eastern Norway, Bø, 3800 Telemark, Norway

Corresponding author: Viet-Ha Nhu (nhuvietha@humg.edu.vn)

This work was supported by the Ministry of Education and Training (MoET), Vietnam, under Grant B2021-MDA-12.

ABSTRACT Groundwater spring plays a crucial role in human life, including water resource management and planning; therefore, developing accurate prediction models for groundwater spring potential mapping is essential. The objective of this research is to introduce and confirm a new modeling approach based on TensorFlow Deep Neural Networks (TF-DNN) and multisource geospatial data for spatial prediction of groundwater spring potential, with a case study in the tropical province in the central highland of Vietnam. For this task, the TF-DNN model structure with three hidden layers with 32 neurons each was established; therein, the Adaptive Moment Estimation (ADAM) algorithm was used as an optimizer, whereas the Rectified Linear Unit (ReLU) was used as the activation function, and the sigmoid was used as the transfer function. A geospatial database for the study area, consisting of 733 groundwater spring locations and 12 influencing factors, was prepared in ArcGIS Pro. Then, it was used to develop and verify the TF-DNN model. Decision Tree, Support Vector Machine, Logistic Regression, Random Forest, and Classification and Regression Trees were used as a benchmark for the model comparison. The results demonstrate that the proposed TF-DNN model (Accuracy = 80.5%, F-score = 0.797, and AUC = 0.864) achieves a high global prediction performance, outperforming the benchmark models. Thus, the TF-DNN represents a novel and effective tool for spatially predicting groundwater spring potential mapping. The groundwater spring potential map generated in this study has the potential to assist provincial authorities in formulating strategies concerning water management and socio-economic development.

INDEX TERMS Groundwater, Tensorflow, deep neural networks, geospatial data, Vietnam.

I. INTRODUCTION

Groundwater is an indispensable resource for societies as it serves various vital purposes, including agricultural cultivation, industrial activities, and drinking water supply [1]; however, the availability of groundwater is not limitless [2], thereby necessitating the need for accurate forecasting, effective management, and responsible utilization. As a result of socio-economic development, population growth, and the adverse impacts of climate change, the rising demand for water supply underscores the urgency of implementing sustainable management practices for groundwater resources [3]. Ensuring the sustainable use of these resources has become a vital priority for safeguarding human well-being. In this regard, the utilization of a potential mapping method for groundwater plays a pivotal role in

The associate editor coordinating the review of this manuscript and approving it for publication was Stefania Bonafoni.

the comprehensive assessment and effective management of groundwater resources. Such a method enables the identification of regions with a high potential for groundwater occurrence and facilitates the estimation of water volumes within targeted areas. By employing this approach, policymakers and resource managers can make informed decisions to optimize the utilization and conservation of groundwater resources.

Different approaches have been applied to map the potential areas of groundwater [4]. First, groundwater potential mapping has frequently been conducted by using statistical approaches, including discriminant analysis and logistic regression [5]. These approaches are commonly simple to implement, but they are based on analyzing a linear relationship between the values of independent and dependent input data, which may contain errors. Although they often require a limited number of samples, they need considerable knowledge from experts to decide on suitable predictors, which may not always provide accurate outputs [6].

Other widely used approaches to groundwater potential mapping are geospatial methods, including weighted overlay and fuzzy logic [7]. These methods utilize various spatial variables, including land use/land cover, soil types, and topography, to create potential maps of groundwater. These variables impact the occurrence of groundwater and can address incomplete or missing data. Specifically, the weighted overlay is based on the relative importance of each independent variable to weigh their impact on predicting the occurrence of groundwater [8]. A statistical analysis or expert knowledge is applied to decide weights, and then the method creates a weighted quantity of the input variables to generate a potential map of groundwater. Meanwhile, fuzzy logic uses fuzzy set theory with various spatial datasets, which can address incomplete or missing data and then predict groundwater potential maps [9]. Nevertheless, the truthfulness of the results varies according to the accuracy and availability of the input data, which are likely challenging to acquire for some regions [10], [11].

In the last decade, machine learning-based approaches have been increasingly applied for mapping potential groundwater areas [12], [13]. These methods offer several advantages, including the ability to address complex and nonlinear relationships between groundwater occurrence and the associated predictors, resulting in more accurate and reliable outcomes. Notable machine learning-based methods for this purpose include support vector machines [14], neural-fuzzy systems [15], artificial neural networks [16], random forests [17], and decision trees [18]. These methods exhibit flexibility, the ability to handle intricate relationships, and adaptability, making them highly suitable for groundwater modeling. As a result, they provide improved results and valuable insights for effective groundwater management and informed decision-making processes.

In recent years, the development of information technology with integrated and ensemble algorithms [19], [20] has offered many new solutions for the field of groundwater research, i.e., Bayesian ensemble [21], Naïve Bayes ensemble [22], multi-model ensemble [23], boosting and bagging [24], cluster ensemble [25], ensemble Kalman filter [26], and iterative ensemble smoother [27], [28]. Overall, the proposed integration and ensemble of multiple models have consistently improved performance in groundwater modeling tasks. The ability to combine models, capture nonlinear relationships, handle uncertainty, generalize well, and adapt to different geo-environmental data sources contribute to the superior performance of integrated and ensemble algorithms in these groundwater modeling applications.

In more recent years, the development of remote sensing technology, geographic information systems (GIS), and deep learning has significantly advanced the spatial modeling of various geo-information science domains [29], [30], [31], including groundwater studies. Remote sensing involves the acquisition of data about the Earth's surface from a distance, typically using satellites or aerial platforms, i.e., Sentinel images [32], Landsat 9 OLI-2 satellite imagery [33], and UAV-LiDAR (Light Detection and Ranging) data [34], provide available valuable information for groundwater modeling. While GIS technology, i.e., ArcGIS Pro and Google Earth Engine, plays a crucial role in managing and analyzing spatial data for groundwater modeling, including geospatial data integration and spatial analysis [35], [36]. Deep learning has shown significant promise in groundwater modeling [37], [38] due to its ability to learn and extract complex patterns from large datasets automatically. More importantly, deep learning can fuse various multisource geospatial data [39], such as remote sensing derived data, soil layer, hydrogeological parameters, and climate data, to improve the accuracy of groundwater models. It enables the integration of diverse information and enhances the representation of groundwater system dynamics. Nevertheless, exploring deep learning and multisource geospatial data for spatial prediction of groundwater spring potential is still limited in the literature.

In this study, we propose and validate a novel modeling approach for groundwater spring potential mapping, addressing the above limitation of literature. Our approach utilizes TensorFlow Deep Neural Networks (TF-DNNs) and multisource geospatial data. TensorFlow, an open-source software library for machine learning and artificial intelligence, provides a flexible framework for constructing deep neural networks. Our model effectively captures the intricate and nonlinear relationships between groundwater occurrence and various influencing factors by employing deep neural networks. We incorporate a range of influential variables, including slope, aspect, elevation, curvature, land use/land cover, distance to fault, distance to river, lithology, rainfall, and indexes (e.g., NDVI, NDMI, NDWI), as input features to calculate the probability of groundwater occurrence in the target area. In order to create groundwater potential maps, we integrate our model outputs with GIS functionalities. This integration allows for the delineation of areas

with high groundwater potential. The utilization of TF-DNNs eliminates the need for expert knowledge and manual intervention, reducing the likelihood of human errors. Moreover, it enhances the accuracy and robustness of the results, overcoming the limitations of existing methods. By adopting this approach, we aim to contribute valuable scientific evidence to support decision-making in the sustainable management and utilization of groundwater resources.

The research aims to introduce and evaluate a novel approach for spatial prediction of groundwater spring potential using TensorFlow Deep Neural Networks and multisource geospatial data. Specifically, it focuses on enhancing the accuracy of groundwater spring potential mapping in tropical areas, such as the Kon Tum province in Vietnam.

The subsequent sections of the paper are structured as follows: Section II delves into the background of the algorithms employed. In Section III, the study area and data are elucidated. Section IV introduces the proposed Tensorflow Deep Neural Network for spatial prediction of groundwater springs potential integrated with GIS. Section V encompasses the results and discussions. The paper concludes with discussions and final remarks in Section VI and Section VII, respectively.

II. BACKGROUND OF THE ALGORITHMS USED

This section describes the background of the deep neural networks used in this research. Additionally, it outlines the key features of five benchmark algorithms, DT (Decision Trees), SVM (Support Vector Machines), LR (Logistic Regression), RF (Random Forest), and CART (Classification and Regression Trees), for the purpose of comparisons.

A. DEEP NEURAL NETWORKS

Deep learning is a subfield of machine learning that has gained significant attention in recent years, especially in geosciences [40], [41], [42], due to its ability to learn highly complex and abstract representations of geo-environmental data automatically. It is based on the use of neural networks, which are composed of interconnected nodes or neurons that mimic the behavior of neurons in the human brain. Herein, deep learning algorithms learn features directly from raw data by utilizing a hierarchical architecture of multiple layers of processing [43], which allows the network to progressively extract and transform features from the input data from simple and local patterns in the lower layers to more abstract and global representations in the higher layers. The most popular deep learning frameworks are shown in Table 1.

Of the various deep learning frameworks, TensorFlow, which was developed by Google [44], was chosen for this research due to its high flexibility in constructing deep neural network models. The main strength of TensorFlow lies in its ability to handle computations with large amounts of data and parameters and to distribute them across multiple devices for parallel processing [52]. This makes it ideal for training deep neural networks on large datasets. Moreover, TensorFlow offers a wide range of high-level APIs and tools that facilitate the development of deep learning models, such as ArcGIS Pro deep learning, for building and training neural networks.

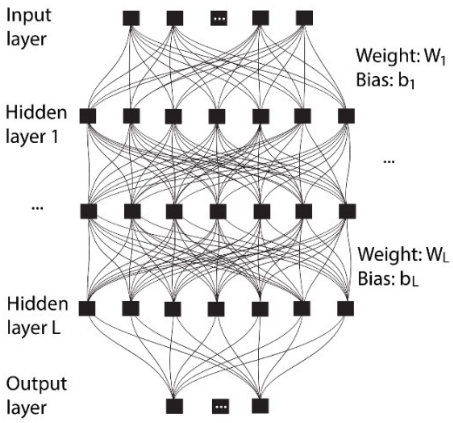

FIGURE 1. A typical structure of a deep neural network.

Let's consider a training dataset DT $\in$ (X, y), with X $\in$ R*m*,y $\in$ [1,0], and m *is the number of input factors.* In this research context, X is a matrix consisting of 12 groundwater spring influencing (GSI) factors (m =12), whereas y $\in$ (0,1) denotes two classes, groundwater spring and non-groundwater spring. The TFDeepNN aims to build an inference model f (X): R$d \rightarrow$ [1,0] to infer 12 GSI factors into groundwater spring indices. Subsequently, these indices are utilized to produce a map indicating the potential locations of groundwater springs using ArcGIS Pro. A typical deep learning structure network is shown in Figure 1, featuring one input layer, one output layer, and some hidden layers.

B. BENCHMARK MACHINE LEARNING

In order to confirm the effectiveness of the proposed TF-DNN model, five benchmark algorithms, DT (Decision Trees), SVM (Support Vector Machines), LR (Logistic Regression), RF (Random Forest), and CART (Classification and Regression Trees) were used for the purpose of performance comparison. They are well-known and widely used machine learning algorithms in the field of environmental modeling. Herein, the selected algorithms represent different types of machine learning techniques to ensure the presence of diversity. DT is based on a series of if-else conditions, SVM finds optimal hyperplanes for classification, LR models probabilities, RF combines multiple decision trees, and CART can handle both classification and regression tasks [53], [54], [55]. Because the detailed description of these benchmark algorithms is overwhelming in literature, only salient features were outlined here.

In this research, the DT algorithm, which is a Java implementation (J48) of the C4.5 algorithm introduced by Quinlan [56], was used. This algorithm has been widely used in environmental modeling applications because it can generate interpretable models and handle various types of multisource geospatial data. The algorithm utilizes the groundwater training dataset DT to construct a decision tree,

TABLE 1. Popular deep-learning frameworks.

No.	Framework	Developer	Language	Reference
1	TensorFlow	Google Inc.	Python	Abadi, et al. [44]
2	Keras	François Chollet	Python	Chollet [45]
3	PyTorch	Facebook	Python	Paszke, et al. [46]
4	Caffe	Berkeley AI Research	Python and Matlab	Jia, et al. [47]
5	Theano	Université de Montréal	Python and C++	Bastien, et al. [48]
6	CNTK	Microsoft Research Lab	Python, Java, and C++	Seide and Agarwal [49]
7	MXNet	Microsoft, Nvidia, Intel, and others	Python, C++, R, and JavaScript	Chen, et al. [50]
8	DeepLearning4j	Skymind	Java	Team [51]

beginning with a root node and progressively expanding to leaf nodes until the tree is fully formed. The minimum number of instances per leaf and the confidence factor are salient parameters controlling the performance of the DT.

SVM, which was introduced by Vapnik [57], is a powerful machine learning algorithm and widely applied to groundwater modeling [58] due to its ability to handle non-linearly separable geospatial data by transforming it into a higher-dimensional space. It is effective in high-dimensional spaces and offers flexibility in capturing complex decision boundaries without explicitly transforming the data [59]. In the SVM with Radial Basis Function (RBF), the hyperparameters C and gamma are crucial in influencing the model's performance. Therefore, they should be carefully selected.

LR, which was developed by Cox [60] is a widely used statistical model for groundwater mapping [59], [61]. It estimates the probability of an input sample belonging to one of two classes and maps it to probabilities by using the sigmoid function. The probability is used as the groundwater spring potential index, with 1– groundwater and 0 – non-groundwater. Thus, the LR model is simple yet effective, making it a popular choice for various environmental applications, including groundwater studies.

Introduced by Breiman [62], RF is a potent ensemble learning method known for its exceptional performance in various environmental modeling domains, including groundwater studies [63], [64]. It constructs an ensemble of decision trees by randomly generating subsets from the groundwater training dataset DT and building a tree for each subset. The final groundwater indices are obtained through majority voting (for classification) or averaging (for regression) of individual tree predictions. This makes RF a versatile and powerful algorithm for spatial prediction in diverse environmental domains. Thus, the number of decision trees in the RF is the most important parameter.

Regarding CART, which was also introduced by Breiman [65], this algorithm constructs a tree-like model by recursively splitting the groundwater data into subsets based on the most significant input factors. CART can be used for classification and regression tasks and is valued for its simplicity, interpretability, and ability to handle discrete and continuous geospatial data.

III. STUDY AREA AND DATA

A. STUDY AREA

The study area is the Kon Tum province, located in the Central Highland of Vietnam, between longitudes 107°19′ E and 108°33′ E and latitudes 13°53′ N and 15°25′ N. The province covers an area of 9691.6 km2 and has a diverse landscape characterized by mountains, hills, and valleys. The altitude ranges from approximately 140.1 m in the Sa Thay area in the south to 2600.3 m above sea level in the Tu Mo Rong in the north (Figure 2), with a mean of 882.2 m and a standard deviation of 391.4 m. The slope in the study area ranges from 0o to 86.8o. Approximately 79.7% of the study area consists of slopes less than 25°, while only 3.9% has slopes greater than 35°. Based on our statistical analysis of land use/land cover, approximately 70.6% of the province is covered by forest land, with deciduous broadleaf forest being the dominant type. Woody crops account for 24.2% of the study area. Scrub/Shrub land, rice paddies, and residential areas comprise 1.6%, 1.4%, and 0.7% of the region.

The province belongs to the tropical monsoon climate, which is characterized by distinct wet and dry seasons. The wet season typically lasts from May to October, whereas the dry season usually from November to April. The average annual rainfall in Kon Tum province varies from 1700 mm to above 3000 mm [66]. The province's mountainous terrain contributes to rainfall variability across different areas, with some regions experiencing more precipitation than others [67].

Regarding the geology, the study area is situated within the Kon Tum uplift block, positioned on the eastern side of the Indosinian orogeny [66]. Over 33 geological formations or complexes have been identified; however, their distribution varies significantly (Figure 3 and Table 2). Ten formations or complexes account for 87.8% of the study area (Table 2). Around 22.8% of the study area is covered by Tac Po formation. It is followed by Hai Van complex (13.5%), Van Canh complex (8.8%), Kham Duc formation (8.5%), BG-QS complex (8.3%), and Mang Yang formation (8.1%) (Table 2).

Particularly 81.1% of the groundwater spring locations observed in the study area are located in the Tac Po formation (15.6%), Kham Duc formation (14.9%), Van Canh complex (12.7%), Mang Yang formation (9.4%), Hai Van complex (7.4%), Dak Long formation (6.2%),

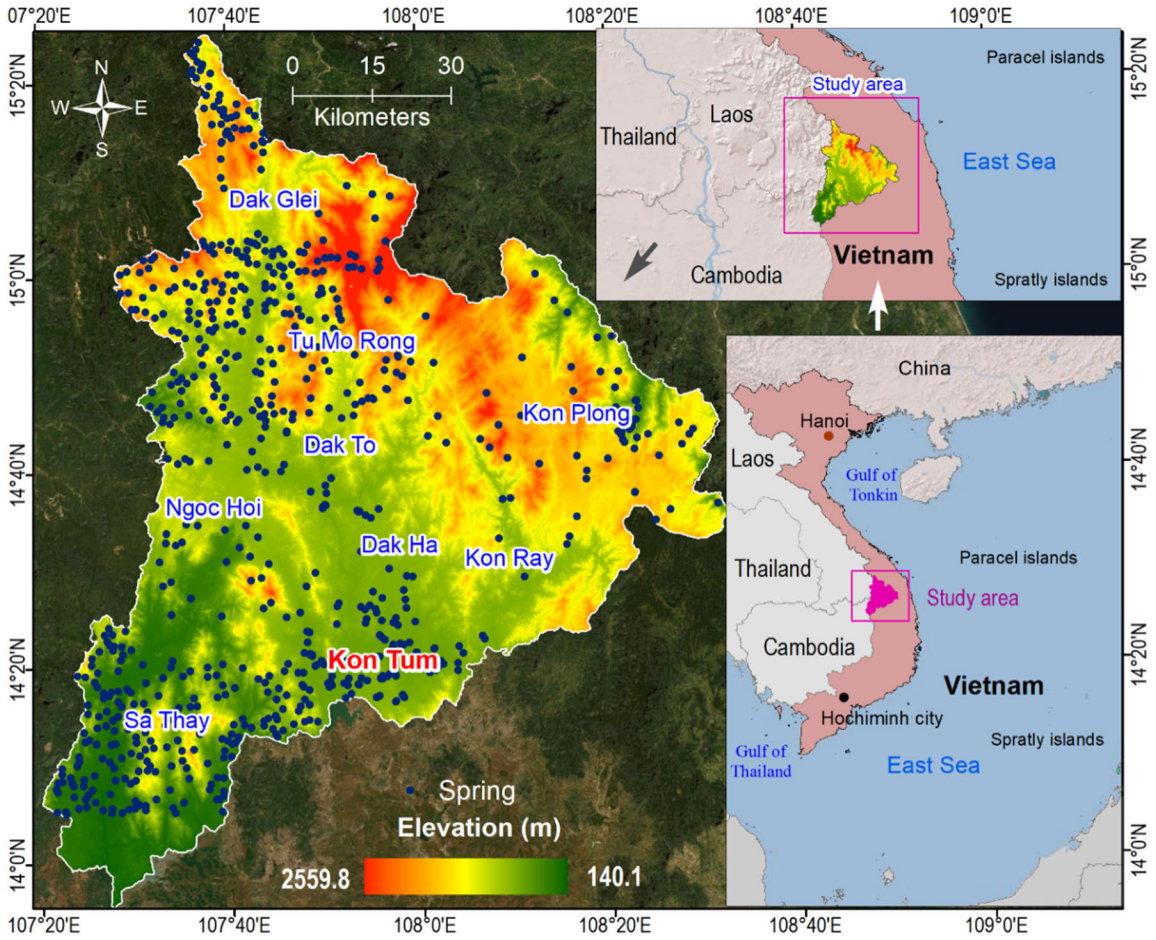

FIGURE 2. Location of Kon Tum province and groundwater spring locations.

BS-QS complex (5.1%), Kon Tum formation (5.1%), and Tuc Trung formation (4.8%) (Table 2). The main lithologies are gneiss, plagiogneiss, schist, granite, granodiorite, conglomerate, sandstone, biotite, quartzite, schist, and tholeiitic.

B. GROUNDWATER SPRING LOCATIONS

The modeling approach in this research uses available groundwater spring as a key for spatial predictions of groundwater spring potential. Therefore, the collection of groundwater spring locations in the Kon Tum province is an essential task. For this task, a total of 733 groundwater spring locations were derived from fieldwork under B2021-MDA-12 project and from the national projects on groundwater in the Central Highlands of Vietnam [68], [69], [70], [71]. Accordingly, each spring location was surveyed and recorded using handheld GPS devices. The coordinates of these points were then imported into ArcGIS Pro to create a comprehensive groundwater spring inventory map. The measured water flow at these locations ranges from 0.01 to 10.89 l/s, with an average of 0.37 l/s. Notably, 54 locations exhibit water flow greater than 1.0 l/s, while 652 locations have water flow ranging from 0.1 l/s to 0.91 l/s. The degree of the mineralization of the water in these locations ranges from 0.01 to 0.99 g/L, which is within the permissible standard for the domestic water supply. Among them, only five locations have a mineralization larger than 0.5 g/L, whereas 412 locations have a mineralization from 0.1 to 0.5 g/L.

C. INFLUENCING FACTORS

Determining the influencing factors for groundwater spring modeling is crucial for accurately understanding and predicting the behavior of groundwater springs. These influencing factors can vary depending on the specific characteristics of the geo-environmental conditions in the Area being considered. Through the analysis of groundwater spring locations and the characteristics of the study area, a total of 12 influencing factors were considered: slope, aspect, elevation, curvature, Landuse/landcover (LULC), NDVI, NDMI, NDWI, distance to fault; distance to river; lithology, and rainfall.

First, a digital elevation model (DEM) for the Kon Tum province is derived from ALOS Global Digital Surface Model "ALOS World 3D - 30m (AW3D30)", measured in degrees

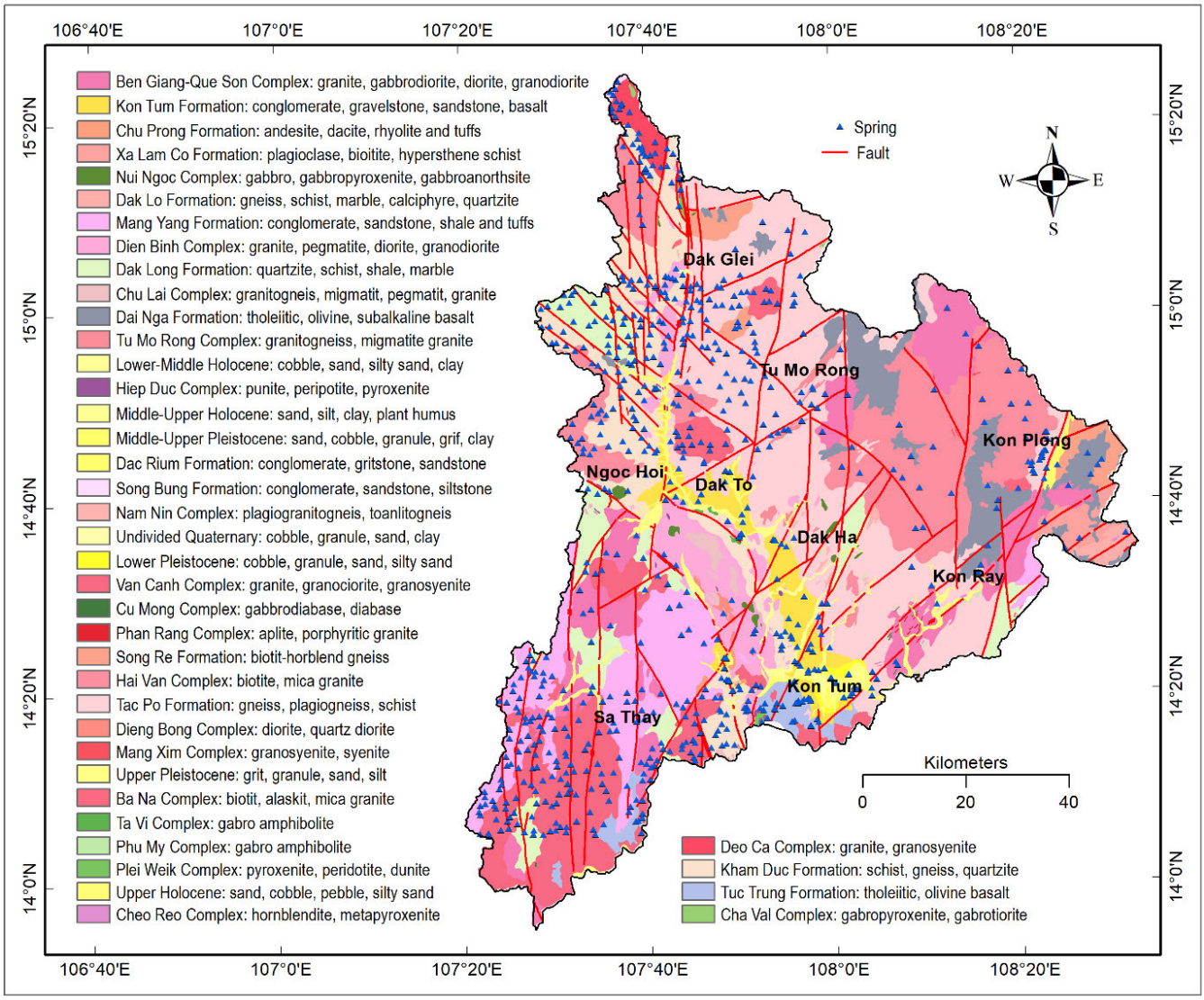

FIGURE 3. Geologic map of the Kon Tum province.

(Figure 2) [72]. Then, four topographic factors, slope, aspect, elevation, and curvature, were extracted and constructed based on the DEM. These topographic factors should be selected because they control groundwater flow and the formation of springs. Slope refers to the steepness or inclination of the land surface [73] that plays a significant role in groundwater flow patterns. Herein, groundwater tends to flow along the path of least resistance [74], typically toward the steepest slope. Aspect can influence the amount of solar radiation and moisture distribution on the land surface, potentially affecting groundwater recharge rates [75]. Elevation influences the gravitational potential energy of groundwater and can determine the direction and speed of groundwater flow [76]. Springs often occur where the water table intersects the land surface, which is more likely to happen in lower-lying regions or at specific elevation thresholds [77].

Curvature refers to the degree of curvature in the land surface [78]. It is essential in groundwater potential mapping as it impacts the direction and rate of water flow, impacting groundwater recharge and storage. Herein, areas of positive curvature, such as hills or ridges, may indicate groundwater recharge zones where water is infiltrating the ground. In contrast, areas of negative curvature, such as valleys, may represent groundwater discharge zones where water is discharged as spring [79].

LULC refers to the types of vegetation, crops, and land use activities in a specific region, which is an essential factor in groundwater potential mapping because it impacts the rate of infiltration, evapotranspiration, and soil moisture [4]. In this study, we use the High-Resolution Land Use and Land Cover (LULC) Map from the Japan Aerospace Exploration Agency (JAXA) Earth Observation Research Center [80]. It was

TABLE 2. Overview of the geological formations and complexes in the Kon Tum province.

No.	Formation/Complex	Area (%)	GS location (%)	Main lithology
1	Tac Po formation	22.8	15.6	Gneiss, plagiogneiss, schist
2	Hai Van complex	13.5	7.4	Biotite and mica granite
3	Van Canh complex	8.8	12.7	Granite, granociorite, and granosyenite
4	Kham Duc formation	8.5	14.9	Schist, gneiss, and quartzite
5	BG-QS complex	8.3	5.1	Granite, gabbrodiorite, and diorite
6	Mang Yang formation	8.1	9.4	Conglomerate, sandstone, and shale
7	Dak Long formation	5.5	6.2	Quartzite, schist, and shale
8	Dai Nga formation	5.0	1.4	Tholeiitic, olivine, and sub-alkaline basalt
9	Dien Binh complex	4.3	4.5	Granite, pegmatite, and diorite
10	Kon Tum formation	3.1	5.1	Conglomerate, gravestone, and sandstone
11	Song Re formation	1.6	1.0	Biotit-horblend gneiss
12	Upper Holocene	1.5	1.2	Sand, cobble, pebble, silty sand
13	Tuc Trung formation	1.5	4.8	Tholeiitic and olivine basalt
14	Ba Na complex	1.3	2.7	Biotite and mica granite
15	Lower-Middle Holocene	1.1	1.0	Cobble, sand, silty sand, and clay
16	Deo Ca complex	0.9	2.3	Granite and granosyenite
17	Xa Lam co formation	0.9	0.4	Plagioclase, biotite, and hypersthene schist
18	Dak Lo formation	0.5	0.1	Gneiss, schist, marble, and caliphate
19	Chu Lai complex	0.4	0.1	Granitogneis, migmatit, and pegmatit
20	Upper Pleistocene	0.4	0.8	Grit, granules, sand, and silt
21	Tu Mo Rong complex	0.4	0.8	Granitogneiss and migmatite granite
22	Others	1.8	2.5	Sand, cobble, gabbro, and conglomerate

created using multiple remotely sensed databases (Landsat, Sentinel, and ALOS.) and field survey data. The LULC of the study area is complex, with a mix of different land types, including forested areas, croplands, and scrub/shrub areas (Figure 4e). However, forest and scrub/shrub lands are the most dominant areas in the region.

Indexes can provide additional information in groundwater potential mapping [81]. For example, the Normalized Difference Vegetation Index (NDVI) [82], Normalized Difference Moisture Index (NDMI), and Normalized Difference Water Index (NDWI) [83] are able to provide information on vegetation cover, moisture content, and water content, respectively. These indexes can be used to identify areas with high vegetation cover, soil moisture, or water content, which can indicate potential groundwater recharge areas. This study uses Landsat 8 OLI to extract these indexes (Figures 4f, 4g, and 4h).

The distance to faults in an area is likely to impact the groundwater potential of the region [84]. Faults can act as conduits for groundwater flow and improve adjacent areas' recharge rates. However, faults can cause water loss through leakage or lead to the depletion of groundwater resources. Here, data on fault locations and their attributes are obtained from the Geological and Mineral Resources Map (1:200,000) created by the Ministry of Natural Resources and Environment (Vietnam) [85]. As shown in Figure 4i, most faults lie in the north-south direction.

The distance to rivers is vital in replenishing groundwater resources. The proximity of a study area to a river can affect its groundwater potential. Rivers can be sources of recharge that increase the water table elevation. In other words, the over-extraction of groundwater near rivers can cause a reduction in river flow, likely leading to environmental issues. In this study, we extract data on river locations and their attributes from the National Topographic Map (1:50,000) developed by the Ministry of Natural Resources and Environment (Vietnam) [86]. We derive river networks by utilizing ArcGIS Pro software. As shown in Figure 4j, the study catchment has a dense river network.

Lithology with the rock types and their structure can impact the groundwater potential [87]. Some rocks are more permeable than others, allowing more effortless water flow and recharge. Impermeable layers can hinder groundwater flow, while faults can act as conduits. Data on lithology are obtained from the Geological and Mineral Resources Map (1:200,000) created by the Ministry of Natural Resources and Environment (Vietnam). The lithology of the study area is diverse, with 40 formations, complexes, and layers (Figure 3), and they were grouped into 28 categories (Figure 4k).

Rainfall is one of the primary sources of groundwater recharge. The amount and distribution of rainfall can impact the groundwater potential of an area. Areas with higher rainfall have a higher recharge rate, leading to higher groundwater potential. We derive rainfall data from the POWER project, National Aeronautics and Space Administration (NASA), USA. In this research, the total rainfall in the last 30 years was used to generate the rainfall map using the Inverse Distance Weight method [88]. As can be seen from Figure 4m, rainfall is highly distributed in the east-north rather than the west-south. It ranges from around 40,000 mm to 60,000 mm over the study area.

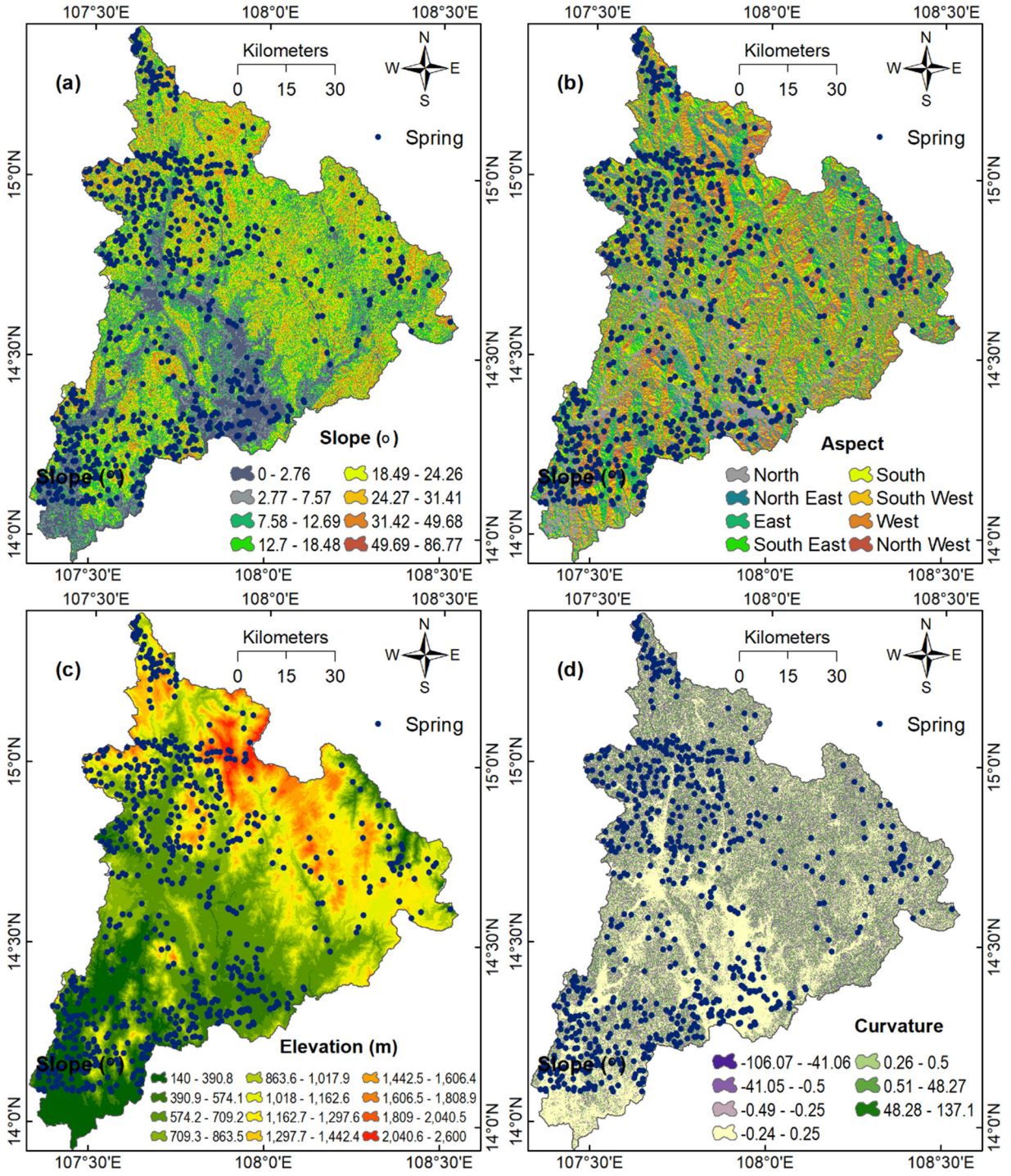

FIGURE 4. Influencing factors: (a) Slope, (b) Aspect, (c) Elevation, and (d) Curvature.

IV. TENSORFLOW DEEP NEURAL NETWORK FOR SPATIAL PREDICTION OF GROUNDWATER SPRINGS POTENTIAL WITH GIS

This section presents the proposed approach (see Figure 5) for predicting the potential locations of groundwater springs using a combination of TensorFlow Deep Neural Network (TF-DNN), multi-geospatial data, and GIS. In this study, ArcGIS Pro 3.0 was utilized to process the groundwater springs inventory and the influencing factors. The Deep Learning Framework within ArcGIS Pro 3.0 was employed, which includes several pre-built modules for TensorFlow, Keras, and PyTorch for training and validating the TF-DNN model. The Python code used to develop the TF-DNN can be accessed at www.tensorflow.org. In addition, the

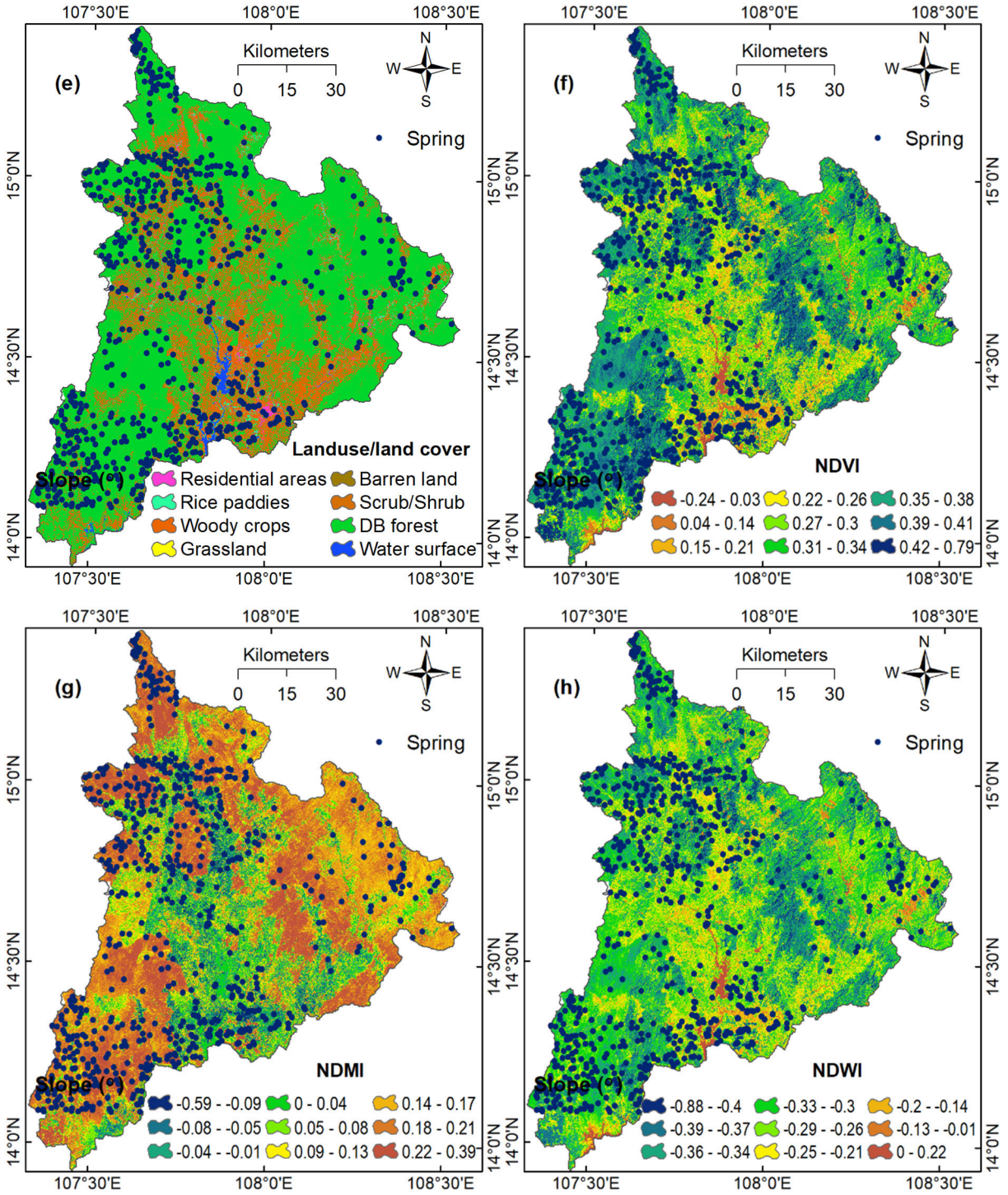

FIGURE 4. *(Continued.)* **Influencing factors: (e) LULC, (f) NDVI, (g)NDMI, and (h) NDWI.**

authors created a Python script that operates within the Spyder 3.9 environment (available at www.spyder-ide.org). The script's purpose is to normalize and encode the raster maps of the influencing factors, extract them into both the training and validation datasets, and ultimately link them to the TF-DNN model. The script used was also used to compute the groundwater springs potential (GSP) index as the output of the TF-DNN model, which was then transformed into a GSP map. Regarding the three benchmarked methods, DT, SVM, LR, RF, and CART, the Python Wrapper for the Weka (https://github.com/fracpete/python-weka-wrapper3) [89] was employed to build these models.

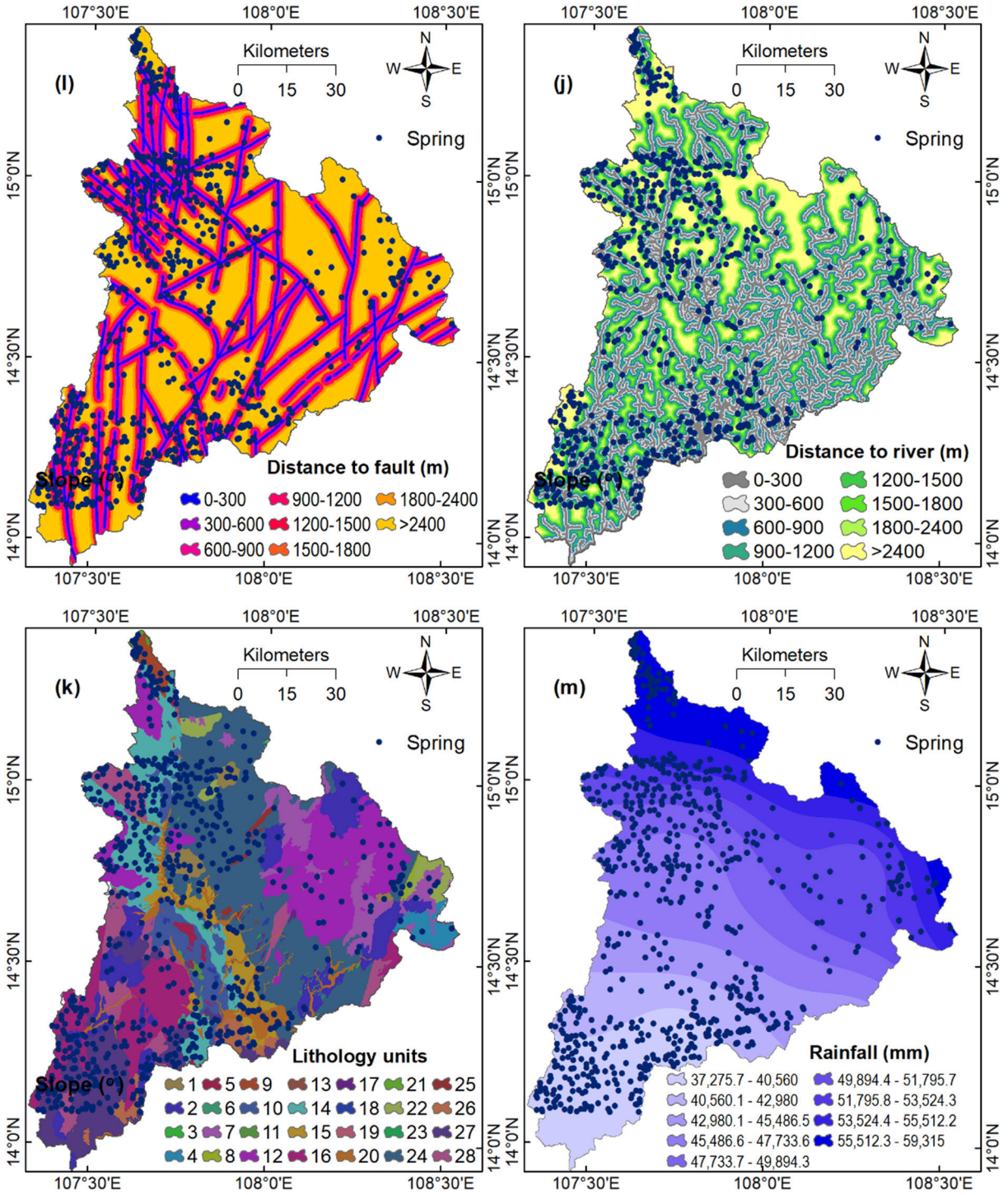

FIGURE 4. *(Continued.)* **Inluencing factors: (i) Distance to fault; (j) Distance to river; (k) Lithology, and (m) Rainfall.**

A. GROUNDWATER SPRING DATABASE

For groundwater springs potential modeling with TF-DNN, creating a groundwater spring database in a GIS is essential because it enables the integration of multisource geospatial data, including geological maps, topographical information, and satellite imagery. Combining these geospatial data allows a more comprehensive analysis of the study area, leading to a better understanding of the factors that influence the location of groundwater springs. For this task, a groundwater database was established for the study area in ArcGIS Pro 3.0, utilizing the ESRI File Geodatabase format. This format is known for its durability and dependability, as it can effectively manage

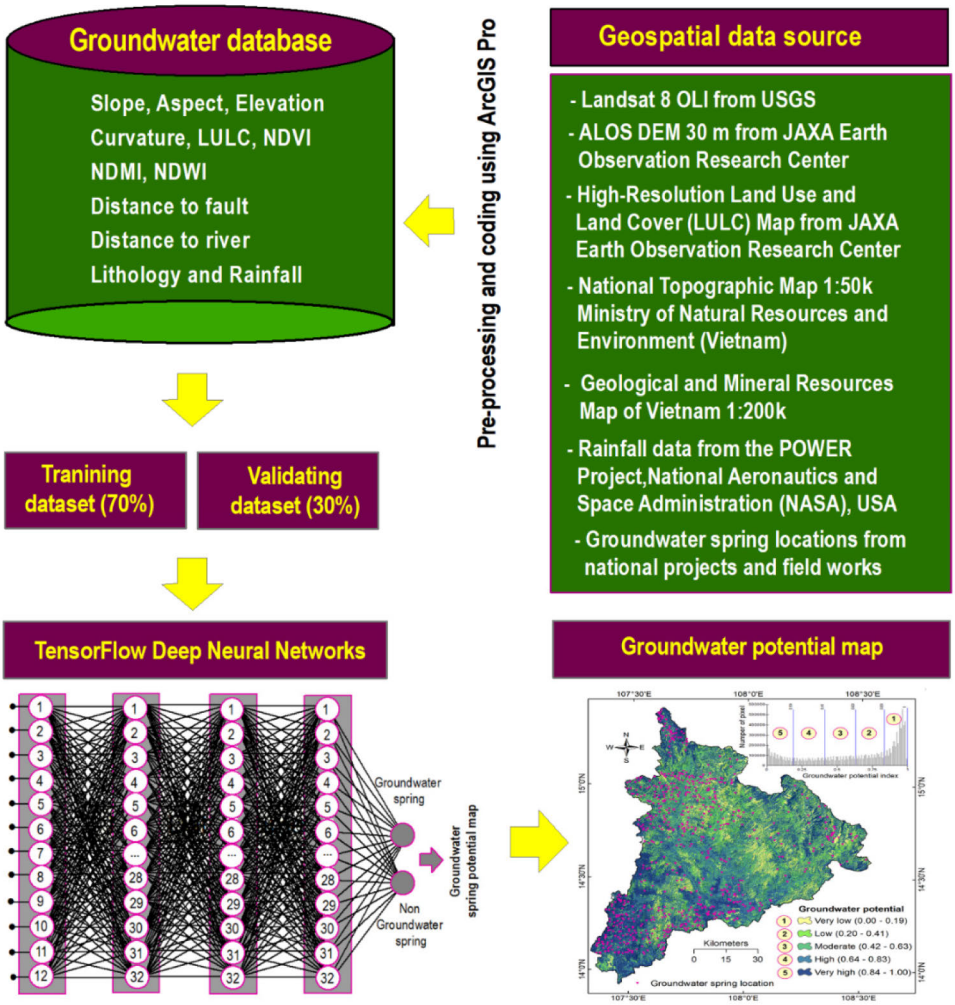

FIGURE 5. Flowchart of the proposed TF-DNN for spatial prediction of groundwater potential.

extensive amounts of data while maintaining high performance [90].

Besides, the ESRI File Geodatabase offers several advantages over other multisource geospatial data formats. Its ability to process multisource data with different coordinate systems and bring them into a unique coordinate system. This feature ensures that all pixels from different GIS layers can be superimposed accurately, which is crucial for precise groundwater spring modeling in this research. It is noted that the ability to manage data from multiple sources with varying coordinate systems is especially important in complex geospatial projects, where datasets may come from various sources and require integration into a unified system.

After completing the multisource geospatial data processing, illustrated in Figure 5, ten influencing factors for groundwater spring potential were derived. These factors included slope, aspect, elevation, curvature, land use and land cover (LULC), normalized difference vegetation index (NDVI), normalized difference moisture index (NDMI), normalized difference water index (NDWI), distance to fault, distance to river, lithology, and rainfall. The resulting data had a spatial resolution of 30 meters and was projected in UTM Zone 48 before being exported into the groundwater spring database. This process allowed for the integration of a range of geospatial data types into the database for further analysis and modeling. In the next step, the ten influencing factors were normalized into the range 0.01 to 0.09 [91] using the following equation.

$$NF_i = \frac{IF_i - Min(IF)}{Max\,(IF) - Min(IF)}[0.99 - 0.01] + 0.01 \qquad (1)$$

where IF_i is the current value of the pixel of the influencing factor, $Max(IF)$ and Min (IF) are the maximum value and the minimum value of that factor, and NF_i is the calculated value for that factor.

Next, the process involved importing a total of 733 groundwater spring locations to the database, which were subsequently randomly split into two parts, following a 70:30 ratio. Part 1, consisting of 513 locations, was used to build the training dataset, while Part 2, containing 220 locations,

was utilized for model validation. In this study, modeling groundwater spring potential was considered a binary classification problem, where the probability of belonging to the groundwater spring class was used as the groundwater spring index. Therefore, an equivalent number of locations were randomly chosen from non-groundwater spring areas within the study area. The groundwater spring locations were assigned a value of "1," while a value of "0" was assigned to the non-groundwater spring locations. Finally, an extraction process was performed to create ten raster values for these locations, and the training and validation datasets were derived.

B. INFLUENCING FACTOR ANALYSIS

Feature selection in groundwater spring potential modeling is crucial due to the analysis of numerous factors associated with geological, topographic, hydrological, and land cover attributes. Therefore, it is imperative to identify the most informative and pertinent factors. This process simplifies the modeling procedure and improves model performance. Incorporating irrelevant or redundant features in a model can result in overfitting, reduced model accuracy, and heightened computational complexity. Thus, careful selection of features becomes essential.

In this analysis, a variable selection procedure using the wrapper random forests classifier (VSP-WrapRFC) [92], [93] was adopted for the Influencing factor analysis and quantifying the most informative and pertinent factors for groundwater spring potential modeling. The VSP-WrapRFC was selected due to its comprehensive evaluation, ability to capture nonlinear relationships and interactions, robustness to noise and outliers, handling of correlated features, and flexibility in different domains [94], [95].

C. SETTING UP THE TENSORFLOW DEEP NEURAL NETWORK

For groundwater spring potential (GSP) modeling, a deep neural network model is designed to be a pattern recognizer to infer the input factors to two classes, groundwater springs and non-groundwater springs. A key feature of designing a deep neural network is the ability to optimize the network's parameters and architecture for the GSP modeling task at hand. This involves determining the number of layers and neurons in the network, selecting appropriate activation functions, and adopting the optimizer function.

The architecture of the TensorFlow deep neural network (TF-DNN) used in this research is depicted in Figure 6, consisting of an input layer, three hidden layers, and an output layer. The input layer receives information related to the 12 GSI factors and passes it on to the three hidden layers for analysis, allowing for discovering meaningful patterns in the collected data. These hidden layers extract useful knowledge, which is then used to generate numerical values within the range of [0, 1]. Finally, the output layer produces numerical values, with a threshold of 0.5 being adopted to divide the output indices into two classes for model performance evaluation. The probability belonging to the groundwater spring class is used as the GSP index and then utilized to generate a GSP map. The rectified linear unit (ReLU) (Eq.2) is selected as the activate function, whereas the sigmoid (Eq. 3) is used as the transfer function [96].

$$fa(x) = max(0, x) \tag{2}$$

$$\sigma(x) = \frac{1}{1 + exp(-x)} \tag{3}$$

where x denotes an input signal to a neuron, fa represents the activation function, and σ represents the transfer function.

A summary of the weights and biases of the TF-DNN model for groundwater spring potential mapping in this research is shown in Table 3. A total of 2561 parameters controlling the performance of the TF-DNN model are determined, including 2464 weights and 97 biases. These parameters are searched and optimized in the training phase.

For optimizing the 2561 parameters of the TF-DNN model, the Adaptive Moment Estimation (ADAM) algorithm introduced by Kingma and Ba [97] was employed due to its effectiveness and efficiency in training deep learning models in various spatial domains [66], [96], [98]. Thus, various combinations of 2561 parameters were checked through training iterations. Besides, Mean Squared Error (MSE) (Eq.4) is used to measure the overall performance of each combination.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(GSPI_i - GSPO_i)^2 \tag{4}$$

where $GSPI_i$ is the groundwater spring potential value in the inventory dataset, whereas, $GSPO_i$ is the groundwater spring potential output computed from the TF-DNN model; n is the total samples used.

TABLE 3. The weights and bias of the TF-DNN model for groundwater spring potential mapping in this research.

No.	Weights and bias of the TF-DNN model	Number of parameters
1	Input Layer	384
2	The bias of the Input Layer	32
3	Hidden Layer 1	1024
4	The bias of the Hidden Layer 1	32
5	Hidden Layer 2	1024
6	The bias of Hidden Layer 2	32
7	Hidden Layer 3	32
8	The bias of Hidden Layer 3	1

D. PERFORMANCE ASSESSMENT

In this study, the GSP modeling is treated as a form of binary pattern recognition. Consequently, we employ well-known statistical metrics such as Kappa, F-score, Receiver Operating Characteristic (ROC) curve, Area under the curve (AUC), classification accuracy (Acc), true positive (TP), true negative (TN), false positive (FP), false negative (FN), Sensitivity (Sens), and specificity (Spec) [99], [100] to evaluate and analyze the performance of the model. Because these metrics have been extensively documented in the literature, especially

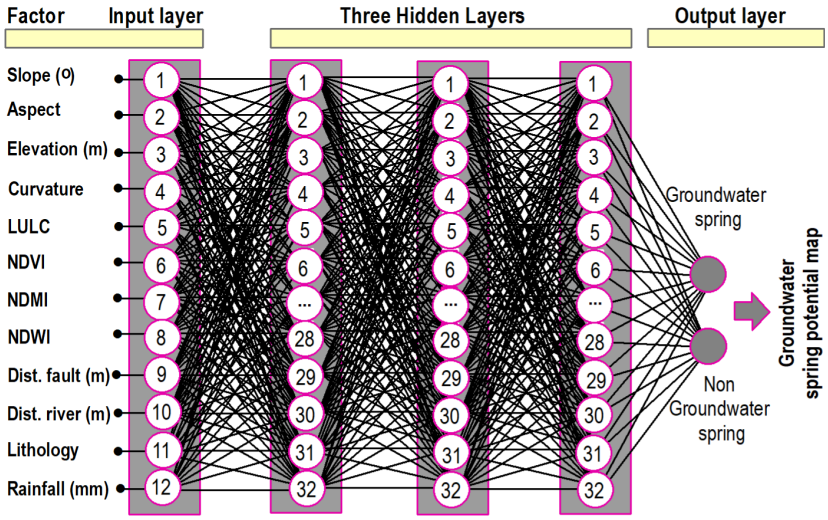

FIGURE 6. TensorFlow Deep Learning (TF-DNN) model for groundwater spring potential mapping in this research.

TABLE 4. Role of the ten groundwater spring influencing factors.

No.	Groundwater spring influencing factor	Merit	Ranking
1	NDWI	0.256	1
2	NDVI	0.186	2
3	Geology	0.155	3
4	Elevation	0.131	4
5	Slope	0.129	5
6	Rainfall	0.124	6
7	Curvature	0.099	7
8	NDMI	0.073	8
9	Landuse	0.066	9
10	Aspect	0.024	10
11	Distance to fault	0.012	11
12	Distance to river	0.006	12

concerning groundwater modeling [101], [102], [103]; therefore, we won't delve into a comprehensive description of each metric in the aforementioned papers.

V. RESULTS AND ANALYSIS

A. FACTOR ANALYSIS

The role of the ten groundwater spring influencing (GSI) factors, which impact or affect the behavior, availability, and characteristics of groundwater springs in the study area, is shown in Table 4. It could be seen that, among the 12 GSI factors, NDWI, NDVI, and geology have the highest importance. The merit value is 0.256, 0.186, and 0.155, respectively. They are followed by elevation (0.131), slope (0.129), and rainfall (0.124). In contrast, distance to river and distance to fault have the lowest contribution to the groundwater spring. The merit value is 0.006 and 0.012, respectively.

B. MODEL TRAINING AND VALIDATION

Using 12 GSI factors and the training dataset with 1026 samples, the TF-DNN model was trained using the ADAM optimizer, and the result is shown in Figure 7. Our observations indicate that the proposed TF-DNN model exhibits a strong level of fitting with the training dataset, where MSE is 0.115, Error mean is 0.026, and Error Standard is 0.338. Besides, the distribution of the error follows the normal distribution (Figure 7). As mentioned in Section IV, the groundwater spring locations were assigned a value of "1," while a value of "0" was assigned to the non-groundwater spring locations; therefore, a threshold of 0.5 was used to separate the two classes, spring and non-spring. As a result, the output values of the TF-DNN model and statistical metrics were computed and shown in Table 5. The results reveal that the TF-DNN model achieved high accuracy, with values of 84.2% for Acc, 0.684 for Kappa, and 0.837 for F-score. These metrics demonstrate the model's effective classification of the training dataset samples. The global performance of the TF-DNN model is evaluated and presented through the ROC curve and AUC (Figure 9). As illustrated in Figure 9, with an AUC value of 0.918, the TF-DNN model demonstrates its ability to effectively distinguish between groundwater spring and non-groundwater spring samples in the training dataset. The remaining statistical performance measures of the TF-DNN model on the training dataset are shown in Table 5.

C. MODEL COMPARISON

In order to draw more definitive conclusions regarding the applicability of the proposed TF-DNN model, it is crucial to validate its efficacy in spatially predicting groundwater spring potential. This entails assessing its predictive capabilities more, including a comparison with other benchmarked methods. Such an analysis will provide valuable insights into the model's performance and suitability for the specific task. For this purpose, five popular machine learning models were selected: Decision Tree (DT), Support Vector Machine

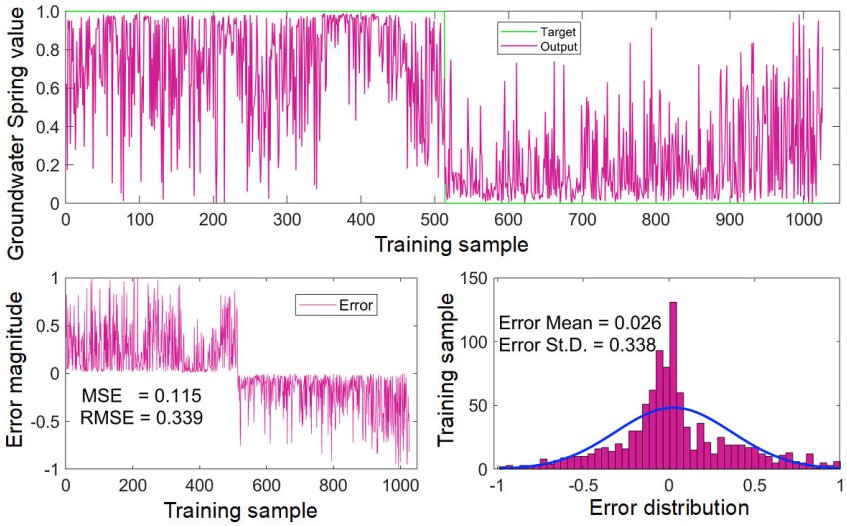

FIGURE 7. Performance of the TF-DNN model in the training dataset.

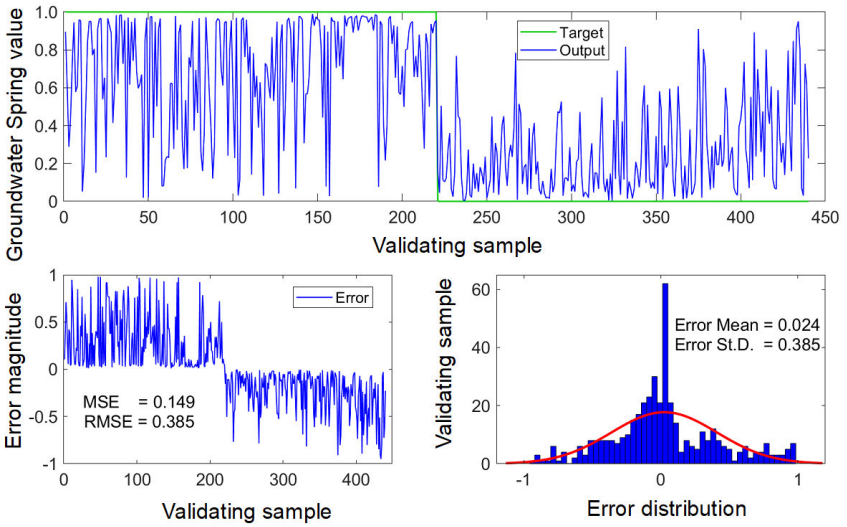

FIGURE 8. Performance of the TF-DNN model in the validating dataset.

(SVM), Logistic Regression (LR), Random Forest (RF), and Classification and Regression Trees (CART). For DT, the C4.5 algorithm was utilized with default values for the confidence factor and the minimum number of instances per leaf. The Radial Basis Function was employed with SVM, and its two parameters, C = 14 and gamma = 0.52, were determined through the grid search method. LR adopted the sigmoid function, while RF utilized 100 trees with the other parameters set to default values. For CART, all parameters were set to their default values.

The training and validating results of the six models are shown in Tables 4 and 5. The results reveal that the RF model (Acc = 80.0%, F-score = 0.800, Kappa = 0.600, and AUC = 0.888) demonstrates a good fit with the training dataset, where the DT model (Acc = 75.0%, F-score = 0.747, Kappa = 0.501, and AUC = 0.779), the SVM model (Acc = 76.0%, F-score = 0.763, Kappa = 0.520, and AUC = 0.822), the LR model (Acc = 68.5%, F-score = 0.696, Kappa = 0.370, and AUC = 0.743), and the CART model (Acc = 76.9%, F-score = 0.767, Kappa = 0.538, and AUC = 0.843) yield poorer performance. Overall, all five models, DT, SVM, LR, RF, and CART, exhibit inferior performance compared to the proposed TF-DNN model.

Regarding the prediction power, among the five models, the RF model (Acc = 79.5%, F-score = 0.791, Kappa = 0.591, and AUC = 0.887) provides the highest prediction. It is followed by the CART model (Acc = 76.8%, F-score = 0.769, Kappa = 0.536, and AUC = 0.847), the SVM model (Acc = 73.9%, F-score 0.738, Kappa = 0.477,

TABLE 5. Performance of the proposed TF-DNN model and the DT, SVM, LR, RF, and CART models in the training dataset.

Groundwater potential model	Measured metrics											
	TP	TN	FP	FN	PPV	NPV	Sens	Spec	Acc	F-score	Kappa	AUC
TF-DNN	416	448	97	65	81.1	87.3	86.5	82.2	84.2	0.837	0.684	0.918
DT	377	393	136	120	73.5	76.6	75.9	74.3	75.0	0.747	0.501	0.779
SVM	395	385	118	128	77.0	75.0	75.5	76.5	76.0	0.763	0.520	0.822
LR	370	333	143	180	72.1	64.9	67.3	70.0	68.5	0.696	0.370	0.743
RF	409	412	104	101	79.7	80.3	80.2	79.8	80.0	0.800	0.600	0.888
CART	390	399	123	114	76.0	77.8	77.4	76.4	76.9	0.767	0.538	0.843

TABLE 6. Prediction power of the proposed TF-DNN model and the DT, SVM, LR, RF, and CART models in the validating dataset.

Groundwater potential model	Measured metrics											
	TP	TN	FP	FN	PPV	NPV	Sens	Spec	Acc	F-score	Kappa	AUC
TF-DNN	169	185	51	35	76.8	84.1	82.8	78.4	80.5	0.797	0.609	0.864
DT	166	161	54	59	75.5	73.2	73.8	74.9	74.3	0.746	0.486	0.788
SVM	162	163	58	57	73.6	74.1	74.0	73.8	73.9	0.738	0.477	0.810
LR	146	133	74	87	66.4	60.5	62.7	64.3	63.4	0.645	0.268	0.683
RF	170	180	50	40	77.3	81.8	81.0	78.3	79.5	0.791	0.591	0.887
CART	170	168	50	52	77.3	76.4	76.6	77.1	76.8	0.769	0.536	0.847

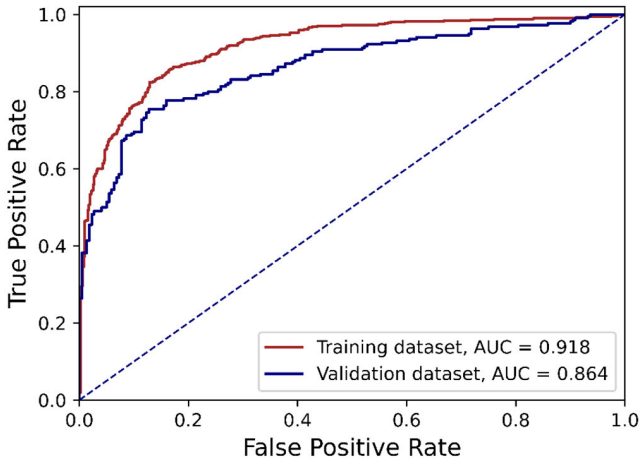

FIGURE 9. ROC curve and Area under the curve (AUC) of the TF-DNN model.

and AUC = 0.810), the DT model (Acc = 74.3%, F-score = 0.746, Kappa = 0.486, and AUC = 0.788), and the LR model (Acc = 63.4%, F-score = 0.645, Kappa = 0.268, and AUC = 0.683). Nevertheless, when compared to the proposed TF-DNN model, all four models (DT, SVM, LR, and CART) demonstrate lower predictive power, as indicated by the results presented in Table 6. For the RF model, although AUC = 0.887 is slightly higher than that of the TF-DNN model (AUC = 0.864). However, The three metrics, Acc, F-score, Kappa of the RF model is lower than those of the TF-DNN model, therefore, the prediction of the TF-DNN model is higher than that of the RF model.

We further conducted a statistical comparison to assess whether the groundwater springs prediction performance of the proposed TF-DNN model is significantly higher than that of the five benchmark models, namely, DT, SVM, LR, RF, and CART. To achieve this, we utilized the paired-samples sign test [104], where the null hypothesis (H0) assumes no significant difference in prediction performance between the TF-DNN model and the aforementioned benchmark models. Subsequently, we computed the Z-value and p-value. If the p-value is less than or equal to 0.05 and the absolute value of the observed test statistic lies outside the range of −1.96 to +1.96, we reject the null hypothesis. As a result, we established that the prediction power of the TF-DNN model is statistically significantly higher than that of the benchmark models.

The comparison results are presented in Table 7. It can be observed that all Z-values for the five pairs (TF-DNN vs. DT, TF-DNN vs. SVM, TF-DNN vs. LR, TF-DNN vs. RF, and TF-DNN vs. CART) are greater than +1.96. As a result, it is concluded that the prediction power of the TF-DNN model is statistically higher than the five benchmark models.

D. GENERATING THE GROUNDWATER SPRINGS POTENTIAL MAP

Using the trained TF-DNN model explained in the previous section, the groundwater springs potential (GSP) index for each pixel in the study area was computed. The GSP indices were converted into the ESRI File Geodatabase format, as described in Section IV-A, to align with the established georeferenced settings (i.e., UTM Zone 48) implemented in this project. Herein, the GSP indices vary from 0.00 to 1.00. Finally, the GSP map was derived by reclassifying the GSP indices into five zones using the natural break tool in ArcGIS Pro: very low (0.00-0.19), low (0.20-0.41), moderate (0.42-0.63), high (0.64-0.83), and very high (0.84-1.00) (Figure 10).

VI. DISCUSSION

Groundwater spring potential is significant for sustainable water management, ecological preservation, environmental

TABLE 7. Paired-samples sign test of the proposed TF-DNN model and the DT, SVM, LR, RF, and CART models.

No.	Pairwise comparison	Z-value	p-value	Significance
1	TF-DNN vs. DT	4.891	0.000	Yes
2	TF-DNN vs. SVM	2.813	0.005	Yes
3	TF-DNN vs. LR	2.908	0.004	Yes
4	TF-DNN vs. RF	2.145	0.032	Yes
5	TF-DNN vs. CART	6.314	0.000	Yes

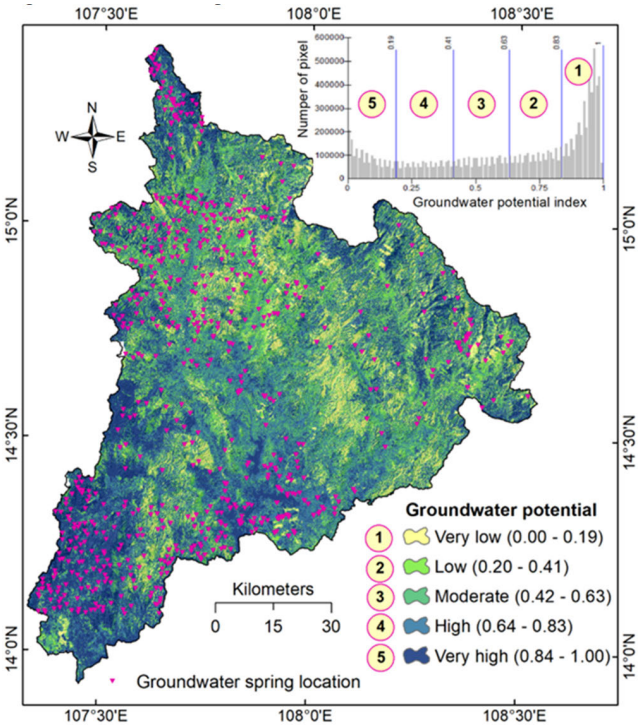

FIGURE 10. Groundwater springs potential map for Kon Tum Province using the TF-DNN model.

monitoring, and overall socio-economic development [105], [106], [107]. Therefore, studying and developing highly accurate models for spatial prediction of groundwater spring potential are essential. These models can serve as fundamental tools for informed decision-making in water management, conservation efforts, and sustainable development, ensuring this valuable natural resource's wise use and protection. In this research paper, we introduce and validate a novel modeling approach that leverages deep neural networks and multisource geospatial data for the spatial prediction of groundwater spring potential, with a case study at the Kon Tum province in the central highland Area of Vietnam.

The performance of the TF-DNN model for the spatial prediction of groundwater spring potential is highly dependent on their architecture or structure, activation function, transfer function, and optimizer; therefore, it is crucial to determine them to ensure optimal performance carefully. The high prediction power of the proposed TF-DNN model substantiates the appropriateness of its structure, featuring three hidden layers with 32 neurons each, for the specific dataset used in this study. Additionally, the use of ReLU as the activation function and sigmoid as the transfer function, along with the ADAM algorithm for optimizing the network weights, proves to be appropriate choices for the model's successful performance. The findings are in line with other recent research on the application of deep learning in various environmental modeling tasks, such as flash floods [96], landslides [98], and forest fires [108]. However, it is challenging to definitively claim that the current structure of the TF-DNN model is optimal. As a result, further research should be conducted to establish a formula or methodology for determining the most suitable TF-DNN architecture.

The effectiveness of the TF-DNN model has been confirmed by its statistically significant higher prediction power compared to the benchmark models, namely DT, SVM, LR, RF, and CART models. Therefore, the TF-DNN model represents a new tool for spatially predicting groundwater spring potential. The result confirms a recent statement in the literature that deep learning can provide more accurate results [109], [110]. This is because the TF-DNN model is capable of automatically learning hierarchical representations from the multisource geospatial data for groundwater spring potential modeling. Thus, its ability to extract intricate features and patterns from 12 GSI factors enables it to handle complex, high-dimensional data more effectively.

The findings in this study also show that the effectiveness and generalization capability of the TF-DNN model heavily depends on the quantity and quality of the GSI factors used. In this analysis, 12 GSI factors were prepared from different geospatial sources with different georeference systems; therefore, it is vital to convert, process, and transfer these data to a unique coordinate system and spatial resolution. Thus, the ESRI File Geodatabase and various spatial analysis tools in ArcGIS Pro 3.0 are suitable.

Among the 12 GSI factors, NDWI, NDVI, and geology are essential for the spatial prediction of groundwater spring potential. This is a reasonable result because the areas with high NDWI and NDVI values in the Kon Tum province are related to higher water content regions (Figures 3f and g), while for the case of geology, the groundwater spring potential (GSP) points distribute mostly in the Tac Po formation, the Kham Duc formation, the Van Canh formation, and the Mang Yang formation, compared to the other formations and complexes. In contrast, distance to river has the lowest contribution to the groundwater spring potential, mainly because

the distribution of the GSP points is not close to the river network of the Kon Tum province.

VII. CONCLUSION

In the present study, we introduced and evaluated a novel modeling approach for spatial prediction of groundwater spring potential using Tensorflow Deep Neural Networks (TF-DNN) and multisource geospatial data. The research focused on the Kon Tum province, Vietnam, and incorporated groundwater spring potential locations along with 12 influencing factors for modeling. Our investigation yielded the following conclusions.

- The TF-DNN model demonstrates the capability to map groundwater spring potential in tropical areas accurately.
- The TF-DNN model exhibits statistically significantly higher global prediction performance than benchmark models, such as DT, SVM, LR, RF, and CART. Thus, the TF-DNN model emerges as a novel and valuable tool for spatially predicting groundwater spring potential.
- NDWI, NDVI, and geology stand out as the most critical factors for spatially predicting groundwater spring potential in the Kon Tum province.
- Future work should investigate further variations in Deep Neural Network architectures to enhance prediction accuracy. Additionally, uncertainty analysis should be included. This analysis should explore uncertainties associated with the predictions, considering the inherent uncertainties in geospatial data and neural network models. Understanding these uncertainties is crucial for decision-makers relying on planning and policy formulation predictions.
- As the final conclusion, the groundwater spring potential map produced in this research may be useful for local authorities in the province for developing measures related to water management and socio-economic development.

REFERENCES

[1] P. Shukla, P. Singh, and R. M. Singh, *Environmental Processes and Management: Tools and Practices for Groundwater*. Berlin, Germany: Springer, 2023.

[2] K. Khosravi, M. Panahi, and D. T. Bui, "Spatial prediction of groundwater spring potential mapping based on an adaptive neuro-fuzzy inference system and metaheuristic optimization," *Hydrol. Earth Syst. Sci.*, vol. 22, no. 9, pp. 4771–4792, Sep. 2018.

[3] A. F. Lutz, W. W. Immerzeel, C. Siderius, R. R. Wijngaard, S. Nepal, A. B. Shrestha, P. Wester, and H. Biemans, "South Asian agriculture increasingly dependent on meltwater and groundwater," *Nature Climate Change*, vol. 12, no. 6, pp. 566–573, Jun. 2022.

[4] S. Díaz-Alcaide and P. Martínez-Santos, "Review: Advances in groundwater potential mapping," *Hydrogeol. J.*, vol. 27, no. 7, pp. 2307–2324, Nov. 2019.

[5] J. Zandi, P. T. Ghazvinei, R. Hashim, K. B. W. Yusof, J. Ariffin, and S. Motamedi, "Mapping of regional potential groundwater springs using logistic regression statistical method," *Water Resour.*, vol. 43, no. 1, pp. 48–57, Jan. 2016.

[6] N. Taghavi, R. K. Niven, D. J. Paull, and M. Kramer, "Groundwater vulnerability assessment: A review including new statistical and hybrid methods," *Sci. Total Environ.*, vol. 822, May 2022, Art. no. 153486.

[7] A.-A. Hussein, V. Govindu, and A. G. M. Nigusse, "Evaluation of groundwater potential using geospatial techniques," *Appl. Water Sci.*, vol. 7, no. 5, pp. 2447–2461, Sep. 2017.

[8] R. Chaudari, D. Lal, S. Dutta, B. Umrikar, and S. Halder, "Weighted overlay analysis for delineation of ground water potential zone: A case study of Pirangut River Basin," *Int. J. Remote Sens. Geosci. (IJRSG)*, vol. 7, no. 1, pp. 1–7, 2018.

[9] I. Ahmad, M. A. Dar, A. H. Teka, M. Teshome, T. G. Andualem, A. Teshome, and T. Shafi, "GIS and fuzzy logic techniques-based demarcation of groundwater potential zones: A case study from jemma river basin, Ethiopia," *J. Afr. Earth Sci.*, vol. 169, Sep. 2020, Art. no. 103860.

[10] C. Chiloane, T. Dube, and C. Shoko, "Impacts of groundwater and climate variability on terrestrial groundwater dependent ecosystems: A review of geospatial assessment approaches and challenges and possible future research directions," *Geocarto Int.*, vol. 37, no. 23, pp. 6755–6779, Dec. 2022.

[11] I. P. Hoyos, N. Krakauer, R. Khanbilvardi, and R. Armstrong, "A review of advances in the identification and characterization of groundwater dependent ecosystems using geospatial technologies," *Geosciences*, vol. 6, no. 2, p. 17, Mar. 2016.

[12] P. Prasad, V. J. Loveson, M. Kotha, and R. Yadav, "Application of machine learning techniques in groundwater potential mapping along the west coast of India," *GIScience Remote Sens.*, vol. 57, no. 6, pp. 735–752, Aug. 2020.

[13] T. X. Bien, A. Jaafari, T. Van Phong, P. T. Trinh, and B. T. Pham, "Groundwater potential mapping in the central highlands of Vietnam using spatially explicit machine learning," *Earth Sci. Informat.*, vol. 16, no. 1, pp. 131–146, Mar. 2023.

[14] S. A. Naghibi, K. Ahmadi, and A. Daneshi, "Application of support vector machine, random forest, and genetic algorithm optimized random forest models in groundwater potential mapping," *Water Resour. Manag.*, vol. 31, no. 9, pp. 2761–2775, Jul. 2017.

[15] V. Gholami, M. R. Khaleghi, S. Pirasteh, and M. J. Booij, "Comparison of self-organizing map, artificial neural network, and co-active neuro-fuzzy inference system methods in simulating groundwater quality: Geospatial artificial intelligence," *Water Resour. Manag.*, vol. 36, no. 2, pp. 451–469, Jan. 2022.

[16] T. Roshni, M. K. Jha, and J. Drisya, "Neural network modeling for groundwater-level forecasting in coastal aquifers," *Neural Comput. Appl.*, vol. 32, no. 16, pp. 12737–12754, Aug. 2020.

[17] Q. B. Pham, D. A. Tran, N. T. Ha, A. R. M. T. Islam, and R. Salam, "Random forest and nature-inspired algorithms for mapping groundwater nitrate concentration in a coastal multi-layer aquifer system," *J. Cleaner Prod.*, vol. 343, Apr. 2022, Art. no. 130900.

[18] M. Khosravi, A. Afshar, and A. Molajou, "Decision tree-based conditional operation rules for optimal conjunctive use of surface and groundwater," *Water Resour. Manag.*, vol. 36, no. 6, pp. 2013–2025, Apr. 2022.

[19] M. S. Jaafarzadeh, N. Tahmasebipour, A. Haghizadeh, H. R. Pourghasemi, and H. Rouhani, "Groundwater recharge potential zonation using an ensemble of machine learning and bivariate statistical models," *Sci. Rep.*, vol. 11, no. 1, p. 5587, Mar. 2021.

[20] H. Morgan, A. Madani, H. M. Hussien, and T. Nassar, "Using an ensemble machine learning model to delineate groundwater potential zones in desert fringes of East Esna-Idfu area, Nile Valley, Upper Egypt," *Geosci. Lett.*, vol. 10, no. 1, p. 9, Feb. 2023.

[21] J. Yin, J. Medellín-Azuara, A. Escriva-Bou, and Z. Liu, "Bayesian machine learning ensemble approach to quantify model uncertainty in predicting groundwater storage change," *Sci. Total Environ.*, vol. 769, May 2021, Art. no. 144715.

[22] B. T. Pham, A. Jaafari, T. V. Phong, D. Mafi-Gholami, M. Amiri, N. Van Tao, V.-H. Duong, and I. Prakash, "Naïve Bayes ensemble models for groundwater potential mapping," *Ecological Informat.*, vol. 64, Sep. 2021, Art. no. 101389.

[23] A. Jafarzadeh, M. Pourreza-Bilondi, A. Akbarpour, A. Khashei-Siuki, and S. Samadi, "Application of multi-model ensemble averaging techniques for groundwater simulation: Synthetic and real-world case studies," *J. Hydroinformatics*, vol. 23, no. 6, pp. 1271–1289, Nov. 2021.

[24] A. Mosavi, F. S. Hosseini, B. Choubin, M. Goodarzi, A. A. Dineva, and E. R. Sardooi, "Ensemble boosting and bagging based machine learning models for groundwater potential prediction," *Water Resour. Manag.*, vol. 35, no. 1, pp. 23–37, Jan. 2021.

[25] E. Sharghi, V. Nourani, Y. Zhang, and P. Ghaneei, "Conjunction of cluster ensemble-model ensemble techniques for spatiotemporal assessment of groundwater depletion in semi-arid plains," *J. Hydrol.*, vol. 610, Jul. 2022, Art. no. 127984.

[26] Z. Wang, W. Lu, Z. Chang, and H. Wang, "Simultaneous identification of groundwater contaminant source and simulation model parameters based on an ensemble Kalman filter—Adaptive step length ant colony optimization algorithm," *J. Hydrol.*, vol. 605, Feb. 2022, Art. no. 127352, doi: 10.1016/j.jhydrol.2021.127352.

[27] A. E. Prieto-Estrada, M. A. Widdowson, and L. D. Stewart, "Numerical modeling and data-worth analysis for characterizing the architecture and dissolution rates of a multicomponent DNAPL source," *Water Resources Res.*, May 2023, Art. no. e2022WR034351.

[28] Z. Pan, W. Lu, and Y. Bai, "Groundwater contaminated source estimation based on adaptive correction iterative ensemble smoother with an auto lightgbm surrogate," *J. Hydrol.*, vol. 620, May 2023, Art. no. 129502, doi: 10.1016/j.jhydrol.2023.129502.

[29] F. Li, T. Yigitcanlar, M. Nepal, K. Nguyen, and F. Dur, "Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework," *Sustain. Cities Soc.*, vol. 96, Sep. 2023, Art. no. 104653.

[30] T. A. Tuan, P. D. Pha, T. T. Tam, and D. T. Bui, "A new approach based on balancing composite motion optimization and deep neural networks for spatial prediction of landslides at tropical cyclone areas," *IEEE Access*, vol. 11, pp. 69495–69511, 2023.

[31] W. Han, X. Zhang, Y. Wang, L. Wang, X. Huang, J. Li, S. Wang, W. Chen, X. Li, R. Feng, and R. Fan, "A survey of machine learning and deep learning in remote sensing of geological environment: Challenges, advances, and opportunities," *ISPRS J. Photogramm. Remote Sens.*, vol. 202, pp. 87–113, Aug. 2023.

[32] S. Awasthi, K. Jain, S. Bhattacharjee, V. Gupta, D. Varade, H. Singh, A. B. Narayan, and A. Budillon, "Analyzing urbanization induced groundwater stress and land deformation using time-series Sentinel-1 datasets applying PSInSAR approach," *Sci. Total Environ.*, vol. 844, Oct. 2022, Art. no. 157103.

[33] J. G. Masek, M. A. Wulder, B. Markham, J. McCorkel, C. J. Crawford, J. Storey, and D. T. Jenstrom, "Landsat 9: Empowering open science and applications through continuity," *Remote Sens. Environ.*, vol. 248, Oct. 2020, Art. no. 111968.

[34] S. García-López, M. Vélez-Nicolás, P. Zarandona-Palacio, A. C. Curcio, V. Ruiz-Ortiz, and L. Barbero, "UAV-borne LiDAR revolutionizing groundwater level mapping," *Sci. Total Environ.*, vol. 859, Feb. 2023, Art. no. 160272.

[35] M. H. Price, *Mastering ArcGIS Pro*. New York, NY, USA: McGraw-Hill, 2023.

[36] M. Sener and M. C. Arslanoglu, "Morphometric analysis in Google Earth engine: An online interactive web-based application for global-scale analysis," *Environ. Model. Softw.*, vol. 162, Apr. 2023, Art. no. 105640.

[37] Z. Wang, J. Wang, and J. Han, "Spatial prediction of groundwater potential and driving factor analysis based on deep learning and geographical detector in an arid endorheic basin," *Ecological Indicators*, vol. 142, Sep. 2022, Art. no. 109256.

[38] S. K. R. Chidepudi, N. Massei, A. Jardani, A. Henriot, D. Allier, and L. Baulon, "A wavelet-assisted deep learning approach for simulating groundwater levels affected by low-frequency variability," *Sci. Total Environ.*, vol. 865, Mar. 2023, Art. no. 161035.

[39] X. Wan, X. Li, X. Wang, X. Yi, Y. Zhao, X. He, R. Wu, and M. Huang, "Water quality prediction model using Gaussian process regression based on deep learning for carbon neutrality in papermaking wastewater treatment system," *Environ. Res.*, vol. 211, Aug. 2022, Art. no. 112942.

[40] R. Zuo, Y. Xiong, J. Wang, and E. J. M. Carranza, "Deep learning and its application in geochemical mapping," *Earth-Sci. Rev.*, vol. 192, pp. 1–14, May 2019.

[41] W. Zhang, X. Gu, L. Tang, Y. Yin, D. Liu, and Y. Zhang, "Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge," *Gondwana Res.*, vol. 109, pp. 1–17, Sep. 2022.

[42] S. M. Mousavi and G. C. Beroza, "Deep-learning seismology," *Science*, vol. 377, no. 6607, Aug. 2022, Art. no. eabm4470.

[43] C. R. Kumar, "Higher level abstraction of deep architecture," in *Prediction and Analysis for Knowledge Representation and Machine Learning*, 2022, pp. 143–161.

[44] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. J. Dean, M. Devin, and S. Ghemawat, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, *arXiv:1603.04467*.

[45] F. Chollet, *Deep Learning Mit Python Und Keras: Das Praxis-Handbuch Vom Entwickler Der Keras-Bibliothek*. MITP-Verlags GmbH & Co. KG, 2018.

[46] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, and A. Desmaison, "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, pp. 8026–8037.

[47] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, Nov. 2014, pp. 675–678.

[48] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley, and Y. Bengio, "Theano: New features and speed improvements," 2012, *arXiv:1211.5590*.

[49] F. Seide and A. Agarwal, "CNTK: Microsoft's open-source deep-learning toolkit," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2016, p. 2135.

[50] T. Chen, M. Li, Y. Li, M. Lin, N. Wang, M. Wang, T. Xiao, B. Xu, C. Zhang, and Z. Zhang, "MXNet: A flexible and efficient machine learning library for heterogeneous distributed systems," 2015, *arXiv:1512.01274*.

[51] D. Team, "Deeplearning4J: Open-source distributed deep learning for the JVM," *Apache Softw. Found. License*, vol. 2, p. 2, Jan. 2016.

[52] J. Verbraeken, M. Wolting, J. Katzy, J. Kloppenburg, T. Verbelen, and J. S. Rellermeyer, "A survey on distributed machine learning," *ACM Comput. Surv.*, vol. 53, no. 2, pp. 1–33, 2020.

[53] S. Sangeetha, G. Sudha Sadasivam, and A. Srikanth, "Differentially private model release for healthcare applications," *Int. J. Comput. Appl.*, vol. 44, no. 10, pp. 953–958, Oct. 2022.

[54] S. Zian, S. A. Kareem, and K. D. Varathan, "An empirical evaluation of stacked ensembles with different meta-learners in imbalanced classification," *IEEE Access*, vol. 9, pp. 87434–87452, 2021.

[55] L. F. Maldaner, J. P. Molin, T. F. Canata, and M. Martello, "A system for plant detection using sensor fusion approach based on machine learning model," *Comput. Electron. Agricult.*, vol. 189, Oct. 2021, Art. no. 106382.

[56] J. R. Quinlan, *C4. 5: Programs for Machine Learning*. Amsterdam, The Netherlands: Elsevier, 2014.

[57] V. Vapnik and O. Chapelle, "Bounds on error expectation for support vector machines," *Neural Comput.*, vol. 12, no. 9, pp. 2013–2036, Sep. 2000.

[58] D. T. Anh, M. Pandey, V. N. Mishra, K. K. Singh, K. Ahmadi, S. Janizadeh, T. T. Tran, N. T. T. Linh, and N. M. Dang, "Assessment of groundwater potential modeling using support vector machine optimization based on Bayesian multi-objective hyperparameter algorithm," *Appl. Soft Comput.*, vol. 132, Jan. 2023, Art. no. 109848, doi: 10.1016/j.asoc.2022.109848.

[59] A. Malekloo, E. Ozer, M. AlHamaydeh, and M. Girolami, "Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights," *Struct. Health Monitor.*, vol. 21, no. 4, pp. 1906–1955, Jul. 2022.

[60] D. R. Cox, "Two further applications of a model for binary regression," *Biometrika*, vol. 45, no. 3, p. 562, Dec. 1958.

[61] W. Chen, H. Li, E. Hou, S. Wang, G. Wang, M. Panahi, T. Li, T. Peng, C. Guo, C. Niu, L. Xiao, J. Wang, X. Xie, and B. B. Ahmad, "GIS-based groundwater potential analysis using novel ensemble weights-of-evidence with logistic regression and functional tree models," *Sci. Total Environ.*, vol. 634, pp. 853–867, Sep. 2018.

[62] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001.

[63] S. He, J. Wu, D. Wang, and X. He, "Predictive modeling of groundwater nitrate pollution and evaluating its main impact factors using random forest," *Chemosphere*, vol. 290, Mar. 2022, Art. no. 133388.

[64] N. N. Thanh, P. Thunyawatcharakul, N. H. Ngu, and S. Chotpantarat, "Global review of groundwater potential models in the last decade: Parameters, model techniques, and validation," *J. Hydrol.*, vol. 614, Nov. 2022, Art. no. 128501.

[65] L. Breiman, "Bagging predictors," *Mach. Learn.*, vol. 24, no. 2, pp. 123–140, Aug. 1996.

[66] D. T. Bui, P. Tsangaratos, V.-T. Nguyen, N. V. Liem, and P. T. Trinh, "Comparing the prediction performance of a deep learning neural network model with conventional machine learning models in landslide susceptibility assessment," *Catena*, vol. 188, May 2020, Art. no. 104426, doi: 10.1016/j.catena.2019.104426.

[67] D. N. D. Phuong, N. T. Huyen, N. D. Liem, N. T. Hong, D. K. Cuong, and N. K. Loi, "On the use of an innovative trend analysis methodology for temporal trend identification in extreme rainfall indices over the central highlands, Vietnam," *Theor. Appl. Climatol.*, vol. 147, nos. 1–2, pp. 835–852, Jan. 2022.

[68] D. V. Canh, "Research to build a scientific basis and propose solutions to protect and rationally use water resources in the central highlands," Hanoi Univ. Mining Geol., Tech. Rep., 2008.

[69] D. V. Canh, N. T. T. Thuy, N. T. Xuan, N. Q. Luat, P. Q. Nhan, D. V. Binh, T. T. Hue, D. D. Nhan, N. T. Tu, and D. D. Long, "Research on scientific basis and develop solutions to store rainwater into the ground for drought prevention and protection of underground water resources in the central highlands," Hanoi Univ. Mining Geol., Tech. Rep., 2010.

[70] H. H. Duong, N. X. Lam, N. T. Tu, H. M. Tho, N. T. Phong, N. X. Tang, H. L. Thuan, N. L. Long, H. V. Hoan, and T. D. Trinh, "Research and propose models and technological solutions to exploit and protect water sources in basalt formations in high-mountainous and water-scarcity areas in the central highlands," Vietnam Academy Water Resour., Tech. Rep., 2018.

[71] P. T. Vinh, D. D. Hai, T. T. Thanh, K. V. Huan, V. N. H. Giang, T. D. Huyen, N. D. Chan, P. C. Nam, N. T. Tu, and N. Luu, "Research and propose models of collection and sustainable exploitation of spring groundwater for high-mountain and water-scarces areas in the central highlands," Vietnam Academy Water Resour., Tech. Rep., 2018.

[72] T. Tadono, H. Nagai, H. Ishida, F. Oda, S. Naito, K. Minakawa, and H. Iwamoto, "Generation of the 30 M-mesh global digital surface model by ALOS PRISM," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 41, pp. 157–162, Jun. 2016.

[73] D. J. Wald and T. I. Allen, "Topographic slope as a proxy for seismic site conditions and amplification," *Bull. Seismological Soc. Amer.*, vol. 97, no. 5, pp. 1379–1395, Oct. 2007.

[74] J. D. Phillips, "Emergence and pseudo-equilibrium in geomorphology," *Geomorphology*, vol. 132, nos. 3–4, pp. 319–326, Sep. 2011.

[75] A. L. Langston, G. E. Tucker, R. S. Anderson, and S. P. Anderson, "Evidence for climatic and hillslope-aspect controls on vadose zone hydrology and implications for saprolite weathering," *Earth Surf. Processes Landforms*, vol. 40, no. 9, pp. 1254–1269, Jul. 2015.

[76] M. Manga and J. W. Kirchner, "Interpreting the temperature of water at cold springs and the importance of gravitational potential energy," *Water Resour. Res.*, vol. 40, no. 5, May 2004.

[77] J. De Waele, "Interaction between a dam site and Karst springs: The case of supramonte (Central-East Sardinia, Italy)," *Eng. Geol.*, vol. 99, nos. 3–4, pp. 128–137, Jun. 2008.

[78] J. Minár, I. S. Evans, and M. Jenco, "A comprehensive system of definitions of land surface (topographic) curvatures, with implications for their application in geoscience modelling and prediction," *Earth-Sci. Rev.*, vol. 211, Dec. 2020, Art. no. 103414.

[79] Y. Zhou and W. Li, "A review of regional groundwater flow modeling," *Geosci. Frontiers*, vol. 2, no. 2, pp. 205–214, Apr. 2011.

[80] D. C. Phan, T. H. Trung, V. T. Truong, and K. N. Nasahara, "Ensemble learning updating classifier for accurate land cover assessment in tropical cloudy areas," *Geocarto Int.*, vol. 37, no. 14, pp. 4053–4070, Jul. 2022.

[81] R. B. Myneni, F. G. Hall, P. J. Sellers, and A. L. Marshak, "The interpretation of spectral vegetation indexes," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 481–486, Mar. 1995.

[82] D. Xu and X. Guo, "Compare NDVI extracted from Landsat 8 imagery with that from Landsat 7 imagery," *Amer. J. Remote Sens.*, vol. 2, no. 2, pp. 10–14, 2014.

[83] E. Özelkan, "Water body detection analysis using NDWI indices derived from Landsat-8 OLI," *Polish J. Environ. Stud.*, vol. 29, no. 2, pp. 1759–1769, 2020.

[84] A. Ozdemir, "Using a binary logistic regression method and GIS for evaluating and mapping the groundwater spring potential in the Sultan Mountains (Aksehir, Turkey)," *J. Hydrol.*, vol. 405, nos. 1–2, pp. 123–136, Jul. 2011.

[85] P.-T.-T. Ngo, T. D. Pham, V.-H. Nhu, T. T. Le, D. A. Tran, D. C. Phan, P. V. Hoa, J. L. Amaro-Mellado, and D. T. Bui, "A novel hybrid quantum-PSO and credal decision tree ensemble for tropical cyclone induced flash flood susceptibility mapping with geospatial data," *J. Hydrol.*, vol. 596, May 2021, Art. no. 125682.

[86] V.-H. Nhu, P.-T. T. Ngo, T. D. Pham, J. Dou, X. Song, N.-D. Hoang, D. A. Tran, D. P. Cao, I. B. Aydilek, M. Amiri, R. Costache, P. V. Hoa, and D. T. Bui, "A new hybrid Firefly–PSO optimized random subspace tree intelligence for torrential rainfall-induced flash flood susceptible mapping," *Remote Sens.*, vol. 12, no. 17, p. 2688, Aug. 2020.

[87] C. Thivya, S. Chidambaram, C. Singaraja, R. Thilagavathi, M. V. Prasanna, P. Anandhan, and I. Jainab, "A study on the significance of lithology in groundwater quality of Madurai district, Tamil Nadu (India)," *Environ., Develop. Sustainability*, vol. 15, no. 5, pp. 1365–1387, Oct. 2013.

[88] A. Benhamrouche, J. Martin-Vide, Q. B. Pham, M. E. Kouachi, and M. C. Moreno-Garcia, "Daily precipitation concentration in Central Coast Vietnam," *Theor. Appl. Climatol.*, vol. 147, nos. 1–2, pp. 37–45, Jan. 2022.

[89] P. Reutemann. (2023). *Python3 Wrapper for the Weka Machine Learning Workbench*. [Online]. Available: https://pypi.org/project/python-weka-wrapper3/

[90] D. W. Allen and J. M. Coffey, *Focus on Geodatabases in ArcGIS Pro*. ESRI Press, 2019.

[91] D. T. Bui, Q.-T. Bui, Q.-P. Nguyen, B. Pradhan, H. Nampak, and P. T. Trinh, "A hybrid artificial intelligence approach using GIS-based neural-fuzzy inference system and particle swarm optimization for forest fire susceptibility modeling at a tropical area," *Agricult. Forest Meteorol.*, vol. 233, pp. 32–44, Feb. 2017.

[92] C. H. Park and S. B. Kim, "Sequential random k-nearest neighbor feature selection for high-dimensional data," *Exp. Syst. Appl.*, vol. 42, no. 5, pp. 2336–2342, Apr. 2015, doi: 10.1016/j.eswa.2014.10.044.

[93] A. S. Rodin, A. Litvinenko, K. Klos, A. C. Morrison, T. Woodage, J. Coresh, and E. Boerwinkle, "Use of wrapper algorithms coupled with a random forests classifier for variable selection in large-scale genomic association studies," *J. Comput. Biol.*, vol. 16, no. 12, pp. 1705–1718, Dec. 2009.

[94] M. Yaqub, M. K. Javaid, C. Cooper, and J. A. Noble, "Investigation of the role of feature selection and weighted voting in random forests for 3-D volumetric segmentation," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 258–271, Feb. 2014.

[95] J. Chong, P. Tjurin, M. Niemelä, T. Jämsä, and V. Farrahi, "Machine-learning models for activity class prediction: A comparative study of feature selection and classification algorithms," *Gait Posture*, vol. 89, pp. 45–53, Sep. 2021.

[96] D. T. Bui, N.-D. Hoang, F. Martínez-Álvarez, P.-T.-T. Ngo, P. V. Hoa, T. D. Pham, P. Samui, and R. Costache, "A novel deep learning neural network approach for predicting flash flood susceptibility: A case study at a high frequency tropical storm area," *Sci. Total Environ.*, vol. 701, Jan. 2020, Art. no. 134413, doi: 10.1016/j.scitotenv.2019.134413.

[97] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. ICLR*, vol. 500, San Diego, CA, USA, 2015, pp. 1–15.

[98] V.-H. Nhu, N.-D. Hoang, H. Nguyen, P. T. T. Ngo, T. T. Bui, P. V. Hoa, P. Samui, and D. T. Bui, "Effectiveness assessment of keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area," *CATENA*, vol. 188, May 2020, Art. no. 104458.

[99] A. R. M. T. Islam, S. C. Pal, I. Chowdhuri, R. Salam, M. S. Islam, M. M. Rahman, A. Zahid, and A. M. Idris, "Application of novel framework approach for prediction of nitrate concentration susceptibility in coastal multi-aquifers, Bangladesh," *Sci. Total Environ.*, vol. 801, Dec. 2021, Art. no. 149811.

[100] O. Rahmati, H. Darabi, M. Panahi, Z. Kalantari, S. A. Naghibi, C. S. S. Ferreira, A. Kornejady, Z. Karimidastenaei, F. Mohammadi, S. Stefanidis, D. T. Bui, and A. T. Haghighi, "Development of novel hybridized models for urban flood susceptibility mapping," *Sci. Rep.*, vol. 10, no. 1, p. 12937, Jul. 2020.

[101] D. Ruidas, S. C. Pal, A. R. M. T. Islam, and A. Saha, "Hydrogeochemical evaluation of groundwater aquifers and associated health hazard risk mapping using ensemble data driven model in a water scares Plateau region of Eastern India," *Exposure Health*, vol. 15, no. 1, pp. 113–131, Mar. 2023.

[102] W. Chen, Z. Wang, G. Wang, Z. Ning, B. Lian, S. Li, P. Tsangaratos, I. Ilia, and W. Xue, "Optimizing rotation forest-based decision tree algorithms for groundwater potential mapping," *Water*, vol. 15, no. 12, p. 2287, Jun. 2023.
[103] S. A. Naghibi, H. Hashemi, R. Berndtsson, and S. Lee, "Application of extreme gradient boosting and parallel random forest algorithms for assessing groundwater spring potential using DEM-derived factors," *J. Hydrol.*, vol. 589, Oct. 2020, Art. no. 125197.
[104] D. W. Zimmerman, "Teacher's corner: A note on interpretation of the paired-samples T test," *J. Educ. Behav. Statist.*, vol. 22, no. 3, pp. 349–360, Sep. 1997.
[105] S. M. Wineland, H. Basagaoglu, J. Fleming, J. Friedman, L. Garza-Diaz, W. Kellogg, and J. Koch, "The environmental flows implementation challenge: Insights and recommendations across water-limited systems," *Wiley Interdiscipl. Rev., Water*, vol. 9, no. 1, p. e1565, 2022.
[106] E. Lee, R. Jayakumar, S. Shrestha, and Z. Han, "Assessment of transboundary aquifer resources in asia: Status and progress towards sustainable groundwater management," *J. Hydrol., Regional Stud.*, vol. 20, pp. 103–115, Dec. 2018.
[107] J. J. KarisAllen, A. A. Mohammed, J. J. Tamborski, R. C. Jamieson, S. Danielescu, and B. L. Kurylyk, "Present and future thermal regimes of intertidal groundwater springs in a threatened coastal ecosystem," *Hydrol. Earth Syst. Sci.*, vol. 26, no. 18, pp. 4721–4740, Sep. 2022.
[108] H. V. Le, D. A. Hoang, C. T. Tran, P. Q. Nguyen, V. H. T. Tran, N. D. Hoang, M. Amiri, T. P. T. Ngo, H. V. Nhu, T. V. Hoang, and D. T. Bui, "A new approach of deep neural computing for spatial prediction of wildfire danger at tropical climate areas," *Ecological Informat.*, vol. 63, Jul. 2021, Art. no. 101300.
[109] R. Haggerty, J. Sun, H. Yu, and Y. Li, "Application of machine learning in groundwater quality modeling—A comprehensive review," *Water Res.*, vol. 233, Apr. 2023, Art. no. 119745, doi: 10.1016/j.watres.2023.119745.
[110] H. Tao, M. M. Hameed, H. A. Marhoon, M. Zounemat-Kermani, S. Heddam, S. Kim, S. O. Sulaiman, M. L. Tan, Z. Sa'adi, A. D. Mehr, and M. F. Allawi, "Groundwater level prediction using machine learning models: A comprehensive review," *Neurocomputing*, vol. 489, pp. 271–308, Jun. 2022, doi: 10.1016/j.neucom.2022.03.014.

VIET-HA NHU received the Ph.D. degree in engineering sciences from the Free University of Brussels, Belgium, in 2016. He is currently a Senior Lecturer/Researcher with the Department of Geological-Geotechnical Engineering, Hanoi University of Mining and Geology, Hanoi, Vietnam. He has more than 20 articles that were published in Science Citation Index (SCI/SCIE) indexed journals. His research interests include natural hazards, geotechnical engineering, applied artificial intelligence, geospatial technology, and machine learning for geo-resources and geo-hazards such as landslide, flood, forest fire, and land subsidence.

DUONG CAO PHAN is currently with the National Centre for Applied Artificial Intelligence (CeADAR), University College Dublin, and NexusUCD, Dublin, Ireland. He is a Trailblazing Postdoctoral Scientist with the CeADAR, School of Computer Science, University College Dublin. With a profound passion for the Earth's ecosystem and cutting-edge technologies, he has made significant contributions to remote sensing, geospatial analysis, and artificial intelligence for sustainable environmental studies. His research interests include disaster monitoring and climate resilience, empowering communities to respond effectively to natural disasters. His work also includes unraveling climate change impacts on land ecosystems through geospatial analysis and remote sensing, aiding conservation efforts, and sustainable land management. His advanced methods for accurate carbon estimation have guided climate change policies, leading to the mitigation of greenhouse gas emissions. As a visionary advocate for sustainability, he inspires the next generation of environmental scientists, pioneering the use of AI and remote sensing for a greener future.

PHAM VIET HOA received the Ph.D. degree in remote sensing from the Hanoi University of Mining and Geology, in 2013.

She is currently an Associate Professor and the Director of the Ho Chi Minh City Institute of Resources Geography, Vietnam Academy of Science and Technology (VASC). Her research interests include remote sensing, GIS, and information science for sustainable and territorial management.

PHAM THE VINH received the Ph.D. degree in water resources engineering from the Vietnam Academy for Water Resources, in 2021. He is currently the Director of the Center of Environmental Technology Science and Ecology, Southern Institute of Water Resources Research, Ho Chi Minh, Vietnam. He has many national and international articles. His research interests include water resources and ecological environment. In these areas, he regularly conducts in-depth research in the fields of natural disasters, floods, droughts, and ecological environment protection.

DIEU TIEN BUI is currently a Full Professor with the GIS Group, Department of Business and IT, University of South-Eastern Norway (USN), Norway. He has more than 200 publications and out of them, more than 180 articles were published in Science Citation Index (SCI/SCIE) indexed journals. His research interests include GIS and geospatial information science, remote sensing, and applied artificial intelligence and machine learning for natural hazards and environmental problems, such as landslide, flood, forest fire, ground biomass, and structural displacement.