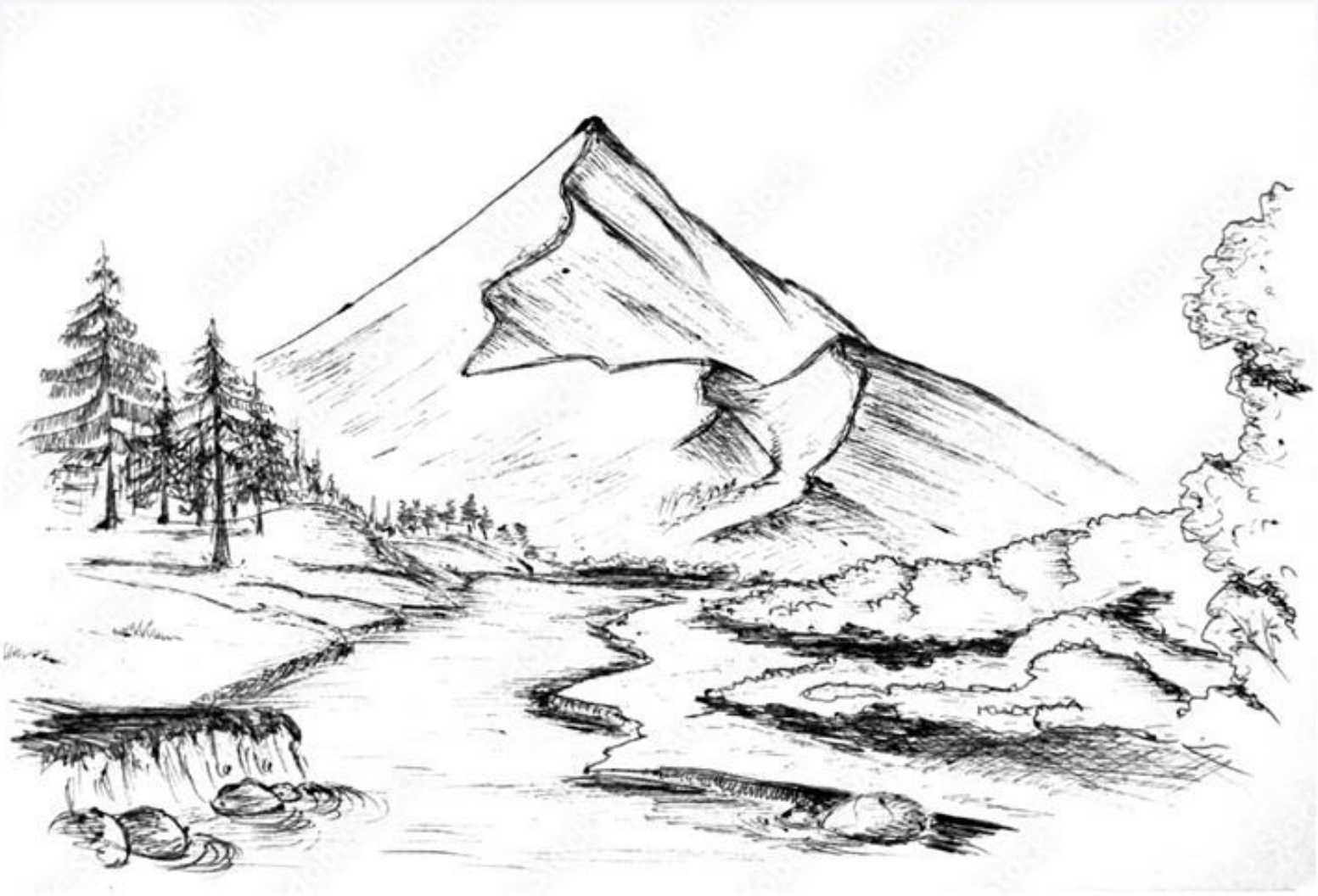
THE PROCEEDINGS OF

Vietnam International Water Conference (VIWC 2022)

19-21 Sept 2022 Online; 09-11 Nov, Onsite, HCM, Vietnam

Ngoc T. NGUYEN, Anh Q. NGUYEN, Toan. D. VU, Phi Q. NGUYEN, Tien L.T. DU,
Nuong T. BUI, Thao. T. P. BUI, Tra T. NGUYEN, Dan D. BUI



SOUTHEAST ASIA UNION FOR WATER, ENVIRONMENT AND GEOSCIENCES (SEAGU)

THE JOINT INNITIATIVE BY ASEAN & US SCIENCES AND TECHNOLOGY FELLOWSHIP & INTERNATIONAL SOCIETY OF GROUNDWATER FOR SUSTAINABLE DEVELOPMENT (ISGSD)

PART II: SELECTED FULL PAPERS

Assessment of environmental pollution levels and forecast of changes in some heavy metals and grease in sediment of Nam Cau Trang Port cluster - Quang Ninh Province
Phuong Nguyen, Cuc Thi Nguyen, Lan Anh Thi Vu, Dong Phuong Nguyen, Hoang Hai Yen 118

Assessment of surface water quality in Cao Bang city using Vietnam water quality index (VN_WQI)
Trung Thanh Dao, Thi Hong Nguyen.....124

Signs of Acid Mine Drainage and Distribution of Heavy Metal into The Environment at Ban Phuc Nickel Mine
Pham Van Chung, Nguyen Van Pho, Nguyen Thi Thuc Anh.....130

Remote sensing-based indices for mapping of water quality in Thai Binh river at Hai Duong province
Nguyen Quoc Phi, Vu Manh Tuong, Nguyen Quang Minh.....138

Researching the Potential Impact of Industrial Stacks of the Bright International Vietnam Company on the Ambient Environment
Tran Anh Quan, Nguyen Thi Hong Ngoc, Do Thi Hai, Do Phuong Nam.....149

Investigating The Causes of The Flood In Ho Chi Minh City Area -Proposing The Solutions
Thiem Quoc Tuan, Do Van Nhuan, Huynh Ngoc Huong.....156

Studying Causes of The Sand-Flow Hazard in Tien Thanh Area, Phan Thiet City, Binh Thuan Province - Proposing Solutions
Thiem Quoc Tuan, Do Van Nhuan, Huynh Ngoc Huong.....165

Characteristics and Exploitation Reserves of Hot Mineral Water at Nkns Borehole, Bao Yen Commune, Thanh Thuy District, Phu Tho Province
Do Van Binh, Ho Van Thuy, Do Thi Hai.....173

A GIS-Based approach for the sustainable management of livestock effluents at Cam Lam district, Khanh Hoa province
Nguyen Quoc Phi, Tran Thi Thu, Phi Truong Thanh.....185

Comparison between ANN & DT to predict landslide susceptibility in Trung Khanh area, Cao Bang province
Trinh Ngoc Nhu Anh, Nguyen Quoc Phi.....191

Surface Modification of Balsa Wood for Functional Oil/Water Separation
Tho D. Le, Duc Anh H. Bui, Bao T.T. Nguyen, Huong T.T. Tong.....196

Collection of domestic solid waste in Hai Duong province and some proposes
Do Van Binh, Vu Manh Hung, Nguyen Thi Hang, Ha Thi Luyen.....204

Monitoring rice growth and predicting rice biomass by Sentinel-2 data at Giao Thuy, Nam Dinh
Hoa Phan Thi Mai, Cuc Nguyen Thi, Phi Nguyen Quoc.....213

Comparison between ANN & DT to predict landslide susceptibility in Trung Khanh area, Cao Bang province

Trinh Ngoc Nhu Anh[1], Nguyen Quoc Phi[1]*

[1]Faculty of Environment, Hanoi University of Mining and Geology

1. Introduction

The main purpose of this study is to establish an effective landslide susceptibility zoning model with 12 landslide condition factors were selected as the evaluation indices to construct the susceptibility assessment model. Two machine learning algorithms for landslide susceptibility prediction (LSP) including: C4.5 Decision Tree and artificial neural network (ANN), which have been used for landslide susceptibility mapping in the Trung Khanh region of Cao Bang province, and the results were compared. Landslide areas were delimited and mapped as landslide inventory after gathering information from historical records, remote sensing detection and periodic field investigations. In the database, 87 landslides were plotted and classified into training (70%, 61 landslides) and testing (30%, 26 landslides) subsets randomly to train and validate the models. The ANN model has a higher correct prediction rate than DT, with a rate of 89.28%, and DT is 83.64%. This research might be useful in landslide studies in mountainous areas, especially in locations with comparable geophysical and climatological characteristics, to aid in decision making for land use planning.

2. Methodology and Research Data

2.1. Study Area and Landslides Inventory

Landslides are one of the most common natural hazards and when they occur, they usually cause loss of life and significant economic losses. The main type of environmental hazard in the study area is landslide. Thus with a total of 87 landslide sites has been mapped by applying the remote sensing (RS), geographic information system (GIS), and spatial data analysis method.

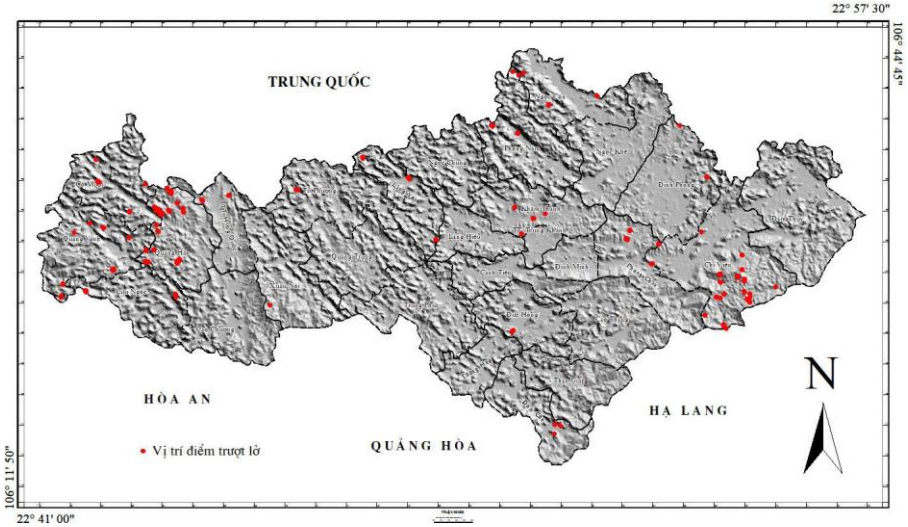


Figure 1. Overview of landslides in Trung Khanh County

The area of Trung Khanh County is about 688.01km^2 and the population is about 70,424. The study area is one of the most landslide-prone areas in Cao Bang province, one of the reasons that is karst landscape with easily leaking surface water and high soil moisture. The annual average precipitation of study area is about 1,500-2,000mm, precipitation is the major inducing factor for landslides.

2.2. Condition Factors

In this study, we collected multi-source data such as field survey data, precipitation data and remote sensing satellite data from Trung Khanh County, Cao Bang province and used the advanced big data models to construct landslide susceptibility map, and compared their performance and applicability in the study area. Landslides in the study areas are affected by many different factors, however, we give 12 most important factors, the data make the most of it to compare the susceptibility assessment landslides in the area.

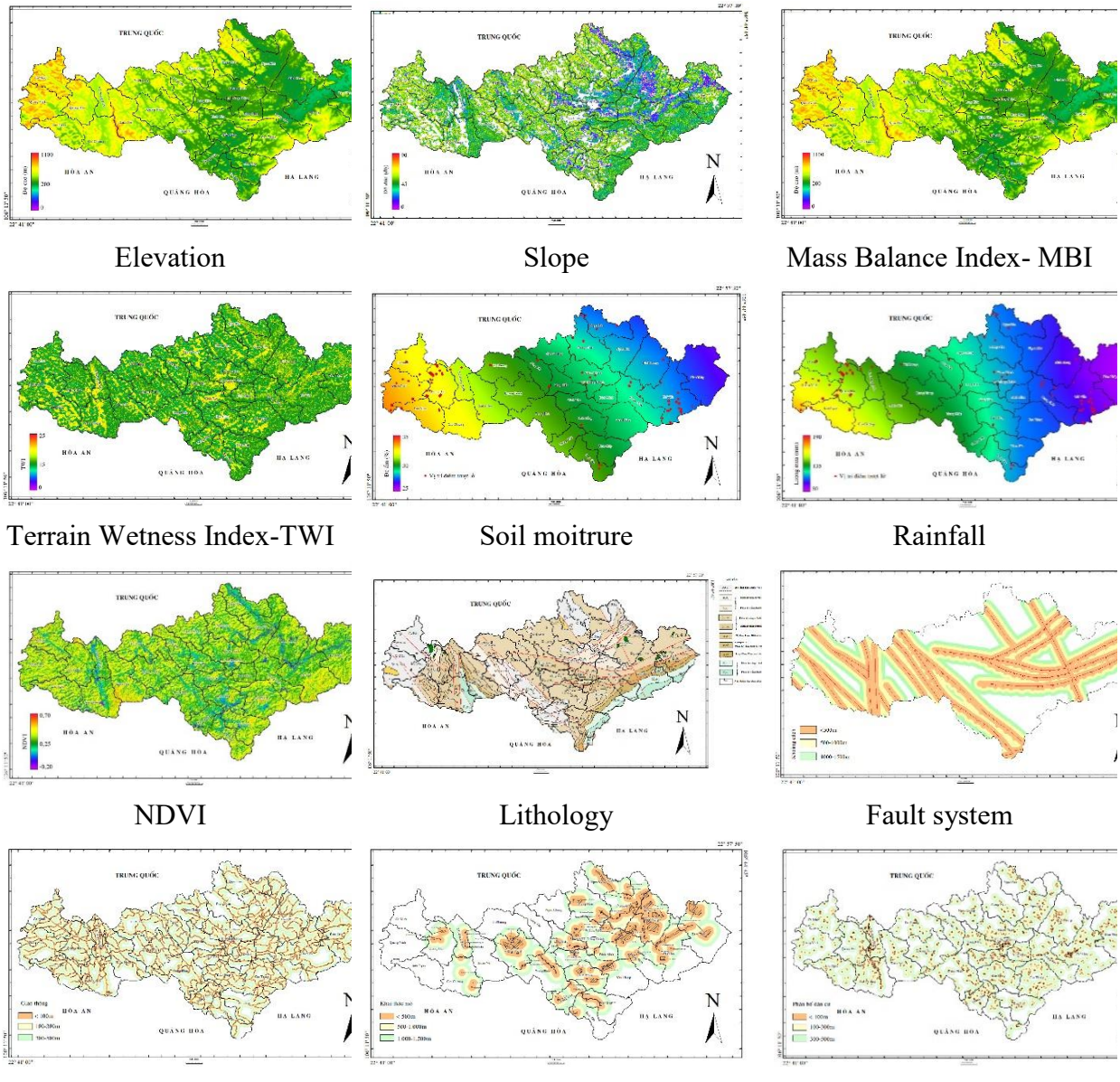


Figure 2: The 12 Landslide condition factors

Combined with multi-source, field and remote sensing data, landslide statistics were collected for 87 landslide sites. A total of 12 conditioning factors were selected based on their impact on the landslides and the data accessibility (in Figure 2). , which including: elevation, slope, TWI, MBI, lithology, fault system, soil moisture, rainfall, NDVI, distance to mined areas, distance to roads, distance to houses. The purpose of analyzing these inputs factors, it is necessary to format the DEM pixel size, they are resized to 20m × 20m. The relationship between landslide hazards and the environment was studied, and its sensitivity was classified and evaluated.

a. Methodology

Evaluation of the models, including their advantages and applicability, is important to obtain a satisfactory landslide susceptibility map. The strongest point of big data analysis methods is that the input data does not need to be normalized or created dummy variables, can use all classes of information collected by checking the statistical independence of the data. The information layer is not required. At the same time, these methods can work with both numerical data (rainfall, slope, etc.) and labeled data (rock groups, land use types, etc.) . The dataset included in the calculation consists of 11 information layers, each layer consists of 1426 pixels, of which 587 pixels of landslide locations and 839 pixels of no landslide are randomly selected over the entire study area. Due to the relatively large data source (total: 1,848,582 pixels), the research data set is divided into three data sets with a separation ratio of 70% for the training dataset to build, 15% for the cross-validation dataset and 15% for the testing dataset to check the efficiency of the computational model.

i.Artificial Neural Network (ANN)

ANN is a learning simulation model identical to the human nervous system. Neural networks consist of a large number of artificial neural connections that can be used to estimate or approximate functions. The back-propagation neural network is the most commonly used neural network architecture.

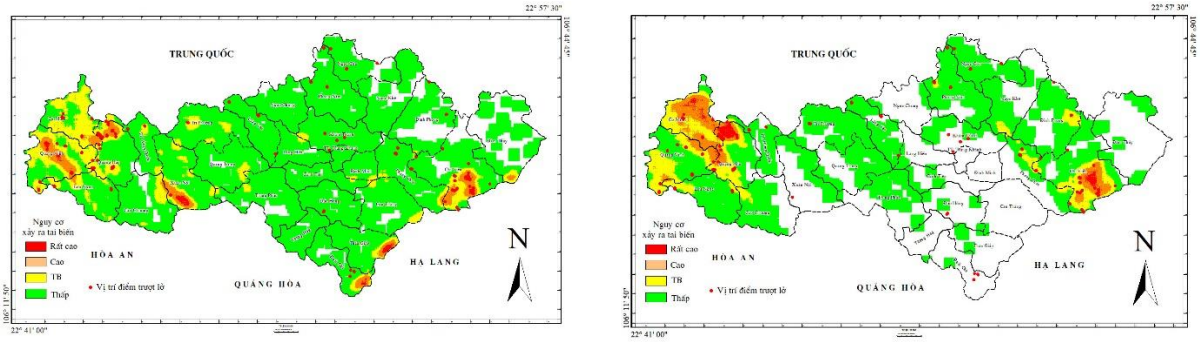
In this study, the ANN model includes 12 input layers, 04 hidden layers and one output layer. Each neuron in the input layer represents different corresponding evaluation factors, the output represents the occurrence of landslides.

ii.Decision Tree Model (DT)

A decision tree is a non-parametric supervised learning method used to generate a model that predicts the value of a target variable by learning the rules, through making inferential decisions. from data features. The tree features a root node, internal nodes, and terminal nodes (leaves). CART (C4.5, J4.8) is a decision tree generation algorithm that divides data with simple predictor variables that can be used for landslide susceptibility analysis.

3. Results and discussion

Since analyzing landslide susceptibility, the landslides are mainly located in the very high and high LSP. ANN has more accuracy but DT model is much easier to understand and interpret.



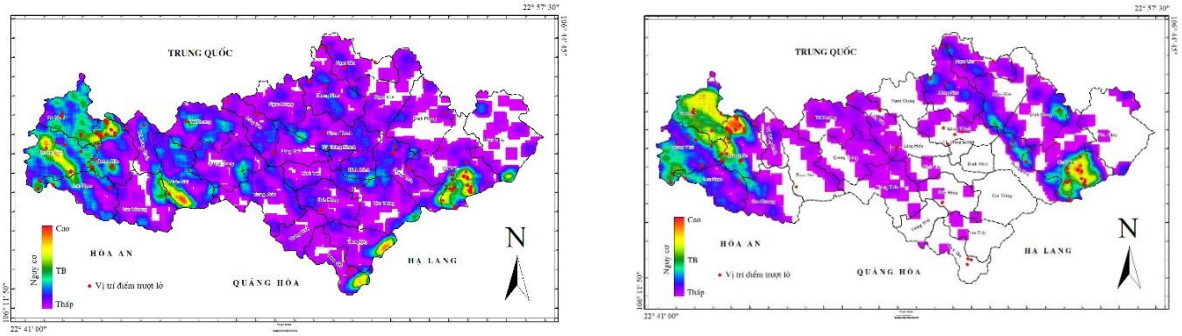
a) b)

Figure 3. Landslide susceptibility maps from a) ANN, b) DT algorithms

The ANN artificial intelligence network model gave the best results with global accuracy of 89.3%, decision tree model (83.6%), followed by the ability to effectively analyze data models. large data for the purpose of predicting the risk of accidents in the study area very well.

To a certain extent, it provides a reference for the link between landslide early warning and extreme precipitation warning. Mining disturbances have also caused extensive damage to the surface, and mining activities have influenced the occurrence of landslides in the area.

For using two models, the training and testing data are both defined as 70/30.



a) b)

Figure 4. Landslide susceptibility maps from a) ANN, b) DT algorithms

In fact, different models have different performance, the different condition factors cannot equally contribute to the development of landslides in the area.

The main cause of landslides in Trung Khanh are rainfall, soil moisture, elevation and the distance to mining area is the conditions of the significant influence contribution in the occurrence of geological hazards.

Compared with the two black box models (ANN), the diagram of the Decision Tree model is quite easy to understand, can create rules (rules) associated with each leaf branch is a rule of the tree. The results indicated ANN, DT have 70 landslide sites (80.75%), 53 landslide sites (60.48%) in the high and very high areas respectively, and the high and very highly prone areas account for 20.97% and 32.46% of the study area. They all have high enough model accuracy to be applied to the prediction of the susceptibility of landslides in study area. These two models also predict that

the landslide sensitivity of the Trung Khanh County is low or moderate, and predict that the highest landslide sus- ceptibility lies in the west section, and the southeast part of the area.

4. Conclusion

This study has contributed to compare 02 machine learning methods (ANN and DT) for landslide susceptibility zoning in Trung Khanh district, Cao Bang province. The study area is affected by 87 recorded landslide sites and 12 condition factors and the results show that ANN is more suitable to assess landslide susceptibility through the analysis of accuracy and characteristics. landslide distribution point. Of all the 12 condition factors, three including rainfall, soil moisture and elevation are the most favorable conditions for landslide susceptibility; land cover (NDVI), mining and other factors contribute less. Last but not least, it is necessary to consider the dangerous support points for management.

Reference

1. Aktas H., San B. T., 2019. Landslide susceptibility mapping using an automatic sampling algorithm based on two level random sampling. Comput. Geosci. 133, 104329.
2. Bragagnolo L., da Silva R. and Grzybowski J., 2020. Artificial neural network en- sembles applied to the mapping of landslide susceptibility. Catena 184, 104240.
3. Dou J., Chang K. T., Chen S., Yunus A. P., Liu J. K., Xia H. and Zhu Z., 2015 Automatic Case-Based Reasoning Approach for Landslide Detection: Integration of Object-Oriented Image Analysis and a Genetic Algorithm. Remote Sens. 7, 4318-4342.
4. Mao Y., Zhang M., Sun P. and Wang G., 2017. Landslide susceptibility assessment using uncertain decision tree model in loess areas. Environ. Earth Sci. 76, 752.
5. Pham B. T., Bui D. T., Pourghasemi H. R., Indra P., Dholakia M. B., 2017. Landslide susceptibility assesssment in the Uttarakhand area (India) using GIS: A comparison study of prediction capability of naïve bayes, multilayer perceptron neural networks, and functional trees methods. Theor. Appl. Clim. 128, 255-273.
6. Qiaomei Su, Weiheng Tao, Shiguang Mei, Xiaoyuan Zhang, Kaixin Li, Xiaoye Su, Jianli Guo and Yonggang Yang. (2021). Landslide Susceptibility Zoning Using C5.0 Decision Tree, Random Forest, Support Vector Machine and Comparison of Their Performance in a Coal Mine Area. Frontiers in Earth Science | www.frontiersin.org
7. Wang Y., Fang Z. and Hong H., 2019. Comparison of convolutional neural networks for landslide susceptibility mapping in Yanshan County, China. Sci. Total Environ. 666, 975-993.
8. Yang J., Song C., Yang Y., Xu C., Guo F. and Xie L., 2019. New method for land- slide susceptibility mapping supported by spatial logistic regression and GeoDetector: A case study of Duwen Highway Basin, Sichuan Province, China. Geomorphology 324, 62-71.
9. Zhou C., Yin K., Cao Y., Ahmed B., Li Y., Catani F. and Pourghasemi H. R., 2018. Landslide susceptibility modeling applying machine learning methods: A case study from Longju in the Three Gorges Reservoir Area, China. Comput. Geosci. 112, 23-37

THANK YOU

FOR BEING A PART OF VIWC2022
WE LOOK FORWARD TO YOUR CONTINUED SUPPORT IN THE NEXT EDITION VIWC2023

NHÀ XUẤT BẢN DÂN TRÍ
Số 9 - Ngõ 26 - Phố Hoàng Cầu - Q. Đống Đa - TP. Hà Nội
VPGD: Số 278 Tôn Đức Thắng, Q. Đống Đa - TP Hà Nội
ĐT: (024). 66860751 - (024). 66860753
Email: nxbdantri@gmail.com
Website: nxbdantri.com.vn

Chịu trách nhiệm xuất bản:
BÙI THỊ HƯƠNG
Chịu trách nhiệm nội dung:
LÊ QUANG KHÔI

Biên tập:	**…….....**
Vẽ bìa:	**Trần Thị Trà Giang**
Sửa bản in:	**…….....**
Trình bày sách:	**Trần Thị Trà Giang**
Chế bản:	**CTCP in cao cấp Quang Minh**

LIÊN KẾT XUẤT BẢN: Công ty cổ phần in cao cấp Quang Minh

In cuốn, khổ cm tại Công ty , Địa chỉ số:
Số xác nhận đăng ký xuất bản số: 1073-2023/CXBIPH/1-37/DT
Quyết định xuất bản số 1473/QĐXB-NXBDT: do Nhà xuất bản Dân Trí cấp ngày 8 tháng 6 năm 2023
Mã ISBN: 978-604-331-965-1
In xong, nộp lưu chiểu ngày... tháng... năm 2023

Vietnam International

WATER

Conference

19-21 Sept 2022 - Online

09-11 Nov - Onsite@HCM

ISBN: 978-604-33-1965-1