

Soil Erosion Assessment and Early Warning Model Based on Big Data and Artificial Intelligence

Yunfeng Li^{1, a}, Ruixiang Song^{2, b*}

¹²Jilin Meteorological Information Network Centre, Changchun, 130062, Jilin, China

^a1285830578@qq.com, ^bsong.rui.xiang@163.com

Abstract: Soil erosion can lead to the pollution of water bodies, further deterioration of the ecological environment, and pose a serious threat to humans and the biodiversity of the earth. In order to cope with this global challenge, it is particularly important to develop soil erosion assessment and early warning systems. With the support of soil erosion model and geographic information system, a new soil erosion assessment and early warning system based on the soil and water environment of Jilin Province was developed in this paper. The system can analyze rainfall data, satellite data and geographic information data more accurately, and then find out the key factors causing soil erosion. The study found that the slope had the greatest correlation with soil erosion, with a correlation of 1.12. The second is the soil type and vegetation type, and in the steep slope of $> 25^\circ$, soil erosion is severe, and the land ecological condition is very poor. These research results not only have important reference value for the land spatial planning of Jilin Province, but also provide valuable experience and reference for the soil erosion control work in other areas. Through scientific planning and early warning system, it is expected to protect valuable land resources, maintain the ecological balance of the earth, and contribute to the sustainable development of mankind.

Keywords: Soil Erosion; Early Warning Model; Artificial Intelligence; Environmental Assessment

1. Introduction

The terrain of China is complex, and soil erosion occurs in many places due to the unique geographical environment. The common erosion types include hydraulic erosion, wind erosion, freeze-thaw erosion, gravity erosion, glacial erosion and plant erosion. The occurrence of soil erosion is mostly a natural phenomenon, but with the increasingly close connection between human production and living activities and the natural environment, human activities may become one of the main factors of soil erosion. The appearance of soil erosion will lead to land degradation, reduce land productivity, destroy farmland and the integrity of cultivated land, so that land resources continue to decrease. At the same time, the soil carried away by the water is deposited in the downstream channel, weakening the flood discharge capacity of the river bed, and silting in the reservoir leads to the reduction of the area of the reservoir, mountain pond and lake, thus aggravating the occurrence of flood disasters. In addition, serious soil erosion will lead to a decline in farmland fertility and significantly reduce crop yields. Therefore, it is necessary to take measures to prevent soil erosion and protect land resources and ecological environment. With the continuous development of science and technology, artificial intelligence and big data technology continue to mature, using artificial intelligence and big data to establish a set of intelligent soil erosion assessment and early warning model can effectively monitor the soil and water environment in various regions, and meet the requirements of ecological protection and soil erosion control.

At present, China is in a period of rapid development, and the territorial space planning system is facing new challenges and opportunities. In this critical period, regional water and soil governance has become the core key to improve the territorial space governance system and governance capacity. Soil and water management is not only related to the protection of ecological environment, but also directly affects the sustainable development of economic society. In order to better cope with this challenge, this study took the soil and water environment of Jilin Province as an example to deeply discuss the construction and application of soil erosion assessment and early warning technology model. Based on the spatial guidance of ecological carrying capacity early warning, the

technical model aims to provide scientific basis for land spatial planning and improve the efficiency and effect of territorial spatial governance. Through this research, we have formed a soil erosion assessment and early warning technology model which has the value of popularization and application. The model integrated use of geographic information system, remote sensing technology, mathematical model and other methods to achieve accurate monitoring and early warning of soil erosion. At the same time, the model also considers the actual situation and characteristics of different regions, and has good universality and operability.

The results of this study not only provide a scientific reference for the land spatial planning of Jilin Province, but also provide a useful reference for soil erosion assessment and early warning in other areas of the country. In the future, we will continue to improve this technology model, promote its wide application in more areas, and make greater contributions to China's land spatial planning and governance.

2. Related Work

In order to better design the soil erosion assessment and early warning system, a large number of literatures were reviewed before the study began, and it was found that there were already studies on soil erosion assessment and early warning. Nie W et al. introduced the latest research advances and methods in natural and engineering slope monitoring, warning, disaster remediation and risk assessment, aiming to study slopes that may become unstable due to natural factors or human activities, and find ways to continuously monitor these slopes to provide disaster warning, so that engineers and emergency services can respond accordingly [1]. In addition, Xiao W et al. emphasized the topic of how soil erosion and climate change exacerbated natural disasters, and researched and developed landslide related models, which had made great contributions to understanding soil erosion and landslide processes and corresponding necessary tools for cultivating resilience [2]. In terms of soil erosion characteristics, Cai W et al. studied the evaluation of the erodibility of geological materials under the action of seepage force based on the determination of erosion characteristics, and established a simplified model to explain the erosion phenomenon of geotechnical materials. The physical model combines Bernoulli's principle and the erosive structure law of internal erosion of ground lines driven by pressure gradients, from which analytical formulas are derived to determine important parameters such as soil particle removal, radial erosion propagation, erosion coefficients, and critical shear stress [3].

For the same early warning model system, Senanayake S et al. planned a deep learning method to predict the probability of soil erosion. They collected daily precipitation data from five agricultural weather stations in the central Highlands of Sri Lanka from 1990 to 2021, and used a long short-term memory neural network model to train these historical precipitation data and learn time series patterns within them. The model can predict precipitation values for each weather station over the next 36 months and assess the corresponding risk of soil erosion. This integrated approach contributes to a better understanding of precipitation and soil erosion dynamics in the central highlands of Sri Lanka and provides a scientific basis for agriculture and soil conservation management [4]. Risaldi M U et al. Planed to build an integrated system with Arduino Uno microcontroller as the core processor and SIM-800LV2 module to realize wireless communication function. In addition, the system includes piezoelectric sensors and soil moisture sensors for collecting environmental data. It is also equipped with a buzzer and LCD display for immediate feedback. The developed system can use the SIM-800LV2 module to send this data to remote users via short messages. The deployment of this system has wide application potential in agricultural monitoring, environmental control, smart home and other fields [5].

Although the above studies have made significant progress in slope monitoring, early warning, disaster rehabilitation and risk assessment, there are still some shortcomings. Many studies rely on historical data, which may limit the predictive power and accuracy of the models. Future research could focus more on the acquisition and processing of real-time data to better reflect the current

environment and predict future changes. In addition, the current research mainly focuses on specific regions or conditions, in order to make the model more general and reliable, it is necessary to further study how to improve the generalization ability of the model, so that it can adapt to various environments and conditions. Since soil erosion problems involve multiple disciplinary fields, such as geography, soil science, meteorology and engineering, etc. [6], research should pay more attention to interdisciplinary cooperation and comprehensively utilize theories and knowledge from different fields to solve these problems. Although there are already some technologies for slope monitoring and early warning [7], there is still room for improvement. For example, more advanced sensors, communication technologies and data analysis methods are used to improve the accuracy and real-time monitoring [8]. Therefore, this paper proposes to develop a soil erosion assessment and early warning model based on big data and artificial intelligence, combined with the soil erosion model and the support of geographic information system, to develop a new soil erosion assessment and early warning system based on the soil and water environment of Jilin Province, and in order to apply the system to the actual experiment.

3. Methods

3.1 Soil Erosion Sensitivity Assessment Methods

Soil erosion sensitivity refers to the potential trend and degree of soil erosion in the natural environment [9]. This indicator reflects the soil erosion sensitivity, which is an important basis for evaluating soil protection capacity and formulating soil and water conservation measures. In this paper, the general soil loss formula is used to estimate soil erosion. The formula is as follows:

$$A = R \times K \times LS \times C \times P \quad (1)$$

In formula (1), A is the actual soil erosion amount in the test area ($t/(km^2 \cdot a)$); R is rainfall erosion factor ($MJ \cdot mm/(hm^2 \cdot h \cdot a)$); K is the soil erosion factor ($t \cdot hm^2 \cdot h/(hm^2 \cdot MJ \cdot mm)$); LS indicates the slope length and slope coefficient; C is the coverage and management factor, and P is the factor of soil and water conservation measures.

3.2 Acquisition and Processing of Remote Sensing Data

Vegetation coverage is the proportion of the projected area of vegetation canopy to the total land area, expressed in percentage [10]. Vegetation coverage is an important factor affecting wind erosion. The higher the vegetation coverage, the weaker the wind erosion, and the worse the wind erosion. The pixel dichotomy method is used to calculate vegetation coverage, and the calculation formula is as follows:

$$VC = \frac{NDVI - NDVI_{\min}}{NDVI_{\max} - NDVI_{\min}} \times 100\% \quad (2)$$

In formula (2), VC represents the vegetation cover rate, and $NDVI$ represents the standardized vegetation index. Among them, $NDVI_{\max}$ is the standardized vegetation index value of the area with completely bare surface and no vegetation cover, and $NDVI_{\min}$ is the standardized vegetation index of the area with completely vegetation cover and no bare land. These parameters are important for evaluating the growth status of vegetation and land cover. The principle of this model is to assume that the reflectance R of pixels in the image can be divided into two parts: pure vegetation reflection R_V and non-vegetation reflection R_S [11]. Therefore, the reflection value of each pixel can be expressed as a linearly weighted sum of R_V (vegetation cover) and R_S (non-vegetation cover).

$$R = R_V + R_S \quad (3)$$

If the proportion of the area covered by vegetation to the total area of the pixel is f_c , then the proportion of the area covered by non-vegetation is $1 - f_c$. In remote sensing images, if an pixel is completely covered by vegetation, its reflectance is usually close to that of pure vegetation R_{vcg} . On the contrary, if an pixel has no vegetation cover at all, its reflectance is close to the background reflectance of the feature R_{soil} . Therefore, the information contributed by the vegetation part of the mixed pixel can be expressed as the product of the reflectivity of the pure vegetation and the

vegetation coverage area in the pixel f_c (formula 4), while the information contributed by the non-vegetation component can be expressed as the product of R_{soil} and $1 - f_c$ (formula 5).

$$R_V = f_c \times R_{vcg} \quad (4)$$

$$R_S = (1 - f_c) \times R_{soil} \quad (5)$$

By solving formulas (3), (4) and (5), the formula for calculating vegetation coverage can be derived, which is as follows:

$$f_c = (R - R_{soil}) / (R_{vcg} - R_{soil}) \quad (6)$$

Among them, R_{soil} and R_{vcg} are the two parameters of the pixel binary model [12]. The vegetation index is calculated by using binary pixel model and represents the reflectance value of the vegetation covered part in pixel. The total pixel count is the number of pixels in a remote sensing image. Through this formula, the vegetation coverage in a specific area can be calculated to further evaluate its ecological status and environmental changes.

4. Results and Discussion

4.1 Data Collection on Factors Causing Soil Erosion

1) Type factor of water source protection area

Different levels of water source protection areas have different impacts on water environment ecology in the study area. The landscape ecology in nature and cultural reserves and buffer zones has been well protected and restored, mainly thanks to strict protection measures and regulations. In contrast, the situation in development and utilization areas is different. As the development and utilization areas mainly meet the needs of agriculture and industry, their requirements for water resources and environment are relatively low [13]. This means that in the development and utilization area, the pressure on the water resources environment may be small, and the damage to the ecological environment is relatively limited. In some cases, the environmental requirements for water resources in the buffer zone may be even more relaxed, because the main function of the buffer zone is to ease the pressure between the protected area and the development and utilization area, providing a transition area.

Therefore, for the water environment ecology in Jilin Province, different levels of water source protection areas have different impacts [14]. This is not only related to the strict protection measures of the protected area, but also closely related to the needs of the development and utilization area and the functional positioning of the buffer zone. In formulating water resources management and protection strategies, various influencing factors need to be fully considered to ensure the sustainable utilization of water resources and the good state of the ecological environment. In this study, the study area is categorized into four types: protected area, conservation area, exploitation area, and buffer area based on the existing planning and actual situation in the study area. This classification helps us to carry out water resources management and protection work in a more targeted manner in order to achieve efficient utilization of water resources and balanced development of the ecological environment. The specific weight classification is shown in Table 1.

Table 1. Resistance coefficient and weight of water source protection area type factors on environmental safety

Molecule of resistance	Grading criteria	Resistance class	Relative resistance value	Resistance factor weights
Type of water source protection area	Conservation area	1	15	25%
	Reserved area	2	25	
	Utilization area	3	35	
	Buffer area	4	45	

2) Terrain factor

Topographic factor, as a key expression factor of local natural environment, provides a comprehensive description of basic landform types and topographic relief. Terrain diversity has a profound impact on all aspects of the entire ecological environment, such as plant growth, animal survival and climate change [15]. This is because the terrain is not only directly related to soil, water and light conditions, but also indirectly affects the migration and reproduction of organisms, as well as the formation and change of climate.

In this paper, we pay special attention to slope as a topographic factor. The main reason for choosing slope factor is that it is closely related to the topographic and geomorphologic characteristics of the study area. The overall terrain slope of the study area is relatively gentle, so compared with other terrain factors, the weight of slope factor in the overall resistance factor system is relatively low [16]. In order to show the coefficients and weights of slope resistance factors more specifically, this paper collates relevant information and makes Table 2. Table 2 will help readers better understand the role and importance of slope factors in the study area. In general, the weight allocation of slope factors is obtained after comprehensive consideration of the topographic characteristics of the study area, the needs of the ecosystem and climate change. Such weight distribution not only reflects the role of slope factor in ecosystem, but also reflects its value in the study of regional environmental change.

Table 2. Resistance coefficient and weight of slope factor on environmental safety

Drag factor	Grading criteria	Resistance class	Relative resistance value	Resistance factor weights
Elevation	0-5	1	5	15%
	5-10	2	15	
	10-15	3	20	
	15-25	4	25	
	25 or more	5	30	

3) Standardization of soil indicators

On this basis, in order to further optimize the early warning results, the research combined soil science knowledge and local agricultural production conditions, while referring to the experience of experts, and carried out detailed quantitative assignment of soil data in Jilin Province. The data are shown in Table 3.

Table 3. Standardization of soil indicators

Metric	10	8	6	4	2
Ph	6.0~7.0	4.5~5.5	4.5~5.0, 6.5~7.5	8.0~9.5	< 4.0, < 8.0
Organic matter	≥ 35	35~25	25~15	15~5	<5
Soil thickness	180.66~210.84	159.86~180.64	141.17~159.86	127.15~141.14	102.19~127.14
Carbon content	24.54~29.45	20.19~24.58	15.48~20.49	10.84~15.47	4.87~10.84

4.2 Put the Model into the Experiment

Through a series of calculation and analysis, the correlation between the soil erosion conditions and the occurrence of soil erosion can be obtained by inputting the data of soil and water conditions in the experimental area of Jilin Province into the newly established soil erosion model. In order to show this result more intuitively, this paper makes Figure 1 based on the calculated data of the model. Figure 1 clearly shows the correlation between each factor and soil erosion.

From Figure 1, we can see that factors such as slope, soil type and vegetation type have a high correlation with soil erosion. This means that in the Jilin Province pilot area, these factors are the main conditions that lead to soil erosion. Based on these results, researchers can develop targeted measures, such as strengthening soil and water conservation, increasing vegetation cover, etc., to reduce the occurrence of soil erosion.

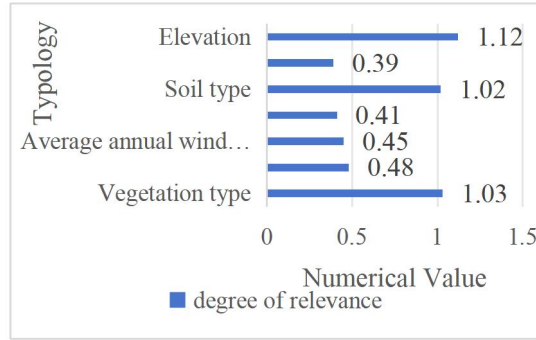


Figure 1. Correlation of each factor with soil erosion

As can be seen from Figure 1, the slope with the greatest correlation with soil erosion is 1.12. This was followed by soil type and vegetation type at 1.02 and 1.03 respectively. Therefore, the model can determine the items that need to be paid attention to in the process of soil erosion prevention to reduce the workload [17].

After that, threshold values were set for the allowable soil loss and warning classification in the experimental area. The allowable soil loss in the experimental area was set to $1000\text{t}\cdot\text{km}^{-2}\cdot\text{a}^{-1}$, and no warning was given when the soil and water carrying capacity was ≤ 1.0 , warning was given when the soil and water carrying capacity was $1\sim 1.5$, and warning was given when the soil and water carrying capacity was > 5.0 . Then, the slope of the study area is divided into 5 grades, the slope $< 5^\circ$ is flat land, the slope of $5^\circ\sim 10^\circ$ is flat land, the slope of $10^\circ\sim 15^\circ$ is gentle slope, the slope of $15^\circ\sim 25^\circ$ is slope land, and the slope $> 25^\circ$ is steep land. Figure 2 shows the early warning results for each slope using the model.

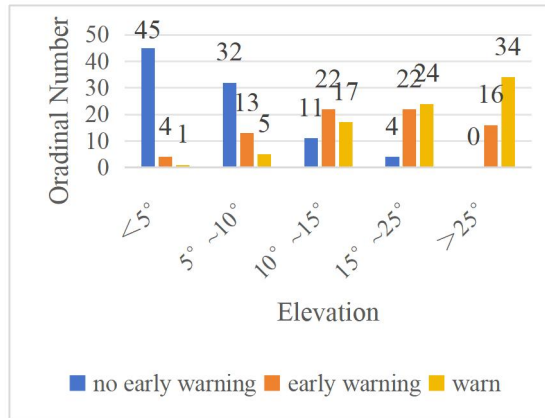


Figure 2. Statistical chart of warning times of each slope

According to the actual soil and water loss on the surface of the test area and the statistics of the warning times of each slope in Figure 2, it can be concluded that the flat land with slope $< 5^\circ$ has flat ground, generally no obvious erosion, and the ecological condition is good. The slope of $5^\circ\sim 10^\circ$ has slight soil erosion, and the ecological condition is better. The gentle slope of $10^\circ\sim 15^\circ$ has the phenomenon of topsoil loss, moderate erosion, and the ecological condition is medium. The slope of $15^\circ\sim 25^\circ$ has serious soil erosion and poor ecological condition. Steep land with slope $> 25^\circ$ has severe soil erosion and poor land ecological condition. The conclusion is consistent with the data of local water resources bureau in Jilin Province, so the early warning model can be put into practical use.

5. Conclusion

Jilin Province is located in northeast China and has a complex physical and geographical environment. The land in this region has formed a large area of loose black soil due to its unique natural geographical location and geological structural background. This soil structure is prone to

geological disasters such as shallow landslides under the erosion of heavy rainfall. This kind of soil structure is prone to geological disasters such as shallow landslides under the erosion of heavy rainfall. Most of the existing early warning methods are based on experience judgment or calculation of infiltration and stability of uniform soil. However, these methods may not provide accurate early warning in the special soil structure and climatic conditions of Jilin Province. In order to solve this problem, a new soil erosion model is proposed in this study. The model takes into account not only the effect of rainfall on soil, but also a variety of factors such as vegetation cover and land use status. Through this model, the relevant personnel can more accurately calculate the amount of soil loss at the soil erosion warning point. Through practical application experiments, it is found that the new model is more reasonable and accurate. This provides a strong scientific basis for the disaster prevention and reduction work in Jilin Province, and helps to better protect the lives and property safety of local residents.

6. Acknowledgements

This work was supported by the Jilin Province Science and Technology Development Plan Project (NO.20230203134SF), Development of Short-term Precipitation Risk Assessment and Early Warning System for Hydraulic Erosion in the Black Soil Region of Jilin Province

References

- [1] Nie W, Qiu H, Asadi A. Monitoring, early warning, and mitigation of natural and engineered slopes[J]. *Frontiers in Earth Science*, 2023, 10(1): 104-180.
- [2] Xiao W, Ren H, Sui T, et al. A drone-and field-based investigation of the land degradation and soil erosion at an opencast coal mine dump after 5 years' evolution of natural processes[J]. *International Journal of Coal Science & Technology*, 2022, 9(1): 42-43.
- [3] Cai W, Bora M J, Pekkat S, et al. A new and simple model for predicting soil erosion based on hole erosion tests[J]. *Acta Geophysica*, 2023, 71(2): 823-836.
- [4] Senanayake S, Pradhan B, Alamri A, et al. A new application of deep neural network (LSTM) and RUSLE models in soil erosion prediction[J]. *Science of The Total Environment*, 2022, 84(5): 157-220.
- [5] Risaldi M U, Baswara A R C. Design of a Soil Erosion Warning System in Watersheds Based on Arduino Uno[J]. *Buletin Ilmiah Sarjana Teknik Elektro*, 2023, 5(1): 167-179.
- [6] Yan R, Gao J. Key factors affecting discharge, soil erosion, nitrogen and phosphorus exports from agricultural polder[J]. *Ecological modelling*, 2021, 45(2): 109-586.
- [7] Borrelli P, Panagos P, Alewell C, et al. Policy implications of multiple concurrent soil erosion processes in European farmland[J]. *Nature Sustainability*, 2023, 6(1): 103-112.
- [8] Masroor M, Sajjad H, Rehman S, et al. Analysing the relationship between drought and soil erosion using vegetation health index and RUSLE models in Godavari middle sub-basin, India[J]. *Geoscience Frontiers*, 2022, 13(2): 101-312.
- [9] Chen H S, Wang X K, Lei S, et al. Soil erosion characteristics in an earthquake-impacted mountainous basin based on RUSLE model: implications for geohazards[J]. *Journal of Soils and Sediments*, 2023, 23(11): 3857-3878.
- [10] Chen C, Wang X, Wu C, et al. Soil erosion prediction based on moth-flame optimizer-evolved kernel extreme learning machine[J]. *Electronics*, 2021, 10(17): 21-25.
- [11] Zhao Z, Huo A, Liu Q, et al. Spatiotemporal Variation in the Coupling Relationship between Human Activities and Soil Erosion—A Case Study in the Weihe River Basin[J]. *Sustainability*, 2023, 15(14): 107-185.
- [12] Egbueri J C. Soil erosion and landslide susceptibility insights based on hierarchical clustering and multilayer perceptron networks: a Nigerian case study[J]. *International Journal of Environmental Science and Technology*, 2023, 20(10): 10763-10786.

- [13] da Silva L C M, Avanzi J C, Peixoto D S, et al. Ecological intensification of cropping systems enhances soil functions, mitigates soil erosion, and promotes crop resilience to dry spells in the Brazilian Cerrado[J]. *International Soil and Water Conservation Research*, 2021, 9(4): 591-604.
- [14] Lee S, Chu M L, Guzman J A, et al. A comprehensive modeling framework to evaluate soil erosion by water and tillage[J]. *Journal of Environmental Management*, 2021, 27(9): 111-131.
- [15] González C A R, Rodríguez-Pérez Á M, Mancera J J C, et al. Applied methodology based on HEC-HMS for reservoir filling estimation due to soil erosion[J]. *Journal of Hydrology and Hydromechanics*, 2022, 70(3): 341-356.
- [16] Haq Y, Shahbaz M, Asif H M S, et al. Spatial Mapping of Soil Salinity Using Machine Learning and Remote Sensing in Kot Addu, Pakistan[J]. *Sustainability*, 2023, 15(17): 129-143.
- [17] Liu C, Deng C, Li Z, et al. Response characteristics of Soil Erosion to Spatial Conflict in the Production-Living-Ecological Space and Their Drivingmechanism: A Case Study of Dongting Lake Basin in China[J]. *Land*, 2022, 11(10): 17-24.