MDPI

Article

Assessment and Prediction of Water Resources Vulnerability Based on a NRS-RF Model: A Case Study of the Song-Liao River Basin, China

Weizhong Chen [1], Yan Chen [1,2,*] and Yazhong Feng [1]

[1] College of Economics and Management, Nanjing Forestry University, Nanjing 210037, China; chenweizhong0713@163.com (W.C.); fengyazhong@njfu.edu.cn (Y.F.)

[2] Academy of Chinese Ecological Progress and Forestry Development Studies, Nanjing Forestry University, Nanjing 210037, China

* Correspondence: sanchen007@njfu.edu.cn; Tel.: +86-025-8542-7377

Abstract: The vulnerability of water resources is an important criterion for evaluating the carrying capacity of water resources systems under the influence of climate change and human activities. Moreover, assessment and prediction of river basins' water resources vulnerability are important means to assess the water resources security state of river basins and identify possible problems in future water resources systems. Based on the constructed indicator system of water resources vulnerability assessment in Song-Liao River Basin, this paper uses the neighborhood rough set (abbreviated as NRS) method to reduce the dimensionality of the original indicator system to remove redundant attributes. Then, assessment indicators' standard values after dimensionality reduction are taken as the evaluation sample, and the random forest regression (abbreviated as RF) model is used to assess the water resources vulnerability of the river basin. Finally, based on data under three different future climate and socio-economic scenarios, scenario predictions are made on the vulnerability of future water resources. The results show that the overall water resources vulnerability of the Song-Liao River Basin has not improved significantly in the past 18 years, and the overall vulnerability of the Song-Liao River Basin is in the level V of moderate to high vulnerability. In the future scenario 1, the overall water resources vulnerability of the river basin will improve, and it is expected to achieve an improvement to the level III of moderate to low vulnerability. At the same time, the natural vulnerability and vulnerability of carrying capacity will increase significantly in the future, and the man-made vulnerability will increase slowly, which will deteriorate to the level V of moderate to high vulnerability under Scenario 3. Therefore, taking active measures can significantly reduce the vulnerability of nature and carrying capacity, but man-made vulnerability will become a bottleneck restricting the fragility of the overall water resources of the river basin in the future.

Keywords: basin water resources; vulnerability assessment; neighborhood rough set; random forest regression; scenario prediction

1. Introduction

As an indispensable natural resource for human survival and social development, water resources play an extremely important role in the sustainable development of a viable socio-economy. However, water is also a limited resource that is unevenly distributed in time and space. Under the influence of increasing human activities and rapid urban development, water environmental pollution and water shortages have always plagued human beings and have made the sustainable development of water resources in river basins appear vulnerable [1]. As one of China's important water systems, although the Song-Liao River Basin has relatively richer water resources than Northwest and North China, there are problems such as uneven temporal and spatial distribution and prominent contradictions between water supply and demand. Also, the distribution of precipitation

in this basin is extremely uneven [2]. Furthermore, water resource utilization has low efficiency and serious waste. At present, rivers in this basin have poor self-purification ability [3]. Also, water pollution has become an important factor that restricts the economic and social development of Song-Liao River Basin [4]. Many of the above factors have contributed to the increase in the vulnerability of water resources in the basin. Therefore, it is necessary to conduct in-depth research on the vulnerability of water resources in the Song-Liao River Basin in order to better manage the water resources to reduce the vulnerability.

The concept of water resources vulnerability originated from the related studies on groundwater vulnerability proposed by the French scholar Margat [5]. After the enrichment and development of Xia et al. [6–8], the connotation and scope of water resources vulnerability research continue to expand, and a rich research system on the vulnerability of water resources systems has been formed. In terms of research object selection, not all scholars chose general water resources system as the research object. Many scholars studied the vulnerability of various specific components of water resources systems, such as surface water [9,10], ground water [11,12], drinking water [13], household water [14], urban runoff [15] and water quantity [16]. What's more, there are also many scholars who have combined water resources with social, economic, ecological and other external factors to conduct research on vulnerability [17,18]. In terms of research perspective selection, many scholars currently choose to study the vulnerability of water resources from the perspective of climate and land, that is, to study the impact of climate change and land use on water resources vulnerability [19–21]. In terms of research frameworks, some frameworks such as the 2014 Intergovernmental Panel for Climate Change (IPCC) framework is applied to the research on vulnerability in the context of climate change [22]. Combining existing relevant research literature [23–25], the authors believe that water resources vulnerability is the degree to which the water resources system is susceptible to and unable to cope with the adverse effects of the external natural climate change and human socio-economic activities. In this paper, the water resources system mentioned in the definition of water resources vulnerability refers to various water bodies (such as rivers, lakes, reservoirs, groundwater, etc.) and related engineering buildings that are related to each other in hydrology, hydraulics, and water conservancy within a basin or region.

The current research on water resources vulnerability includes quantitative research and qualitative research. There are few qualitative studies, which mainly focus on the theoretical framework and tools related to water resources vulnerability [22,26]. At present, most of the research related to the vulnerability of water resources focuses on the quantitative model research, which mainly includes the evaluation and prediction of the vulnerability of the water resources system. In terms of evaluation, the functional evaluation method and indicator evaluation method are used a lot. Among them, the function evaluation method [27] has a clear operation mechanism and is easy to evaluate multiple regions, but there are problems such as difficulty in interpreting the results. Therefore, at present, scholars have published relatively few studies on this method. Compared with the function method, the index evaluation method is flexible, easy to explain, clear in structure, and compatible with multiple model methods, so most of the current studies are carried out on the basis of indicator evaluations. Regarding the selection of water resources vulnerability indicators, there is no unified objective standard. Scholars have combined different evaluation objects and theoretical connotations to construct corresponding indicators, and they used a variety of methods to evaluate water resources vulnerability. In the application of evaluation methods, weight evaluation methods, such as entropy weight method [28], projection pursuit method [29], analytic hierarchy process (AHP) method [30] are used more frequently. In addition, principal component analysis (PCA) [31,32], technique for order preference by similarity to an ideal solution (TOPSIS) [33], rough set cloud [34], connection entropy approach [35], expert-based model [36] and other quantitative methods are also used in the assessment of water resources vulnerability. What's more, Geographical Information System (GIS) is often used in the study of water resources vulnerability [37]. Many scholars have improved the GIS model and combined it with AHP [37,38] and F-Hydra [39].

Among them, GIS and the integrated climatic modelling framework (geographical spatially distributed water balance model (GEO-CWB)) model developed and verified by Trinity College (Dublin, Ireland) are used together in the study of water resources vulnerability [40]. Furthermore, some indicator models are used for the vulnerability assessment and mapping. Among them, the indicator models such as DRASTIC (depth to groundwater, recharge rate, aquifer, soil, topography, vadose zone's impact, aquifer's hydraulic conductivity) model [41–44], SINTACS (Soggiacence, actual infiltration, non-saturated zone, typology of overburden, hydrogeological characteristics of the aquifer, hydraulic conductivity, topographic Slope) model [45], GALDIT model (an abbreviation for the six parameters of groundwater occurrence, aquifer hydraulic conductivity, level of groundwater above sea level, distance from the shoreline, impact of the existing status of seawater intrusion, and thickness of the aquifer) [46] and the susceptibility index (SI) method [47] are the most used ones. There are also some methods for evaluating the vulnerability of karst water, such as epikarst, protective cover, infiltration conditions, karst network development (EPIK), protective cover, infiltration conditions (PI), concentration of flow, overlaying layers, precipitation (COP) and altitude, slope, lithology, infiltration, soil (APLIS) methods [48,49]. For the prediction of water resources vulnerability, the number of evaluation models is also considerable. In addition to the indicator method and function method related to water resources vulnerability evaluation used in the prediction, there are other efficient prediction methods and technologies, such as the Monte Carlo model [50], soil and water assessment tool-water evaluation and planning system (SWAT-WEAP) coupling method [51], Arc-Wofe system simulation method [52], enriching the existing prediction method system of water resources vulnerability.

At present, the research on the vulnerability of water resources in river basins has achieved fruitful results. The connotation of water resources has been continuously enriched and developed, and the corresponding evaluation methods have become increasingly complete. Specifically, the evaluation content has realized the transformation from a single water quality study to a comprehensive system evaluation combining natural climate, ecological, and socio-economic systems. In the past, the indicator evaluation and prediction of water resources vulnerability had the following problems. First, in terms of the construction of the indicator system, some researched indicator settings were insufficiently targeted. What's more, there are many and complex indicators involved in constructing a water resources vulnerability indicator system, and the interrelationships between indicators are closely related. Therefore, problems such as multi-index data redundancy, inter-index correlation and duplicate information are easy to appear. If they are not handled properly, these problems will cause a certain interference to the accuracy of the evaluation results objectively. Second, most of the vulnerability research of water resources is in terms of evaluation and relatively little research is in terms of prediction. As a result, management control has become a post-control, and it is difficult to make timely preparations for adjustment and control in advance. Third, for the data to be predicted in the future scenario, the accuracy of the model is a problem worthy of attention. In order to more accurately predict the vulnerability of water resources under future scenarios, the comparison and selection of models is crucial.

The Song-Liao River Basin is a densely populated area, a food production base and an important heavy industry base. The vulnerability of water resources in the Song-Liao River Basin deeply affects people's lives, food and energy production in the basin. In addition, compared with the other five major rivers in China, the rivers in the Song-Liao Basin in the Northeast China have the following unique characteristics: a long dry season, many tributaries, small water flow and long-term freezing on the water surface. These characteristics make the Songhua River and Liao Rivers have poor self-purification ability and be more susceptible to pollution. However, there are few studies on the overall vulnerability of water resources in it. The existing research focuses on the vulnerability of groundwater resources in the Songhua and Liao River Basin, which mainly uses the indicator evaluation method of the DRASTIC model [53,54]. Other studies on the vulnerability of water

resources in the Song-Liao River Basin include the risk assessment of storm flood disasters in the mainstream of the Songhua River [55], the study on the vulnerability of the Songhua River Basin landscape pattern and its driving forces [56], risk assessment of groundwater pollution caused by agricultural activities in the Lower Liao Plain [57] and so on.

In order to better analyze the vulnerability of water resources to better manage the water resources in the Song-Liao River basin, we will do the following tasks: combining the actual conditions of the Song-Liao River Basin, this paper constructs an indicator system for water resources vulnerability assessment in the Song-Liao River Basin. In order to solve the problems of data redundancy and duplicate information in the index system, this paper uses the neighborhood rough set attribute reduction method to reduce the dimensionality of the original index system and remove redundant indicators. Then, this paper uses the standard value of the original indicator level threshold as the input vector and the vulnerability level value as the output vector to construct a random forest model. The assessment indicator system after dimensionality reduction is used as the basis of the random forest model to evaluate and predict the water resources vulnerability of the Song-Liao River Basin. The indicator data under different scenarios in 2025 and 2030 is substituted into the random forest model as input data to obtain the predicted value of water resources vulnerability in 2025 and 2030, so as to provide a reference for future water resources planning and adaptive management.

2. Data and Methodology

2.1. Research Area

The Song-Liao River Basin includes two sub-basins, the Liao River Basin and the Songhua River Basin, whose total area is 1,248,948 km^2. The basin is surrounded by mountains to the west, north and east, and there is a vast plain in the middle of the basin. The specific distribution of rivers is shown in Figure 1. In the Song-Liao River Basin, that is rich in black soil resources, the plains are flat, the soil is fertile, the rainfall is abundant, the climate is suitable, the sunlight is sufficient, and it has good conditions for agricultural development [58]. What's more, the basin is also a major industrial hub for coal, steel, petrochemical and equipment manufacturing [59–61]. In terms of precipitation, the annual average precipitation in the Song-Liao River Basin shows a spatial distribution characteristic of less precipitation in the northwest and more precipitation in the southeast [2]. The spatial distribution of annual average precipitation in Song-Liao Basin is shown in Figure 1. At present, the Song-Liao River Basin is in a stage of extreme precipitation, and the probability of extreme precipitation has increased [62]. What's more, due to uneven seasonal precipitation and the high-water level of the drainage channel, the internal water cannot be discharged. The mountainous area has serious soil erosion, and a large amount of sediment is discharged and deposited on the drainage channel, which causes that the riverbed has been silted up year by year, the channel is blocked, and the water is urgent in the flood season. Also, poor water leakage in the waterlogged area and low infiltration capacity have led to frequent occurrence of waterlogging in the watershed. As the northwestern part of the basin is in the westerly zone, when a large precipitation front and low-pressure system pass through this area, the rainfall will be significantly reduced or no precipitation. The normal annual precipitation is below 400 mm, and the water surface evaporation capacity is 1100–1200 mm, which caused many droughts [2]. With the development of industrialization and urbanization and the massive discharge of industrial wastewater and domestic sewage, the Song-Liao River Basin is facing a severe situation of aggravated pollution. According to relevant statistical data such as environmental yearbook, since 2014, the chemical oxygen demand (COD) emissions of the Song-Liao River Basin are about 830,400 tons, and the ammonia and nitrogen emissions are about 12.85 tons. Compared with the previous level, the water pollution problem still cannot be ignored. This objectively aggravates the vulnerability of water resources, so research on the vulnerability of water resources is timely.
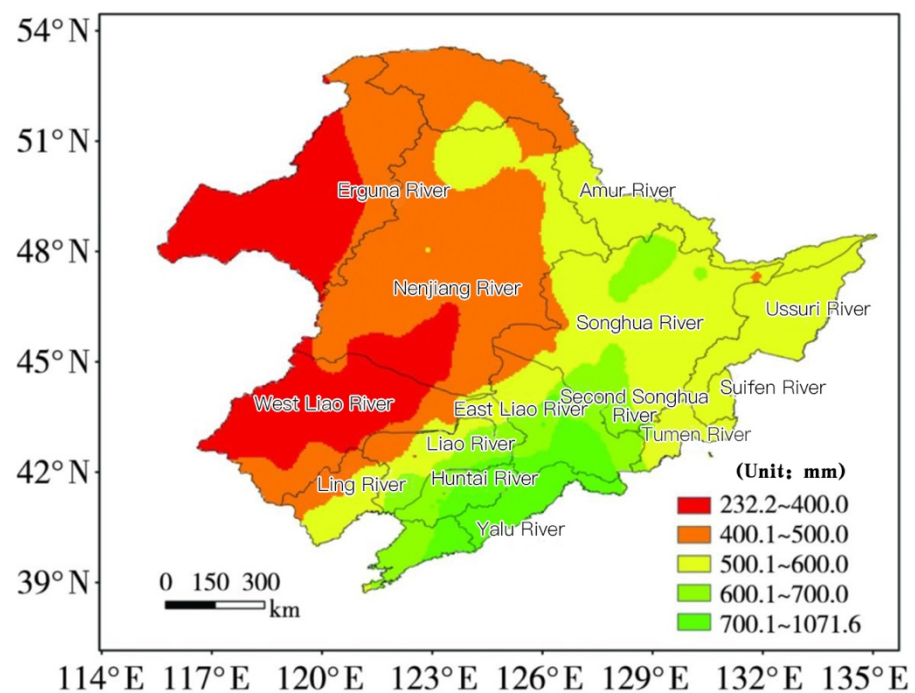
Figure 1. Research area.

2.2. Data Sources

In the empirical research part of this article, the water resource vulnerability assessment period is 2000–2017. In the scenario prediction stage, we mainly study the vulnerability of water resources in 2025 and 2030. Also, the data format of all indicators is time series. Among them, some data on indicators related to water resources vulnerability of the Song-Liao River Basin from 2000 to 2017 is derived from the "Song-Liao River Basin Water Resources Bulletin" (2000–2017), "China Statistical Yearbook" (2000–2018), and "China Environmental Statistical Yearbook" (2000–2018) and "China Water Conservancy Yearbook" (2000–2018). What's more, some indicators need to be calculated by the data directly obtained. For example, the reclamation index is obtained by dividing the cultivated land area by the total land area.

2.3. Methodology

In this paper, we construct an indicator system for water resources vulnerability assessment in the Song-Liao River Basin. Then, we use the neighborhood rough set method to reduce the dimensionality of the original index system and remove redundant indicators. Finally, we construct a random forest model to evaluate and predict the water resources vulnerability.

2.3.1. Assessment Indicator System

Water resources refer to available resources or the water source that may be used, which should have sufficient quantity and suitable quality and meet the specific needs of a certain place in a period of time. What's more, the drainage basin refers to the entire area through which the mainstream and tributary of the river flow. Compared with static areas, its coverage is wider. Therefore, the composition of the vulnerability of water resources in a river basin is more complex, which includes the vulnerability generated in the degradation of the water resources system itself and the recycling process, as well as the vulnerability generated in the development and utilization of water resources. In this paper, starting from the characteristics, manifestations and causes of water vulnerability, combining the situation of the Song-Liao River Basin itself, we construct an assessment indicator system on the framework of nature, man-made and carrying capacity. Therefore, we construct the

first-level indicators from the three aspects of natural vulnerability, man-made vulnerability and vulnerability of carrying capacity. In particular, the indicators of natural vulnerability mainly characterize the impact of natural conditions on the quantity and quality of water resources. The indicators of man-made vulnerability mainly characterize the impact of human activities on the quantity and quality of water resources. The indicators of the vulnerability of carrying capacity mainly characterize the carrying capacity of water resources for human social and economic development from the perspective of the demand for water by the human socio-economic system. Then, we construct third-level indicators under each of second-level indicators of water quantity, water quality and disasters. According to the construction principles of the indicator system and the availability and completeness of data, 24 assessment indicators are selected to construct an assessment indicator system for water resources vulnerability finally, which is shown in Table 1.

Table 1. Assessment indicator system of water resources vulnerability in Song-Liao River Basin.

First-Level	Second-Level	Number	Indicator	Attribute [1]
Natural vulnerability	Water quantity	A_1	Water production modulus	Negative
		A_2	Variation coefficient of annual precipitation	Positive
		A_3	Change rate of annual precipitation	Positive
	Water quality	A_4	Water quality examination pass rate in water function area	Negative
		A_5	Qualified ratio of water quality of river basin	Negative
		A_6	Decline rate of water quality examination pass rate	Positive
	Disasters	A_7	Proportion of area affected by flood and drought	Positive
		A_8	Water production coefficient	Negative
Man-made vulnerability	Water quantity	B_1	Proportion of surface water resources being utilized	Positive
		B_2	Proportion of groundwater resources being utilized	Positive
	Water quality	B_3	Total COD emission per 10,000 people	Positive
		B_4	Total ammonia and nitrogen emission per 10,000 people	Positive
	Disasters	B_5	Proportion of farmland area being the effectively irrigated	Negative
		B_6	Proportion of population under levee protection	Negative
		B_7	Proportion of soil erosion being controlled	Negative
		B_8	Water conservancy project storage capacity	Negative
Vulnerability of carrying capacity	Water quantity	C_1	Ratio of groundwater supply to total water supply	Positive
		C_2	Per capita water consumption	Positive
		C_3	Water consumption for irrigation per mu	Positive
	Water quality	C_4	Population density	Positive
		C_5	Wastewater generation per 10,000-yuan GDP	Positive
		C_6	Ecosystem water consumption rate	Negative
	Disasters	C_7	Population per 10,000 cubic meters of water	Positive
		C_8	Reclamation index	Positive

[1] The positive indicator indicates that the larger the value of the assessment index is, the higher the water resources vulnerability is; the negative indicator indicates that the smaller the value of the assessment index is, the lower the water resources vulnerability is.

There is currently no unified classification of the criteria for evaluating the vulnerability of water resources in river basins. Combining the situation of the Song-Liao River Basin and other river basins with similar environmental conditions, such as the Huai River Basin and Hai River Basin, a scale and standard suitable for evaluating the vulnerability of the Song-Liao River Basin has been established by referring to previous studies [3,15,63–65] and the current domestic and foreign water resources management standards, such as the evaluation standards of water scarcity established by organizations such as the World Meteorological Organization, the World Food and Agriculture Organization, the United Nations Educational, Scientific, Cultural and Health Organization, and the United Nations Commission for Sustainable Development. This paper divides the vulnerability of water

resources into seven levels, ranging from the level I of no vulnerability to the level VII of extreme vulnerability. The specific grade standard is shown in Table 2.

Table 2. Grade standard for assessment of water resources vulnerability [1].

	Level I	Level II	Level III	Level IV	Level V	Level VI	Level VII
A1	(60, 100]	(50, 60]	(40, 50]	(30, 40]	(20, 30]	(10, 20]	(0, 10]
A2	(0, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]	(0.5, 0.6]	(0.6, 0.8]
A3	(0, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]	(0.5, 0.6]	(0.6, 0.7]
A4	(0.9, 1]	(0.8, 0.9]	(0.7, 0.8]	(0.6, 0.7]	(0.5, 0.6]	(0.4, 0.5]	(0.2, 0.4]
A5	(0.8, 1]	(0.7, 0.8]	(0.6, 0.7]	(0.5, 0.6]	(0.4, 0.5]	(0.25, 0.4]	(0.15, 0.25]
A6	(−0.8, 0]	(0, 0.05]	(0.05, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]
A7	(0, 0.05]	(0.05, 0.1]	(0.1, 0.15]	(0.15, 0.2]	(0.2, 0.25]	(0.25, 0.3]	(0.3, 1]
A8	(0.7, 0.8]	(0.6, 0.7]	(0.5, 0.6]	(0.4, 0.5]	(0.3, 0.4]	(0.2, 0.3]	(0.1, 0.2]
B1	(0, 0.2]	(0.2, 0.25]	(0.25, 0.4]	(0.4, 0.55]	(0.55, 0.7]	(0.7, 0.85]	(0.85, 1]
B2	(0, 0.2]	(0.2, 0.25]	(0.25, 0.4]	(0.4, 0.55]	(0.55, 0.7]	(0.7, 0.85]	(0.85, 2]
B3	(15, 30]	(30, 45]	(45, 60]	(60, 75]	(75, 90]	(90, 105]	(105, 160]
B4	(4, 6]	(6, 8]	(8, 10]	(10, 12]	(12, 14]	(14, 16]	(16, 18]
B5	(0.95, 1]	(0.9, 0.95]	(0.85, 0.9]	(0.8, 0.85]	(0.75, 0.8]	(0.7, 0.75]	(0.65, 0.7]
B6	(0.9, 1]	(0.8, 0.9]	(0.7, 0.8]	(0.6, 0.7]	(0.5, 0.6]	(0.4, 0.5]	(0.2, 0.4]
B7	(0.9, 1]	(0.75, 0.9]	(0.6, 0.75]	(0.45, 0.6]	(0.3, 0.45]	(0.15, 0.3]	(0, 0.15]
B8	(0.9, 1.2]	(0.75, 0.9]	(0.6, 0.75]	(0.45, 0.6]	(0.3, 0.45]	(0.15, 0.3]	(0, 0.15]
C1	(0, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]	(0.5, 0.6]	(0.6, 0.7]
C2	(180, 200]	(200, 300]	(300, 400]	(400, 450]	(450, 500]	(500, 550]	(550, 600]
C3	(180, 200]	(200, 300]	(300, 400]	(400, 450]	(450, 500]	(500, 550]	(550, 600]
C4	(0.008, 0.01]	(0.01, 0.02]	(0.02, 0.03]	(0.03, 0.04]	(0.04, 0.05]	(0.05, 0.06]	(0.06, 0.09]
C5	(2, 4]	(4, 6]	(6, 8]	(8, 10]	(10, 15]	(15, 25]	(25, 75]
C6	(0.06, 0.09]	(0.05, 0.06]	(0.04, 0.05]	(0.03, 0.04]	(0.02, 0.03]	(0.01, 0.02]	(0, 0.01]
C7	(0, 3.34]	(3.34, 5]	(5, 10]	(10, 15]	(15, 20]	(20, 30]	(30, 90]
C8	(0, 0.1]	(0.1, 0.2]	(0.2, 0.3]	(0.3, 0.4]	(0.4, 0.5]	(0.5, 0.6]	(0.6, 0.7]

[1] The level I is the level of no vulnerability; The level II is the level of mild vulnerability; The level III is the level of moderate to low vulnerability; The level IV is the level of moderate vulnerability; The level V is the level of moderate to high vulnerability; The level VI is the level of highly vulnerability; The level VII is the level of extreme vulnerability.

2.3.2. Attribute Reduction Method

The rough set concept was originally proposed by Polish scientist Pawlak [66] in 1982 to deal with data duplication and redundancy. It is a mathematical tool for processing all kinds of data by solving the upper and lower approximate relations. The attribute reduction of a rough set refers to the identification of key information and the removal of redundant information by specific calculation methods without changing the existing decision-making attributes.

The traditional rough set needs to use the discretization method to divide the data before the attribute reduction, so as to achieve the discretization state. However, there are many discretization methods at present, and there is no recognized best processing method for discretization. At the same time, improper processing or external interference in the process of discretization of data may cause data distortion and change the nature of the original data, thereby bringing wrong judgments to decision makers. In order to make up for the shortcomings of the classic rough set, Hu [67] introduced the neighborhood model into rough set to form the neighborhood rough set theory. Neighborhood rough set is proposed to overcome the problem that traditional rough set reduction methods rely on discretized data. Therefore, the method of neighborhood rough set is used to reduce the dimensionality of the original index system in this paper. The related definitions are as follows [68]:

Definition 1 (*Measurement*)**.** *Given the universe* $U = \{x_1, x_2, \cdots, x_n\}$, *if for any* $x_i, x_j, x_k \in U$, *there are unique real functions* Δ *corresponding to them, and* Δ *satisfies:*

1. *Non-negative.* $\Delta(x_i, x_j) \geq 0$, *if and only if* $x_i = x_j$, $\Delta(x_i, x_j) = 0$.

2. 　Symmetry. $\Delta(x_i, x_j) = \Delta(x_j, x_i)$.
3. 　Triangle inequality. $\Delta(x_i, x_k) \leq \Delta(x_i, x_j) + \Delta(x_j, x_k)$.

　　Δ is called the metric or distance function, and $< U, \Delta >$ is called the metric space.

Definition 2 *(Neighborhood). For any metric space $< U, \Delta >$ and any U, the neighborhood δ of any x_i can be expressed as:*

$$\delta(x_i) = \left\{ x \middle| x \in U, \Delta(x_i, x_j) \leq \delta \right\}, \delta \geq 0 \tag{1}$$

　　Here the formula of the neighborhood radius is $\delta(x_i) = Std(a_i)/\lambda$; $Std(a_i)$ is the standard deviation of the indicator attribute a_i; λ is a characteristic parameter, which is set according to the difference in the accuracy of the final reduction. By adjusting and setting the size of λ, the size of the neighborhood $\delta(x_i)$ of each indicator attribute can be changed, and the value is generally [2,4].

Definition 3 *(Upper and lower approximation). In the neighborhood rough set decision-making system NDS = $(U, A \cup D)$, given the universe $U = \{x_1, x_2, \ldots, x_n\}$ and the neighborhood relationship N on it, that is, two-tuple NS = (U, N), for any $X \subseteq U$, the lower approximation of X in the neighborhood approximation space NS = (U, N) is:*

$$NX = \{x_i | \delta(x_i)X, x_i \in U\} \tag{2}$$

　　At the same time, the lower approximation NX of X is also defined as the positive domain, denoted as $Pos(X) = NX$.

Definition 4 *(Neighborhood dependence). In the decision system $S = (U, C \cup D)$, $c \in C$, the importance of conditional attribute (indicator) c is defined as: $Sig(c) = \gamma_C(D) - \gamma_{C-\{c\}}(D)$, where the calculation formula of the dependency γ_C of the conditional attribute set C relative to the decision attribute D is:*

$$\gamma_C(D) = \frac{CARD(POS_C D)}{CARD(U)} \tag{3}$$

Definition 5 *(Knowledge reduction). Given a knowledge base $S = (U, R)$ and a set of equivalent relations $P \subseteq R$ on it, for any $G \subseteq P$, if G satisfies the following conditions:*
(1)　G is independent, that is, every element in G is indispensable.
(2)　$IND(G) = IND(P)$, so that the positive domains of G and P relative to the decision attribute are the same, that is, $pos_G D = pos_P D$, then G is a reduction of P, denoted as $G \in Red(P)$.

　　At present, scholars have gradually increased their research on the attribute reduction algorithm of neighborhood rough set. The attribute reduction algorithm has been optimized to form a rich system, and it has also been extended to the integration with other machine learning methods such as Bayesian networks, neural networks, and support vector machines. The attribute reduction mechanism of the neighborhood rough set is similar to the traditional rough set attribute reduction method. The commonly used method is the attribute reduction method based on importance, which is calculated essentially based on the positive domain of the decision attribute relative to the conditional attribute (that is, the lower approximation). In the process of attribute reduction, if the positive range of an indicator in the assessment indicator system is higher, the removal of the indicator from the constructed assessment indicator system will have a greater impact on the entire assessment system, and the indicator will be given a higher degree of importance. What's

more, the limit of the importance calculated in the attribute reduction of the assessment indicator is 0. In fact, the importance of 0 does not exist, because the constructed indicators have their specific theoretical basis and will have some impact on the overall indicator system to a certain extent. Therefore, choosing a suitable attribute reduction method is very important for the overall assessment of the indicator.

In this paper, the forward greedy algorithm of neighborhood rough set is used for attribute reduction. It retains the indicator with the largest attribute importance and starts to select backwards. Also, it ensures that the core is not reduced. The main algorithm steps are as follows:

Step 1: Generate decision tables and determine the values of various model parameters, according to the data set decision system $NDS = (U, A \cup D)$, including discretized decision attributes.

Step 2: Calculate and search the neighborhood radius. The optimal neighborhood radius of the conditional attributes of each subsystem is determined respectively, and the neighborhood set of the sample is obtained separately according to the definition of the neighborhood radius. The size of the neighborhood radius is calculated based on the standard deviation of the attribute samples and the relative number of neighborhood parameters. The samples under the same neighborhood radius as sought are regarded as the same-attribute neighborhood set.

Step 3: Calculate the upper and lower approximate set, that is, calculate the upper and lower approximations of the decision attribute set relative to the condition attribute set. The lower approximation set is also the positive domain of the neighborhood rough set.

Step 4: Calculate the dependency of the decision attribute in each condition attribute subset by the positive domain and calculate the importance of each decision attribute relative to each condition attribute according to the importance solving formula.

Step 5: Get the reduction set. The samples whose attribute importance degree exceeds the set appropriate lower limit of importance degree are taken as the final attribute reduction set, so as to obtain a satisfactory solution.

2.3.3. Random Forest Models

Random forest is a relatively novel machine learning model, which has certain applications in many fields [69,70]. The algorithm is a classification regression model composed of multiple decision trees proposed by Breiman [71]. It is constructed by a certain sampling method such as bootstrap sampling method, based on the classification decision tree without pruning constructed by the categorical and regression trees (CART) algorithm, which has higher model accuracy and adaptability compared with the decision tree. The final output of random forest regression uses the average of the output results of a single tree (for the regression model) or the classification result with the most votes obtained by the "simple majority voting method" (for the classification model).

The randomness of random forest is mainly reflected in the randomness of sample extraction and variable extraction. In terms of the randomness of sample extraction, the random forest model randomly selects a certain number of samples from the samples to be trained. Then, it takes the extracted sample as the root node of the tree and construct a corresponding number of decision tree models based on this, and then model training is conducted on this basis. The randomness of variable extraction refers to using the random sampling method to select variables to be extracted that will be used as the split node of the tree when each decision tree model is constructed. When the random forest model is used for data regression prediction, it is necessary to construct a regression model and obtain the output result of a single tree by averaging. Specific steps are as follows:

Step 1: Randomly select a decision tree with a number of k. The bootstrap method is used to resample the original samples, thereby randomly generating k training sets $\theta_1, \theta_2, \theta_3 \dots \theta_k$, that is, the number of trees generated is k (that is, the value of parameter *ntree*). At the same time, each training set trained is used to generate the corresponding decision tree $\{Tr(x, \theta_1), Tr(x, \theta_2), Tr(x, \theta_3), \dots, Tr(x, \theta_k)\}$.

Step 2: Randomly extract the dimensional feature set with the number *m*, that is, randomly extract *m* features from the dimensional features with the indicator feature number *M* as the split feature set of the current node (that is, the value of parameter *mtry*), and use the standardized mean-square error as the standard to judge whether these *m* features follow the most appropriate split method to carry out splitting, so that the whole after splitting has the best stability.

Step 3: Calculate the observation value of a single tree. The prediction of a single decision tree is obtained by the weighted average of the dependent variables' observed values $Y_i(i = 1, 2, 3, \ldots n)$.

Step 4: Calculate the predicted value of the random forest. According to the weight of each decision tree $\omega_i(x, \theta_t)(t = 1, 2, 3 \ldots k)$, the mean value of the observation value of each decision tree is taken as the final result.

2.3.4. Integration of Neighborhood Rough Set and Random Forest Algorithm

Aiming at the solving the problems of multi-index data redundancy, inter-index correlation and duplicate information in the water resources vulnerability assessment system constructed in this paper, this paper proposes to use the neighborhood rough set method to reduce the dimensionality of the original indicator system to remove redundant attributes. For methods that can remove redundant attributes, there are some methods that can remove redundant attributes such as principal component analysis [31,32] and projection pursuit [29]. Although the use of these methods helps to achieve the dimensionality reduction effect of the indicators, they are not good for the subsequent prediction work. The reasons are as follows. These methods synthesize new indicator values with quantitative methods by combining certain indicators. Also, they do not consider the decision attributes and do not use conditional attributes to judge decision attributes, which may neglect the important information, further reduce the space [72,73]. Therefore, they are difficult to interpret the newly synthesized indicators and to integrate the data in different situations in the future. Also, it will also increase the workload of calculation, so it is not suitable for the study. The neighborhood rough set can effectively make up for these shortcomings [73]. Moreover, the random forest regression algorithm proposed in this paper has further advantages over multiple linear regression, support vector machine regression, neural network, decision tree and other methods, which can effectively improve the accuracy of model prediction and the effect of model construction. The following comparison of these machine learning methods in Section 3.3.2 also confirms this point well.

Therefore, this paper integrates neighborhood rough set and random forest model for the assessment and prediction of water resources vulnerability. Using the neighborhood rough set theory to reduce the dimensionality of the original assessment indicator system can objectively improve the efficiency of the random forest algorithm and the speed of model establishment, which is the basis for the training of the random forest regression model. What's more, the integration of the two methods can reduce the interference caused by unfavorable assessment indicators that may appear in the research of the paper, thereby improving the effect of the assessment and prediction of the water resources vulnerability.

The integration process of neighborhood rough set and random forest model is shown in Figure 2 and specific steps are as follows:

Step 1: Attribute reduction. After constructing the assessment indicator system of water resources vulnerability, we use the forward greedy algorithm of neighborhood rough set to reduce the dimensionality of the original indicator system to remove redundant attributes. It retains the indicators with the greatest attribute importance, and then starts to select it backwards, while also ensuring that the core is not reduced.

Step 2: Construction of water resources vulnerability assessment model. We take the standard value of the indicator level threshold of the assessment indicator system after dimensionality reduction as the input vector and the vulnerability level value as the output vector to construct a random forest model.

Step 3: Assessment of water resources vulnerability. The indicator data of the Song-Liao River Basin from 2000 to 2017 is substituted into the model to obtain the water resources vulnerability evaluation value in the past few years.

Step 4: Testing of the assessment model. We use the 10-fold cross-validation method to test the trained model to judge the reliability of the results. In order to verify whether the accuracy of the random forest regression model is better than other models in this paper, the neural network model, decision tree and support vector machine regression model with excellent nonlinear regression function are compared with it.

Step 5: Scenario prediction of water resources vulnerability. The indicator data under different scenarios in 2025 and 2030 is substituted into the random forest model as input data to obtain the predicted value of water resources vulnerability, so as to provide reference for future water resources planning and adaptive management.
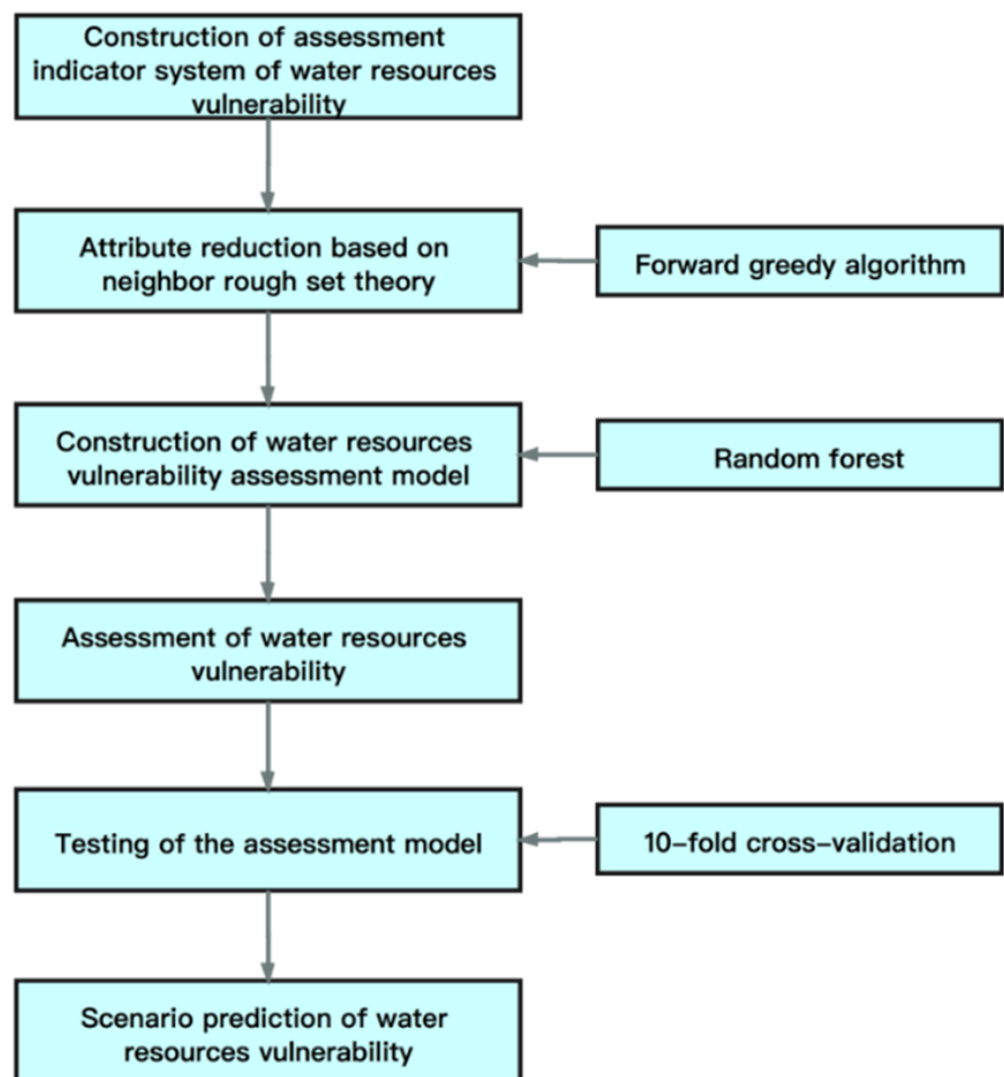

Figure 2. Water resources vulnerability assessment model based on the integration of neighborhood rough set and random forest model.

3. Results and Discussion

3.1. Attribute Reduction of Evaluation Indicators

3.1.1. Correlation Analysis of Evaluation Indicators

From the above process of establishing the water resources vulnerability assessment indicator system, it can be seen that there are many factors affecting the vulnerability of water resources, and some indicators have greater redundancy and repetition. In order

to facilitate the analysis of the repeatability of the indicators, this paper analyzes the correlation between the indicators under each first-level indicator. Each indicator includes 18 years of data in the Song-Liao River Basin from 2000 to 2017. See Figure 3 for the analysis results. Specifically, the correlation coefficient diagram of natural vulnerability is shown in Figure 3a. Most absolute values of the correlation coefficient between indicators are not higher than 0.5. The correlation coefficient between A_1 and A_8 is up to 0.9, and the repeatability between the two is relatively large. The correlation coefficient diagram of man-made vulnerability is shown in Figure 3b. Among them, the correlation coefficient between B_5 and B_7 is 0.8, which is the strongest. In addition, the correlations between indicator B_3 and other indicators are relatively large. The absolute value of the correlation coefficient between B_3 and B_5, B_3 and B_6 are all greater than 0.7. The correlations between other indicators are not so strong. The correlation coefficient diagram of vulnerability of carrying capacity is shown in Figure 3c. Compared with the indicators in the first two aspects, the redundancy between the indicators in the aspect of the vulnerability of carrying capacity is significantly greater. Among them, the absolute values of the correlation coefficient between C_2 and other indicators are all not less than 0.3. In summary, there are more or less correlations among various indicators. At the same time, the definition and connotation of each specific indicator and the actual situation of the evaluation industry explain that there is duplication of information in the indicators initially constructed in the article. This information may have an impact on the final prediction of water resources vulnerability of the river basin. Therefore, it is very important to choose a suitable method to eliminate the impact of the constructed index as much as possible. On the basis of the indicator structure and classification ability unchanged, this article uses the neighborhood rough set theory to reduce the attributes of the remaining indicators for the important indicators selected by the qualitative analysis, which is more conducive to the subsequent assessment and prediction work.
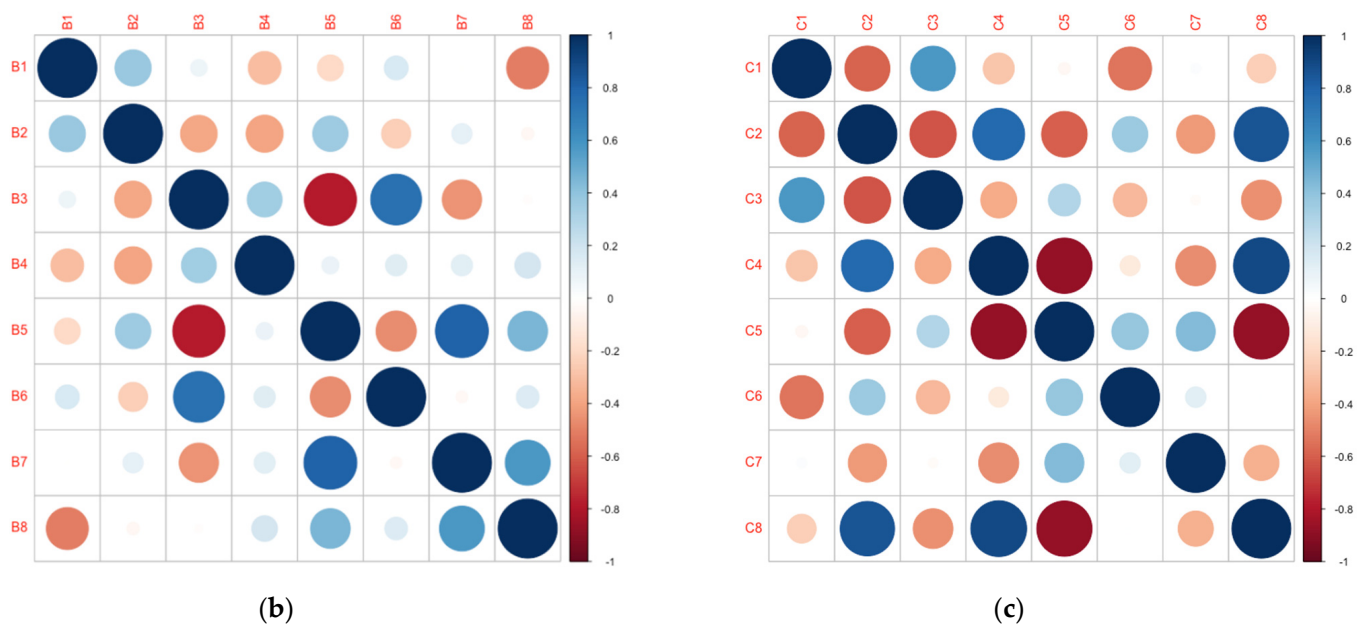

(**a**)

Figure 3. *Cont.*

(b)

(c)
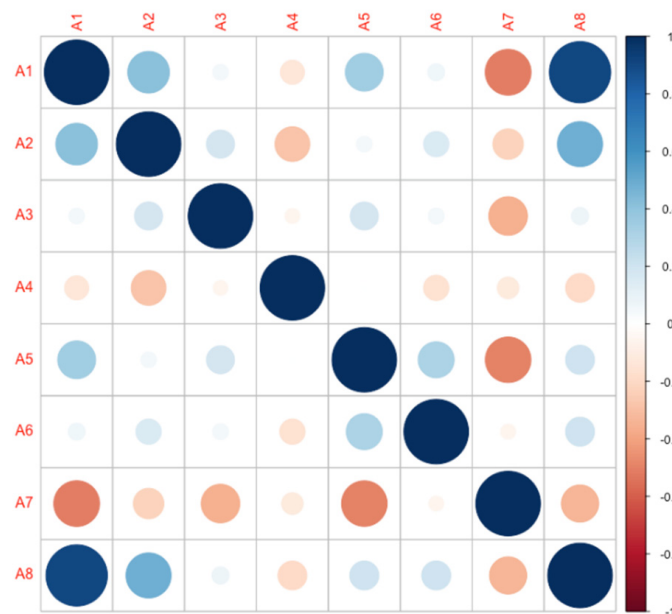
Figure 3. Correlation coefficient diagram of the original index system: (**a**) correlation coefficient diagram of natural vulnerability; (**b**) correlation coefficient diagram of man-made vulnerability; (**c**) correlation coefficient diagram of vulnerability of carrying capacity.

3.1.2. Determination of Decision-Making Attributes

When we use the neighborhood rough set to reduce the attributes of an assessment indicator system, the collected raw data does not need to be discretized, and can be directly used as conditional attributes. Due to the lack of decision-making attributes, we need to use linear weighted method to determine the comprehensive evaluation value of natural vulnerability, man-made vulnerability and vulnerability of carrying capacity. And then, the discretized comprehensive evaluation values are used as decision attributes to construct a neighborhood rough set decision system NDS. The discretization process is to transform the original continuous data into several "segments". In this way, continuous data is converted into discrete data to reduce the complexity of continuous data and reduce the influence of extreme values on the data structure. The commonly used discretization methods currently used in rough set methods include equal frequency division discrete method, equal distance division discrete method, self-organizing neural network method, discrete method based on information entropy, etc. This paper uses k-means clustering discretization method for discretization, which is an unsupervised discretization method. It is carried out by the k-means clustering method, which can overcome the shortcomings of the isometric method and the equal frequency method that do not consider the data characteristics and related attributes of the index to a certain extent. The calculation is relatively simple, and it is also suitable for the actual situation with a small amount of sample data, so it is more suitable for this paper.

Among them, this paper uses the game weighting method [74] to determine the constructed water resource vulnerability indicator weights and uses the calculated weights to calculate the comprehensive evaluation value of natural vulnerability, man-made vulnerability and vulnerability of carrying capacity. The calculation steps are as follows:

Step 1: Perform non-dimensional standardization on the original data. The non-dimensional standardization processing formula of the positive indicator is as follows:

$$X_{ij} = \frac{x_{ij} - minx_{ij}}{maxx_{ij} - minx_{ij}} \tag{4}$$

The non-dimensional standardization processing formula of the negative indicator is as follows:

$$X_{ij} = \frac{maxx_{ij} - x_{ij}}{maxx_{ij} - minx_{ij}} \tag{5}$$

In the formula, x_{ij} represents the original value of the *j*-th indicator in the *i*-th year; $maxx_{ij}$ represents the maximum value of the *j*-th indicator; $minx_{ij}$ represents the minimum value of the *j*-th indicator.

Step 2: Determine the weight of the entropy method [75]. First of all, we calculate the information entropy E_j of the *j*-th index:

$$E_j = -\frac{1}{\ln(n)} \sum_{i=1}^{j} (f_{ij} ln f_{ij}) \tag{6}$$

In this formula, $f_{ij} = X_{ij} / \sum_{i=1}^{j} X_{ij}$; when $f_{ij} = 0$, $f_{ij} ln f_{ij} = 0$. Then, the entropy weight of the *j*-th index $w1_j$ is:

$$w1_j = \frac{1 - E_j}{\sum_{j=1}^{n} (1 - E_j)} = \frac{D_j}{\sum_{j=1}^{n} (D_j)} \tag{7}$$

In this formula, D_j is the indicator difference degree, and $D_j = 1 - E_j$.

Step 3: Determine the weight of the CRITIC (criteria importance though intercriteria correlation) law [76]. First of all, we quantitatively calculate the information amount C_i of the assessment indicator:

$$C_i = S_j * R_j = \sqrt{\frac{\sum_{i=1}^{n} \left(X_{ij} - \frac{1}{n} \sum_{i=1}^{n} X_{ij}\right)}{n - 1}} * \sum_{t=1}^{n} (1 - r_{rj}) \tag{8}$$

Here, r_{rj} represents the correlation coefficient between the assessment indicator *i* and *j*; S_j is the contrast strength between the *j*-th indicator and other indicators, that is, the standard deviation of the *j*-th indicator; $R_j = \sum_{t=1}^{n} (1 - r_{rj})$ is the conflicting quantitative indicator between the *j*-th indicator and other indicators. Then, the weight of the CRITIC law of the *j*-th index $w2_j$ is:

$$w2_j = \frac{C_j}{\sum_{j=1}^{m} C_i} \tag{9}$$

Step 4: Determine the comprehensive weight of the game theory method. This paper uses the game theory method to integrate the entropy method and the CRITIC method to comprehensively determine the weight of the indicator. At the same time, the relevance, dispersion and relative strength of the indicator data information have also been fully tested. Thereby, the weighting result tends to be balanced, and the scientificity of the indicator weight is improved. The steps for determining the comprehensive weight of the game theory method are as follows:

Firstly, the entropy method and CRITIC method are used to weight the indicators respectively, and a basic weight vector set $w_k = \{w_{k1}, w_{k2}\}$ is constructed. The arbitrary linear combination between the above 2 different vectors is:

$$w = \sum_{k=1}^{2} \alpha_k w_k^T \left(\alpha_k > 0, \sum_{k=1}^{2} \alpha_k = 1\right) \tag{10}$$

where w is a possible weight vector in the basic weight vector set; α_k is the linear combination coefficient. Then, we use game theory to optimize the 2 linear combination coefficients α_k, so that the deviation between w and each w_{km} is the smallest, namely:

$$min\|\sum_{j=1}^{2}\alpha_j w_j^T - w_i\|_2(i = 1, 2) \tag{11}$$

The optimal first derivative condition of the above formula can be converted into the following formula:

$$\begin{bmatrix} w_1 w_1^T & w_1 w_2^T \\ w_2 w_1^T & w_2 w_2^T \end{bmatrix}\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} w_1 w_1^T \\ w_2 w_2^T \end{bmatrix} \tag{12}$$

Finally, according to the above process, (α_1, α_2) is obtained and then normalized, that is, the objective weight w_j of the game of the j-th index is calculated:

$$w_j = \frac{|\alpha_k|}{\sum_{k=1}^{2}|\alpha_k|} \tag{13}$$

Step 5: Use the standardized values of related indicators and the weights of each indicator in the water resources vulnerability assessment indicator system to calculate the final comprehensive evaluation value of vulnerability in the three aspects of natural vulnerability, man-made vulnerability and vulnerability of carrying capacity. The calculation formula is as follows:

$$Z_i = \sum_{j=1}^{n}(w_j X_{ij}) \tag{14}$$

3.1.3. Attribute Reduction of Indicators

Through the abovementioned forward greedy algorithm, the original data is reduced by attributes. After many parameter adjustments, the calculated 3-dimensional set is the smallest reduction of the original decision table, that is, the reduction set of natural vulnerability $A = \{A_3, A_4, A_7\}$, the reduction set of man-made vulnerability $B = \{B_2, B_6, B_7, B_8\}$, the reduction set of vulnerability of carrying capacity $C = \{C_1, C_3\}$. After inspection, they all meet the definition requirements of the above reduction.

The result calculated by the forward greedy algorithm satisfies Definition 5. The above minimum reduction set represents the processing result of the collected indicator data from the perspective of data mining of the neighborhood rough set, but it cannot fully consider the actual situation of the area, and some indicators that are important for the assessment of water resource vulnerability may be deleted by mistake. This article's assessment of the water resources vulnerability of the river basin not only includes the overall status of the river basin, but also includes the three aspects of natural vulnerability, man-made vulnerability and vulnerability of carrying capacity. In order to facilitate the assessment of the vulnerability of water resources in the three aspects, A_1 is added to the set A, and C_5, C_6 are added to the set C. Finally, a reduced indicator system is formed, as shown in Table 3 below, which serves as the basis for the random forest regression model. In addition, according to the correlation analysis of evaluation indicators in the Section 3.1.1, the absolute values of the correlation coefficient between indicators are controlled below 0.6 after the reduction, that is, the redundancy of the indicator system is further reduced.

Table 3. Indicator collection after reduction.

First-Level Indicator	Third Level Indicator
Natural vulnerability	Water production modulus A_1 Change rate of annual precipitation A_3 Water quality examination pass rate in water function area A_4 Proportion of area affected by flood and drought A_7
Man-made vulnerability	Proportion of groundwater resources being utilized B_2 Proportion of population under levee protection B_6 Proportion of soil erosion being controlled B_7 Water conservancy project storage capacity B_8
Vulnerability of carrying capacity	Ratio of groundwater supply to total water supply C_1 Water consumption for irrigation per mu C_3 Wastewater generation per 10,000-yuan GDP C_5 Ecosystem water consumption rate C_6

3.2. Construction of Water Resources Vulnerability Assessment Model

3.2.1. Interpolation of Regression Samples

We take the standard values of the indicator level threshold of the assessment indicator system after dimensionality reduction as the input vector and the vulnerability level values as the output vector to construct a random forest model. Among them, the average value of the thresholds on both sides of the level range is used as the standard value of each indicator. In this way, the standard value matrix of each indicator corresponding to seven levels is calculated, which has dimensions of 7 rows (levels) × 12 attributes. Then, we use the standard value matrix as the input data for the random forest regression model calculation. Furthermore, a vector composed of seven vulnerability level values is used as the output vector.

However, considering that the size of standard value matrix is small, it is likely to have an adverse effect on the accuracy of the regression. Therefore, it is necessary to eliminate this effect and expand the matrix size. Then, we use the two-dimensional cubic interpolation method to complete by the function *interp2* of the MATLAB software. Finally, the standard value matrix which has dimensions of 7 rows (levels) × 12 attributes is expanded to a matrix whose dimensions are 100 rows (levels) × 12 attributes. Also, the output vector is expanded to a 100-dimensional vector. And then, the interpolated matrix is used as the input data and the interpolated vector is used as output data of the random forest model regression model.

3.2.2. Construction of Random Forest Regression

This paper uses the random forest (randomForest package) in the R language software for sample training. In order to better determine the tree and other parameters of the random forest, the existing sample set is randomly divided into two categories. The first category is used as the training set, and the second category is used as the test set. Then, we test the accuracy of a random forest with *n* trees, where *n* takes all integers from 0 to 500. It can be seen from Figure 4a that the higher the number of trees is, the smaller the error between the test set and the training set is. Therefore, the number of *ntree* in the random forest is set to 500. Figure 4b shows the relationship between the number of trees trained in the random forest and the error outside the bag. It can be seen that the higher the number of *ntree* is, the smaller the error of the model is and the more stable it becomes.

For the number of *mtry*, the traversal method is adopted. By setting the number of *mtry* in the interval [2,8], the NMSE of the training set and the test set is obtained as the *mtry* selection criterion. The calculation results are shown in Table 4. It can be seen that the accuracy of the random forest model is the best when the number of *mtry* reaches 4. Therefore, the number of parameters *mtry* is set to 4, and the random forest model is trained by the parameters set above.
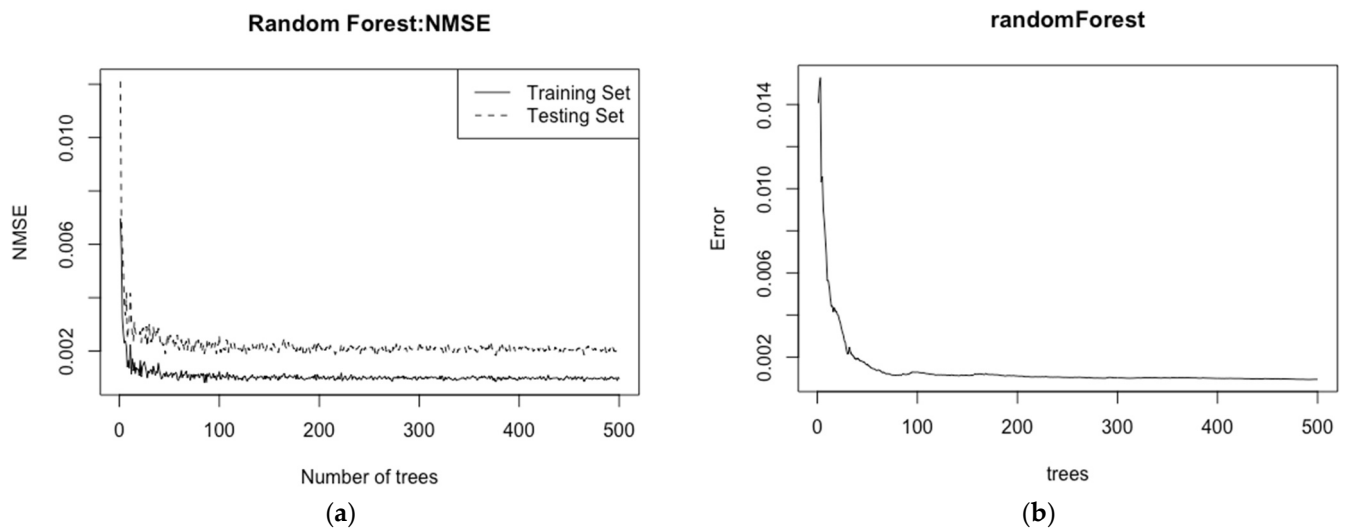
Figure 4. Random forest model training image: (**a**) relationship graph between model NMSE and *ntree*; (**b**) relationship graph between model error and *ntree*.

Table 4. The accuracy of the test set and the training set under each different number of *mtry*.

Number of *mtry*	*NMSE* of Training Set	*NMSE* of Test Set
2	0.001004	0.005345
3	0.000937	0.005332
4	0.000896	0.005255
5	0.000977	0.005369
6	0.000927	0.005399
7	0.000927	0.005318
8	0.00096	0.00506

In order to verify whether the random forest regression model is better than other models in the accuracy of the research in this article, we will compare the neural network model (net package), decision tree (rpart package) and support vector machine regression model (e1071 package) with excellent nonlinear regression functions. After setting and debugging the appropriate parameters of each model, the training of the model is carried out.

For the test of the trained model, this paper uses the 10-fold cross-validation method to judge the reliability of the result. We will use 10 randomly selected samples as 10 training sets, and take turns using nine of them as training data and one as test data for testing, so that each sample can be used as a test set. Finally, the average value of the mean square error (*MSE*), normalized mean square error (*NMSE*) and coefficient of determination (*R-squared* or R^2) of all test results is obtained. The related error test formula is as follows [77]:

$$MSE = \frac{1}{n} \sum \left(\hat{y}_i - y_i \right)^2 \tag{15}$$

$$NMSE = \frac{\sum \left(y_i - \hat{y}_i \right)^2}{\sum \left(y_i - \overline{y}_i \right)^2} \tag{16}$$

$$R^2 = 1 - \frac{\sum_i \left(\hat{y}_i - y_i \right)^2}{\sum_i \left(\hat{y}_i - \overline{y}_i \right)^2} \tag{17}$$

Results of each error test obtained by the above method are shown in Table 5 above. It can be seen from the table that the selected regression models have good training accuracy, and the accuracy of random forest regression is higher than that of neural network models,

support vector machine and decision trees. Therefore, this paper uses the random forest regression model as the water resource vulnerability assessment and prediction model of the Song-Liao River Basin.

Table 5. Regression training accuracy of each model after attribute reduction.

Methods	MSE	NMSE	R-Squared [1]
Random forest	0.0001529336	1.738314×10^{-8}	0.9999968
Decision tree	0.00013895	0.01205421	0.9999924
Support vector machine	0.001653784	1.620705×10^{-6}	0.9999991
Neural network	0.8765589	1.414752	0.9969101

[1] The greater the R-squared is, the greater the precision of model is.

3.3. Assessment of Water Resources Vulnerability

3.3.1. Calculation of Current Situation of Water Resources Vulnerability

We use all indicator data on the natural vulnerability of the Song-Liao River Basin from 2000 to 2017 as input data. The random forest model trained by the above software package is used to obtain the comprehensive evaluation value of natural vulnerability in the past few years. The same operation is applied to the calculation of the vulnerability value of man-made vulnerability and vulnerability of carrying capacity. Then, we use the original indicator data of the Song-Liao River Basin from 2000 to 2017 as the input data and use the random forest model trained by the above software package to obtain the comprehensive evaluation value of the water resource vulnerability, as well as comprehensive evaluation value of natural vulnerability, man-made vulnerability and vulnerability of carrying capacity in the past few years. Then we rounded the relevant vulnerability data to get the relevant grade value. The calculation results are shown in Table 6 below.

Table 6. Evaluation results of water resources vulnerability in the Song-Liao River Basin.

	Vulnerability Value (0–7)				Vulnerability Grade (I–VII)			
	Basin Vulnerability	Natural Vulnerability	Man-Made Vulnerability	Vulnerability of Carrying Capacity	Basin Vulnerability	Natural Vulnerability	Man-Made Vulnerability	Vulnerability of Carrying Capacity
2000	4.6439	3.8538	5.0428	4.0412	V	IV	V	IV
2001	4.7345	3.8520	5.1176	4.0910	V	IV	V	IV
2002	4.7350	3.8148	4.9121	4.2335	V	IV	V	IV
2003	4.9053	3.7021	4.6849	5.7085	V	IV	V	VI
2004	4.9896	3.8451	4.7396	5.6695	V	IV	V	VI
2005	4.8879	3.8350	4.6024	5.4638	V	IV	V	V
2006	4.8516	3.7702	4.5788	5.4150	V	IV	V	V
2007	4.8796	3.9211	5.1552	5.2339	V	IV	V	V
2008	4.8247	3.8551	4.9260	5.1380	V	IV	V	V
2009	4.9444	3.8665	4.9004	5.3324	V	IV	V	V
2010	4.5139	3.7642	4.5418	4.8695	V	IV	V	V
2011	4.6983	4.0618	5.2211	4.2512	V	IV	V	IV
2012	4.5862	4.4266	4.6821	4.2025	V	IV	V	IV
2013	4.6497	3.5852	5.1762	4.5860	V	IV	V	V
2014	4.8027	4.0528	5.2035	4.6551	V	IV	V	IV
2015	4.6008	3.7439	5.0740	4.1847	V	IV	V	IV
2016	4.5366	3.7731	4.8101	4.1949	V	IV	V	IV
2017	4.7030	3.8582	5.1625	4.3442	V	IV	V	IV

3.3.2. Current Situation Evaluation of River Basin's Water Resources Vulnerability

As can be seen from Figure 5 below, the Song-Liao River Basin is relatively fragile as a whole, basically in the level of moderate to high vulnerability. From a numerical point of view, the water resources vulnerability of the river basin has experienced a trend of first deterioration and then improvement from 2000 to 2017. However, its vulnerability level

has been maintained at the level V, which shows that the overall water vulnerability of the river basin is still at a relatively poor level.
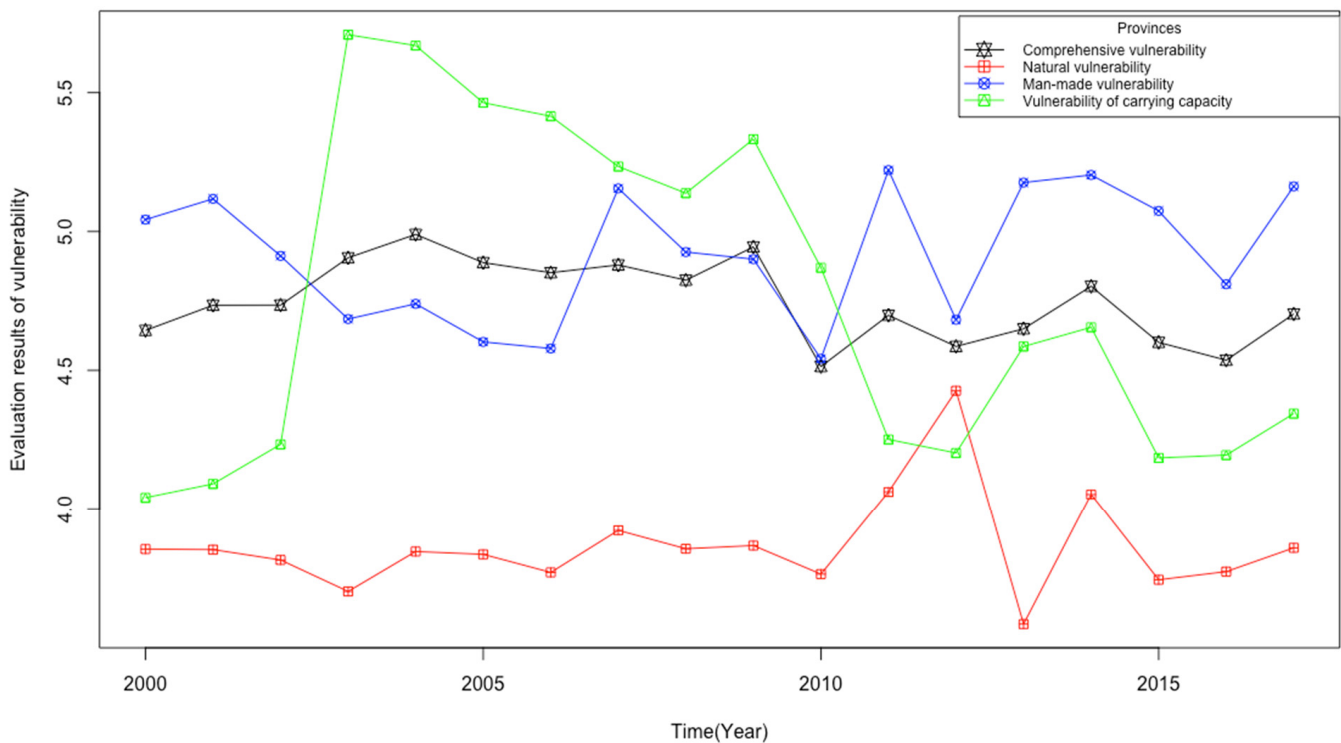

Figure 5. Evaluation results of water resources vulnerability in Song-Liao River Basin.

In terms of natural vulnerability, the water resource vulnerability of the Song-Liao River Basin is more serious, but the vulnerability is better compared with other aspects of the basin. According to relevant data, the water production modulus of the Song-Liao River Basin has been at a relatively poor level since 2000, with an average value of only 14.36 m^3/km^2, and some years have large fluctuations. There is a big gap between the Song-Liao River Basin and other water-producing areas, such as the Huai River Basin whose average water production modulus is 320,000 m^3/km^2. In the Song-Liao River Basin, basin itself has a vast area, the total amount of water resources is relatively abundant and there is a large gap between regions, which is an important factor that is likely to cause the uncertainty of water resources' natural vulnerability. At the same time, in some areas, precipitation fluctuates greatly in some years, causing the basin to be affected by external natural climatic conditions. In addition, the imbalance of precipitation in the basin, especially the water shortage in inland areas, cannot be ignored, which will objectively aggravate the natural vulnerability of the Song-Liao basin. Additionally, in terms of water quality, the situation of regional water resources pollution has improved recently, but the extent is not very large. From the water quality qualified rate of the river basin, the relevant rate of Song-Liao River Basin in 2017 has increased by 2.96% compared with the previous year. Therefore, there is room for improvement in the future.

In terms of man-made vulnerability, the vulnerability of water resources in the Song-Liao River Basin is extremely high. From 2000 to 2017, the grade of water vulnerability basically stagnated to the level V of moderate to high vulnerability. Due to the lack of water resources in parts of the Song-Liao River Basin, there has been overexploitation in water resources exploitation, especially groundwater. The average proportion of groundwater resources being utilized from 2011 to 2017 was 48.49%, and it reached 51.98% in some years, which is a problem that cannot be ignored. The ratio of effectively irrigated farmland area is also steadily increasing in view of the transformation from extensive water use to intensive water use. In terms of environmental pollution emissions, the Song-Liao River Basin has a

trend of improvement compared with the previous year. According to the "China Environmental Statistics Yearbook", we can see that the main indicators of industrial pollution include per capita COD emissions and per capita ammonia and nitrogen emissions are gradually decreasing. In particular, COD emissions per capita in 2017 were reduced by about 32% compared to 2000, the initial year of the evaluation. In terms of water conservancy project construction, although there is a large investment every year, the effect is not very obvious. It can be seen from the three aspects of water conservancy regulation and storage capacity, water and soil erosion control, and embankment protection population construction that the relevant indicators are basically the same as those of the original year, which shows that the man-made vulnerability of water resources in the Song-Liao River Basin is difficult to manage, and the governance effect is not obvious. What's more, the discharge of agricultural and domestic sewage is also a factor that cannot be ignored in the Song-Liao River Basin in terms of water quality. Also, this is an aspect needed to be paid attention to in the future to reduce the water vulnerability of the river basin.

In terms of vulnerability of carrying capacity, the overall water resource vulnerability of the Song-Liao River Basin has deteriorated first and then improved, and it is developing in a better direction. At the same time, the vulnerability of carrying capacity has improved the most compared to other vulnerabilities, whose grade has been changed from the past level V of moderate to high vulnerability to the level IV of moderate vulnerability. As the population in the Song-Liao River Basin has grown at a steady rate each year in recent years, the carrying pressure in the basin has also increased. On the other hand, with the advancement of active and effective measures in the basin, the pressure on the Song-Liao basin in other fields is also decreasing. Regarding the pressure on the amount of water resources, the ratio of the supply of groundwater resources to the total water supply is slowly decreasing, and it is expected to decrease to 45% in the next few years with the current trend. Due to the uncertainty of annual precipitation, the water carrying pressure in each year is also fluctuating. But from the trend of the population per 10,000 cubic meters of water, it is likely to further decline in the future. Therefore, there is much room for improvement in the vulnerability of river basin water resources.

3.4. Scenario Prediction and Analysis of River Basin's Water Resources Vulnerability

3.4.1. Raw Data under Different Scenarios in the Future

Setting different scenarios according to different levels of goals can more effectively control resources in advance, so as to prepare for advance control measures [78]. We set three scenarios—Scenario 1, Scenario 2, and Scenario 3—in the years of 2025 and 2030, respectively. At the same time, we collect original data under the three future scenarios and predict the future water vulnerability state of Song-Liao River Basin under these three scenarios. Recently, affected by human economic activities and lifestyles, forests have been continuously being cut down, cities have been expanding, and the stability of forest ecosystems have been declining [79]. This has further led to man-made greenhouse gas emissions reach the highest in current history, which has affected humanity and climate change in natural systems. All of the above factors in turn increase the uncertainty of water security and the vulnerability of water resources [80]. Therefore, the scenario data in scenario predictions that are greatly affected by the climate are collected based on the RCP4.5, RCP6.0 and RCP8.5 concentration emission scenarios set by the fifth IPCC assessment report to predict the future river runoff, precipitation and other data of the Song-Liao River Basin. We use this as the input data for the random forest prediction model of Scenario 1, Scenario 2, and Scenario 3 in the future forecast years.

In order to obtain other scenario data related to economic and social development, such as water efficiency, water conservancy projects, water pollution control, we need to first set the water resources management goals corresponding with the scenario indicators. The target data are based on the "Opinions of the State Council on Implementing the Strictest Water Resources Management System", "Comprehensive Planning of the Song-Liao River Basin" (2012–2030) and the forecast results of the future socioeconomic water of the Song-

Liao River Basin. We assume that Scenario 1, Scenario 2, and Scenario 3 respectively follow the principles of pessimism, mediocrity, and optimism, and they complete the progress according to the goal set based on a certain proportion. Among them, Scenario 1 represents that the most stringent water resources management system can be implemented in the water resources management of the river basin, and it can basically reach or exceed the target value of the relevant plan; Scenario 2 represents a medium management scenario, which can make the target variable reach 2/3 of the target value; Scenario 3 represents a weak management situation, which can make the target variable reach only 1/3 of the target value. Then, we use these as scenario data under 3 different management efforts in the future.

3.4.2. Forecast Result Analysis

The water resources vulnerability prediction results and related grades of the Song-Liao River Basin under different scenarios in the future can be obtained by inputting the raw data under the above 3 different scenarios into the trained random forest regression model, which is shown in Table 7 below.

Table 7. Vulnerability value and grade of water resources in different scenarios in the future.

	Vulnerability Value (0–7)				Vulnerability Grade (I–VII)			
	Basin Vulnerability	**Natural Vulnerability**	**Man-Made Vulnerability**	**Vulnerability of Carrying Capacity**	**Basin Vulnerability**	**Natural Vulnerability**	**Man-Made Vulnerability**	**Vulnerability of Carrying Capacity**
2025 S1 [1]	3.9930	2.6677	4.4839	4.1512	IV	III	IV	IV
2025 S2 [1]	4.5857	3.6436	4.8763	4.6821	V	IV	V	V
2025 S3 [1]	5.0388	3.8908	5.1832	5.4807	V	IV	V	V
2030 S1	3.0448	2.3793	3.5221	2.3116	III	III	IV	III
2030 S2	3.7156	3.3894	4.0384	3.4934	IV	IV	IV	III
2030 S3	4.3642	3.8460	4.2620	4.2809	IV	IV	IV	IV

[1] S1, S2 and S3 are short for Scenario 1, Scenario 2 and Scenario 3.

The above prediction results show that the vulnerability of water resources in the future will be greatly improved compared with the previous evaluation years, and the vulnerability prediction results of the three scenarios in 2030 are better than those in 2025. Additionally, the water resources vulnerability situation under Scenario 1 is the best, and the water resource vulnerability grade will reach the level III is the level of moderate to low vulnerability in 2030. Also, Scenario 2 is only in Scenario 1, and the prediction result of water vulnerability under Scenario 3 is the worst. This shows that if the Song-Liao River Basin can implement water resources management and control measures in accordance with relevant water resources management systems in the future. Scenario 1 can effectively change the current deterioration situation of the water resources vulnerability in the Song-Liao River Basin, while water resources vulnerability in Scenarios 2 and 3 increases slowly. The governance effect of Scenario 2 and Scenario 3 in 2030 is not as good as the governance effect of Scenario 1 in 2025, and the vulnerability of water resources in the basin will further deteriorate.

In terms of natural vulnerability, the natural vulnerability under scenario 1 will be raised to the level III of moderate to low vulnerability, which has a trend of further improvement compared to previous years. In terms of man-made vulnerability, improvement under the three different scenarios is relatively slow. In 2025, the vulnerability of the three scenarios will be level IV and V. It will be stabilized in a level IV of moderate vulnerability until 2030. This will also become one of the difficulties in managing the vulnerability of

water resources in the Song-Liao River Basin in the future. In terms of vulnerability of carrying capacity, the water resource vulnerability state of Scenario 1 will be significantly faster than the state of Scenario 2 and Scenario 3. It can be seen from the above that if favorable measures can be taken, such as the construction of water conservancy projects, it can effectively reduce the risk of future disasters and further enhance the security of water resources in the river basin.

4. Conclusions

This paper uses the forward greedy algorithm of the neighborhood rough set method to reduce the dimensionality of the original indicator system to remove redundant attributes, which is based on the constructed indicator system of water resources vulnerability assessment in Song-Liao River Basin. Then, assessment indicators' standard values after dimensionality reduction are taken as the evaluation sample, and the random forest regression (abbreviated as RF) model is used to assess the water resources vulnerability of the river basin. Finally, based on data under three different climate and socio-economic scenarios in the future, scenario predictions are made on the vulnerability of future water resources.

We apply the above methods to carry out an empirical analysis of the Song-Liao River Basin and obtain the current situation and prediction results of the water resources vulnerability assessment. First of all, from the perspective of the current assessment, the water resource vulnerability of the Song-Liao River Basin is relatively poor, and it is currently at the level V of moderate to high vulnerability. In particular, natural vulnerability and man-made vulnerability improve relatively slowly, which are at the level IV of moderate vulnerability and the level V of moderate to high vulnerability, respectively. The vulnerability of carrying capacity improves relatively more obviously, and it has improved from the level V of moderate to high vulnerability to the level IV of moderate vulnerability. Secondly, judging from the forecast results, the vulnerability of water resources in Scenario 1 will be significantly improved in the future. However, the progress of water vulnerability in Scenario 2 and Scenario 3 is very slow and may even deteriorate. It can be seen that whether the most stringent water resources management system and related regulations can be well implemented in the future is an important factor in determining whether the vulnerability of water resources can be improved.

This article has certain innovations and characteristics in indicator selection, assessment model and scenario prediction. First, in terms of indicator selection, some research indicators are not well-targeted, and the influence of correlation between the indicators and information redundancy on the accuracy of the model are not well considered, which will cause interference with the accuracy of the assessment results. On the basis of the indicator structure and classification ability unchanged, this article uses the neighborhood rough set theory to reduce the attributes of the remaining indicators, which is more conducive to the subsequent assessment and prediction work. In addition, compared with the traditional rough set, the neighborhood rough set does not depend on the discretized data, and there is no need to use the discretization method to divide the data. Moreover, as a basis for the training of the random forest regression model, it can objectively improve the efficiency of the random forest algorithm and the speed of model establishment. Second, in the assessment model, this paper compares multiple regression models with excellent nonlinear regression functions. We train the random forest, neural network model, decision tree and support vector machine regression model after setting and debugging the parameters properly. After judging the reliability of the results by using the 10-fold cross-validation method, it is found that the accuracy of the random forest regression model is higher than other models in this paper. Therefore, this paper selects the random forest regression model to assess the water vulnerability of Song-Liao River Basin. Third, the current prediction-related research is relatively small, which has caused the adjustment and control not to be prepared in time for early. Therefore, the future prediction of the vulnerability of water resources in the Song-Liao River Basin is worthy of attention. We refer to some

important documents and set up three different scenarios in accordance with the principles of pessimism, mediocrity and optimism. Then, we input the raw data under three different scenarios into the random forest regression model with better training effect to study the vulnerability of water resources.

Restricted by data search, the assessment and prediction of the water vulnerability of the Song-Liao River Basin is limited to the basin's overall situation in a certain year. In the internal system of the Song-Liao River Basin, due to the different geography, climate, economic and social conditions of each internal zone, the assessment indicators applicable to each internal system will be different. In future studies, we will focus on exploring the vulnerability of water resources at a smaller scale. Also, we will discuss the uncertainty of each sub-system of water resources and a dynamic adaptive forecasting management plan made in response to the predicted water vulnerability in the future. In addition, in view of the previous studies on the water-energy-food nexus [76,81], subsequent studies on water vulnerability should not only combine water resources with social, economic, and ecological external factors, but also should put water resources into the water-energy-food nexus to study the vulnerability of water-energy-food coupling.

Author Contributions: Conceptualization, Y.C.; methodology, Y.C. and W.C.; validation, Y.C. and W.C.; investigation, Y.C., Y.F. and W.C.; writing—original draft preparation, Y.C. and W.C.; writing—review and editing, Y.C., Y.F. and W.C.; visualization, W.C. and Y.F.; funding acquisition, Y.C. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhong, S.; Wang, H. Retrospective vulnerability evaluation and sensitivity analysis of watershed water resources sustainability. *J. Water Resour. Water Eng.* **2021**, *32*, 43–50.
2. Xi, Z.; Ji, L.; Yang, X.; Liu, Y.; Du, G. Temporal and spatial variation characteristics of precipitation concentration index from 1961 to 2017 in Songliao Basin. *J. Arid Meteorol.* **2019**, *37*, 885–891.
3. Yu, G.; Xiu, C.; Zhao, C.; Ding, Z. Strategic Cross-Border Water Pollution in Songliao Basin. *Sustainability* **2018**, *10*, 4713. [CrossRef]
4. Wei, W.; Zhang, Y.; Wan, G. Research on Surface Water Quality in China-based on Observation of 148 Automatic Monitoring Stations. *J. Coast. Res.* **2020**, *115*, 413–416. [CrossRef]
5. Albinet, M.; Margat, J. Cartographie de la vulnérabilité à la pollution des nappes d'eau souterraine. *Bull. BRGM* **1970**, *2*, 4.
6. Weng, J.; Xia, J.; Chen, J. Water Resources Vulnerability Assessment of the Upper Yellow River. *Yellow River* **2013**, *35*, 15–20.
7. Liu, Q.; Chen, Y. Vulnerability prediction of basin water resources on rough set and BP neural network: A case of Huai River basin. *Resour. Environ. Yangtze Basin* **2016**, *25*, 1317–1327.
8. Pan, Z.; Jin, J.; Liu, X.; Zhou, R. Mechanism and Evaluation Method of Water Resources Utilization System Vulnerability. *J. Nat. Resour.* **2016**, *31*, 1599–1609.
9. Qin, J.; Ding, Y.; Zhao, Q.; Wang, S.-P.; Chang, Y.-P. Assessments on surface water resources and their vulnerability and adaptability in China. *Adv. Clim. Chang. Res.* **2020**, *11*, 381–391. [CrossRef]
10. Hang, N.; Phung, N. Vulnerability assessment of surface water resource of Dong Nai River basin in Lam Dong province. *Vietnam J. Hydrometeorol.* **2020**, *4*, 31–40. [CrossRef]
11. Muthusankar, G.; Lakshumanan, C.; Roy, P. Ground water vulnerability assessment of the South East India in the context of climate change stressors. In Proceedings of the 2nd International Water Conference on Water Resources in Arid Areas (WRAA2020), Muscat, Oman, 16–19 March 2020; p. 80.
12. Jesudhas, C.; Chinnasamy, A.; Muniraj, K.; Sundaram, A. Assessment of vulnerability in the aquifers of rapidly growing sub-urban: A case study with special reference to land use. *Arab. J. Geosci.* **2021**, *14*, 1–20. [CrossRef]
13. Hayward, J.; Johnston, L.; Jackson, A.; Jamieson, R. Hydrological Analysis of Municipal Source Water Availability in the Canadian Arctic Territory of Nunavut. *ARCTIC* **2021**, *74*, 30–41. [CrossRef]
14. Sohns, A.; Ford, J.; Adamowski, J.; Robinson, B.E. Participatory Modeling of Water Vulnerability in Remote Alaskan Households Using Causal Loop Diagrams. *Environ. Manag.* **2021**, *67*, 26–42. [CrossRef]
15. Ghazavi, R.; Imani, R. Development and application of a new index-overlay method to assess urban runoff vulnerability to contamination (evaluation in the Ardabil city, Iran). *Arab. J. Geosci.* **2020**, *13*, 1–12. [CrossRef]

16. Nguyen, T.; Ngo, H.; Guo, W.; Nguyen, H.Q.; Luu, C.; Dang, K.B.; Zhang, X. New approach of water quantity vulnerability assessment using satellite images and GIS-based model: An application to a case study in Vietnam. *Sci. Total Environ.* **2020**, *737*, 139784. [CrossRef] [PubMed]

17. Orozco, I.; Martínez, A.; Ortega, V. Assessment of the Water, Environmental, Economic and Social Vulnerability of a Watershed to the Potential Effects of Climate Change and Land Use Change. *Water* **2020**, *12*, 1682. [CrossRef]

18. Sahana, M.; Rehman, S.; Paul, A.; Sajjad, H. Assessing socio-economic vulnerability to climate change-induced disasters: Evidence from Sundarban Biosphere Reserve, India. *Geol. Ecol. Landsc.* **2021**, *5*, 40–52. [CrossRef]

19. Chen, Y.; Feng, Y.; Zhang, F.; Yang, F.; Wang, L. Assessing and Predicting the Water Resources Vulnerability under Various Climate-Change Scenarios: A Case Study of Huang-Huai-Hai River Basin, China. *Entropy* **2020**, *22*, 333. [CrossRef] [PubMed]

20. Chishugi, D.U.; Sonwa, D.J.; Kahindo, J.-M.; Itunda, D.; Chishugi, J.B.; Félix, F.L.; Sahani, M. How Climate Change and Land Use/Land Cover Change Affect Domestic Water Vulnerability in Yangambi Watersheds (D. R. Congo). *Land* **2021**, *10*, 165. [CrossRef]

21. Papadopoulou, M.P.; Charchousi, D.; Spanoudaki, K.; Karali, A.; Varotsos, K.V.; Giannakopoulos, C.; Markou, M.; Loizidou, M. Agricultural Water Vulnerability under Climate Change in Cyprus. *Atmosphere* **2020**, *11*, 648. [CrossRef]

22. Sharma, J.; Ravindranath, N. Applying IPCC 2014 framework for hazard-specific vulnerability assessment under climate change. *Environ. Res. Commun.* **2019**, *1*, 051004. [CrossRef]

23. Cutter, S. Vulnerability to environmental hazards. *Progr. Hum. Geogr.* **1996**, *20*, 529–539. [CrossRef]

24. Blaikie, P.; Cannon, T.; Davis, I.; Wisner, B. *At Risk: Natural Hazards, People's Vulnerability, and Disasters*; Routledge: London, UK, 1994.

25. Patwardhan, A.; Semenov, S.; Schnieder, S. Assessing key vulnerabilities and the risk from climate change. *Clim. Chang.* **2007**, 779–810.

26. Miguel, A.; Snethlage, J.; Daniels, E.; Froebrich, J. Water Vulnerability Assessment Framework. Available online: http://amsacta. unibo.it/6309/# (accessed on 17 June 2021).

27. Xia, J.; Shi, W.; Chen, J.; Hong, S. Study on vulnerability of water resources and its adaptive regulation under changing environment: A case of Haihe River basin. *Water Resour. Hydropower Eng.* **2015**, *6*, 27–33.

28. Wang, B.; Teng, Y.; Wang, H.; Zuo, R.; Zhai, Y.; Yue, W.; Yang, J. Entropy weight method coupled with an improved DRASTIC model to evaluate the special vulnerability of groundwater in Songnen Plain, Northeastern China. *Hydrol. Res.* **2020**, *51*, 1184–1200. [CrossRef]

29. Xiang, X.; Li, Q. Water Resources Vulnerability Assessment and Adaptive Management Based on Projection Pursuit Model. *J. Coast. Res.* **2020**, *103*, 431–435. [CrossRef]

30. Nasri, G.; Hajji, S.; Aydi, W.; Boughariou, E.; Allouche, N.; Bouri, S. Water vulnerability of coastal aquifers using AHP and parametric models: Methodological overview and a case study assessment. *Arab. J. Geosci.* **2021**, *14*, 1–19. [CrossRef]

31. Haak, L.; Pagilla, K. The water-economy nexus: A composite index approach to evaluate urban water vulnerability. *Water Resour. Manag.* **2020**, *34*, 409–423. [CrossRef]

32. Ioele, G.; De Luca, M.; Grande, F.; Durante, G.; Trozzo, R.; Crupi, C.; Ragno, G. Assessment of Surface Water Quality Using Multivariate Analysis: Case Study of the Crati River, Italy. *Water* **2020**, *12*, 2214. [CrossRef]

33. Tornyeviadzi, H.; Neba, F.; Mohammed, H.; Seidu, R. Nodal vulnerability assessment of water distribution networks: An integrated Fuzzy AHP-TOPSIS approach. *Int. J. Crit. Infrastruct. Prot.* **2021**, *34*, 100434. [CrossRef]

34. Chen, Y.; Feng, Y.; Zhang, F.; Wang, L. Assessing Water Resources Vulnerability by Using a Rough Set Cloud Model: A Case Study of the Huai River Basin, China. *Entropy* **2019**, *21*, 14. [CrossRef]

35. Pan, Z.; Jin, J.; Li, C.; Ning, S.; Zhou, R. A Connection Entropy Approach to Water Resources Vulnerability Analysis in a Changing Environment. *Entropy* **2017**, *19*, 591. [CrossRef]

36. Masocha, M.; Dube, T.; Owen, R. Using an expert-based model to develop a groundwater pollution vulnerability assessment framework for Zimbabwe. *Phys. Chem. Earth* **2020**, *115*, 102826. [CrossRef]

37. Saha, D.; Talukdar, D.; Senapati, U. Exploring Vulnerability of Groundwater Using AHP and GIS Techniques: A Study in Cooch Behar District, West Bengal, India. *Groundw. Soc. Appl. Geospat. Technol.* **2021**, 445–472.

38. Ma, S.; Xu, X.; Chui, J.; Ni, J.; Ma, H.; Tian, H. A New Model to Evaluate Groundwater Vulnerability that Uses a Hybrid AHP-GIS Approach. *J. Geol. Soc. India* **2021**, *97*, 94–103. [CrossRef]

39. Agyare, A.; Kabo-Bah, A.; Bayel, B.; Jalloh, A. Integration of GIS and F-Hydra Model for Aquifer Vulnerability Monitoring in the Afram Plains, Ghana. *J. Geosci. Environ. Prot.* **2021**, *9*, 222.

40. Salimbene, O.; Gharbia, S.; Pilla, F. GIS-based integrated analysis for water resources vulnerability: A case study in the north-west of Italy. *Int. J. Environ. Impacts* **2020**, *3*, 260–271. [CrossRef]

41. Aller, L.; Bennet, T.; Lehr, J.H.; Petty, R.J. *DRASTIC: Standardized System for Evaluating Ground Water Pollution Potential Using Hydrogeologic Settings*; Office of Research Development, US EPA: Ada, OK, USA, 1985.

42. Li, P.; Zhang, C.; Li, W.; Li, Y. Groundwater Vulnerability Assessment of Pingtan Island in Fuzhou City, China, Based on DRASLI-QUE. *J. Hydrol. Eng.* **2021**, *26*, 05020050. [CrossRef]

43. Taheri Tizro, A.; Qaleban Tgmedash, M.; Farajy, F. Vulnerability evaluation of Qazvin Plain Using DRASTIC-AHP Based on Geographic Information System. *J. Nat. Environ. Hazards* **2020**. [CrossRef]

44. Karan, S.; Samadder, S.; Singh, V. Groundwater vulnerability assessment in degraded coal mining areas using the AHP–Modified DRASTIC model. *Land Degrad. Dev.* **2018**, *29*, 2351–2365. [CrossRef]

45. Eftekhari, M.; Akbari, M. Evaluation of the SINTACS-LU model capability in the analysis of aquifer vulnerability potential in semi-arid regions. *J. Appl. Res. Water Wastewater* **2020**, *7*, 111–119.

46. Mirzavand, M.; Ghasemieh, H.; Sadatinejad, S.; Bagheri, R.; Clark, I. Saltwater intrusion vulnerability assessment using AHP-GALDIT model in Kashan plain aquifer as critical aquifer in a semi-arid region. *Desert* **2018**, *23*, 255–264.

47. Ghouili, N.; Jarraya-Horriche, F.; Hamzaoui-Azaza, F.; Zaghrarni, M.; Ribeiro, L.; Zammouri, M. Groundwater vulnerability mapping using the Susceptibility Index (SI) method: Case study of Takelsa aquifer, Northeastern Tunisia. *J. Afr. Earth. Sci.* **2021**, *173*, 104035. [CrossRef]

48. Petrović, B. Intrinsic groundwater vulnerability assessment by multiparameter methods, a case study of Suva Planina Mountain (SE Serbia). *Environ. Earth Sci.* **2020**, *79*, 1–14. [CrossRef]

49. Yogafanny, E.; Anasstasia, T. Utama V F.Zonation of ground water vulnerability using COP and APLIS methods in Gremeng Watershed, Umbulrejo Village, Ponjong, Gunungkidul. *J. Watershed Manag. Res.* **2020**, *4*, 103–120.

50. Sun, C.; Xi, X. Assessment of groundwater intrinsic vulnerability in the Lower Reaches of Liaohe River Plain under uncertain conditions. *Adv. Sci. Technol. Water Resour.* **2014**, *34*, 1–7.

51. Hao, L.; Wang, J. Evaluate to Water Resources Vulnerability Using SWAT-WEAP Model in Tributary of Xiliaohe River. *J. Nat. Resour.* **2012**, *27*, 468–479.

52. Sun, C.; Wang, Y. The Study on Intrinsic Vulnerability of Nitrate in Groundwater in the Lower Reach of Liaohe River Based on WOE Method. *Res. Soil Water Conserv.* **2009**, *16*, 80–84.

53. Ye, X.; You, C.; Du, C.; Wang, L. Groundwater vulnerability to contamination near the river: A case study of the second Songhua River. *Sci. Technol. Rev.* **2015**, *33*, 78–83.

54. Wang, L.; Du, X.; Dong, L.; Zhong, W.; Ding, S. Time-varying Characteristics of Shallow Groundwater Vulnerability in the Plain Area of the Second Songhua River Watershed. *J. Yangtze River Sci. Res. Inst.* **2017**, *34*, 29–34.

55. Fan, J.; Zhang, J.; Tong, Z.; Liu, X. GIS-based Risk Assessment on Rain and Flood Disasters of Songhua River. *J. Catastrophol.* **2009**, *24*, 51–56.

56. Sun, H.; Zhang, Z. Change of Landscape Pattern Vulnerability in the Songhua River Basin in Jilin Province and Its Driving Forces. *Arid Zone Res.* **2019**, *36*, 1005–1014.

57. Liu, M.; Su, X.; Lin, G. Assessment of Groundwater Contamination Caused by Agricultural Activities in Lower Liaohe River Plain. *Water Saving Irrig.* **2013**, *8*, 54–59.

58. Songliao Water Conservancy Network. Overview of Songliao River Basin. Available online: http://www.slwr.gov.cn/slzx/slwlyjj/slwllgk/ (accessed on 17 June 2021).

59. Sigman, H. International Spillovers and Water Quality in Rivers: Do Countries Free Ride? *Am. Econ. Rev.* **2002**, *92*, 1152–1159. [CrossRef]

60. Fang, H.; Sun, L.; Qi, D. Using 137 Cs technique to quantify soil erosion and deposition rates in an agricultural catchment in the black soil region, Northeast China. *Geomorphology* **2012**, *169*, 142–150. [CrossRef]

61. Jia, X.; Lee, H.F.; Zhang, W. Human-environment interactions within the West Liao River Basin in Northeastern China during the Holocene Optimum. *Quat. Int.* **2016**, *426*, 10–17. [CrossRef]

62. Xi, Z.; Yang, X.; Liu, Y.; Ji, L.; Rao, W. Characteristics of Extreme Precipitation Change from 1961 to 2017 in Songliao Basin. *Res. Soil Water Conserv.* **2019**, *26*, 199–203, 212.

63. Lv, C.; Chou, Y.; Jia, Y.; Niu, C.; Ding, X.; Han, C. Water resources vulnerability and its assessment of Haihe River Basin. *South-to-North Water Divers. Water Sci. Technol.* **2012**, *10*, 55–59.

64. Cui, D. Comprehensive assessment of the vulnerability of water resources by improved BP neural network model. *J. Yangtze River Sci. Res. Inst.* **2013**, *30*, 1–7.

65. Cao, Y.; Gao, L.; Zhu, M.; Zhang, L.; Ma, J. Vulnerability assessment method of water resources system and its applications. *Yellow River* **2016**, *38*, 46–49.

66. Pawlak, Z. Rough set approach to knowledge-based decision support. *Eur. J. Oper. Res.* **1997**, *99*, 48–57. [CrossRef]

67. Hu, Q.; Yu, D.; Xie, Z. Numerical Attribute Reduction Based on Neighborhood Granulation and Rough Approximation. *J. Softw.* **2008**, *3*, 640–649. [CrossRef]

68. Zhang, Y.; Wang, Y. Research on Classification Model based on Neighborhood Rough Set and Evidence Theory. *J. Phys. Conf. Ser.* **2021**, *1746*, 012018. [CrossRef]

69. Mansour, K.; Mutanga, O.; Adam, E.; Abdel-Rahman, E. Multispectral remote sensing for mapping grassland degradation using the key indicators of grass species and edaphic factors. *Geocarto Int.* **2015**, *31*, 1–28. [CrossRef]

70. Yan, D.; De Beurs, K. Mapping the distributions of C3 and C4 grasses in the mixed-grass prairies of southwest Oklahoma using the Random Forest classification algorithm. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *47*, 125–138. [CrossRef]

71. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

72. Wu, M.; Liu, J.; Gao, Y.; Kong, X.; Feng, C. Feature selection and clustering via robust graph-laplacian PCA based on capped L 1-norm. In Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA, 13–16 November 2017; pp. 1741–1745.

73. Zong, Y.; Huang, G. A feature dimension reduction technology for predicting DDoS intrusion behavior in multimedia internet of things. *Multimedia Tools Appl.* **2019**, 1–14. [CrossRef]

74. Chen, W.; Chen, Y. Two-Step Measurement of Water–Energy–Food Symbiotic Coordination and Identification of Key Influencing Factors in the Yangtze River Basin. *Entropy* **2021**, *23*, 798. [CrossRef]

75. Zhu, J.; Sun, H.; Liu, N.; Zhou, D.; Taghizadeh-Hesary, F. Measuring Carbon Market Transaction Efficiency in the Power Industry: An Entropy-Weighted TOPSIS Approach. *Entropy* **2020**, *22*, 973. [CrossRef] [PubMed]

76. Chen, W.; Chen, Y. Pre-Warning Measurement of Water Resources Security in the Yangtze River Basin from the Perspective of Water-Energy-Food Symbiosis. *Water* **2021**, *13*, 475. [CrossRef]

77. Chen, Y.; Xu, L. Evaluation and Scenario Prediction of the Water-Energy-Food System Security in the Yangtze River Economic Belt Based on the RF-Haken Model. *Water* **2021**, *13*, 695. [CrossRef]

78. Zhang, L.; Yuan, W.; Jiang, S.; Wu, H.; Chen, T.; Liu, X. Modelling the generation of household automobiles scrap in the context of urban-rural disparity: A case study of Nanjing, China. *J. Clean. Prod.* **2020**, *268*, 122237. [CrossRef]

79. Chen, S.; Jiang, H.; Chen, Y.; Cai, Z. Spatial-temporal patterns of net primary production in Anji (China) between 1984 and 2014. *Ecol. Indic.* **2020**, *110*, 105954. [CrossRef]

80. IPCC. *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Core Writing Team, Pachauri, R.K., Meyer, L.A., Eds.; IPCC: Geneva, Switzerland, 2014; p. 151.

81. Chen, Y.; Chen, W. Simulation Study on the Different Policies of Jiangsu Province for a Dynamic Balance of Water Resources under the Water–Energy–Food Nexus. *Water* **2020**, *12*, 1666. [CrossRef]