Based on DPSIR Model Environmental Planning and Risk Evaluation in the Early Stage of State-Level New Areas Development: In China

Yanyu Chen¹,*, Huaqi Chai¹, Lili Ke², Yimiao Huang³ and Chenglin Xiao¹

¹School of Management, Northwestern Polytechnical University, Xi'an Shaanxi710072, China
²HEC Montreal, H3T 2A7, Canada
³School of Public Administration, Changchun University of Technology, Changchun Jilin130012, China

*Corresponding author e-mail: cheny628@mail.nwpu.edu.cn

Abstract. The development of state-level new areas is a requirement for China's in-depth reform and sustainable development. The strategic objectives of sustainable development put forward higher requirements for the accuracy and scientificity of environmental planning in state-level new areas. Therefore, identifying potential risks through pre-environmental evaluation is the basis for achieving sustainable green development in state-level new areas. According to the characteristics of the early stage of development period of the state-level new Area, this paper applies the DPSIR conceptual model to summarize and extract, constructs the state-level new Area risk evaluation index system, and uses the adaptability and reliability and validity to test the reliability of the index system, and introduces the random forest algorithm. The general process and feasibility of state-level new Area development risk evaluation based on random forest algorithm. This paper identifies 23 risk indicators such as “growth of GDP”, “unbalanced proportion of three-product structure” and “greening coverage”, and judges the risk level and risk evaluation value of the development of 11 state-level-level new Areas. The relevant conclusions provide strong support for the environmental planning and evaluation of the state-level new Area, and also promote the improvement of the environmental evaluation system with sustainable development as the goal.

1. Introduction

The establishment of a State-level New Area is a new strategic deployment of the international development in the future, under the background of global changes in the new era and economic society. Accurate and scientific assessments of their development risks and anticipation of its potential risks are the basis for the effective implementation of this policy and its benefits to the people. The Drivers, Pressures, State, Impact, Response model of intervention (DPSIR) was organized by the Organisation for Economic Cooperation and Development (OECD) and Development Organization in the late 20th century. Based on the Pressure-State-Response (PSR) model, the response model framework for describing the interaction between society and the environment proposed by the dimensions of driving force and influence is added. The core mechanism is Human influence on the environment, and vice
versa, the structural subsystems are interdependent, mutually influential, and mutually integral. Because of its rigorous and well-organized hierarchy of system problems, it has been widely adopted by government organizations and academic institutions, and has been applied to research on ecological security, landscape models, public management, environmental development, and biodiversity, and has become One of the standard methods of environmental assessment and economic and social policy evaluation by the European Environment Agency (EEA). At present, the methods for environmental planning and risk assessment can be divided into traditional statistical methods, high-order mathematical models, and intelligent models.

In comparison, traditional statistical methods and high-order mathematical models are simple to operate and intuitive to visualize the model, but the evaluation process and results are susceptible to data noise; while the higher-order mathematical models are complex in structure, redundant in operation, and low in model robustness. And the calculation scheme is poorly portable. The intelligent model can have good adaptive ability to high-dimensional nonlinear and unstructured problems. It can exploit a large amount of hidden information through autonomous learning, high-dimensional mapping, fault tolerance, memory and other means to obtain the inherent law of risk assessment change. It is regarded as one of the paradigms of data mining. Various intelligent algorithms have the characteristics of easy dimension expansion, high degree of fitting, and no dependence on samples, and have certain advantages in method application.

Throughout the relevant series of studies, most of the research only stays at the level of result simulation, that is, the risk evaluation value of the relevant object is given by experts or traditional methods in advance, and then the original value of the risk evaluation index system is taken as the input variable, thereby establishing The neural network and machine learning algorithm model are used to analyze the precision of the simulation results. However, the application of the algorithm is not ok. Few scholars have given a universal case based on this method.

In view of this, this paper uses the DPSIR model as a tool to construct a national-level new district risk assessment index system, and identifies 23 risk indicators such as “growth rate of production”, and uses the adaptability and reliability and validity to test the reliability of the indicator system. The general process of risk assessment in the development environment planning of State-level New Areas based on this model is explained. The random forest algorithm is used as a tool to explain the general process of State-level New Areas development risk assessment based on the algorithm.[1~3]

2. Model and Algorithm Introduction

2.1. The connotation of DPSIR conceptual framework
DPSIR model (Drivers, Pressures, State, Impact, Response model of intervention, DPSIR) conceptual framework consists of five dimensions, driving force, pressure, state, intervention, response, each the dimension chain contains information about the different elements. The driving force is the change of the basic background environment, which directly and indirectly exerts pressure on the system condition; the pressure is the human factor that leads to the evolution of the system; the state is driven by various factors, the state of the system after the pressure is affected; the influence is the system environment The impact of changes in function on the environment of the system and the outcome factors supporting its development; response is a measure of policy action that is directly or indirectly triggered by the perception of impact and attempts to prevent, eliminate, compensate or reduce its consequences. In order to visually describe the structure of the model system, Figure 1 shows the corresponding schematic, as shown below:
2.2. Mechanism of DPSIR Model for Risk Assessment System in Environmental Planning of State-level New Area

Regarding the application of risk assessment in environmental planning, foreign scholars have more research and mature systems. Domestic scholars began to introduce this concept in 2003, and expounded its applicability and measurement methods, and then developed into ecology and human geography. One of the key elements of the science, sociology, management, systems science, and environmental science departments. The research similar to this paper mostly focuses on economics, finance, etc., and uses PSR and DPSIR models to explore the risks of investment and operation. In view of its application in the evaluation of State-level New Areas development, this paper will use the concept method to combine the characteristics of the target object and give a relatively reasonable model explanation.

According to the classification of economic society and its related statistical indicators given by the Economic Commission for Europe, based on the risk characteristics of the new national districts, the driving factors can be defined as “national social form, population and economic development, as well as lifestyle, overall consumption level and production mode. Corresponding changes”. According to the driving force, it is recommended to classify the four non-hierarchical but interactive drivers that influence the structure and relationship between social, economic, political and environmental systems; among them, the first is the “main driving force” and economic management. Levels of pressure (such as industry, tourism) are directly related to socio-economic activities. The “second driving force” appears at the policy level (eg, waste policy, law). In the long run, with the expansion of the broader spatial sphere of influence, there are levels of “five-level driving force”, ideology and lifestyle (eg, media, consumption patterns). Finally, “basic drivers” include basic trends (population or culture) that are only affected by long-term social decisions (such as climate change, demographics); The pressure is the release of human-made risk in the State-level New Areas, including the ability of decision makers, the quality of practitioners, and the concentration of key personnel. Although different methods for defining them can be found in other literatures, given the object characteristics, some pressure indicators should be given reference (Figure 2); the state can reflect the quality of the natural system of the country alone, or the level of operation of a natural socio-economic system. The indicators may vary from qualitative to quantitative, from quantity to quality, and the state reflects the interpretation of the target object's pressure; The impact is mainly the redistribution of the resources required or operated by the target system of the State-level New Area, which is mutually dynamic with the response.

The benefits of the development of State-level New Area are extensive and their responses can come from different social levels. They have complex and dynamic relationships with the development and operation of new districts, and correspondingly take different measures.

Figure 1. Structure of DPSIR model

Note: The missing line in the figure is a solid line, indicating the positive effect of this dimension on another dimension; the dotted line is a negative effect.
Figure 2. Structural model of risk evaluation system for state-level new areas based on DPSIR framework

Note: The vector line in the figure is a solid line, indicating the positive effect of this dimension on another dimension; the dotted line is a negative effect.

According to Figure 2, the development risk of the State-level New Areas is the result of the combined effects of the five dimensions of driving, pressure, state, impact and response. Among them, the driving and pressure are the source of risk factors affecting the development of the State-level New Areas, and the impact and response dimensions are fundamental. It affects the operation quality of the State-level New Areas, and the state dimension has simple independence.

3. In the Early Stage of State-level New Areas Development environmental planning and risk assessment based on DPSIR model

3.1. Selection of risk indicators
Firstly, the “three-step” method is used to search for the risk assessment index factors in the environmental planning of the State-level New Areas, and then it is included in the DPSIR concept framework to identify the risk indicator factors in the State-level New Areas development environment planning. The table shows. There are 23 risk indicators, among which there are 5 index factors belonging to the driving force D dimension, 4 index factors belonging to the pressure P dimension, 4 index factors belonging to the state S dimension, and 4 index factors belonging to the influence I dimension. Item; there are 5 factors belonging to the response R dimension.
In the current State-level New Areas development environmental planning risk assessment framework, the interpretation of the relationship between the risk index and the source dimension of risk factors is diversified and complex, and a mature indicator framework with universal significance has not yet been formed.

Based on the previous studies, this paper introduces the DPSIR conceptual model into the State-level New Areas environmental planning, and extracts the new district development environment on the basis of in-depth analysis of the State-level New Areas's basic environment, development potential, policy orientation, location significance, and economic level. The dimensional characteristics of risk are incorporated into the DPSIR framework; in order to describe the multivariate structure and internal differences of risk in the development planning of State-level New Areas with causality and systematic multiple feedback. Therefore, the risk assessment index system in environmental planning is well considered, which considers the basic principles of systemic, scientific, operability and hierarchy. It can accurately describe the risk characteristics in environmental planning. In order to test the rationality of the system, it still needs to be strengthened. Expert judgement and mathematical test, and the scientific verification of its dimensions and indicators will be elaborated later.

In addition, the model framework decomposes the connotation of development risk with new ideas and perspectives, and contributes to a comprehensive understanding of the development planning of State-level New Areas. It has positive significance for understanding, identifying risk factors, controlling and predicting risks; In other words, the model has the characteristics of easy expansion and operation, and has good application potential.

### Table 1. Risk Assessment Index System of State-level New Areas

| aims | Dimension | Indicator number | Risk index |
|------|-----------|------------------|------------|
| Driver (D) | D 1 | GDP growth rate | |
| | D 2 | General public budget revenue growth rate | |
| | D 3 | GDP as a percentage of province (city) | |
| | D 4 | Pct of GDP in State-level New Areas higher than that in its province (city) | |
| | D 5 | Import and Export Total Growth Rate | |
| Pressure (P) | P 1 | Unbalanced proportion of three-product structure | |
| | P 2 | High-tech industrial output value as a percentage of total industrial output value | |
| | P 3 | Proportion of employees in the tertiary industry | |
| | P 4 | Residents’ disposable income | |
| State (S) | S 1 | Actual foreign investment rate | |
| | S 2 | Domestic actual occupancy rate | |
| | S 3 | Infrastructure investment | |
| | S 4 | Number of important projects attracting investment | |
| | S 5 | People's livelihood investment accounts for public finance expenditure | |
| Impact (I) | I 1 | Number of Patent Applications | |
| | I 2 | patent authorization number | |
| | I 3 | The number of invention patent applications accounted for the total proportion of patent applications | |
| | I 4 | Number of people engaged in scientific and technological activities | |
| Response (R) | R 1 | Green coverage rate | |
| | R 2 | Air quality rate | |
| | R 3 | sulfur dioxide emissions exceeded the standard | |
| | R 4 | Excessive nitrogen oxide emissions | |
| | R 5 | Industrial solid waste production exceeds the standard | |
3.2. Risk Evaluation Thought in the Early Stage of State-level New Areas Development Environment Planning

According to the foregoing, the State-level New Areas development environment planning based on the sustainable development strategy is a hierarchical, complex, random and structural nonlinear system. The following basic processes should be followed to complete the risk assessment in environmental planning: Risk identification - indicator factor determination - comprehensive evaluation. The comprehensive evaluation process is based on the attribute of each index factor to construct its affiliation relationship with the risk system, and the influence of the single factor is included in the risk index, and the risk index is determined according to the size of the risk index in the environmental plan. It can be seen that the process of environmental assessment in the development of State-level New Areas is essentially a comprehensive decision-making problem, and this solution can be divided into two types of methods, one is the conventional mathematical statistics method, and the other is the pattern recognition scheme. The former obtains the comprehensive index size by supervising or objectively empowering each other, and then by weighting the weighted relationship between the multivariate statistical linear programming weights and the normalized values of the indicators; the latter does not obtain the weight of the index, and mainly determines the modulus rule based on the size of the index. The nonlinear statistical model is realized by the relationship between the index hierarchy and the magnitude of the risk index, so that the inherent value of the risk assessment in the environmental planning is obtained. Law.

3.3. Indicator data source

There are four main ways to obtain data for each risk indicator in environmental planning: one is the relevant statistical yearbook published by the national statistical department; the other is the academic report formed by some research scholars and research institutes. The third is to obtain data through the Internet, including Internet engine search, literature search, official website query, etc. Fourth, by consulting the teachers around, the new district administrators and other parts to obtain some data. Since the development and construction of China's State-level New Areas is a long-term large-scale project, the establishment of the first new district, Shanghai Pudong New Area, was also established after the 1990s. Most of the new districts were established after 2010, so it can be said that China All new districts are currently in the early stages of development and construction. Considering the availability of data, the data of the 2015 new district was mainly collected, and the new district was established before 2015, with 11 new districts. The remaining 8 new districts were established after 2015, and the data could not be obtained.

3.4. System Dynamics Mechanism of Risk Assessment System in Environmental Planning under DPSIR Framework

3.4.1. System dynamics model. System Dynamics (SD) is a quantitative analysis method developed by Professor JW Forrester, a professor of the Massachusetts Institute of Technology (MIT) in 1956, based on feedback control theory and system simulation technology. The system is an object that resolves information transmission and feedback mechanisms between substructures. For half a century, the method has been applied to economic and social development, automation, ecology, industrial control, and economic modeling. The key basis for applying the system dynamics method to assist the establishment of the risk assessment model in the environmental planning of the national new district is to follow the systematic scientific thinking of “the system must have structure, the system structure determines the system function”, determine the feedback characteristics of each dimension relationship, and then search for it. The transmission path, on the basis of the mathematical basis of system structural integrity and vector directionality, solves the behavioral nature and influence of internal structural factors. The fitting of the elementary flow and its path in the system dynamics constitutes a complete structural equation modeling (SEM). The dimensions are multidimensional nonlinear relations, so the SEM is used to solve the model. The structural equation model formula is as follows:
In the formula: \( \eta, \xi \) represent the internal and external potential variables of the system; \( \beta \) represents the matrix of effect coefficients between potential variables in the system and potential variables outside the system; \( \lambda \) represents the matrix of effect coefficients between potential variables outside the system and potential variables within the system; \( a_x, a_y \) represents the matrix of effect coefficients of \( x \) versus \( \xi \) and \( y \) versus \( \eta \); where \( \zeta \) is the error term of the structural model; \( \varepsilon \) and \( \delta \) are the error terms of the measurement model.

The initial modeling is performed after all variables and their linkage relationship are determined. The value of the observed variable has been given, and the latent variable, the target variable, is pending and cannot be directly estimated. If the system structure model is correctly defined, its overall covariance matrix and the model covariance matrix are equal. The property and covariance matrix between the observed variables is as follows:

\[
\begin{bmatrix}
\text{Cov}(\eta, \eta) + \text{Cov}(\xi, \eta) \\
\text{Cov}(\xi, \eta) \cdot \alpha_y \\
\text{Cov}(\xi, \eta) \cdot \alpha_x
\end{bmatrix}
\]

This formula is the covariance equation. By formula (1), it can be obtained as follows:

\[
\begin{bmatrix}
\alpha_y \text{Cov}(\eta) \alpha_y' + \varepsilon_x y + \theta_e \\
\alpha_x \text{Cov}(\xi, \eta) \cdot \alpha_y \\
\alpha_y \text{Cov}(\xi, \eta) \cdot \alpha_x
\end{bmatrix}
\]

There are 8 parameters to be estimated in the matrix, and the least variable method is usually used to obtain the matrix of potential variable coefficients, as follows:

\[
B = \begin{bmatrix}
a_{11} & a_{21} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

Then its structural equation model is expressed as follows:

\[
\begin{cases}
\eta_1 = a_{11} \xi + a_{21} \xi + a_{31} \xi + \cdots + a_{m1} \xi \\
\eta_2 = a_{21} \xi + a_{22} \xi + a_{31} \xi + \cdots + a_{m2} \xi \\
\vdots \\
\eta_m = a_{11} \xi + a_{m1} \xi + a_{m2} \xi + \cdots + a_{mm} \xi
\end{cases}
\]

(5)
3.4.2. System Dynamics Parameters of Risk Assessment Model in Environmental Planning of State-level New Areas under DPSIR Framework. The DPSIR model of risk assessment in the national new district environmental planning shown in Figure 2 is simple and clear, and clearly reflects the feedback correlation between the various dimensions. Based on this structure, we use AMOS 24.0 simulation software to simulate the dynamic process to calculate the path coefficient between the observed variable and the latent variable, latent variable and latent variable. The results are shown in Table 2.

Table 2. results of path coefficient of structural equation model based on risk evaluation of DPSIR state-level new area

| Path     | Vector Direction | Coefficient | P    |
|----------|------------------|-------------|------|
| D → P    | Positive         | 0.891       | 0.000|
| P → S    | Negative         | -0.887      | 0.001|
| S → I    | Negative         | -0.673      | 0.001|
| I → R    | Negative         | -0.623      | 0.001|
| R → D    | Positive         | 0.865       | 0.000|
| R → P    | Positive         | 0.921       | 0.000|
| R → S    | Positive         | 0.804       | 0.000|
| D → T    | Positive         | 0.236       | 0.010|
| P → T    | Positive         | 0.218       | 0.010|
| S → T    | Positive         | 0.166       | 0.010|
| R → T    | Positive         | 0.198       | 0.010|
| I → T    | Positive         | 0.182       | 0.010|

Note: → indicates the path. T in the table indicates the potential dependent variable outside the system, that is, the target variable—the risk index in the early stage of state-level new areas environment planning.

It can be seen that in the early stage of state-level new areas environment planning risk assessment model structure system, there are three path coefficients with negative values, and correspondingly the vector direction is negative; the other nine path coefficients are positive and the vector direction is positive. For the national dependent area development risk index, the target dependent variable, the driving D dimension has the greatest impact on it, the path coefficient is 0.236, which indicates that it is the key content of risk prevention and control; the impact of pressure PP on external potential dependent variable I The path coefficient is 0.218; the influence of state SS on the national new zone risk is the smallest, the path coefficient is only 0.166; the effect of the influence dimension II is also small, the path coefficient is 0.182; the response RI dimension is the influence of the target variable. I force up to 0.198. The weighted sum of the path coefficients of the five dimensions is 1, and the coefficient value is also a representation of its weight.

In the structure of state-level new areas environmental planning risk assessment system based on the DPSIR framework, the driving force D has a significant positive impact on the pressure P, and its path coefficient is 0.891, indicating that the risk factor of the national new district increases by one unit. Correspondingly, a pressure of 0.891 units will be produced.

The pressure dimension P has a negative correlation with the state S dimension, and the former has a path coefficient of -0.887 to the latter information, indicating that the forward vector information of the pressure evolves into negative information after being passed to the state dimension, and the information energy is weakened. That is, only 1 unit of energy is transferred from the pressure to the state, leaving only 0.887 of stock.

There is a negative correlation between the state dimension S and the influence dimension. The former has a path coefficient of -0.673 to the latter, indicating that when the state dimension is increased by 1 unit, the response influence dimension will be reduced by 0.673 units.
The information transmission path coefficient that affects the dimension I to the response dimension R is -0.623, which means that when the energy of the dimension changes by one unit, there is a change in the energy of -0.623 units.

The influence of the response dimension R on the driving force dimension is very obvious. The path coefficient of the former to the latter information transmission is 0.865, which indicates that when the response dimension is increased by 1 unit, the driving force dimension will increase by 0.865 units. The transmission path coefficient of the response dimension R to the pressure dimension is 0.924, which reflects a positive correlation between the two. The response dimension changes by 1 unit, and the pressure dimension will change by 0.921 units. There is also a positive correlation between the response dimension R and the state dimension S. The information transmission path coefficient of the former to the latter is 0.804, indicating that the state dimension will change by 0.804 units when the response dimension changes by one unit.

It is not difficult to find that when the information of each dimension is transmitted to another dimension information, the absolute value of the path coefficient is less than 1, which is due to information loss and energy reduction in the dynamic structure of the system.

It can be seen that the absolute values of the path coefficients of the vector direction such as the driving force dimension--pressure dimension, the pressure dimension--state dimension, the response dimension--driving force dimension, the response dimension--pressure dimension, etc. are relatively large, which are 0.891, -0.887, 0.865, respectively. 0.821, indicating that the loss of information in the direction of the path is small, and also reflects that the former has a greater influence on the latter; on the contrary, its influence is smaller.

3.5. Verification of Risk Assessment Index System in the early stage of state-level new areas environment planning Based on DPSIR Framework

3.5.1. Verification of Index System Adaptability Based on DPSIR Framework of System Structure Equation. In order to verify the rigor of the model and the rationality of the structure, different methods are needed for testing. This chapter proposes a number of detection methods, in order to master its scientific nature, first based on system dynamics and structural equations to verify. It is usually judged by three parameters: absolute fit, value-added fit, and simple fit, used for the relationship between latent variables and the part of the variance that cannot be explained by other variables in the model. Based on the above-mentioned schematic diagram of the power flow system of the risk assessment index system in the national new district development environment planning based on the DPSIR module framework, the AMOS 24.0 software modeling is used to statistically obtain the structural correlation model of the building structure. As shown in Table 3.

| Measurement method       | Statistical detection amount | Adaptation standard | Test result | Adaptation result     |
|--------------------------|-----------------------------|---------------------|-------------|-----------------------|
| Absolute fit             | Chi-square value            | > 0.05              | 0.061       | Ideal fit             |
|                          | Residual mean square and    | < 0.05              | 0.046       | Ideal fit             |
|                          | square root                |                     |             |                       |
|                          | Adaptability index         | > 0.90              | 0.896       | Proximity adaptation  |
| Value-added adaptation   | Standard fit index         | > 0.90              | 0.912       | Ideal fit             |
|                          | Relative fitness index     | > 0.90              | 0.907       | Ideal fit             |
|                          | Value-added fitness index  | > 0.90              | 0.918       | Ideal fit             |
| Simple adaptation        | Adjusted dominance index   | > 0.90              | 0.936       | Ideal fit             |
|                          | Simple adaptation index    | > 0.90              | 0.908       | Proximity adaptation  |
From the measurement method of absolute fitness, there are three specific statistical indicators, and the chi-square test value of the value is 0.061, which satisfies the adaptation threshold requirement of more than 0.05. Therefore, from the point of view, it is ideal adaptation type.

From the sub-indicator of the mean square of the residual and the square root, the test value is 0.046, which satisfies the threshold requirement of less than 0.05. From this point of view, it is an ideal adaptation type; the fitness index indicates that the test value of the sub-indicator is 0.896. Although it is close to the horizontal flat powder with a threshold of 0.90, it still does not reach the threshold line, so it belongs to the near-adaptive type;

From the perspective of the absolute adaptation measurement method, two of the three sub-indicators exceed the threshold line, which is an ideal adaptation type, and only one of them is close to the threshold, which is the near-the adaptation type. From the measurement method of value-added adaptation degree, there are also three sub-indicator metric values, wherein the standard adaptation degree index test value is 0.912, which reaches the threshold requirement that should be greater than 0.90, so it is determined that it belongs to the ideal adaptation type;

From the point of view of the value-added fitness index, the test value of the sub-indicator is 0.907, which is also above the threshold line of 0.90, which satisfies the inspection requirements, and therefore belongs to the ideal adaptation type;

From the relative fitness index, the test value is 0.918, which exceeds the equivalent of 0.18 of the threshold line, so it is determined that it belongs to the ideal fit type. The test values of the three sub-measurement indicators in the measurement method all meet the threshold requirement, and the value-added adaptation degree is determined to be an ideal adaptation type.

From the measurement method of the simple adaptation degree, there are two measurement sub-indicators, wherein the adjusted value of the adjusted fitness index is 0.936, which satisfies the threshold requirement greater than 0.90, according to which it is considered to be an ideal adaptation type; From another sub-indicator, the test value of the simple fit index is 0.908, which also is the ideal adaptation type. Although the three measurement methods consider different angles, they can help us determine the stability of the system structure. The test values of 7 sub-indicators of the 8 measurement sub-indicators meet the threshold requirements, and the judgment result is ideally adapted. There is a sub-indicator that fails to reach the threshold level and is the near-adaptive type.

It should be pointed out that although there is a measure sub-indicator that fails the ideal adaptation test, but comprehensively meets the ideal adaptation criteria, it can be known that the risk assessment index of the State-level New Areas development environment planning based on the DPSIR framework that construction is basically reasonable. The dynamic system of the structure is rigorous and scientific, and the system simulation accuracy is high, which has certain applicability. However, due to the lack of analysis, induction and dimension determination of specific indicators, there are individual indicators that fail to reach the threshold level, which indicates that this aspect should be strengthened in the future.

3.5.2 Reliability and validity analysis. Reliability and validity analysis can rationalize the construction of macro-measurement index system, and it is also an important basis for index factor optimization and dimension re-induction. Reliability is a measure of the consistency, consistency and stability of indicators. It is usually related to the heterogeneity of indicators. Generally, the better the heterogeneity of indicators, the higher the reliability, which also requires indicators when constructing the indicator system. The identification of factors ensures their independence and dimensionality. Validity is a measure of the effectiveness of the indicator system. Reliability is a prerequisite for validity. It mainly measures the effectiveness of the indicator system, namely availability.

In view of this, this paper is based on the above-mentioned dimensions of risk assessment in the State-level New Areas environmental planning based on the DPSIR model framework. The SPSS21 statistical software is used for reliability and validity analysis. The analyzed indicator statistics mainly include Cronbach's $\alpha$ coefficient, half-fold reliability, convergence validity, KMO test value and Bartlett ball test value, are shown in Table 4.
Table 4. Reliability and validity of the test results of the adaptability of the preset model for the index system of the development risk evaluation of the state-level new area based on DPSIR

| Latent variable | Cronbach's $\alpha$ coefficient | half-fold reliability | Convergence validity | KMO test | Bartlett ball test |
|-----------------|----------------------------------|-----------------------|----------------------|----------|-------------------|
| D               | 0.972                            | 0.961                 | 0.927                | 0.761    | 120.3 (0.000)     |
| P               | 0.918                            | 0.908                 | 0.886                | 0.752    | 91.28 (0.000)     |
| S               | 0.856                            | 0.829                 | 0.813                | 0.737    | 89.65 (0.000)     |
| I               | 0.874                            | 0.851                 | 0.837                | 0.718    | 43.57 (0.000)     |
| R               | 0.838                            | 0.823                 | 0.808                | 0.720    | 68.69 (0.000)     |

According to the table, there are differences in the reliability and validity test values of each dimension of the risk assessment index system based on the State-level New Areas development environment planning under the DPSIR framework.

The Cronbach's $\alpha$ coefficient of the driving force dimension is the largest, 0.972, followed by the pressure dimension, which is 0.918. The values of Cronbach's $\alpha$ coefficient of state, response and influence dimension are not much different, which are 0.856, 0.838 and 0.874 respectively.

From the perspective of half-fold reliability, the half-fold reliability value of the driving force dimension is higher than the correlation value of other dimensions, which is 0.961; the half-fold reliability of the pressure dimension is 0.908; the half-fold reliability value of the influence dimension is centered, which is 0.851; the response dimension and the half-factor reliability values of the state dimension are 0.829 and 0.823, which are the minimum values in each dimension.

From the point of view of convergence efficiency, the convergence efficiency of the driving force dimension is 0.927, which is larger than other dimensions; the convergence efficiency of the pressure dimension is second, 0.886; the convergence validity value of the influence dimension is relatively small, 0.837; state and response The dimensional convergence validity values are not much different, in order of 0.813 and 0.808.

The KMO test value shows that the test values of the driving force dimension and the pressure dimension are the largest, reaching 0.761 and 0.752 respectively; the KMO test value of the state dimension is centered at 0.737; the KMO test values of the influence dimension and the response dimension are not much different, respectively 0.718 and 0.720.

The Bartlett autumn test value of each dimension spans a large extent, and the Bartlett autumn test value of the driving force dimension is 120.36; the Bartlett autumn test value of the pressure dimension is similar to the state test value, which are 91.28 and 89.65 respectively; the Bartlett autumn test value of the response dimension and the Bartlett autumn test value of the impact dimension is the smallest, only 43.57.

According to the overall situation, the Cronbach's $\alpha$ coefficient, the half-fold reliability and the convergence validity of the five dimensions are all greater than 0.8; and the KMO test values of each dimension are greater than 0.7; the Bartlett ball test values of each dimension of DPSIR are greater than 40. And passed the reliability test of 0.001 level, indicating that the reliability and validity test have statistical significance; The construction of the risk assessment index system of the State-level New Areas development environment planning based on the DPSIR model framework is basically reasonable, and can be applied to subsequent empirical analysis and data research.

4. Conclusion
This paper takes the pre-development environmental planning of the State-level New Areas as the theme, takes the sustainable development as the goal, applies the DPSIR conceptual model to summarize and extract, and constructs the risk assessment index system of the national new district environmental planning, and determines the growth rate of GDP. "Consider 23 risk indicators, and use the adaptability and reliability and validity to test the reliability of the indicator system. The random forest algorithm is used to transform the comprehensive decision problem of risk assessment of national new district
development environment planning into a nonlinear mapping problem. The model relationship between index factor input and risk level identification is established by the idea of classification and regression, and the construction is completed through network training. Model prediction. This conclusion can provide data support for the proposed pre-development environmental planning of the national new district, and provide a basis for the subsequent proposed more targeted countermeasures.

References

[1] Wenfa Hu. Political Risk Evaluation Model of International Engineering Project Based on BP Algorithm [J]. Civil Engineering and Environmental Engineering, 2006, 28 (4): 98-100.

[2] ShiBin Zhang, Chunxiang Xu, Yujun An. Research on risk assessment method based on cloud model [J]. Journal of University of Electronic Science and Technology of China, 2013, 42 (1): 92-97.

[3] Yi Li. Risk Evaluation and Application of Real Estate Investment Based on SVM [J]. Statistics and Decision, 2012 (1): 70-72.

[4] Breiman L. Mach. Learn. [J]. Machine Learning, 2001, 45: 5-32.

[5] Breiman L, Last M, Rice J. Random Forests: Finding Quasars [M] // Statistical Challenges in Astronomy. Springer New York, 2003: 243-254.

[6] Breiman L. RANDOM FORESTS—RANDOM FEATURES [J]. Machine Learning, 1999, 45 (1): 5-32.

[7] Lawrence R L, Wood S D, Shley R L. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest) [J]. Remote Sensing of Environment, 2006, 100 (3): 356-362.

[8] Han H, Guo X, Yu H. Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest [C] // IEEE Interstate-level Conference on Software Engineering and Service Science. IEEE, 2017: 219-224.

[9] Jun Wang. How to evaluate the national new district [J]. China Economic Report, 2017 (11).

[10] Fangshu Ni, Jia Wang, Qingfeng Cao, et al. Construction of National New District Evaluation Index System and Its Implications for the Development of Xiong'an New District in Hebei Province——Based on the Perspective of Five Development Concepts [J]. City, 2017(6): 3-8.

[11] Lu Jiao, RuiYang, Lin Guo. Research on Evaluation of Resource and Environment Carrying Capacity of National New District——Taking Gui’an New Area as an Example [J]. Journal of Sichuan University of Science and Technology (Social Science Edition), 2017, 32 (5): 87-100.

[12] Mingkui Li, Lei Shi, Xue Tan. Preliminary Study on Construction and Application of Environmental Performance Evaluation Index System in National New District [J]. Environmental Protection, 2016, 44 (23): 31-34.