Evaluation of regional industrial cluster innovation capability based on particle swarm clustering algorithm and multi-objective optimization

Yongcai Yan¹ · Mengxue He¹ · Lifang Song²

Received: 19 June 2021 / Accepted: 31 August 2021
© The Author(s) 2021

Abstract

With the progress of the times and the development of science, industrial clusters have been regarded by all countries in the world as one of the important ways to enhance regional competitiveness, and become an inevitable trend of industrial development. The research on the innovation ability of industrial clusters can not only maintain sustainable development of industrial clusters and obtain sustained competitive advantages, but also provide reference for the government's policy formulation of industrial clusters. This paper aims to study the evaluation of regional industrial clusters' innovation capability based on particle swarm clustering and multi-objective optimization. This paper uses the theory of industrial cluster innovation and takes regional industrial system as the empirical research object to establish a regional industrial system capability evaluation system, which is based on the selection of indicators, combined with analytic hierarchy process and factor analysis to evaluate industrial innovation capability. On this basis, the particle swarm clustering theory is used to verify the innovation ability and evaluation index system of industrial clusters, and provide a reference for the evaluation of the innovation ability of industrial clusters. This paper divides the regional cluster innovation capability into four aspects: innovation input capability, environment support capability, self-development capability and innovation output capability, and systematically analyzes the key elements and in the composition of innovation elements and their relationships. It then constructs the evaluation index system of regional cluster innovation capability. At the same time, this paper introduces clustering analysis algorithm and swarm intelligence algorithm into regional innovation evaluation, combines particle swarm optimization algorithm and K-means clustering algorithm, and optimizes particle swarm clustering algorithm by adjusting adaptive parameters and adding fitness variance. The experimental results of this paper show that from the results of the tested innovation potential of the three industrial clusters, industrial cluster F has the strongest innovation ability, with an evaluation coefficient of 0.851, followed by industrial cluster F, which has a value of 0.623. This result is consistent with the actual innovation status of the selected industry. From this point of view, the established particle swarm clustering model for evaluating the innovation capability of regional industrial clusters is reliable and can be used to evaluate the innovation capability of different industrial clusters.

Keywords Machine learning · Particle swarm clustering · Multi-objective optimization · Regional industrial clusters · Innovation capability evaluation

✉ Lifang Song
Slifang@ruc.edu.cn

Yongcai Yan
yanyc@hbnu.edu.cn

Mengxue He
hbjince@163.com

¹ School of Economics Management and Law, Hubei Normal University, Huangshi 435002, Hubei, China

² Labor Union, Renmin University of China, Beijing 100872, China

Introduction

From an international point of view, with the deepening of economic globalization, international division of labor and cooperation is becoming more and more obvious. The world economy presents the development trend of regional industrial clusters. The limited resources in the world have begun to transfer and gather to regions with strong regional innovation ability and obvious industrial clusters, which promotes the rapid growth of these regional economies and

also drives the common development of national economy and world economy. Under the same conditions, through the optimization technology, the system efficiency is improved, the energy consumption is reduced, and the rational use of resources and the economic benefits are improved. Natural organisms adapt to the surrounding environment through their own evolution, so as to continue to move forward. Evolutionary algorithm is a kind of random search technology based on this idea, which is a mathematical simulation of the biological evolution process. They simulate the collective learning process of groups composed of individuals.

How to establish enterprise clusters with innovative advantages and how to effectively exert one's own competitive advantages are urgent s to be solved. From the perspective of clusters, research and evaluation of cluster innovation capabilities can help clusters create a good innovation environment, effectively promote the transfer and absorption of knowledge in the dissemination, innovation system, and encourage cluster enterprises to produce active innovation interaction processes. The transformation of enterprise clusters can improve the overall competitiveness of the enterprise group, evaluate the innovation capability of the business system, and find the best way to improve the innovation performance of the business system.

Jean N showed how to train a convolutional neural network to recognize image features that can explain up to 75% of local economic changes. It also shows how to apply powerful machine learning techniques with limited training data, showing that there are broad potential applications in many scientific fields. But the recognition of the image is not very high [1]. Unlike ordinary computational fluid dynamics prediction methods, Ren T has developed a back-propagation neural network method to quickly obtain the receiver temperature, such as the peak temperature of the inner and outer surfaces, the average outlet temperature, and the molten salt with the highest outlet temperature. Numerical simulations verify the feasibility and effectiveness of the proposed method. In addition, in the proposed method, the temperature of the tube wall and molten salt can be quickly predicted without considering the thermophysical parameters of the material, the boundary conditions or initial conditions, and the solution of complex control equations. However, the practicality of this method is not very high [2]. Liu Q provides distributed, but tightly integrated services that have rich functions in large-scale management, reliability and fault tolerance. As far as big data processing is concerned, newly built cloud clusters are facing performance optimization challenges, which are focused on faster task execution and more efficient use of computing resources. The currently proposed methods focus on time improvement, that is, shorten the MapReduce time, but pay little attention to storage usage. However, the unbalanced cloud storage strategy may exhaust the heavy nodes of the MapReduce

cycle and further challenge the security and stability of the entire cluster [3].

The innovation of this paper is (1) Combining the specific conditions of China's industry, designed and developed a Chinese enterprise cluster innovation capability evaluation model, which is used to evaluate whether the enterprise is suitable for clusters and mergers. The percentage of the number of science and technology service organizations in the total number of enterprises is used as an indicator to evaluate the interactive capabilities of the enterprise merger. (2) Based on the standard PSO, the algorithm uses the decreasing inertia weight strategy to punish the unconstrained particles. In the later stage of the algorithm, the mutation operator is introduced to enhance the diversity of the population and make the algorithm better for global optimization.

Evaluation method of regional industrial clusters' innovation capability based on particle swarm clustering and multi-objective optimization

Learning algorithm of particle swarm clustering

Particle swarm optimization (PSO) is a global optimization algorithm based on the individual population. PSO is composed of a certain number of individuals. These individuals (i.e., particles) are randomly generated. These particles search in a certain range. By calculating the value of the fitness function, the position of particles is determined and the quality of possible solutions is evaluated. In the iterative process, particles update their own speed [4]. According to the criteria in the group optimal solution and the individual optimal solution, a new position of the particle is found in the feasible region to find a new possible solution. In the local pattern particle swarm optimization algorithm, the direction and position of particles are only affected by their own cognition and the state of adjacent particles, and they no longer follow the global optimal particles, but are determined by the local optimal position of particles [5, 6].

1. Variable definition.

The group optimal position makes the particles converge quickly to form a particle swarm and searches the neighborhood of the global extremum. The individual's own experiential optimal position ensures that the particles will not converge to the group optimal too quickly and fall into the local minimum, so that the particles can search the region between the individual extremum and the global extremum in one iteration [7]. The m -th neuron in the input layer is denoted as x_p , the m -th neuron in the hidden layer is denoted

as k_m , and the n th neuron in the output layer is denoted as y_n . The connection weight from x_p to k_m is w_{pm} , and the connection weight from k_m to y_n is w_{mn} .

The input and output of each layer are represented by u and v , respectively, and q is the number of iterations [8, 9]. The formula of particle update speed is as follows:

$$Y(q) = [v_N^1, v_N^2, \dots, v_N^N]. \quad (1)$$

The expected output is:

$$d(q) = [d_1, d_2, \dots, d_N]. \quad (2)$$

The error of the q -th iteration is defined as:

$$e_n(q) = d_n(q) - Y_n(q). \quad (3)$$

The error energy is defined as:

$$e(n) = 1/2 \sum_{n=1}^N e_n^2(q). \quad (4)$$

2. Positive working signal propagation.

Fitness function value of each particle:

$$v_p^p(q) = x(q). \quad (5)$$

The input of the m -th neuron in the hidden layer is equal to the weighted sum of $v_p^p(q)$, namely:

$$u_M^m(q) = \sum_{p=1}^P w_{pm}(q) v_p^p(q). \quad (6)$$

Assuming $f(\cdot)$ is the sigmoid function, the output of the q -th neuron in the hidden layer is:

$$v_M^m(q) = f(u_M^m(q)). \quad (7)$$

In the linear decreasing strategy of inertia weight, with the increase of iteration times, the inertia weight decreases linearly:

$$u_N^N(q) = \sum_{m=1}^M w_{mn}(q) v_M^m(q). \quad (8)$$

The output of the n th neuron in the output layer is equal to:

$$v_N^n(q) = g(u_N^n(q)). \quad (9)$$

The function expression of the penalty term is as follows:

$$e_n(q) = d_n(q) - v_N^n(q). \quad (10)$$

The total error of the network is:

$$e(q) = 1/2 \sum_{n=1}^N e_n^2(q). \quad (11)$$

3. Multi-objective particle swarm optimization process.

In the algorithm, a total of three sets are set to save the particle swarm, non-dominated set and external set. Particle swarm optimization (PSO) is the main body to perform the search, while non-dominated set and external set are the main body to save the search results. When the algorithm starts to run, it first initializes the PSO and related parameters randomly, then finds out all the non-dominated particles in the PSO and inserts the non-dominated set [10]. The non-dominated set represents the best part of the particles searched by the algorithm in this generation, but it must also separate the best part of the particles found by the algorithm so far; therefore, it is necessary to insert the non-dominated particles from each generation into the external set, and the global extremum should be the representative of the solution found by the algorithm. Obviously, it is very appropriate to take the external set as the candidate set of the global extremum, and the particle swarm optimization constantly searches for the better solution under the guidance of the extremum and then enters the next cycle [11, 12]. The position change of particles in the search space is based on the individual's successful surpassing other individuals' social psychological intention. Therefore, the change of particles in a group is influenced by the experience or knowledge of its neighboring particles. The search behavior of one particle is affected by the search behavior of other particles in the swarm.

Multi-objective optimization solution method

1. Converting multiple goals into single goals.

The basic idea of this method is to combine all objective functions into a total objective function form, called AOF. Common methods include evaluation function method, constraint method and objective programming method. These methods are briefly introduced below. These methods are also often the method used [13].

For a multi-objective programming problem, if the real function can be constructed according to the preference information provided by the designer, the solution of the most satisfactory design solution is equivalent to the solution of the optimal solution with the actual function as the new objective function, then it is called a multi-objective problem and it is scalable and measurable. Multi-objective utility theory studies the conditions for the existence of such real functions and how to construct problems [14].

The basis of the utility theory is that the preferences of the designer are considered to be expressed by a real function called the utility function. Once the utility function can be constructed, the final choice of the design must be determined according to the value of the utility function.

In some cases, the design with the highest utility value is selected. In case of uncertainty, the design with the highest expected value of the function is chosen. Although utility theory provides tools for analyzing multi-objective optimization, in many cases, the preferred information provided by the designer is insufficient to determine this utility function. For a practical problem, it is difficult or even impossible to evaluate or construct a utility function [15, 16]. To help designers choose satisfactory solutions and overcome the difficulties of usefulness theory, the concept of evaluation function came into being. The evaluation function is used to approximate the utility function of multiple features that are often fuzzy and difficult to construct in the designer's mind and to evaluate the quality of the design. The basic idea is to establish an evaluation function $h(f(x))$ for the multi-objective optimization problem and then solve:

$$\min h(f(x)), \text{ s.t. } x \in X. \quad (12)$$

We use the optimal solution of the problem as the optimal solution of the multi-objective optimization problem [17].

Through the weighted summation of the objective, the MOP is converted into SOP through the linear aggregation of the objective function, that is, the multi-objective optimization problem is converted into a single-objective optimization problem:

$$\min \sum_{i=1}^m \rho_i \cdot f_i, \text{ s.t. } x_i^z \leq x_i \leq x_i^u. \quad (13)$$

This method is the simplest method to solve the multi-objective optimization problem and the optimal Pareto frontier bending problem. If all the weights are positive, in theory, this method can ensure that any optimal Pareto solution set is found. But when the mapping relationship between the decision area and the target area is nonlinear, a set of uniformly distributed weights cannot find a set of uniformly distributed solution sets. In addition, using different weighting factors does not guarantee finding different optimal Pareto solutions. The biggest disadvantage of this method is that it cannot find all the solutions, and the optimal Pareto front is a non-curve average surface [18, 19].

2. Normal constraint algorithm.

It can generate a set of evenly spaced solutions on the Pareto frontier of the multi-objective optimization problem. A new representation is that it combines the critical curve mapping of the design objective and a Pareto solution set filter. This is in NC and other solutions. The method is very effective in practical application. The normal constraint method is a Pareto frontier generator, which is generated by the following characteristic normal constraint method. Pareto points are evenly distributed along the Pareto front.

They are not sensitive to the severity of the design goals. It is effective and relatively easy to achieve any number of design goals. The normal constraint method will ensure that any Pareto points are generated in the feasible design space, as well as Pareto points uniformly distributed along the complete front edge [20].

By performing a series of optimizations, the normal constraint method can obtain a uniformly distributed Pareto solution set of the genetic multi-objective problem. Each optimization execution in this series is constrained by the feasible design space. With the reduction of each design space, a single Pareto solution can be obtained. By changing the original multi-objective problem to a single-objective problem, the single-objective problem affected by the design space reduction can be minimized [21]. Starting from the initial feasible design space, and then reducing until the entire design space has been explored, the normal constraint generates a Pareto solution through the Pareto front. In this sense, the normal constraint method is based on design space reduction, which is similar to other Pareto frontier generators. However, the normal constraint method is different in solution and efficiency [22, 23].

Evaluation method of industrial cluster innovation ability

1. Chromatography analysis method.

The basic idea of the analytic hierarchy process (AHP) is that the decision maker divides the problem into different levels and different elements and performs simple comparisons. The characteristic of the analytic hierarchy process is that on the basis of in-depth analysis of nature, influencing factors and internal relations, it constructs a hierarchical model of complex decision-making problems and then uses less quantitative information to mathematicize the decision-making process, thus having multiple goals. Multi-criteria or unstructured feature complex decision-making problems provide a simple decision-making method [24]. It is to mathematicize and systematize people's thinking process to facilitate acceptance. This method requires less quantitative information, but requires decision-makers to fully understand the nature of the decision-making problem, the elements contained and the logical relationship between them.

The basic steps of the AHP method are:

1. Create a multi-level progressive structure hierarchy model for various elements in the decision-making problem.
2. Establish a crisis matrix.

According to the hierarchical structure, for the upper-level element, compare the importance of the relationship between all the elements related to the next level. The comparison scale should be used to quantify the comparison results, and the crisis table should be compiled at the bottom.

3. Calculate the weight distribution through a specific matrix calculation.

Find the maximum value $\kappa_{\max}$ of the crisis matrix and the corresponding manifold, and normalize the manifold to one of the related elements of the previous layer with the weight of the specific related component. The calculation method generally adopts the numerical repetition method [25].

4. Consistency check and correction of judgment matrix.

The calculation procedure of the consistency test: using the analytic hierarchy process, consulting experts, qualitative and quantitative analysis of the crisis matrix, classification of the level and consistency test, the separator corresponding to the highest efficiency of the crisis matrix should be standardized and characterized as W [26]. The W element is the relative importance classification weight of the coefficient at the same level as the upper level factor. This process is called level classification. To check the consistency of the classification, the consistency index should be calculated:

$$CI = \frac{\kappa_{\max} - n}{n - 1} \quad (14)$$

In the above formula, CI is the consistency index, n represents the dimension of the judgment matrix, and the value of the average random consistency index RI is shown in Table 1.

When the random consistency ratio is $CR = CI/RI < 0.1$, it indicates that the result of the single-level sorting has satisfactory consistency; otherwise, the value of the element of the judgment matrix needs to be adjusted.

5. Determine the overall priority of the relevant factors of the multi-level merger and check the overall consistency.

The maceAHP software is a very effective software, which helps calculate the level of the analysis process. It only needs to input the evaluation data into the software, and the software will calculate the weight results and check the consistency. This article will use maceAHP software to conduct empirical research, using detailed analytic hierarchy process.

2. Factor analysis method.

Factor analysis is a technique that can simplify complex and difficult data. It analyzes the internal connections between multiple data, reduces the data dimension, and finds various hypothetical variables that can reflect the main information of the original multivariate. Because these hypothetical variables are intangible, they are called factors in academia. The mathematical model of factor analysis is as follows:

Suppose $X_i (i = 1, 2, \dots, q)$ q variables, there are two expressions, respectively, as in the following formula:

$$X_i = u_i + a_{i1}f_1 + \dots + a_{in}f_n + c_i, \quad (15)$$

$$\begin{bmatrix} X_1 \\ X_2 \\ \dots \\ X_q \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_q \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{q1} & a_{q1} & \dots & a_{qn} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_q \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \\ \dots \\ c_q \end{bmatrix} \quad (16)$$

$f_1, f_2, \dots, f_n$ are called common factors, and their coefficients are called factor loadings. c_i is a special factor, which is not included in the first n factors and satisfies $\text{Cov}(f, c) = 0$, that is, f and c are not related;

$$B(f) = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \dots & \\ & & & 1 \end{bmatrix} = I \quad (17)$$

It can be seen from the formula that the $f_1, f_2, \dots, f_n$ common factors are not correlated with each other, and the variance is equal to 1.

Normally, the interpretation effect of the main factor is not obvious, and the factor load needs to be rotated. By rotating the factor loading, the square of the factor loading in the matrix tends to the poles from 0 to 1. Finally, the

Table 1 The value of the average random consistency index RI

n	1	2	3	4	5
RI	0	0	0.63	0.91	1.11
n	6	7	8	9	10
RI	1.25	1.33	1.43	1.53	1.64

factor score is calculated according to the rotation matrix, and the evaluation conclusion is drawn.

The basic steps of factor analysis:

1. Calculate the KMO value of the original data: $KMO > 0.5$ is more suitable for factor analysis, while $KMO < 0.5$ is not suitable for factor analysis.
2. Construct factor variables that can reflect most of the original data.
3. Use matrix rotation to make factor variables interpretable [27].
4. Calculate the scores of each factor variable and draw relevant conclusions.

Evaluation experiment of regional industrial cluster innovation ability based on particle swarm clustering

Particle swarm clustering model of innovation ability evaluation

Particle swarm optimization is an optimization algorithm based on iterative thought, similar to the genetic algorithm. Firstly, by evaluating the adaptability of each particle, the current individual optimal solution of each particle and the current group optimization of the whole particle swarm are determined after t iteration. Particles in the community adjust their speed and direction according to their best position and the best position of all particles, so as to get closer to "food" better and faster. After the initial particles are obtained, Gaussian variation is carried out on the particles, which enhances the diversity of the particles. The original velocity update formula is changed when the particle is updated, and random terms are added to the formula to ensure the global search of particles. The probability variation of particles with similar degree of advantage and disadvantage can prevent the algorithm from falling into the local optima. Particle swarm optimization (PSO) is the main body to perform the search, while non-dominated set and external set are the main bodies to save the search results. So far, we have to find out which part of the particle swarm algorithm is not the best. Its functional framework is shown in Fig. 1.

Establishment of evaluation index system

1. Principles of index selection.

Scientific principle The selection of the index system should combine theory and practice. It should be able to reflect the actual situation of the pharmaceutical industry cluster appreciably, and the theory should be selected reasonably. In the design of the index system, we must first use

Fig. 1 Particle swarm optimization model of innovation ability evaluation

scientific theories as a guide to ensure the rationality and rigor of index selection, accurately grasp the nature of the evaluation object, and solve problems in a targeted manner. Secondly, it is necessary to be able to find out the most essential and most original things about the innovation capability of industrial clusters and to choose the most accurate and comprehensive method to describe the subject of evaluation. The more realistic the index system, the stronger is its scientific nature.

Principle of operability Due to the different external conditions and internal mechanisms of industrial clusters in various provinces and cities, on the basis of ensuring the principle of scientificity, try to choose identity indicators as far as possible. And the data must be maneuverable: first, the indicators should be easy to obtain, and the data source should be accurate and reliable, So try to choose the data in the statistical yearbook; second, the data can be quantified and convenient for calculation; finally, the selection of indicators should not be too many, so try to use the indicator system and simplify and evaluate comprehensively.

The principle of universal comparability The principle of universal comparability compares the evaluation subjects in various periods and different objects, that is, it compares the regional industrial clusters vertically and horizontally. Horizontal comparison is to compare regional industrial clusters with the pharmaceutical industries of other provinces, finds out their common points according to the situation of each industry, and designs an index system based on the common points.

Guiding principle The evaluation of regional industrial clusters is not simply to find out their position in the country, but to find out the problems that exist in the emergence stage and guide the industrial clusters to develop in the correct goal and direction.

2. Selection of evaluation indicators.

According to the above-mentioned index selection principle, in view of the actual situation of regional industrial clusters, the selection of indexes in this paper is as follows:

Number of companies with R&D activities: the number of companies engaged in research and development activities in the cluster reflects the scale of innovative research in the cluster.

R&D personnel Refers to personnel engaged in three types of activities including basic research, applied research, and experimental development among the personnel of enterprise science and technology activities.

Full-time equivalent of R&D personnel Refers to the sum of the number of full-time personnel plus the workload of part-time personnel converted into the number of full-time personnel.

Personnel and labor service fees: income earned by individuals when they engage in research and development activities.

Instruments and equipment Expenditures for purchasing instruments and equipment in research and development.

Government capital Usually refers to the government's use of financial tools to promote R&D investment in specific industries, through joint investment between the government and enterprises and institutions, to accelerate the process of industrialization, promote the rapid development of related industries, and achieve national-level macroeconomic goals and scientific research goals, even national defense goals.

Working capital Refers to the sum of various assets and materials disposed of by an enterprise in the form of currency and the currency used to pay workers' labor compensation.

External expenditure R&D Refers to the funds allocated to another party by an institution that rewards or cooperates with another party.

Number of new products developed The number of new products developed in the province this year.

New product development expenditure Refers to the expenditure for the research and development of new products from the internal expenditures of the enterprise's scientific and technological activities during the reporting year, including new product research, design, model development, testing, testing and other expenses.

Evaluation of innovation capability of regional industrial clusters based on particle swarm optimization and multi-objective optimization

Evaluation of the innovation ability of different industrial clusters based on particle swarm optimization

This article conducted field interviews and questionnaire surveys on industries in the Yangtze River Delta, the Pearl River Delta, and asked for help from government statistics department. The above evaluation model was used to evaluate and analyze the innovation ability, conduct a questionnaire survey, obtain the evaluation parameters of the relevant industrial clusters, and use $f(x)$ and $g(x)$ to normalize each indicator. The processed data is shown in Table 2 and Fig. 2.

We divide the data in the chart into two parts and select the first four groups as learning sample training weights. Learning accuracy $e = 0.001$, and the other three groups are used as test samples. After 5000 classes, the training results are shown in Table 3 and Fig. 3. After the training is completed, use the trained three-level BP network to input three sets of data samples for testing.

Among them, the output results of the innovation capabilities of the three industrial clusters are shown in Table 4.

Through the above analysis, it is concluded that the evaluation of the innovation capability of industrial clusters based on neural network can achieve satisfactory results. From the test results of the innovation capability of the three

Table 2 Normalized results of industrial cluster evaluation parameters

Industrial cluster name	Industrial clusters A	Industrial clusters B	Industrial clusters C	Industrial clusters D	Industrial clusters E	Industrial clusters F	Industrial clusters G
Knowledge flow	0.43	0.46	0.36	0.70	0.73	0.87	0.32
Technological innovation ability	0.64	0.56	0.52	0.69	0.72	0.89	0.44
Innovation environment	0.73	0.68	0.57	0.83	0.80	0.92	0.46
Innovation economic performance	0.52	0.50	0.35	0.53	0.45	0.79	0.31

Fig. 2 Innovation data of industrial clusters **Table 3** Data on the training results of the innovation capability network of industrial clusters

Industrial cluster name	Industrial clusters A	Industrial clusters B	Industrial clusters C	Industrial clusters D
Knowledge flow	0.44	0.60	0.70	0.47
Technological innovation ability	0.65	0.72	0.62	0.65
Innovation environment	0.35	0.50	0.51	0.38
Innovation economic performance	0.69	0.66	0.79	0.57

Fig. 3 Data on the training results of the innovation capability network of industrial clusters **Table 4** Output data of the innovation capability of the three industrial clusters

Grade	Industrial clusters E	Industrial clusters F	Industrial clusters G
0.8–1.0	/	0.851	/
0.6–0.8	0.623	/	/
0.4–0.6	/	/	/
0.2–0.4	/	/	0.393
0–0.2	/	/	/

industrial clusters, it can be seen that the cluster with the strongest innovation capability is industrial cluster F, followed by industrial cluster E, and the innovation capability of industrial cluster G is the weakest among the three tested industrial clusters.

Weight distribution of the evaluation index of industrial cluster innovation ability

The index weight is obtained through the calculation of the analytic hierarchy process. Make a judgment matrix questionnaire for each level of indicators of enterprise cluster innovation ability and distribute it to experts who are engaged in the research and management of enterprise cluster innovation ability, so as to obtain more objective, fair and authoritative original evaluation data. Use software to calculate the statistical evaluation data to get the weight results and consistency test of each questionnaire. Calculate the weight of the sub-indices to the overall goal according to the formula. Among them, the first-level weights of the major categories of indicators and the second-level weights of the sub-indices and the weights of the sub-indices to the overall goal are shown in Table 5 and Fig. 4.

Among the three major categories of indicators of enterprise cluster innovation capability, the output capability indicator occupies the most important position, with a weight of 0.4972. This is mainly because the purpose of enterprise cluster innovation is to promote the

commercialization of high-tech, and the level of innovation capability is ultimately important, reflected in product output, measured by economic level. The strength of the innovation output capacity of enterprise clusters is mainly manifested in the proportion of export value to the total industrial output value, and the weight of the indicator of the proportion of export value to the total industrial output value to the total goal of enterprise cluster innovation capacity is 0.3369, that is, it is through the product competitive advantage reflected by sales in foreign markets. Secondly, the number of authorized patents per 1,000 scientific and technical personnel also has a large weight in the evaluation system. Its weight for the overall goal of the innovation capability of the enterprise cluster is 0.1342, which is much higher than the ratio of the number of patents granted to the number of patent applications.

The data are standardized by calculating the score of the secondary index, and the score of the secondary index is a number between 0 and 1, so that the data can be normalized. According to the above method, the scores of the nine sub-indicators of the regional industrial cluster innovation

Table 5 Weight distribution of industrial cluster innovation capability evaluation index

Primary indicator name	First-level weight	Secondary indicator name	Secondary indicator number	Secondary weight	Percentage of sub-indicators (%)
Innovative R&D	0.3785	Scientific and techni	1	0.1834	7.07
		Government investment	2	0.0691	2.75
		High-tech	3	0.1991	7.32
Innovation output	0.4972	Patent professor	4	0.1572	7.69
		Characteristic economy	5	0.2118	10.77
		Export value	6	0.3369	16.71
Innovative interaction	0.1249	University cooperation	7	0.1342	1.66
		Institute cooperation	8	0.1212	1.54
		Technology Service Organization	9	0.1161	1.45

Fig. 4 Weight distribution of industrial cluster innovation capability evaluation index

Table 6 Standardized scores of regional industrial cluster innovation capability data

Secondary indicator name	Secondary indicator number	2015	2016	2017	2018	2019
Science technology	1	0.55	0.60	0.60	0.74	0.78
Government investment	2	0.44	0.50	0.52	0.48	0.23
High-tech	3	0.11	0.08	0.05	0.46	0.56
Patent professor	4	0.05	0.58	0.71	0.25	0.26
Characteristic economy	5	0.20	0.04	0.63	0.23	0.74
Export value	6	0.43	0.22	0.03	0.60	0.07
University cooperation	7	0.43	0.70	0.06	0.04	0.05
Institute cooperation	8	0.05	0.26	0.31	0.07	0.39
Technology Service Organization	9	0.20	0.27	0.31	0.39	0.50

Fig. 5 Standardized scores of regional industrial cluster innovation capability data **Table 7** 2015–2019 Regional industrial cluster innovation capability evaluation index

Evaluation standard	First-level weight	2015	2016	2017	2018	2019
Innovative R&D capability index	0.39	0.231	0.240	0.257	0.290	0.357
Innovation output capability index	0.49	0.274	0.328	0.381	0.354	0.445
Innovation and interaction capability index	0.13	0.162	0.201	0.228	0.277	0.318
Innovation capability index	1	0.242	0.278	0.313	0.313	0.395

Fig. 6 2015–2019 regional industrial cluster innovation capability evaluation index

capability evaluation system from 2015 to 2019 are shown in Table 6 and Fig. 5:

According to the above chart, the scores and weights of each indicator of the innovation capability of regional enterprise clusters in 2015–2019 are obtained and then summed to obtain the category index and evaluation index. We multiply the evaluation index of the major category index by the corresponding first-level weight and then sum up to obtain the total evaluation index of regional enterprise cluster innovation ability. The specific index is shown in Table 7 and Fig. 6.

Judging from the data trend in the chart, although the innovative R&D capabilities of regional industrial clusters continued to rise from 2015 to 2019, the growth was slow, and the growth level of innovative R&D capabilities could not keep up with the development of innovative capabilities. In 2015, the innovative R&D capability index of regional industrial clusters was only 0.231, which was at a low level. This was mainly because the regional operation time was not long in 2015, and the government and specialized towns themselves attach great importance to R&D investment and R&D personnel investment. The output capacity index of this region is very high. Thanks to the continuous increase in the proportion of export value in the total industrial output value, professional towns have the investment in research and development funds and manpower has been fruitful. The new products and new technologies of specialized town enterprises have been recognized in the fierce market competition, and the market share and the industrial output value of high-tech enterprises have also increased significantly. In general, the innovation ability of regional industrial clusters has received more and more attention, and there is still much room for improvement in the innovation ability of specialized towns.

Conclusions

This paper decomposes the evaluation system of enterprise cluster innovation capability into three categories: innovation R&D capability, innovation output capability, and innovation interaction capability. These three categories of indicators represent the different functional positioning of enterprise cluster innovation. Then according to the relevant principles of establishing a scientific evaluation index system, nine sub-indices were selected to establish an evaluation index system of enterprise cluster innovation capability, which is used to quantify the innovation capability of enterprise clusters, aiming to find out the problems and gaps in the innovation capability of enterprise clusters.

Industrial cluster innovation ability is network innovation ability; it is the ability of industrial cluster network to identify opportunities, decision-making, and resource allocation.

This paper further discusses the ability of industrial cluster network to identify opportunities, decision-making innovation capabilities, and innovative resource allocation capabilities. We put forward the relevant ways to increase the innovation ability of industrial clusters.

On the basis of combining k-means algorithm in clustering analysis algorithm and particle swarm algorithm in swarm intelligence algorithm, the particle swarm clustering algorithm is formed. According to the change of population fitness function value, the cut in time of K-means algorithm is determined to enhance the local accurate search ability of the clustering algorithm and shorten the convergence time. The algorithm is optimized from four aspects: the selection of fitness function variance, the adjustment of inertia weight factor, the adjustment of learning factor, and the setting of contraction factor. The improved model is simulated.

Acknowledgements This work was supported by the Research Center for Transformation and Development of Resource Exhausted Cities (RCTDREC) of Hubei Normal University, the Research Center for Collaboration Innovation of Airport Economy (RCCIAE) of South China Business College of Guangdong University of Foreign Studies, the Special Projects in Key Areas (New Generation of Information Technology) of Colleges and Universities in Guangdong Province (No. 2020ZDZX3046), and the Characteristics innovation project of colleges and universities of Guangdong Province (Natural Science, No. 2019KTSCX235). This work was supported by Major project of Beijing Social Science Foundation "Research on Financial Support System Adapting to the Coordinated Development of Strategic Emerging Industries in Beijing-Tianjin-Hebei", (No. 20ZDA11)

Declarations

Conflict of interests The author(s) declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Jean N, Burke M, Xie M et al (2016) Combining satellite imagery and machine learning to predict poverty. *Science* 353(6301):790–794
2. Ren T, Liu S, Yan G et al (2016) Temperature prediction of the molten salt collector tube using BP neural network. *Renew Power Gener IET* 10(2):212–220

3. Liu Q, Cai W, Shen J et al (2016) A speculative approach to spatial-temporal efficiency with multi-objective optimization in a heterogeneous cloud environment. *Secur Commun Netw* 9(17):4002–4012
4. Sidiropoulos ND, De Lathauwer L, Fu X et al (2017) Tensor decomposition for signal processing and machine learning. *IEEE Trans Signal Process* 13:3551–3582
5. Helma C, Cramer T, Kramer S et al (2018) Data mining and machine learning techniques for the identification of mutagenicity inducing substructures and structure activity relationships of noncongeneric compounds. *J Chem Inf Comput* 35(4):1402–1411
6. Cai J, Luo J, Wang S et al (2018) Feature selection in machine learning: a new perspective. *Neurocomputing* 300(26):70–79
7. Ruske S, Topping DO, Foot VE et al (2017) Evaluation of machine learning algorithms for classification of primary biological aerosol using a new UV-LIF spectrometer. *Atmos Meas Tech* 10(2):1–18
8. Benjamin SL, Aspuru-Guzik A (2018) Inverse molecular design using machine learning: generative models for matter engineering. *Science* 361(6400):360–365
9. Janani R, Vijayarani S (2019) Text document clustering using Spectral Clustering algorithm with Particle Swarm Optimization. *Expert Syst Appl* 134:192–200
10. Zhu X, Shang J, Sun Y et al (2019) PSO-CFDP: a particle swarm optimization-based automatic density peaks clustering method for cancer subtyping. *Hum Hered* 84(1):1–12
11. Assouline D, Mohajeri N, Scartezzini JL (2017) Quantifying rooftop photovoltaic solar energy potential: a machine learning approach. *Sol Energy* 141:278–296
12. Feng Z, Rong P, Cao P et al (2018) Machine learning-based quantitative texture analysis of CT images of small renal masses: Differentiation of angiomyolipoma without visible fat from renal cell carcinoma. *Eur Radiol* 28(4):1625–1633
13. Luo X, Liu J, Zhang D et al (2016) A large-scale web QoS prediction scheme for the Industrial Internet of Things based on a kernel machine learning algorithm. *Comput Netw* 101(4):81–89
14. Dev S, Wen B, Lee YH et al (2016) Ground-based image analysis: a tutorial on machine-learning techniques and applications. *IEEE Geosci Remote Sens Mag* 4(2):79–93
15. Fei Z, Li B, Yang S et al (2017) A survey of multi-objective optimization in wireless sensor networks: metrics, algorithms, and open problems. *IEEE Commun Surv Tutor* 19(1):550–586
16. Hamdy M, Nguyen AT, Hensen JLM (2016) A performance comparison of multi-objective optimization algorithms for solving nearly-zero-energy-building design problems. *Energy Build* 121(6):57–71
17. Rubén S, Ruiz AB, José DB et al (2016) Evolutionary multi-objective optimization algorithms for fuzzy portfolio selection. *Appl Soft Comput* 39(C):48–63
18. Boada Y, Reynoso-Meza G, Picó J et al (2016) Multi-objective optimization framework to obtain model-based guidelines for tuning biological synthetic devices: an adaptive network case. *BMC Syst Biol* 10(1):1–19
19. Gupta N, Bari A (2017) Fuzzy multi-objective optimization for optimum allocation in multivariate stratified sampling with quadratic cost and parabolic fuzzy numbers. *J Stat Comput Simul* 87(10–12):1–12
20. Kang HS, Kim YJ (2016) A study on the multi-objective optimization of impeller for high-power centrifugal compressor. *Int J Fluid Mach Syst* 9(2):143–149
21. Yu L, Yang Z, Tang L (2016) Prediction-based multi-objective optimization for oil purchasing and distribution with the NSGA-II algorithm. *Int J Inf Technol Dec Mak* 15(02):423–451
22. Chen T, Li K, Bahsoon R et al (2018) FEMOSAA: feature guided and knee driven multi-objective optimization for self-adaptive software at runtime. *ACM Trans Softw Eng Methodol* 27(2):5.1–5.50
23. Prakash C, Kansal HK, Pabla BS et al (2016) Multi-objective optimization of powder mixed electric discharge machining parameters for fabrication of biocompatible layer on β -Ti alloy using NSGA-II coupled with Taguchi based response surface methodology. *J Mech Sci Technol* 30(9):4195–4204
24. Chakraborty D, Guha D, Dutta B (2016) Multi-objective optimization problem under fuzzy rule constraints using particle swarm optimization. *Soft Comput* 20(6):2245–2259
25. Donoso V, Martin V (2016) Product relatedness and economic diversification in the USA: an analysis at the state level. *Ann Reg Sci* 56(2):449–471
26. Li S, Park MK, Jo CO et al (2017) Emission estimates of methyl chloride from industrial sources in China based on high frequency atmospheric observations. *J Atmos Chem* 74(2):227–243
27. Parker R (2016) Explaining regional departures from national patterns of industry specialization: regional institutions, policies and state coordination. *Org Stud* 26(12):1787–1807

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.