



OPEN Spatial distribution characteristics and influencing factors of youth soccer specialty schools in China

Yujin Zhou¹, Qiuhan Chen¹, Dong He¹, Yu Cheng² & Xiaowei Peng¹✉

Youth soccer specialty schools are the primary project to strengthen the foundation of soccer talents, improve the level of soccer development, and achieve the dream of Chinese soccer. Taking 32,775 soccer specialty schools in China as the research sample, ArcGIS spatial analysis technology is used to explore the spatial distribution characteristics of soccer specialty schools and their influencing factors. The analysis shows that the overall distribution of the country forms five quantity concentration area and eight high-density core areas, with a high degree of spatial concentration and significant differences between the East and West. The factors influencing soccer specialty schools include nine factors in four dimensions: historical development, geographical environment, socio-economic, and educational resources. It was found that geographic environment and socio-economy provide the foundation and background for developing educational resources. However, the direct role of educational resources is more significant in a specific educational program, such as soccer specialty schools, followed by socio-economy and geographic environment. This study comprehensively analyzes and interprets soccer specialty schools' spatial distribution characteristics and influencing factors. These results enrich the research on the spatial layout of sports, providing necessary theoretical support and practical references for the layout and high-quality development of related industries in China and other countries or regions.

As the global economy flourishes and people's living standards rise, sports development has become a significant concern worldwide¹. As an essential part of sports, the revitalization and development of soccer have become a common expectation of Chinese sports and the general public. Since 2012, the Chinese government has listed the revitalization of soccer as a strategic task for developing sports and constructing an intense sports country². Today, soccer has a pivotal role in developing sports in China.

In this background, accelerating the development of youth campus soccer is not only an important measure to implement the government's education policy and promote the physical and mental health of young people but also a basic project to consolidate the foundation of soccer talents, improve the level of development of soccer and achieve the dream of China's soccer. As an essential way to promote the popularization and enhancement of campus soccer, the construction of the Soccer Specialty School has excellent significance in cultivating youth soccer talents and integrating educational resources. The construction of the Soccer Specialty School is a critical way to promote popularizing and enhancing school soccer, which is significant in cultivating youth soccer talents and integrating educational resources³.

Since 2015, the General Office of the State Council, the Ministry of Education, the General Administration of Sport, and other departments have issued several essential documents, clearly put forward: in order to build in line with the law of growth of talents, extensive participation of young people, the level of sports continued to improve, the system and mechanism is full of vitality, the primary conditions of protection is strong, and the cultural atmosphere of the soccer specialty schools development system, there is an urgent need to strengthen the promotion and popularization of school soccer.

At the same time, it is emphasized that the requirements of the regional layout of soccer specialty schools, explicitly including the integration of urban and rural regional layout, the integration of all levels and types of schools, the integration of various kinds of social resources, and encouraging the localities and schools with a foundation to explore the practice, and increase support for rural schools, and strongly support the development of school soccer in the profoundly impoverished areas and central and western regions. Promote high-quality and balanced youth soccer development for both men and women in urban and rural areas. Make efforts to expand school soccer coverage, with a plan to reach 20,000 schools by 2020 and 50,000 or so by 2025. By the end of 2021, 32,775 schools with soccer characteristics were already in the country. However, due to the lack of

¹College of Physical Education, Wuhan Sports University, Wuhan 430079, China. ²College of Journalism and Communication, Wuhan Sports University, Wuhan 430079, China. ✉email: pengxiaowei027@126.com

scientific spatial planning and layout, the distribution of schools in some areas was uneven. There were problems such as unreasonable resource allocation and insufficient or wasted educational resources^{4–6}.

The construction and development of specialty schools will help satisfy the enthusiasm and demand for soccer among the general youth and inject new vitality and vigor into Chinese soccer. Globally, soccer as a sport and cultural phenomenon has significant social and economic potential^{7–9}. Its development can cultivate more excellent soccer players, promote the sports industry's prosperity, enhance the country's image, and promote international exchanges and cooperation^{10–12}. Therefore, the successful construction of soccer specialty schools will provide a new way for quality soccer education and positively contribute to advancing world soccer.

Literature review

Currently, Research on Chinese soccer specialty schools (i.e., school soccer) has primarily focused on the following aspects:

- ① **Policy Analysis.** The Chinese government has recently implemented various policies to promote school soccer development. These policies include initiatives to expand soccer programs through improving facilities, enhancing teacher training, and developing national soccer curriculum standards. However, the effectiveness of these policies varies significantly across different regions. Xu et al. noted that while these national policies have generally fostered the growth of school soccer, substantial challenges persist due to uneven resource allocation. Specifically, the more developed eastern regions have benefitted from better facilities and qualified teachers, leading to more effective policy implementation. In contrast, despite policy intentions, the central and western regions face significant barriers due to insufficient funding and resources¹³. Peng et al. further analyzed the dynamics of policy formulation in China's soccer reforms, emphasizing the critical role of local government participation and the need for policies to be adapted to local economic, cultural, and educational contexts¹⁴. International studies, such as those by Michalis et al., reinforce that uniform policies across diverse regions often fail to achieve desired outcomes, underscoring the necessity for context-specific adjustments¹⁵.
- ② **Developmental Challenges:** Despite increasing policy support and societal attention, the development of school soccer in China still faces numerous challenges. A significant issue is the imbalance in regional development. Research indicates that the eastern regions, with relatively abundant educational resources, have a distinct advantage. In contrast, the central and western regions lag due to limited economic and social resources. This disparity is evident in the availability of facilities and teaching staff and the cultivation and delivery of soccer talent¹⁶. Another critical concern is the health risks associated with intensive professional training. Liu et al. demonstrated that while professional training significantly enhances soccer skills, it also increases the risk of acute injuries. These health concerns are crucial to address in the future development of school soccer, particularly in ensuring that training programs balance intensity with student health and safety¹⁷.
- ③ **Case studies.** On a more localized level, some scholars have conducted in-depth case studies to explore the development experiences of specific schools or regions within the context of school soccer. For instance, Fan and Sungkawadee developed an evaluation index system to assess the sustainable development of school soccer in Shaanxi Province, using empirical research on primary schools in Xi'an. Their study found that leadership in policy implementation and establishing a solid school soccer culture are pivotal for sustainable development¹⁸. Similarly, Zhang and Gu conducted a study on school soccer activities in Yunnan Province's border areas, using interviews with local stakeholders. They identified that limited resources, inadequate facilities, and weak teaching staff are significant constraints on soccer development in these regions. However, by integrating soccer into secondary education, these challenges can be partially mitigated, promoting the popularization and development of the sport¹⁹. These case studies provide valuable insights into addressing specific regional challenges in the development of school soccer.
- ④ **Development Pathways.** Research on the development pathways has predominantly focused on strategies to improve soccer programs' quality, including policy adjustments, resource optimization, and collaboration among various societal sectors. Peng et al. emphasize the importance of stakeholder consensus and the need for policies tailored to regional and cultural specifics to ensure sustainable development²⁰. Jonathan et al. discussed the significance of coordinated efforts between national and local governments in promoting school soccer and made some suggestions for the sustainable development of Chinese soccer in the future²¹.
- ⑤ **Spatial Distribution Analysis.** Despite the growing body of research on school soccer, the spatial distribution of Chinese Soccer Specialty Schools has received relatively little academic attention. The spatial distribution characteristics of these schools and the key factors influencing their distribution are underexplored. Most existing research has concentrated on policy analysis, development pathways, and case studies, with spatial distribution research remaining supplementary and lacking systematic and comprehensive analysis. While some studies have mentioned the uneven distribution of educational resources across regions, they have not thoroughly investigated how these spatial patterns affect the effectiveness and reach of soccer education. Moreover, although research has been conducted on the spatial distribution of sports facilities^{22–25}, these studies are not directly applicable to schools, as the roles of facilities and schools differ significantly.

Overall, while existing research has provided valuable insights into the development of Chinese soccer from perspectives such as policy orientation, developmental challenges, pathway exploration, and localized case studies, systematic studies on soccer specialty schools' spatial distribution characteristics and influencing factors remain relatively scarce. This research gap limits our comprehensive understanding of the nationwide dynamics of school soccer programs. This study examines the spatial distribution characteristics and key influencing factors of Chinese Youth Soccer Specialty Schools to address this gap. By employing various spatial analysis

methods, including ArcGIS, Kernel Density Analysis, Pearson Correlation Coefficient, Grey Relation Analysis, and GeoDetector, this research aims to uncover the distribution patterns across different regions and identify the primary driving factors. The findings of this study not only provide a new theoretical perspective on school soccer research and offer practical guidance for the planning and development of similar projects. Furthermore, this work serves as a valuable reference for future policy formulation and the strategic allocation of educational resources.

Results

Soccer specialty schools spatial distribution status

According to the Ministry of Education's official data, China aims to establish 52,227 soccer specialty schools between 2015 and 2025. By December 2021, 32,775 schools had been completed nationwide, reflecting a completion rate of 62.75% and representing 14.31% of all primary and secondary schools in the country (Fig. 1).

In order to identify regional differences in soccer specialty schools, we conducted a statistical analysis of seven geographic regions in China (east, central, southwest, north, northwest, south, and northeast). The proportion of soccer specialty schools in ordinary schools (development rate) in each region shows significant differences in development. The proportion of soccer specialty schools in the northeast is significantly higher than the national average. At the same time, the southern and central regions are relatively lagging. The proportions in the remaining regions are close to the national average.

By analyzing the development rate of soccer specialty schools in each province, it can be found that (Fig. 2) The Inner Mongolia Autonomous Region, with a proportion of 44.54%, leads the nation, more than 2.6 times the national average. Provinces like Jiangsu, Shanghai, and Liaoning also surpass the national average significantly, with proportions exceeding 20%. In contrast, provinces such as Henan, Hebei, and Guizhou report proportions below 10%, indicating a considerable gap from the national average of 14.31%. The Guangxi Zhuang Autonomous

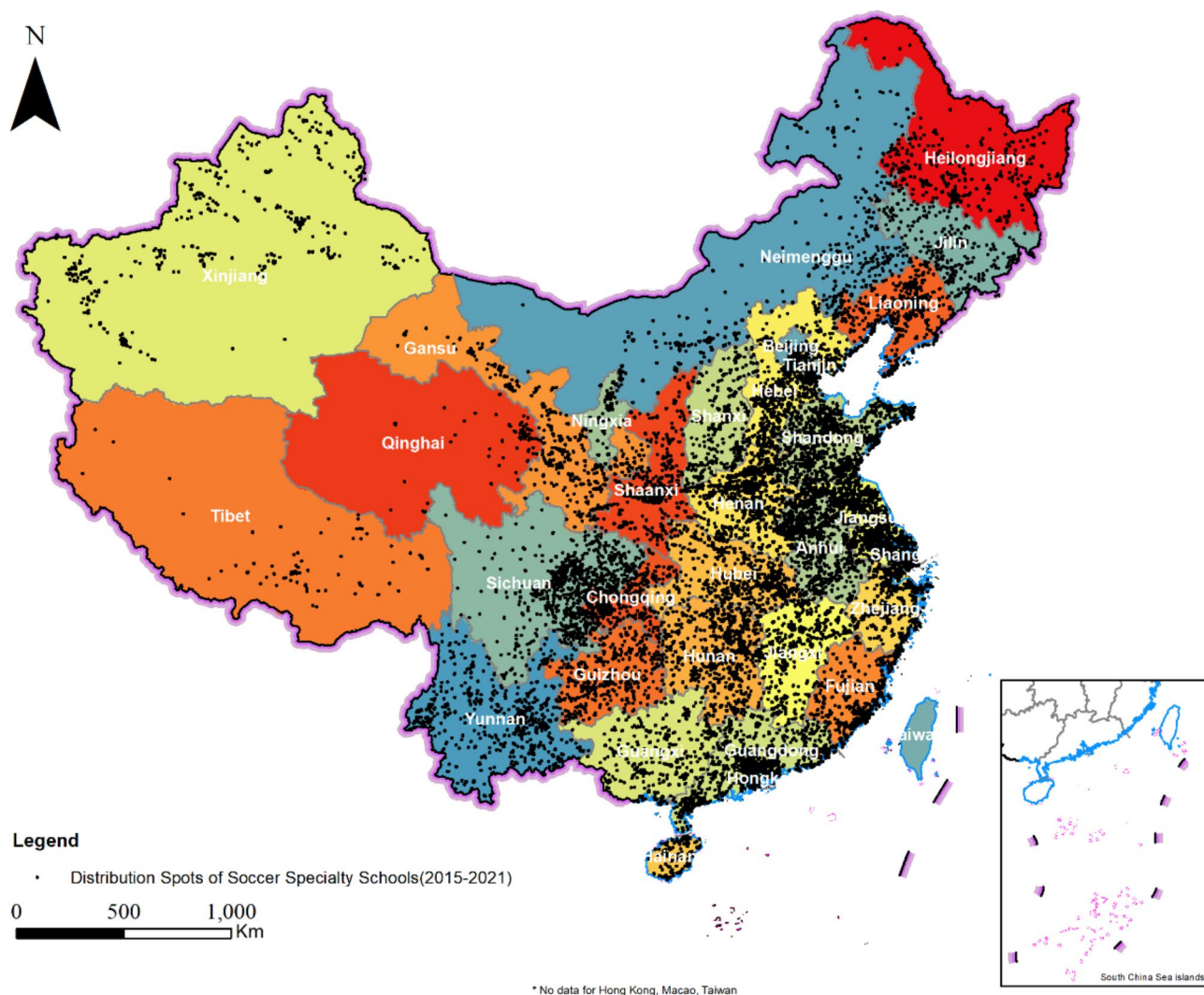


Fig. 1. Spot map of national soccer specialty school distribution from 2015–2021. Maps were drawn by authors, using ArcGIS 10.7 (Environmental Systems Research Institute, USA. <https://www.esri.com/>).

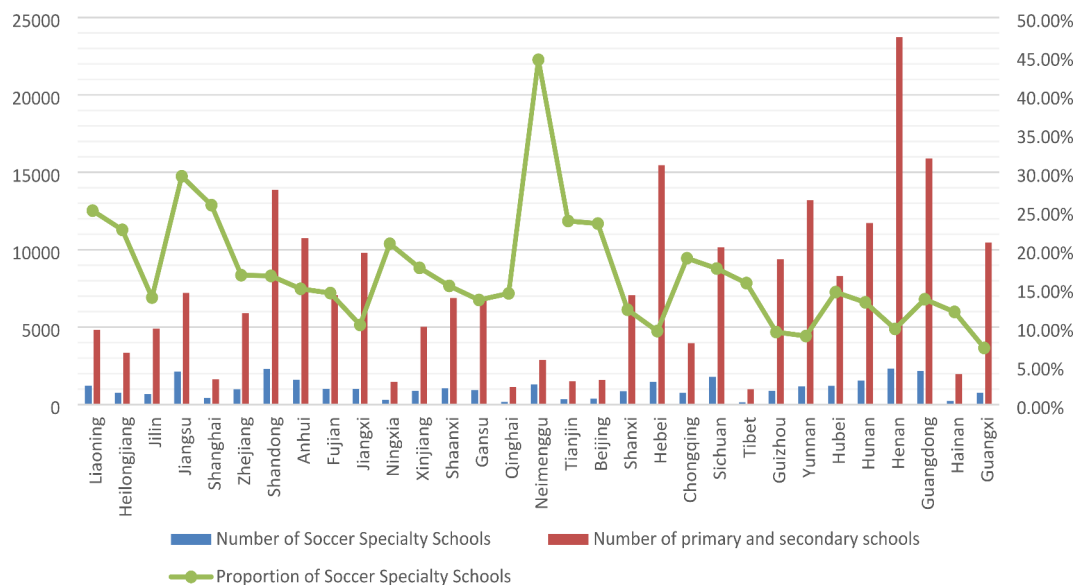


Fig. 2. Spatial distribution data for soccer specialty schools by region.

Region and Yunnan Province show the lowest proportions at 7.31% and 8.85%, respectively. Despite most provinces meeting the national standard of 6–8% for the proportion of specialty schools, there are pronounced disparities in their distribution across provinces.

Characteristics of spatial distribution of soccer specialty schools

In terms of the overall distribution of soccer specialty schools in each province (Fig. 3), there is a clear difference in the east-west distribution, using the “Hu Huanyong Line” as a boundary (the line is a stable boundary for the distribution of China’s population, with the east side of the line accounting for 43.18% of the country’s total land area, while the west side accounts for 56.82%), the east-west distribution shows apparent differences. In terms of the actual number of soccer specialty schools developed, five large concentration areas have been formed, including the concentration area in the northeast region centered on Heilongjiang Province; The Bohai Rim Concentration Area centered in Beijing, together with Tianjin, Hebei, and Shandong Provinces; The Southeast Coastal Concentration Area centered on Shanghai, together with Jiangsu and Zhejiang Provinces; The Central China Centralized Area centered on Anhui Province and constructed together with Henan, Hunan and Hubei Provinces; The Southwest China Centralized Area centered on Chongqing and created with Sichuan and Guizhou Provinces.

Subsequently, the kernel density distribution map of soccer specialty schools was obtained using the ArcGIS 10.7 spatial analysis tool for kernel density analysis (Fig. 4). The map presents the distribution status of soccer specialty school density nationwide, with several core areas of different layers, i.e., areas with a higher number of soccer specialty schools and higher density. In the high-density core area, there are eight regions with higher density and more extensive coverage, including Beijing, Tianjin, Shanghai, Chongqing, central Henan Province, central and southern Guangdong Province, central Sichuan Province, and southern Jiangsu Province, where schools with unique characteristics have formed an apparent concentration of distribution; in addition, Liaoning Province, Shandong Province, and Hunan Province have initially developed a high-density core area. However, the area is too small, and the coverage is limited. In the sub-density core areas, the six areas with higher density and broader coverage include southern Hebei Province, northern Jiangsu Province, south-central Shandong Province, northern Anhui Province, northern Zhejiang Province, eastern Hubei Province, eastern Yunnan Province, and north Jiangxi Province. In addition, the eastern part of Yunnan Province, the northern part of Jiangxi Province, the southern part of Heilongjiang Province, and the central part of coastal Fujian Province also formed sub-density zones. The coverage area was small. However, these sub-density core zones further enriched the spatial distribution pattern of the national soccer specialty schools.

Factors influencing the spatial distribution of Soccer Specialty Schools

Selection of influencing factors

Sports development is inevitably constrained by geographical and infrastructural limitations²⁶, and the spatial distribution is shaped by the interplay of critical factors such as natural geography, economic conditions, political context, and societal influences^{27–29}. Within this framework, this study will analyze and discuss the influencing factors of the spatial distribution of soccer specialty schools through a combination of qualitative and quantitative analyses, mainly from the dimensions of historical development, geographic environment, socio-economics, and educational resources, based on the whole territory of China’s 31 provincial-level administrative districts (Hong Kong, Macao, and Taiwan are not included).



Fig. 3. The trend in spatial distribution of soccer specialty schools by city. Maps were drawn by authors, using ArcGIS 10.7 (Environmental Systems Research Institute, USA. <https://www.esri.com/>).

The selection of the factor of historical development is mainly based on its crucial role in shaping the emergence and development path of schools with unique characteristics, reflecting the impact of policy evolution on schools, revealing the differences in geographic culture and social background, and providing valuable lessons learned^{30,31}. Through an in-depth analysis of historical development, the spatial distribution characteristics of schools and their influencing factors can be comprehensively understood, providing theoretical support and empirical analysis as a basis for the research.

Among the indicators in the dimension of geographic environment, a total of three indicators, namely, average altitude, average annual temperature, and annual rainfall, were selected for each province. First, the average altitude and the terrain and landscape complexity may restrict training conditions or place higher demands on students' physical fitness^{32,33}. Secondly, the average annual temperature and rainfall reflect the differences in regional climatic conditions directly related to soccer training and the regulation of students' physical condition^{34–36}. Therefore, these three indicators in the geographic environment dimension can comprehensively reflect the natural conditions of different regions, thus affecting the construction and operation of special soccer schools and having an essential impact on the spatial layout of specialty schools.

Socio-economic factors are a vital consideration in characterizing the spatial distribution of soccer specialty schools. They include three indicators: the student population, the urban population ratio, and the total GDP of each province. First, the student population reflects the size and age structure of the local population to a certain extent³⁷, directly impacting the choice of student source and scale expansion of specialty schools. Secondly, the urban population ratio reflects the degree of urbanization, and the increase in urbanization level is usually accompanied by better infrastructure and resource allocation, which is conducive to developing schools^{38,39}. Finally, the total GDP is an essential indicator of the economic strength of a region. Regions with high economic levels can usually provide better soccer training facilities, coaching teams, and financial support conducive to developing soccer specialty schools^{40,41}. Therefore, these three indicators in the socio-economic dimension can

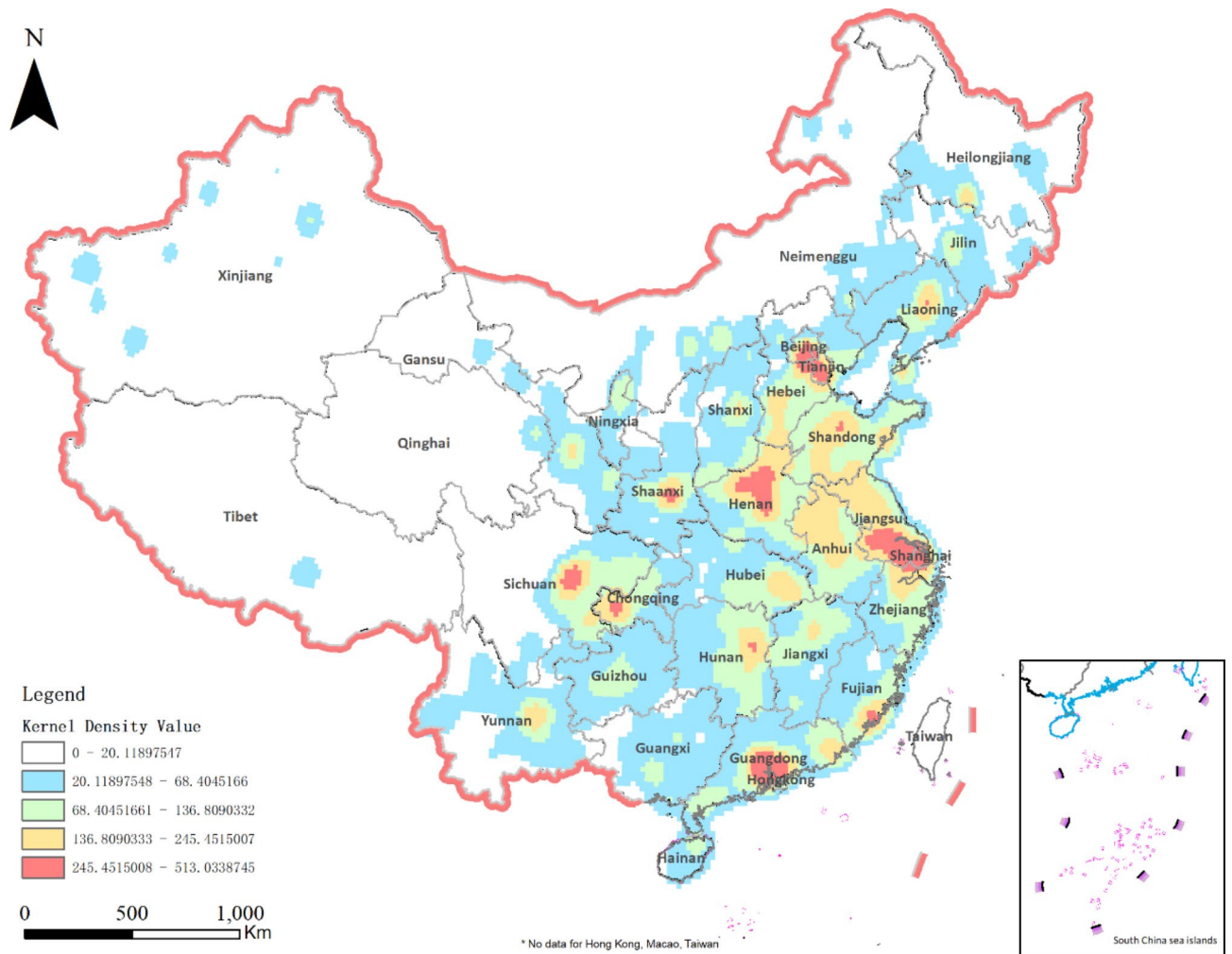


Fig. 4. Kernel density distribution of soccer specialty schools. Maps were drawn by authors, using ArcGIS 10.7 (Environmental Systems Research Institute, USA. <https://www.esri.com/>).

comprehensively reflect the economic strength and development level of different regions, which directly affects the spatial distribution pattern of specialty schools.

Educational resources are also a crucial factor. This dimension contains three indicators: the total number of school soccer fields, the area of school sports fields, and the total number of schools in each province. First, the total number of school soccer fields directly reflects the supply of campus soccer training facilities in each province, and the quantity and quality of soccer fields are crucial to the development of student soccer exercises^{42,43}. Secondly, the area of school sports fields reflects the importance schools attach to sports activities. Adequate sports grounds can help enrich students' physical education programs and campus soccer activities and promote the improvement of students' physical fitness^{44,45}. Finally, the total number of schools reflects the overall scale of educational resources in each province, and the adequacy of educational resources directly affects the construction and development of schools with soccer characteristics. Therefore, these three indicators in the dimension of educational resources can comprehensively reflect the foundation and potential of school soccer development in each province and provide an essential reference for the planning and layout of specialty schools.

Historical development

In terms of historical development, firstly, there are significant differences in the level of economic development of different regions in China, which directly impacts the spatial distribution of soccer specialty schools. Coastal regions have a higher level of economic development due to the earlier implementation of the open-door policy and industrialization and, therefore, have a clear advantage in the investment and construction of soccer specialty schools. In contrast, the economic development of inland and western regions lagged, leading to a corresponding backwardness in related resource input and school construction. Secondly, sports policies have undergone significant adjustments in different historical periods. Early policies favored elite sports development and paid insufficient attention to universal physical education.

However, there has been a gradual policy shift towards a more balanced universal access and education model since China's reform and opening up. The case of Sichuan Province stands out in this regard. Although

Sichuan Province is not a national leader in natural geography and economic conditions⁴⁶, the province has implemented in-depth reforms under the Overall Program for the Reform and Development of Soccer in China, relying on a straightforward top-level design and policy orientation. Among other things, the introduction of the Implementation Opinions on the Reform and Development of Soccer in Sichuan Province and related planning documents has significantly boosted the development of soccer in schools by incorporating it into school physical education curricula, building provincial bases, improving enrollment policies, and developing standards for reserve talents and a system for advancement to higher education. Especially in Chengdu and its neighboring regions, the number and density of specialty schools have caught up with or even surpassed some economically developed regions in several aspects.

Thirdly, the problem of unbalanced distribution of educational resources persists. Schools in provincial capitals and developed coastal areas usually have richer educational resources, including high-quality sports facilities and coaches, which provide favorable conditions for developing specialty schools in these areas.

Finally, differences in different regions' sports culture and traditions also affect the construction of specialty schools. For example, the Northeast region has been more active in constructing specialty schools because of its strong sports tradition, especially in ice and snow sports and soccer. In contrast, provinces such as Tibet and Qinghai are relatively weak in relying on sports culture to support the development of specialty schools due to their shorter exposure to soccer.

In summary, the foundation of the spatial distribution pattern of soccer specialty schools results from the intertwined effects of various historical, economic, political, and cultural factors, which collectively have shaped each region's unique development trajectories and characteristics.

Pearson correlation coefficient

In order to improve the precision of the analysis and the visualization of the data, this study used SPSS software and GeoDetector tools to quantitatively analyze the three dimensions of geographic environment, socio-economics, and educational resources. Specifically, the relationship between the average altitude, average annual temperature, annual rainfall, student population, urban population ratio, GDP, total number of school soccer fields, area of school sports fields, total number of schools, and soccer specialty schools in each province were analyzed. This analysis set the number of soccer specialty schools in each province as the dependent variable. At the same time, the rest of the indicators were used as independent variables.

The Pearson correlation coefficient method was first used to explore the linear correlation between the number of soccer specialty schools and these nine independent variables. The comprehensive analysis of Table 1; Fig. 5 found that the correlations of the four indicators, namely, average altitude, average annual temperature, urban population ratio, and total number of school soccer fields, failed to pass the 5% significance level test. The degree of linear correlation between the independent variables, from strongest to weakest, is as follows: area of school sports fields > student population > total number of schools > GDP > annual rainfall. It is worth noting that the correlation coefficients of the area of school sports fields, student population, and total number of schools are all over 0.8, showing a solid correlation. The correlation coefficients of the GDP are between 0.6 and 0.8, indicating a strong correlation. In comparison, the correlation coefficients for annual rainfall ranged between 0.4 and 0.59, showing a moderate correlation. Therefore, it can be inferred that there is a positive linear correlation between these five independent variables and the number of soccer specialty schools.

Grey relation analysis

However, there are limitations in applying the Pearson correlation coefficient, which is limited to assessing linear correlations between variables and may not be able to capture more complex interaction patterns fully. Because of this, gray correlation analysis was introduced in this study to compensate for these limitations. Grey correlation analysis is not limited to linear relationships. However, it quantifies the links between variables by assessing the geometric similarity of the data series, which can reveal the degree of association between two sets of data more comprehensively, especially in the case of non-linear relationships. Therefore, this study continues

Dimension	Indicator	Correlation coefficient		
		Pearson	Gray	Q-value (factor detection)
Geographic environment	Average altitude	− 0.276	0.807	0.150
	Average annual temperature	0.288	0.966	0.395
	Annual rainfall	0.515**	0.972	0.505
Socio-economic	Student population	0.895**	0.975	0.840
	Urban population ratio	− 0.063	0.961	0.042
	Total GDP	0.778**	0.965	0.611
Educational resources	Number of school soccer fields	0.105	0.957	0.650
	Area of school sports fields	0.907**	0.984	0.739
	Number of schools	0.814**	0.964	0.692

Table 1. The results of the analysis of Pearson's correlation coefficient, gray correlation and factor detection.

*Denotes $p < 0.05$, significant correlation; ** denotes $p < 0.01$, significant correlation; bold denotes $p > 0.05$, not significant correlation.

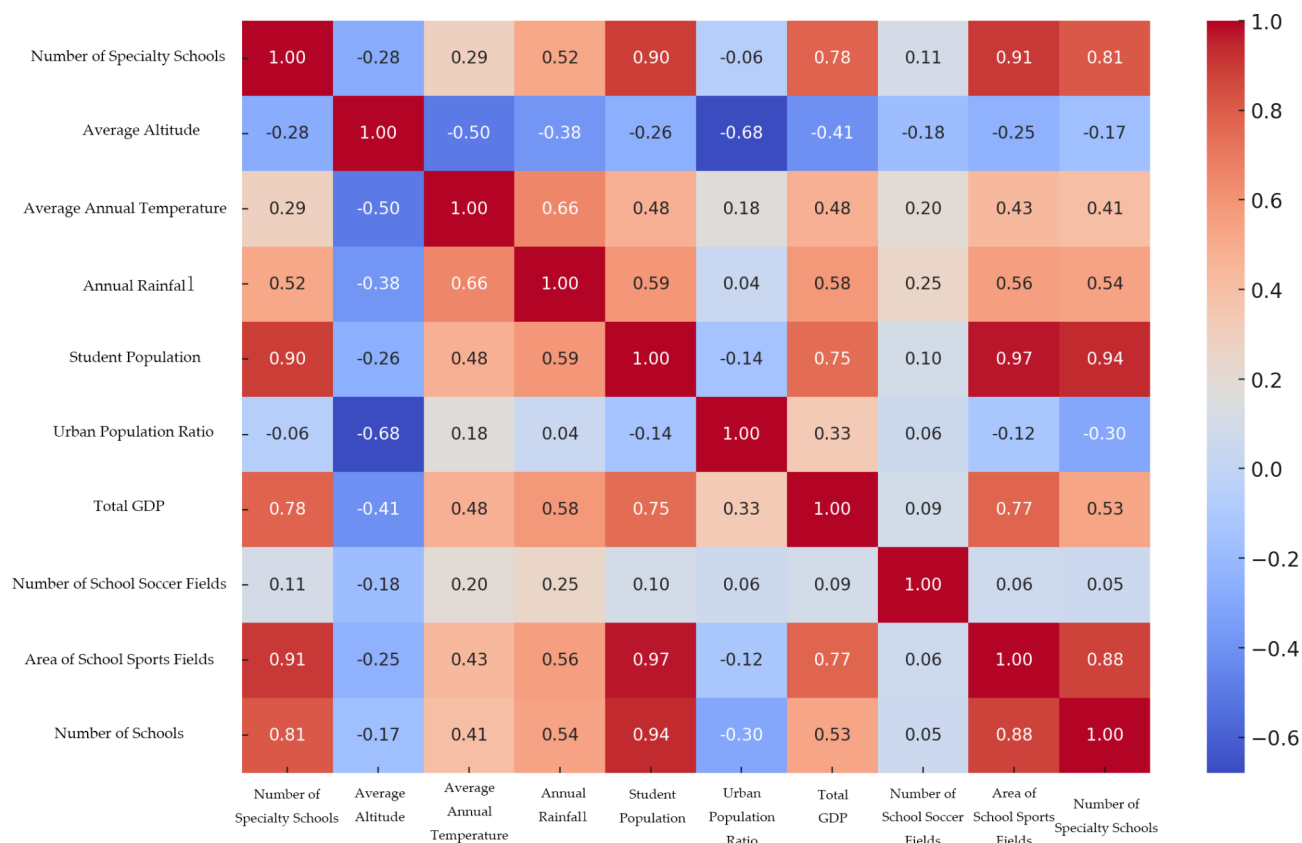


Fig. 5. Matrix of Pearson correlation coefficients for factors influencing soccer specialty schools.

to use gray correlation analysis in order to provide a more comprehensive and nuanced measure of association and, thus, to determine the interactions between variables more accurately.

According to the analysis results in Table 1, the gray correlations between the nine independent and dependent variables are generally high. The gray correlation values were similar for all indicators except for the average elevation. Compared with the Pearson correlation coefficient results, the gray correlation analysis indicates a higher correlation between these independent and dependent variables. The magnitude of correlation was in the following order: Area of school sports fields > student population > annual rainfall > average annual temperature > GDP > total number of schools > urban population ratio > total number of school soccer fields > average altitude.

Factor detection

Combining the Pearson correlation coefficient and gray correlation analysis can provide a more comprehensive data view. However, both methods ignore the issue of spatial stratified heterogeneity and fail to reflect the spatial differences characterizing the data. Therefore, a new statistical analysis method with a spatially stratified heterogeneity perspective, i.e., geo-detector, is used to analyze further the factors influencing the spatial distribution of the number of soccer specialty schools in each province.

The results of the factor probes in Table 1 show that among the nine independent variables, average altitude and urban population ratio did not pass the 5% significance test, which implies that the remaining seven independent variables are essential drivers of the spatial distribution of soccer specialty schools. Among them, the level of q value indicates to what extent these indicators influence soccer specialty schools. The explanatory power of these seven influences on the spatial distribution of soccer specialty schools in descending order is student population > area of school sports field > total number of schools > total number of school soccer fields > GDP > annual rainfall > average annual temperature. Regarding the magnitude of the coefficients, the q-values were generally lower than the scores of the Pearson correlation coefficient and the gray correlation after taking spatial heterogeneity into account, which indicated that the degree of correlation between the variables was corrected after full consideration of the heterogeneity.

Although the coefficients are different, the closeness of the relationship between these seven independent variables and the dependent variable has not changed significantly. From the perspective of factor explanatory power, educational resources can be regarded as the most critical factor in soccer specialty school selection, determining the number of soccer specialty schools in the region. Next is the socio-economic level. Population and GDP are important indicators of the economic level of a region, and the number of soccer specialty schools in the region depends mainly on the local socio-economic level. On the other hand, geography has a relatively low degree of influence on soccer specialty school selection compared to the first two.

Interaction detector

Based on the results of the factor detection, this study further used Geodector's Interaction Detector method to systematically analyze the interaction of different factors on the spatial distribution of soccer specialty schools. The results show that the interaction between the nine factors forms 36 pairs (see Table 2). Twenty-five pairs of factors show a double factor enhancement effect, that is, the explanatory power of the two factors on the spatial distribution is significantly improved after interaction; the remaining 11 pairs of factors show a non-linear enhancement effect, although there is an enhancement effect, the amplitude is relatively limited. This shows that the explanatory power of any two independent variables on the spatial distribution of the number of soccer specialty schools after interaction exceeds that of their individual effects, further revealing the complexity of the spatial distribution of soccer specialty schools being affected by multiple factors.

From the average q value of the three dimensions, educational resources ($q=0.87$) > socio-economic ($q=0.85$) > geographical environment ($q=0.80$), which is consistent with the previous analysis. However, socio-economic factors also play an essential role. The Interaction Detector shows that the interaction between GDP and the area of sports fields is the strongest ($q=0.9662$), followed by the interaction between the student population and GDP ($q=0.9573$), and the interaction between the student population and the urban population ratio ($q=0.9445$). In addition, the average q value of the interaction between the student population and other factors is also the highest, indicating the significant role of socio-economic factors in interacting with other factors, reflecting its critical position in influencing the spatial distribution of soccer specialty schools.

Based on the above factor detection results, all nine influencing factors have significantly impacted the spatial distribution of soccer specialty schools, especially when interacting with the student population factor, significantly enhancing this influence. This further indicates that differences in student populations in different regions mainly drive the spatial distribution of soccer specialty schools. However, overall, the number of soccer specialty schools in different regions still highly depends on the availability and level of development of educational resources. In other words, although the student population is an essential factor in determining the distribution of soccer specialty schools, the availability and level of development of educational resources are more critical in determining the final distribution of soccer specialty schools.

Discussion

This study comprehensively analyzes the spatial distribution characteristics of China's campus soccer specialty schools from 2015 to 2021. The results show that as the number of soccer specialty schools increases, their spatial distribution shows significant regional differences. Using the "Hu Line" as a boundary, the eastern region has formed a dense distribution pattern. In contrast, the western region is relatively sparse. At the same time, soccer specialty schools have further concentrated in key cities and principal economic belts, eventually forming eight high-density core areas centered in Beijing and Henan Province. In addition, this study systematically explores the multiple factors that influence the spatial distribution of these schools, revealing the critical roles played by educational resources, socio-economic conditions, and geographical environment. These findings enrich existing sports geography research and provide an essential reference for formulating relevant policies.

The study used a multi-dimensional analytical approach, including Pearson correlation coefficient analysis, grey relational analysis, and the GeoDetector tool, to identify interactions between complex factors that may be overlooked by a single analytical method. A comparison with existing literature shows that traditional studies often focus on the impact of geographical environment and socio-economic conditions on the distribution of educational resources^{47–49}. In contrast, this study systematically reveals for the first time the central role of educational resources in determining the number and distribution of soccer specialty schools. Based on the results, educational resources are more influential in distributing unique educational resources, such as soccer specialty schools, than the traditionally considered socio-economic and geographical factors. Although socio-economic conditions and geographical environment play a fundamental role in allocating educational resources, the availability and quality of educational resources directly impact the final spatial distribution. This finding provides empirical evidence for policymakers to prioritize the construction of educational infrastructure in resource allocation.

Factors	Average altitude (m)	Average annual temperature (°C)	Rainfall (mm)	Student population	Urban population ratio	Total GDP (billions of yuan)	Number of soccer fields in schools	Area of sports fields (square meters)	Total number of schools
Average altitude (m)	0.15								
Average annual temperature (°C)	0.906432	0.395954							
Rainfall (mm)	0.820025	0.666419	0.505451						
Student population	0.904519	0.89452	0.910141	0.840194					
Urban population ratio	0.367774	0.783882	0.637252	0.944534	0.042765				
Total GDP (billions of yuan)	0.900688	0.751546	0.875394	0.957341	0.872221	0.611484			
Number of soccer fields in schools	0.838972	0.824428	0.787278	0.908504	0.78973	0.829963	0.650478		
Area of sports fields (square meters)	0.842304	0.795387	0.922384	0.904729	0.870784	0.966205	0.933844	0.73903	
Total number of schools	0.876781	0.836642	0.88877	0.889818	0.91658	0.929466	0.919175	0.838882	0.692742

Table 2. The results of the interaction detector. Bold denotes non-linear enhancement.

In addition, the study re-evaluates the influence of geographical factors. Although geographical factors are considered an essential factor affecting the distribution of educational facilities in many cases^{50–53}, their role in the context of soccer specialty schools appears to be more complex and multifaceted. Some previous studies have shown that factors such as topography, climate, and natural resources significantly affect the construction of educational infrastructure, especially in areas with complex topography or harsh climate conditions^{54–56}. For instance, establishing schools or soccer fields presents significant challenges in mountainous areas like Qinghai Province and Tibet or regions with frequent extreme weather. These areas require additional infrastructure investments and specialized planning to overcome natural obstacles. However, this study reveals that, although geographical challenges exist, they are not insurmountable. On the contrary, through strategic policy interventions and targeted resource allocation, the impact of these challenges can be mitigated. The case of Sichuan Province exemplifies this approach. Despite its less favorable natural geography⁴⁶—characterized by rugged terrain and underdeveloped economic conditions—Sichuan has successfully advanced the development of soccer specialty schools through comprehensive reforms and effective policy implementation, achieving a leading position nationwide in the number and prevalence of these schools. Such success stories underscore the importance of top-level design and policy orientation, which can effectively alleviate geographical disadvantages by fostering an environment conducive to developing sports education.

By analyzing the interaction of various factors, the study revealed the importance of the interaction between educational resources and socio-economic factors in explaining the distribution of soccer specialty schools. In particular, the combination of student population and educational resources significantly enhanced the explanatory power of the distribution pattern of soccer specialty schools. The spatial distribution of soccer specialty schools results from a single factor and the product of multiple factors' complex interaction. Therefore, policymakers should consider these factors when planning and allocating educational resources to achieve a more balanced and effective allocation of resources.

Based on the research results, policymakers should pay more attention to the fair distribution of educational resources and consider the impact of socio-economic conditions to reduce the imbalance of resources between regions and promote the balanced development of soccer specialty schools. Future research should further explore the long-term impact of these factors, especially in the context of educational resource allocation and policy changes. In addition, comparing the results of the Chinese study with similar projects in other countries or regions can also provide a valuable reference for educational policy on a global scale.

In summary, this study reveals the spatial distribution characteristics of soccer specialty schools in China and their key influencing factors through multi-dimensional analysis. It further emphasizes educational resources' vital role in this process compared with previous studies. The research results provide a practical reference framework for policy formulation and make the following suggestions: First, adhere to policy orientation, strengthen top-level design, adjust and optimize policies promptly, and control the spatial distribution of soccer specialty schools nationwide from a macro perspective; Second, strengthen spatial awareness, deepen the construction of soccer specialty schools in high-density areas, strengthen regional and inter-school cooperation, form more core areas and development zones, and build a sustainable development model for special schools; Third, governments at all levels need to pay attention to the rational allocation of educational resources, especially in the relatively scarce inland and western regions, and should appropriately tilt policies to ensure that special schools move towards more comprehensive and balanced development.

Limitation

In the study of the spatial distribution of specialty schools, as the project is scheduled to end in 2025, the data of this study is not complete, which will require us to continue the study in the future by adding data on an ongoing basis.

In addition, the factors contributing to the distribution of specialty schools may be complex and multifaceted. In analyzing the factors influencing the spatial distribution of specialty schools, we selected nine factors from the three dimensions of geography, socio-economics, and educational resources to analyze, such as the average altitude, the school population, and the number of soccer fields. Meanwhile, factors such as policy-making, spatial accessibility, financial support, and teacher resources in the region are also important but were not included in this study due to the lack of detailed official data and the unavailability of some of the data. We should conduct continuous and in-depth research on these issues in future studies.

Conclusion

This study systematically analyzed the spatial distribution characteristics of China's campus soccer specialty schools from 2015 to 2021, revealing significant regional differences and concentration trends in these schools nationwide. The study shows that the distribution of soccer specialty schools is significantly uneven on both sides of the Hu Huanyong Line, with a dense distribution pattern in the eastern region and a relatively sparse distribution in the western region. At the same time, soccer specialty schools are further concentrated around critical cities and principal economic belts, eventually forming five major concentration areas in the Bohai Rim, the southeast coast, etc., as well as eight high-density core areas centered on Beijing, Henan Province, etc. Through multi-dimensional analysis, this study confirms the decisive role of educational resources, socio-economic conditions, and geographical environment in the spatial distribution of schools, especially the availability and quality of educational resources that play a central role in the distribution of special projects such as soccer specialty schools. This finding provides a new perspective on the importance of educational resources in promoting educational equity and optimizing spatial distribution. In addition, the study also shows that the interaction between educational resources and socio-economic conditions is critical to explaining the distribution pattern of soccer specialty schools. These findings provide a new theoretical perspective for understanding the spatial distribution of soccer specialty schools in China and empirical evidence for policymakers in allocating

resources and policy design for similar projects. Future research should further explore the role of these factors in other educational projects and reveal general laws and particularities in different policy environments through international comparative research to provide scientific support for education policy formulation worldwide.

Methods

Data source

The data sources for this study are extensive, covering data on the distribution of soccer specialty schools in China, geographic information, natural conditions, and socio-economic background to ensure the analysis's comprehensiveness and the data's accuracy. The specific data and sources are as follows: ① Data on Chinese soccer specialty schools and related data, including the names, numbers, and soccer field numbers of soccer specialty schools, as well as the area of sports fields on campus, are all from the official website of the Ministry of Education of the People's Republic of China, covering the period from 2015 to 2021. The data collection time was December 2022, and these data provide a solid foundation for studying the spatial distribution and resource allocation of youth soccer specialty schools across the country. ② The administrative division data used by the "Administrative Divisions of the People's Republic of China" (2019 edition) published on the official website of the Central People's Government of the People's Republic of China. This data ensures the accuracy and consistency of the study's administrative divisions of provinces, cities, and counties. ③ Map of China and natural geographical data, including map data of China (standard map review number GS (2019) 1823), average elevation, average annual temperature, and annual precipitation, are all from the Ministry of Natural Resources of the People's Republic of China. The data has not been modified in any way to ensure the accuracy of the geographical information and is suitable for spatial analysis and visualization. ④ The geospatial coordinate data of soccer specialty schools was obtained through Baidu Maps and uniformly converted to WGS-84 format to ensure the consistency of the coordinate system and the applicability of international standards. Baidu Maps was selected based on its wide application and high accuracy in China, especially regarding the accuracy of geographic coordinates and spatial locations. Compared with other international map services, Baidu Maps better reflects local conditions and complies with the relevant regulations on geographic information services in China. ⑤ Socio-economic data, including critical indicators such as GDP, population, and urban population ratio, are from each province and city's official Statistical Yearbook (2021). These data support studying and analyzing the relationship between educational resources and socio-economic conditions.

Number	Data name	Data source
①	Chinese soccer specialty schools and related data	Ministry of Education of the People's Republic of China (http://www.moe.gov.cn/)
②	Administrative division data for each region	Central People's Government of the People's Republic of China (https://www.gov.cn/)
③	Map and physical geography data	Ministry of Natural Resources of the People's Republic of China (https://www.mnr.gov.cn/)
④	Geospatial coordinate data for soccer specialty schools	Baidu Map (https://map.baidu.com/)
⑤	Socio-economic data	China Statistical Yearbook (2021), etc.

Spatial data analysis

GIS spatial analysis is a research method based on geographical location and spatial characteristics. It is widely used in spatial data calculation and thematic map compilation^{57–59}. This study used GIS technology to process and analyze the spatial data of China's youth soccer specialty schools to ensure the scientific and accurate nature of the analysis. All spatial data processing and analysis were completed using ArcGIS 10.7, and the application of GIS technology in the study covered data standardization, geocoding, and spatial visualization.

First, we used GIS technology to standardize the geographic coordinates of the soccer specialty schools from Baidu Maps. All coordinates were converted to the WGS-84 coordinate system to ensure the data's universality and the spatial analysis's accuracy. In addition, geocoding technology was used to process complex address data, converting the text addresses of the schools into precise spatial coordinates, allowing the study to be analyzed at a more detailed geographical scale.

In terms of spatial visualization, we used GIS technology to generate multi-level spatial distribution maps. These maps show the distribution patterns of soccer specialty schools in different regions of the country, providing a wealth of spatial background information. These visualizations allow the study to present regional differences and concentration trends, laying a solid foundation for subsequent quantitative analysis. By applying GIS technology, this study was able to efficiently process large amounts of complex spatial data, accurately describe the geographical distribution characteristics of soccer specialty schools, and provide strong support for further spatial analysis. The application of these technologies ensures the scientific nature and reproducibility of the research results, providing a solid basis for formulating relevant policies.

Kernel density analysis (KDA)

The kernel density estimation method is a commonly used method to assess the characteristics of the density distribution of the study elements in the overall space, which can reflect the degree of agglomeration and the location of the agglomeration of the study elements from the layers and its mathematical expression is^{60–62}:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x-x_i}{h}\right)$$

where $k(x)$ is the kernel function; h is the bandwidth; and $(x-x_i)$ denotes the distance from the valuation point x to the featured school x_i . The larger the value of $f(x)$, the denser the distribution of featured schools.

Pearson correlation coefficient

This study employed the Pearson Correlation Coefficient analysis to explore the relationship between the spatial distribution of Chinese Soccer Specialty Schools and various socio-economic and geographical factors. The Pearson Correlation Coefficient is a statistical method used to measure the linear correlation between two continuous variables, calculated using the following formula^{63–66}:

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}}$$

where r represents the correlation coefficient, x_i and y_i are the observed values of the two variables, \bar{x} and \bar{y} are the means of the two variables. The value of the correlation coefficient r ranges from $[-1, 1]$, with $r > 0$ indicating a positive correlation, $r < 0$ indicating a negative correlation, and $r = 0$ indicating no linear correlation.

This study first standardized all variables to ensure comparability across different scales. The Pearson Correlation Coefficients between the number of soccer specialty schools and various factors were then calculated using SPSS 26.0 software to quantify the influence of these factors on the distribution of schools. The results are presented in a correlation matrix, with significance levels (p-values) used to determine the strength of the correlations. Correlations with a significance level below 0.05 were considered statistically significant.

By integrating Pearson Correlation Coefficient analysis with GIS technology, this study provides a comprehensive exploration of the complex factors influencing the spatial distribution of soccer specialty schools, offering scientific insights for policy-making and resource allocation.

Grey relation analysis (GRA)

To thoroughly investigate the relationship between the spatial distribution of Chinese Youth Soccer Specialty Schools and various socio-economic and natural geographical factors, this study employed the Grey Relation Analysis (GRA) method. GRA is a quantitative analysis tool based on systems theory, capable of effectively revealing the degree of association between factors within a multi-dimensional and complex system^{67–69}.

In this study, GRA was used to measure the impact of various influencing factors, such as GDP, population density, average altitude, and annual temperature, on the spatial distribution of soccer specialty schools. First, all variables were standardized to eliminate the influence of different units and scales on the analysis results. Subsequently, the grey relational coefficients and grey relational degrees were calculated to quantify the impact of these factors on the distribution of schools.

The grey relational coefficient is calculated using the following formula⁷⁰:

$$\xi(x_0, x_i) = \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \cdot \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \cdot \max_i \max_k |x_0(k) - x_i(k)|}$$

where $\xi(x_0, x_i)$ represents the grey relational coefficient between the reference sequence x_0 and the comparative sequence x_i at the k th sample, and ρ is the distinguishing coefficient, typically set at 0.5. The grey relational degree, the average of all grey relational coefficients, reflects the overall association between each factor and the system's behavior.

By applying grey relational analysis, the factors were effectively ranked, allowing the identification of those most significantly influencing the spatial distribution of soccer specialty schools. The results of the grey relational analysis provided detailed quantitative insights, revealing the comprehensive effects of multiple factors on the spatial distribution of soccer specialty schools and offering a robust scientific foundation for policy-making.

GeoDetector

Leveraging the spatial stratified heterogeneity detection technique, GeoDetector emerges as a robust spatial analytical tool designed to assess the concordance between the spatial distributions of dependent variables and their respective explanatory factors^{71,72}. The foundational principle of GeoDetector is its ability to dissect spatial variance without being confounded by multicollinearity among explanatory variables. This makes it an indispensable methodology widely employed across various disciplines, including regional economics, spatial planning, and environmental science. Its application facilitates a nuanced understanding of geospatial dynamics, contributing significantly to the precision of regional analyses and the strategic development of context-specific interventions^{72,73}.

In this study, we used the Factor Detection function of GeoDetector to explore the factors affecting the spatial distribution of soccer specialty schools. The q-statistic, denoted as q, represents the explanatory power of each factor. It is calculated as^{74,75}:

$$q = 1 - \frac{\sum_{h=1}^L N_h \sigma_h^2}{N \sigma^2}$$

where h represents each strata or subregion within the study area. N_h and N are the number of samples in strata h and the entire study area, respectively. σ_h^2 is the variance within strata h . σ^2 is the variance of the entire study area. The q-statistic ranges from 0 to 1, with higher q-values indicating that the factor explains a more significant proportion of the spatial variance in the distribution of soccer specialty schools. For example, a q-value of 0.8 suggests that 80% of the spatial variation can be attributed to the specific factor under consideration.

To analyze the combined effects of multiple factors on the spatial distribution of soccer specialty schools, this study employed the Interaction Detection method within the GeoDetector framework. Interaction Detection assesses how the interaction between two factors influences their combined contribution to spatial heterogeneity⁷⁶. By comparing the q-value of the interaction between two factors to the individual q-values of each factor, the interaction can be categorized as Bivariate Enhancement, Non-linear Enhancement, Independence, or Bivariate Weakening. These interactions are defined as follows: Bivariate Enhancement^{76–78}: Occurs when the q-value of the interaction between two factors exceeds the sum of their q-values, indicating that their combined effect on spatial distribution is synergistic. Non-linear Enhancement: Identified when the q-value of the interaction is greater than the higher of the individual q-values but less than their sum, suggesting a significant but not fully additive interaction. Independence: Observed when the combined q-value equals the sum of the individual q-values, implying that the factors do not influence each other's effects. Bivariate Weakening occurs when the combined q-value is less than the sum of the individual q-values, indicating that the interaction reduces the overall impact of the factors.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Received: 3 April 2024; Accepted: 7 October 2024

Published online: 15 October 2024

References

- Liu, Y., Dong, B. & Zeng, X. Healthy and sustainable development of sports economy based on artificial intelligence and mental model. *Front. Psychol.* **13**, 1 (2022).
- State Council Office of the People's Republic of China. Overall Program for the Reform and Development of Chinese Soccer. (2015). https://www.gov.cn/zhengce/content/2015-03/16/content_9537.htm
- Li, G. *Analysis on the Current Situation of the Development of Chibi's Primary and Secondary School Football*, 90–104. <https://doi.org/10.2991/erms-18.2018.20> (Atlantis Press, 2018).
- Deng, X., Liang, L., Wu, F., Wang, Z. & He, S. A review of the balance of regional development in China from the perspective of development geography. *J. Geogr. Sci.* **32**, 3–22 (2022).
- Long, H., Tu, S., Ge, D., Li, T. & Liu, Y. The allocation and management of critical resources in rural China under restructuring: problems and prospects. *J. Rural Stud.* **47**, 392–412 (2016).
- Gao, Z., Tan, N., Geddes, R. R. & Ma, T. Population distribution characteristics and spatial planning response analysis in Metropolises. *Int. Rev. Spat. Plann. Sustain. Dev.* **7**, 134–154 (2019).
- Rensmann, L. Sports, global politics, and social value change: A research Agenda. *Altre modernità/ otras modernidades/ Autres modernités/ Other Mod.* **14**, 114–133 (2014).
- Waalkes, S. Does soccer explain the world or does the world explain soccer? Soccer and globalization. In *The State of the Field* (Routledge, 2018).
- Hoffmann, R., Ging, L. C. & Ramasamy, B. The socio-economic determinants of international soccer performance. *J. Appl. Econ.* **5**, 253–272 (2002).
- Garcia-del-Barrio, P. & Pujol, F. Hidden monopsony rents in winner-take-all markets—sport and economic contribution of Spanish soccer players. *Manag. Decis. Econ.* **28**, 57–70 (2007).
- Dolci, F. et al. Movement economy in soccer: current data and limitations. *Sports.* **6**, 124 (2018).
- Giulianotti, R. Playing an aerial game: the new political economy of soccer. In *The Political Economy of Sport* (eds. Nauright, J. & Schimmel, K. S.) 19–37. https://doi.org/10.1057/9780230524057_2 (Palgrave Macmillan UK, 2005).
- Xu, J., Gao, C. & Zhao, J. National football promotion in China: opportunities and challenges in public health. *J. Sport Health Sci.* **5**, 250–251 (2016).
- Peng, Q., Skinner, J. & Houlihan, B. An analysis of the Chinese Football Reform of 2015: why then and not earlier? *Int. J. Sport Policy Politics* (2019).
- Stylianou, M., Hogan, A. & Enright, E. Youth sport policy: the enactment and possibilities of 'soft policy' in schools. *Sport Educ. Soc.* (2019).
- Li, X. Education in China-Unbalanced Educational Development caused by Regional Differences: Taking Gansu and Beijing (2010–2015) from economic perspective as examples. *J. Educ. Humanit. Soc. Sci.* **8**, 1441–1448 (2023).
- Liu, H. et al. Epidemiology of campus football injuries in Ningxia, China: occurrence, causes, and management. *Front. Public Health* **10**, (2022).
- Fan, X. & Sungkawadee, P. Construction of the evaluation index system for the sustainable development of primary school campus football in Shaanxi Province, China. *Int. J. Sociol. Anthropol. Sci. Rev.* **3**, 201–214 (2023).
- Zhang, J. & Gu, W. Study on the present situation and countermeasure of campus football development in Yunnan Border Area. *Acad. J. Sci. Technol.* **6**, 139–142 (2023).
- Peng, Q., Chen, Z., Li, J., Houlihan, B. & Scelles, N. The new hope of Chinese football? Youth football reforms and policy conflicts in the implementation process. *Eur. Sport Manag. Q.* **23**, 1928–1950 (2023).

21. Sullivan, J., Chadwick, S. & Gow, M. China's football dream: sport, citizenship, symbolic power, and civic spaces. *J. Sport Soc. Issues*. <https://doi.org/10.1177/0193723519867588> (2019).
22. Zhang, Y., Ming, Y., Shi, B. & Y. & Spatial distribution characteristics and causes of public sports venues in China. *Sci. Rep.* **13**, 15056 (2023).
23. Chen, Y. et al. Spatial equity in the layout of urban public sports facilities in Hangzhou. *PLoS One*. **16**, e0256174 (2021).
24. Jing, Y. et al. Decoding the street-based spatiality of urban gyms: implications for healthy city planning. *Land*. **10**, 175 (2021).
25. Chen, B. et al. Research on the spatial distribution characteristics and influencing factors of Educational facilities based on POI Data: a case study of the Guangdong–Hong Kong–Macao Greater Bay Area. *ISPRS Int. J. Geo-Inf.* **13**, 225 (2024).
26. Mason, D. Sport cities, and communities. In *The Oxford Handbook of Sport and Society* (ed. Wenner, L. A.) <https://doi.org/10.1093/oxfordhb/9780197519011.013.15> (Oxford University Press, 2022).
27. Jiang, L., Chen, J., Tian, Y. & Luo, J. Spatial pattern and influencing factors of basic education resources in rural areas around metropolises—A case study of Wuhan City's New Urban districts. *ISPRS Int. J. Geo-Inf.* **11**, 576 (2022).
28. Zhang, Z., Luan, W., Tian, C., Su, M. & Li, Z. Spatial distribution equilibrium and relationship between construction land expansion and basic education schools in Shanghai based on POI data. *Land*. **10**, 1059 (2021).
29. Li, Y., Yuan, Y. & Cheng, N. Analysis on spatial difference and spillover effect of development resilience index of sports industry: a case study of 285 cities in China. *PLoS One*. **18**, e0295313 (2023).
30. Buhaş, S., Buhaş, R., Grama, V., Herman, G. & Drago, P. The impact of local public policies on sports performance. Oradea—An example of success. *Baltic J. Health Phys. Act.* **13**, 25 (2022).
31. Malchrowicz-Moško, E., Pocza, J. A. Small-scale event and a big impact—Is this relationship possible in the World of Sport? The meaning of heritage sporting events for sustainable development of tourism—experiences from Poland. *Sustainability*. **10**, 4289 (2018).
32. Fan, C. et al. The cardiorespiratory fitness of children and adolescents in Tibet at altitudes over 3,500 meters. *PLoS One*. **16**, e0256258 (2021).
33. Chang, Z., Ye, H., Huang, B. & Xu, J. Physical fitness of tibetan adolescents at different high altitudes in Tibet, China: a comparative study. *Am. J. Hum. Biol.* **36**, e23975 (2024).
34. Beghin, L. et al. Influence of meteorological conditions on physical activity in adolescents. *J. Epidemiol. Community Health*. **74**, 395–400 (2020).
35. Zheng, C., Feng, J., Huang, W. & Wong, S. H. S. Associations between weather conditions and physical activity and sedentary time in children and adolescents: a systematic review and meta-analysis. *Health Place*. **69**, 102546 (2021).
36. Draper, G. et al. Do environmental temperatures and altitudes affect physical outputs of elite football athletes in match conditions? A systematic review of the 'real world' studies. *Sci. Med. Footb.* (2023).
37. Apatov, E. & Grimes, A. Impacts of higher education institutions on local population and employment growth. *Int. Reg. Sci. Rev.* **42**, 31–64 (2019).
38. Lan, F., Gong, X., Da, H. & Wen, H. How do population inflow and social infrastructure affect urban vitality? Evidence from 35 large- and medium-sized cities in China. *Cities*. **100**, 102454 (2020).
39. Tripathi, S. Does smart city development promote urbanization in India? In *Smart Cities—Their Framework and Applications*. <https://doi.org/10.5772/intechopen.94568> (IntechOpen, 2021).
40. Black, N., Johnston, D. W., Propper, C. & Shields, M. A. The effect of school sports facilities on physical activity, health and socio-economic status in adulthood. *Soc. Sci. Med.* **220**, 120–128 (2019).
41. Thompson, F., Rongen, F., Cowburn, I. & Till, K. A case study of the features and holistic athlete impacts of a UK sports-friendly school: Student-athlete, coach and teacher perspectives. *PLoS One*. **17**, e0278401 (2022).
42. Wang, Y., Wu, K. & Gao, W. Analysis on spatial distribution pattern of football fields: a case study in Wuhan, China. *E3S Web Conf.* **293**, 02051 (2021).
43. Gao, A., Chen, Z., Li, S., Zhang, C. & Quan, Z. Exploring the development of Chinese campus football from the perspective of health. *Front. Sport Res.* **3** (2021).
44. Peralta, M. et al. Promoting health-related cardiorespiratory fitness in physical education: the role of lesson context and teacher behavior in an observational longitudinal study. (2021).
45. Cocca, A., Espino Verdugo, F., Ródenas Cuenca, L. T. & Cocca, M. Effect of a game-based physical education program on physical fitness and mental health in elementary school children. *Int. J. Environ. Res. Public Health*. **17**, 4883 (2020).
46. He, X., Mai, X. & Shen, G. Poverty and physical geographic factors: an empirical analysis of Sichuan Province using the GWR model. *Sustainability*. **13**, 100 (2021).
47. Brownie, S., Yan, A. R., Broman, P., Comer, L. & Blanchard, D. Geographic location of students and course choice, completion, and achievement in higher education: a scoping review. *Equity Educ. Soc.* <https://doi.org/10.1177/27526461231200280> (2023).
48. Chand, D. D. & Mohan, P. Impact of school locality on teaching and learning: a qualitative inquiry. *Waikato J. Educ.* **24**, 65–72 (2019).
49. Banerjee, S. et al. Transforming social determinants to educational outcomes: geospatial considerations. *Healthcare*. **10**, 1974 (2022).
50. Yan, Q., Li, C., Zhang, J., Ma, Z. & Luo, F. Uneven distribution of primary and middle schools: a case study of Changchun, China. *Front. Educ. China*. **13**, 633–655 (2018).
51. Huang, Q., Cui, X. & Ma, L. The equity of basic educational facilities from the perspective of space. *Sustainability*. **15**, 12031 (2023).
52. Murad, A. A., Dalhat, A. I. & Naji, A. A. Using geographical information system for mapping public schools distribution in Jeddah City. *Int. J. Adv. Comput. Sci. Appl. (IJACSA)* **11** (2020).
53. Xu, Y., Song, W. & Liu, C. Social-spatial accessibility to urban educational resources under the school district system: a case study of public primary schools in Nanjing, China. *Sustainability*. **10**, 2305 (2018).
54. Zhu, Y. et al. Accessibility of primary schools in rural areas and the impact of topography: a case study in Nanjiang County, China. *Land*. **12**, 1134 (2023).
55. Yao, L., Lv, M., Li, T., Wang, D. & Cao, X. Exploring the evolution of the accessibility of Educational Facilities and its influencing factors in mountainous areas: a case study of the Rocky Desertification Area in Yunnan, Guangxi, and Guizhou. *ISPRS Int. J. Geo-Inf.* **11**, 296 (2022).
56. Yang, Z., Hong, Y., Guo, Q., Yu, X. & Zhao, M. The impact of topographic relief on population and economy in the Southern Anhui Mountainous Area, China. *Sustainability*. **14**, 14332 (2022).
57. Awawdeh, M. M. et al. Urban expansion in greater irbid municipality, Jordan: the spatial patterns and the driving factors. *GeoJournal*. **89**, 43 (2024).
58. Keroglou, I. & Tsoutsos, T. Optimal siting of solar desalination plants in Crete, Greece employing a GIS/MCDM approach. *Renew. Energy*. **224**, 120168 (2024).
59. Dudar, I. et al. Enhancing urban solid waste management through an integrated geographic information system and multicriteria decision analysis: a case study in postwar reconstruction. In *Handbook on Post-War Reconstruction and Development Economics of Ukraine: Catalyzing Progress* (ed. Cirella, G. T.) 377–392. https://doi.org/10.1007/978-3-031-48735-4_21 (Springer International Publishing, 2024).
60. Xiong, S., Wang, H., Liao, Z. & Hashim, R. Exploring the factors and spatial patterns of national night cultural tourism consumption agglomeration zones in China. *Heliyon*. **10**, e24132 (2024).

61. Guo, M., Han, C., Guan, Q., Huang, Y. & Xie, Z. A universal parallel scheduling approach to polyline and polygon vector data buffer analysis on conventional GIS platforms. *Trans. GIS*. **24**, 1630–1654 (2020).
62. Tan, R. & Zhou, K. Examining the causal effects of road networks on landscape ecological risk: the case of Wuhan, China. *J. Urban. Plan. Dev.* **149**, 05023004 (2023).
63. Kalogirou, S. Testing local versions of correlation coefficients. *Jahrb Reg. Wiss.* **32**, 45–61 (2012).
64. Chen, Y. A new methodology of spatial cross-correlation analysis. *PLoS One*. **10**, e0126158 (2015).
65. Chatterjee, S. A new coefficient of correlation. *J. Am. Stat. Assoc.* (2021).
66. Dufera, A. G., Liu, T. & Xu, J. Regression models of Pearson correlation coefficient. *Stat. Theory Relat. Fields* (2023).
67. Wu, H. H. A comparative study of using grey relational analysis in multiple attribute decision making problems. *Qual. Eng.* **15**, 209–217 (2002).
68. He, D. et al. Temporal and spatial characteristics of EVI and its response to climatic factors in recent 16 years based on grey relational analysis in Inner Mongolia Autonomous Region, China. *Remote Sens.* **10**, 961 (2018).
69. Guo, F., Zhang, L., Wang, Z. & Ji, S. Research on determining the critical influencing factors of carbon emission integrating GRA with an improved STIRPAT model: taking the Yangtze River Delta as an example. *Int. J. Environ. Res. Public Health*. **19**, 8791 (2022).
70. Liu, S., Yang, Y., Cao, Y. & Xie, N. A summary on the research of GRA models. *Grey Syst. Theory Appl.* **3**, 7–15 (2013).
71. Yang, J. et al. New method for landslide susceptibility mapping supported by spatial logistic regression and GeoDetector: a case study of Duwen Highway Basin, Sichuan Province, China. *Geomorphology*. **324**, 62–71 (2019).
72. Zhang, Z., Song, Y. & Wu, P. Robust geographical detector. *Int. J. Appl. Earth Obs. Geoinf.* **109**, 102782 (2022).
73. Wang, H. et al. Evaluating the suitability of urban development land with a Geodetector. *Ecol. Ind.* **123**, 107339 (2021).
74. Wang, J. F., Zhang, T. L. & Fu, B.-J. A measure of spatial stratified heterogeneity. *Ecol. Ind.* **67**, 250–256 (2016).
75. Wang, L. et al. Spatiotemporal heterogeneity and its determinants of COVID-19 transmission in typical labor export provinces of China. *BMC Infect. Dis.* **21**, 242 (2021).
76. Wang, J., Xu, C. & Geodetector principle and prospective. *Acta Geogr. Sin.* **72**, 116–134 (2017).
77. Wang, J. F. et al. Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. *Int. J. Geogr. Inf. Sci.* <https://doi.org/10.1080/13658810802443457> (2010).
78. Zhang, Z., Song, Y., Karunaratne, L. & Wu, P. Robust interaction detector: a case of road life expectancy analysis. *Spat. Stat.* **59**, 100814 (2024).

Acknowledgements

We acknowledge the support from the National Natural Science Foundation of China (No. 42271226).

Author contributions

Conceptualization, X.P. and Y.Z.; methodology, Y.Z.; software, Y.Z.; validation, X.P., Q.C., and D.H.; formal analysis, Y.Z.; investigation, Q.C.; resources, X.P.; writing—original draft preparation, Y.Z.; writing—review and editing, X.P. and Y.C.; visualization, Q.C. and D.H.; supervision, X.P.; project administration, X.P.; funding acquisition, X.P. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the National Natural Science Foundation of China, Grant Number 42271226.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-75489-9>.

Correspondence and requests for materials should be addressed to X.P.

Reprints and permissions information is available at www.nature.com/reprints.

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024