GISCIENCE & REMOTE SENSING
2025, VOL. 62, NO. 1, 2533487
https://doi.org/10.1080/15481603.2025.2533487

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Identifying climate and environmental determinants of spatial disparities in wheat production using a geospatial machine learning model

Kai Ren[a,b,c], Yongze Song [c], Linchao Li[a,d], Francesco Mancini[c], Zhuoyao Xiao[c], Xueyuan Zhang[c,e,f], Rui Qu [f,g,h] and Qiang Yu[a,b]

[a]College of Natural Resources and Environment, Northwest A&F University, Yangling, Shaanxi, China; [b]State Key Laboratory of Soil and Water Conservation and Desertification Control, Northwest A&F University, Yangling, Shaanxi, China; [c]School of Design and the Built Environment, Curtin University, Perth, WA, Australia; [d]College of Agronomy, Inner Mongolia Agricultural University, Hohhot, China; [e]College of Earth and Environmental Sciences, Lanzhou University, Lanzhou, Gansu, China; [f]Spatial Sciences Discipline, School of Earth and Planetary Sciences, Curtin University, Perth, Australia; [g]State Key Laboratory of Geohazard Prevention and Geoenvironment Protection, Chengdu University of Technology, Chengdu, China; [h]College of Geography and Planning, Chengdu University of Technology, Chengdu, China

ABSTRACT
Wheat production is crucial in global food security and sustainable development, especially in severe global climate change, frequent extreme weather events, and significant population growth worldwide. A deeper understanding of spatial variation in wheat production and its determining factors is essential for implementing different cultivation practices, water and fertilizer management, and adaptive variety selection across different regions. However, existing methods primarily focused on identifying single-variable factors while lacking geographical spatial characteristics, which may lead to an incomplete exploration of spatial disparities in wheat production, predictions, and responses to changes in determining factors. This study develops a geospatial machine learning model by integrating spatial autocorrelation, spatial stratified heterogeneity, and decision tree to identify spatial disparities and their determinants of wheat production. The model is applied to wheat production analysis in Australia, the world's 5th (2022) wheat-producing country. First, a spatial autocorrelation method is employed to identify the hotspot area of wheat production in Australia. Next, the geographically optimal zones-based heterogeneity (GOZH) model, an integration of spatial stratified heterogeneity and decision tree learning models, is used to identify determinants and their interactions on spatial disparities of wheat production. Finally, the developed geospatial machine learning model is evaluated by comparing its effectiveness with the commonly used geographical detector model. The results demonstrate pronounced spatial heterogeneity in Australian wheat production driven by environmental, climatic, and soil factors and their interactions. Identifying these spatial determinants enables more efficient crop management – such as targeted sub – regional practices, climate – adaptive variety selection, and soil health strategies – thereby supporting food security and sustainable agricultural systems.

ARTICLE HISTORY
Received 28 March 2025
Accepted 9 July 2025

KEYWORDS
Agricultural remote sensing; spatial heterogeneity; spatial statistics; geographical detector; geospatial intelligence

1. Introduction

Global crop demand is critically increasing due to the rapid expansion of the global economy and population, leading to the urgent need to enhance crop production (Guarin et al. 2022; Langridge et al. 2022). Wheat is one of the most important crops in the world, providing 20% of the calories and nutrients humans rely on for survival (Gutierrez and Braunstein 2017). Wheat production essentially affects food security in regions with dense agricultural and climate-sensitive regions, leading to the increased challenges in developing climate adaptation strategies and effective regional management in crop production (Chakraborty et al. 2022; Dong et al. 2020; Sharps et al. 2021).

Understanding the spatial disparities of wheat production provides evidence for the decision-making and adaptation strategies (Hernandez-Ochoa et al. 2018). For instance, studies in South Australia and Mediterranean countries demonstrated that distinct weather patterns, soil variability, and climatic conditions can significantly impact crop attributes, production, and quality (Diacono et al. 2012). Therefore, accurate

CONTACT Yongze Song yongze.song@curtin.edu.au; Qiang Yu yuq@nwafu.edu.cn

factor exploration for assessing the spatial disparity of wheat production provides quantitative solutions for optimizing resource allocation and formulating targeted strategies to enhance regional and cooperative cross-regional wheat production and improve their stability and efficiency (X. Lv et al. 2023; Y. Zhang and Li 2022).

A variety of methodologies have been developed to investigate spatial disparities in wheat production from different perspectives, including statistical econometric models, kriging interpolation, and machine – learning algorithms, which capture spatial dependencies, heterogeneity, and complex factor interactions (Min et al. 2014). Stratified heterogeneity models further improve upon these approaches by dividing spatial data into strata and examining variation within each group (P. Luo et al. 2022; Z. Zhang, Li, and Song 2024; Zhang et al. 2024). The geographical detector (GD) method represents a prominent stratified – heterogeneity technique; it partitions the study area based on explanatory variables and evaluates the resulting spatial variance to quantify each factor's explanatory power (J. F. Wang et al. 2010, 2014). GD has been widely applied to explore spatial differences in crop production and to assess relationships between yield and environmental drivers (Chu et al. 2019; Hou et al. 2023; Zeng et al. 2023).

Building on GD, the optimal parameters – based geographical detector (OPGD) model optimizes discretization breakpoints and scale effects to improve the extraction of geographic features from spatial explanatory variables, yielding more precise spatial analyses (Song et al. 2020). OPGD has been used to analyze soil organic matter distribution in the Huangshui River Basin and to examine driving – factor interactions in vegetation net primary productivity (Liu et al. 2023; W. Zhang et al. 2023). However, these approaches do not fully elucidate the spatial distribution characteristics of the response variable during the discretization process, resulting in incomplete exploration of the relationships between explanatory and response variables. To address this gap, the geographically optimal zones – based heterogeneity (GOZH) model identifies regions that maximize inter – region differences and minimize intra – region variation, thereby enhancing the reliability of interaction estimates among multiple explanatory variables (P. Luo et al. 2022).

Factors affecting the spatial disparity of wheat production can be classified into four categories. First, climate is an essential factor, as changes in local climate impact wheat production per hectare, alter the distribution of climatically suitable areas for wheat growth, and affect the availability of land suitable for wheat cultivation, leading to regional variations in wheat production potential and trends (Fan et al. 2018; B. Wang et al. 2018). Second, geographical factor significantly shapes wheat production's spatial disparities by interacting with climate, location, and soil, impacting production patterns and agricultural sustainability. Third, soil attributes influence spatial disparities in wheat production due to their impact on factors such as soil quality, moisture distribution, and nutrient availability, which affect crop production in various regions and environmental change scenarios (Ajami et al. 2020; Chu et al. 2019; Q. Luo et al. 2005). Finally, environmental factors play a pivotal role in influencing the variation in wheat production across different areas. Remote sensing and geospatial technologies are effective approaches to monitor and analyze production environments, climate patterns, and biotic stress variations and identify complex relationships between environment and wheat production (Hodson and White 2007; Sbahi, Ziboon, and Hassoon 2018). In addition to the individual impacts of variables on wheat production, it is essential to explore complex interactions among these variables (Ajami et al. 2020; Chu et al. 2019; Jiu-Jiang et al. 2022). Furthermore, the spatial disparities in wheat production are also influenced by spatial scales, including the field scale (Mao et al. 2021; Kravchenko, catchment scale (Mao et al. 2021; Richter et al. 1998), regional scale (Jin et al. 2022; Z. Lv et al. 2017; Xiong et al. 2008), administrative scale (Fu et al. 2021), and global scale (He et al. 2022). Remote sensing and geospatial technologies are effective for monitoring and analyzing production environments, climate patterns, and biotic stress variations, which helps identify complex relationships between the environment and wheat production (Hodson and White 2007; Sbahi, Ziboon, and Hassoon 2018). Integrating remote sensing and geospatial technologies facilitates a comprehensive understanding of the impacts of environmental factors on wheat production. This approach not only enhances the ability to monitor and analyze spatial disparities but also informs better agricultural practices and policy-making.

However, there are still several challenges to examining the spatial disparities of wheat production using geographical technologies. Given the vast expanse of Australia and the spatial distribution of factors, such as climate, soil, and environmental conditions, wheat production exhibits significant spatial disparities (Fletcher et al. 2020; Orton et al. 2018). To address these complexities, we employed the GOZH model, which is well suited for correcting single-variable underestimation, mitigating multi-variable overestimation, and delineating fine-scale regional patterns (Hu, Song, and Zhang 2025; P. Luo et al. 2022), thereby enabling a more accurate identification of the key factors driving wheat production and their spatially heterogeneous impacts.

This study aims to explore the spatial disparities and identify the determinants of wheat production through the development of a geospatial machine learning model that incorporates spatial autocorrelation, spatial stratified heterogeneity, and decision tree analysis. Given Australia's extensive territory and diverse climate, coupled with its position as the 5th largest wheat producer globally in 2022 and the 2nd largest wheat exporter (FAO 2024), it serves as an ideal region for examining the spatial heterogeneity of wheat production. Wheat production data from Local Government Areas (LGAs) across Australia for the years 2016 and 2021 were collected as a research dataset, sourced from official Australian government websites. Data on 20 critical factors influencing wheat production over the past year were collected, including four distinct categories: geography, climate, soil, and environment. First, the study employs spatial autocorrelation methods to explore the spatial clustering and hotspot distribution characteristics of wheat production. Second, the GOZH model is utilized to identify geographically optimal regions for wheat production and analyze their spatial variations. Third, the effects of univariate and multivariate interactions on wheat production are examined to identify determinants. Finally, the efficacy of the model developed in this study is validated by comparing it with another geospatial heterogeneity model named OPGD.

The paper is organized as follows: Section 2 introduces the study area, wheat production data and explanatory variables used in this study. Section 3 presents the theory, methods, and formulas of the GOZH model. Section 4 presents the results, including the spatial autocorrelation analysis of wheat production, the delineation of geographically optimal zones, the identification of primary determinants of spatial disparities, the effects of individual and multiple variables, and a comparison with another spatial stratified heterogeneity model. Section 5 and 6 present the discussion and conclusions of this study.

2. Study area and data

2.1. Study area and wheat production

Wheat is the most important cereal crop produced in Australia, accounting for approximately 3.5% of global annual wheat production, with an average national output of around 25 million tonnes (Australian Export Grains Innovation Centre 2022; Kingwell 2020). Australia's wheat is mainly exported, accounting for 65–75% of the national wheat production, exported to more than 50 countries around the world, of which Western Australia (WA) is the largest wheat exporting state (Australian Export Grains Innovation Centre 2022; Kingwell 2020). Since Australia has a large proportion of wheat production and exports, we will conduct research on spatial disparities in Australian wheat production.

In this study, the wheat production data were retrieved from the Australian Bureau of Statistics (ABS) for 2015 to 2016 and 2020 to 2021 (Australian Bureau of Statistics 2017, 2022). In 2021, Australia's total national wheat production was 31.9 Mt. Among them, New South Wales (NSW) had the most wheat production, accounting for 44.06% of the total. The spatial distribution of wheat production is shown in Figure 1, which shows the total and average wheat production within the Australian wheatbelt in 2016 and 2021, respectively. Figure 1 shows that total wheat production in Australia is predominantly concentrated in the central and eastern regions, the southeastern coastal areas, as well as in the western and southwestern regions. Total wheat production is relatively higher in the southeastern part of Queensland (QLD), central NSW, western Victoria (VIC), eastern South Australia (SA), and southeastern WA.

To ensure a more accurate spatial representation of wheat production, we constructed a refined spatial unit by overlaying the Local Government Areas (LGAs) with the wheatbelt boundaries, as defined by the ABS. This modified LGA approach eliminates the bias introduced by including large non-producing areas, thereby improving spatial accuracy (Feng et al. 2022). Based on this method, 186 LGAs were identified within the wheatbelt in 2016, while in 2021, the number slightly decreased to 179 (Australian Bureau of Statistics 2020, 2021).

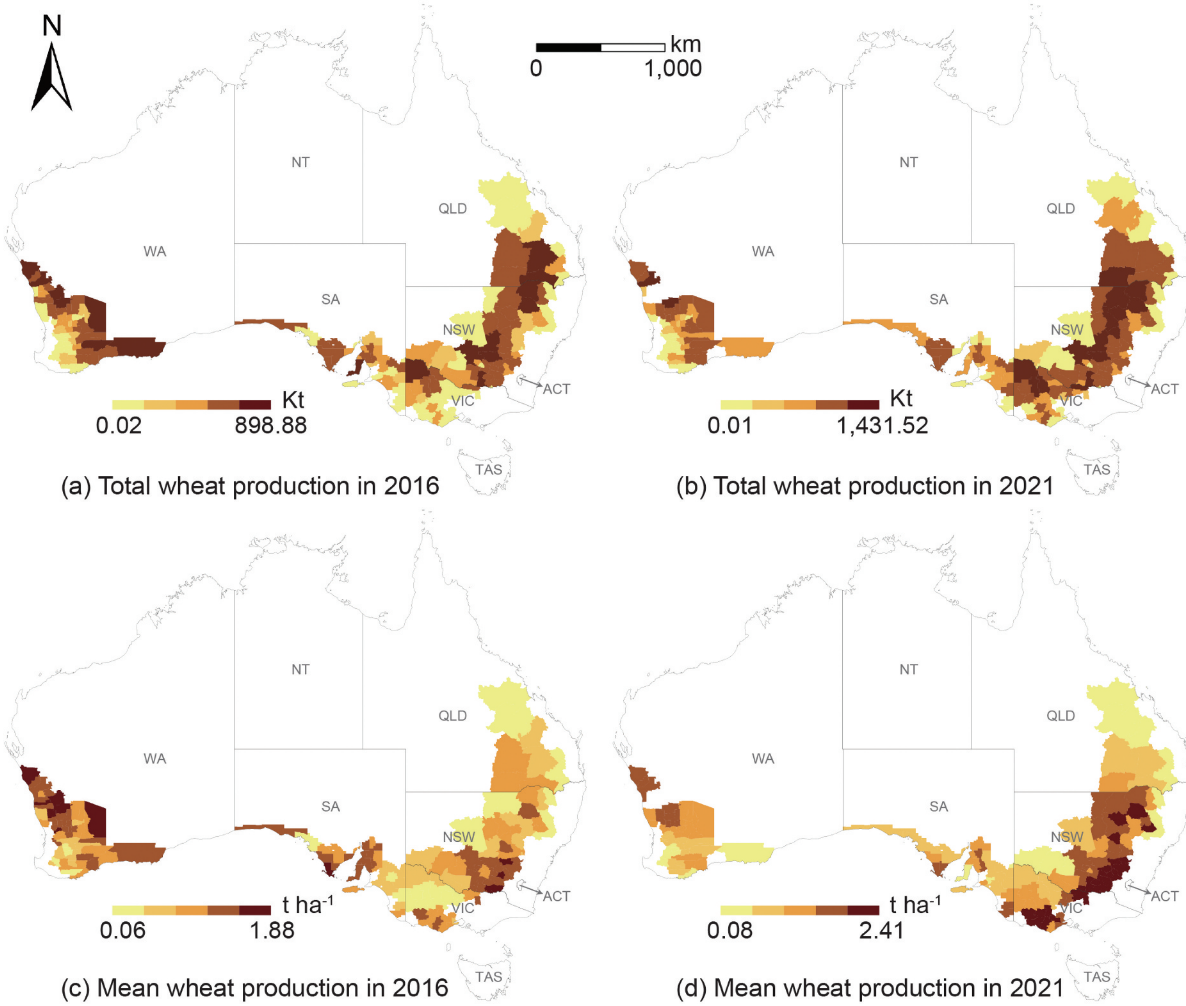

Figure 1. Distribution of Australian total wheat production and mean wheat production in 2016 and 2021.

Table 1. The wheat growing seasons in various states of Australia.

State	Wheat growing sesaons
NSW	From April to December (Gomez-Macpherson and Richards 1995; Pang, Chang, and Chen 2022; B. Wang et al. 2015).
VIC	From April to January of the following year (Pang, Chang, and Chen 2022; E. Wang et al. 2009).
QLD	From March to November (Cammarano et al. 2012; Obanor et al. 2013).
SA	From April to December (Pang, Chang, and Chen 2022; E. Wang et al. 2009).
WA	From May to January of the following year (Duncan et al. 2017; Shen and Evans 2021).

Due to differences in climate conditions and policies, major Australian states have different wheat planting, farming, and growth and development periods. Table 1 shows the wheat growing seasons in various Australian states. This study accounts for the wheat growing season by using data corresponding to the appropriate time periods across LGAs in different states. To minimize the influence of regional size on the outcomes of geospatial analysis, the average wheat production per LGA is used in the analysis.

2.2. Explanatory variables

We selected 20 explanatory variables in four categories – geographic, climatic, soil, and environmental – to capture the main drivers of wheat – yield heterogeneity in Australia (Table 2). Variables were selected based on three criteria: empirical relevance in temperate cropping systems (Han et al. 2020; Millar et al. 2018), representation of key growth processes (temperature, precipitation, soil moisture and texture, vegetation vigor, and water use), and data completeness at the LGA level using

Table 2. A summary of explanatory variables that potentially affect spatial disparities of wheat production.

Category	Variable	Code	Product	Resolution
Geography	Elevation	EL	DEM-S	30 m
	Slope	SLP	DEM-S	30 m
	Aspect	ASP	DEM-S	30 m
Climate	Air temperature	AT	ERA5_land	0.25°
	Total precipitation	TP	ERA5_land	0.25°
	Surface pressure	SP	ERA5_land	0.25°
	Wind speed	WS	ERA5_land	0.25°
Soil data	Available water capacity	AWC	CSIRO/SLGA	92.77 m
	Bulk density (Whole earth)	BDW	CSIRO/SLGA	92.77 m
	Clay	CLY	CSIRO/SLGA	92.77 m
	Total Nitrogen	NTO	CSIRO/SLGA	92.77 m
	Total Phosphorus	PTO	CSIRO/SLGA	92.77 m
	Silt	SLT	CSIRO/SLGA	92.77 m
	Sand	SND	CSIRO/SLGA	92.77 m
	Soil organic carbon	SOC	CSIRO/SLGA	92.77 m
	pH ($CaCl_2$)	pHc	CSIRO/SLGA	92.77 m
Environment	Evapotranspiration	ETa	CMRSET Landsat V2.2	30 m
	Net primary production	NPP	MOD17A3HGF V6.1	500 m
	Normalized difference vegetation index	NDVI	MOD13A2 V6.1	1000 m
	Enhanced vegetation index	EVI	MOD13A2 V6.1	1000 m

harmonized products from Google Earth Engine (MODIS EVI and NDVI) and institutional archives (ECMWF reanalysis, national soil surveys). Although climate modes such as El Niño – Southern Oscillation and the Indian Ocean Dipole affect Australia's rainfall and temperature, their local impacts are captured by our precipitation, temperature, soil moisture, and vegetation indices, so we did not include them as separate variables. Figure 2 shows the spatial distributions of these variables and their relation to regional yield patterns.

2.2.1. Geographical variables

Terrain will directly affect various factors such as sunlight, solar radiation, temperature, precipitation, and air pressure that crops receive (Ajami et al. 2020; Kitchen et al. 2003). The most intuitive model to describe terrain is the digital elevation model (DEM), which is elevation data as our geographical variable. The DEM dataset of Australia was collected from the GEE data catalog provided by Geoscience Australia (Geoscience Australia 2015), which was derived from the SRTM data acquired by the National Aeronautics and Space Administration (NASA) in February 2000. This dataset has been smoothed to reduce noise and improve the representation of surface shape, and it has a high resolution of 30 m. Then, the terrain data include slope and aspect data calculated based on DEM.

2.2.2. Climate variables

Temperature and precipitation are extremely important climate-influencing factors for the spatial disparity of wheat production (Song et al. 2019; H. Zhang et al. 2022). Therefore, we use the air temperature (AT), total precipitation (TP), surface pressure (SP), and wind speed (WS) as climate variables, which are collected from the ECMWF Reanalysis v5 (ERA5) dataset produced by Copernicus Climate Change Service (C3S), the fifth generation ECMWF reanalysis for the global climate and weather for the past eight decades (Hersbach et al. 2023). The resolution of the climatic variables dataset is 0.25°.

2.2.3. Soil attributes

Soil attributes influence crop growth and development by affecting nutrient availability, water retention capacity, soil texture, and pH levels (Kitchen et al. 2003; Nabiollahi et al. 2020; Miller, Singer, and Nielsen 1988). The Soil and Landscape Grid of Australia (SLGA) is a comprehensive dataset of soil attributes across Australia at 3 arc-second resolution (~90 m pixels) which can be obtained from GEE (Rossel et al. 2015). This study uses nine variables of soil attributes from this dataset as follows: Available water capacity (AWC), Bulk density (BDW), Clay (CLY), Total nitrogen (NTO), Total phosphorus (PTO), Silt (SLT), Sand (SND), Soil organic carbon (SOC) and pH ($CaCl_2$) (pHc). These variables have six depths representing the relative soil attributes, such as 0–15 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm, and 100–200 cm. The resolution of soil attributes variables is 92.77 m.

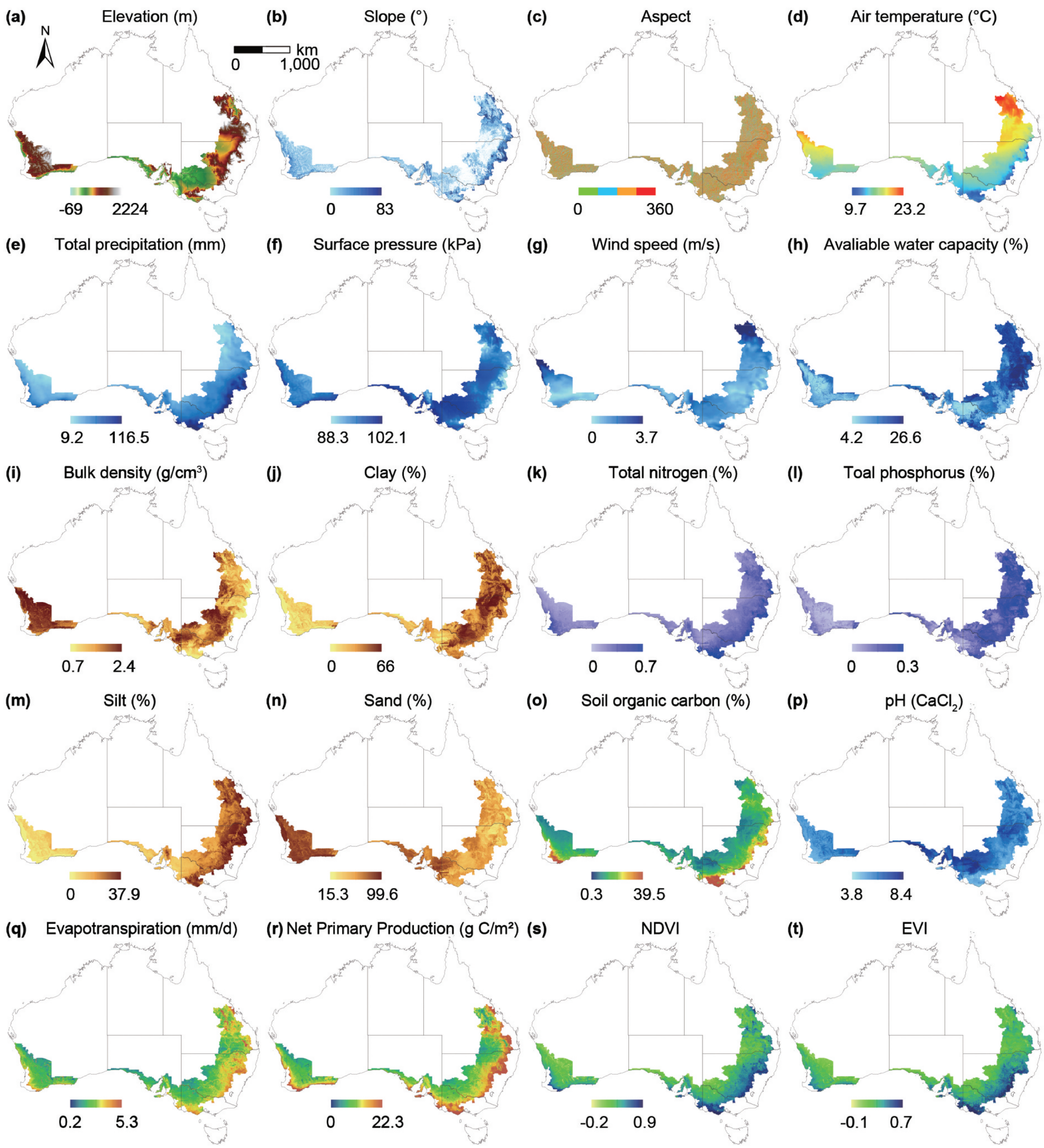

Figure 2. Spatial distributions of explanatory variables in 2021. (a) - (c) geographical variable, (d) - (g) climate variables, (h) - (p) soil attributes, (q) - (t) environmental variables.

2.2.4. Environmental variables

Environmental and ecological conditions of LGAs were characterized using evapotranspiration (ETa), net primary production (NPP), normalized difference vegetation index (NDVI), and enhanced vegetation index (EVI) data (Qader, Dash, and Atkinson 2018; Y. Wang et al. 2019; Xu et al. 2019). The ETa data were collected from GEE provided by the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Land and Water with a 30 m resolution (Guerschman et al. 2022). Similarly, data for NPP, NDVI, and EVI were also obtained from GEE, ensuring consistency in the data quality (Didan 2021; Running and Zhao 2021). These datasets were instrumental in assessing the environmental and vegetative health of the regions, providing key indicators for the study.

3. Methods

This study develops a geospatial machine learning model for analyzing spatial disparities and identifying the determinants in Australian wheat production. The geospatial machine learning model includes spatial autocorrelation, spatial stratified heterogeneity, and decision tree analysis. Figure 3 illustrates the workflow of this study, which includes six main steps. First, the collected variable data undergoes preprocessing to align with the input requirements of the model within this study. Second, a spatial autocorrelation analysis of the response variable will be performed, examining information such as hotspot distribution and clustering patterns. Third, the contribution of individual explanatory variables to wheat production will be determined. The fourth step involves the identification of geographically optimal regions for wheat production in Australia, followed by grouping optimal regions. The fifth step is to identify the interaction impacts among explanatory variables and compute the model's power of determinants (PD) values for hotspot regions and individual states. Finally, the developed model is validated by comparing it with another geographic detector model.

3.1. Data pre-processing

In the first stage, it is essential to subject the collected raw data to data pre-processing to meet the requirements of our research. The data pre-processing comprises four primary components. First, a consistent spatial unit is maintained for all explanatory variables. Second, data standardization is performed. Mean values for each LGA region are computed, and the temperature data undergo unit conversion from Kelvin to Celsius. Third, an outlier detection process is implemented. The BDW and SOC soil attributes exhibit pixel values that surpass the surrounding values. A threshold is set at 2.5 standard deviations from the mean to identify outliers, and values exceeding this threshold are removed. Finally, given that there are six soil attribute datasets corresponding to six geological layers, principal component analysis (PCA) is employed to extract the most significant components and reduce the dimensionality of the soil attribute data. This approach facilitates the condensation of the data into a more analyzable form while preserving the most pertinent information. Table 3 illustrates the pre-processing operations conducted on the explanatory variable data. Moreover, the GOZH model employed in this study has checked the input variables during the calculation process, so no collinearity adjustments are required in the data pre-processing phase.

3.2. Spatial pattern analysis

This study employed spatial autocorrelation methods, including Moran's I and local indicators of spatial association (LISA), to identify global and local clusters and significance levels of wheat production.

Moran's I is a spatial autocorrelation measure in geostatistics, utilized to quantify the spatial clustering, dispersion tendencies, and significance of spatial data within the study area (Li, Calder, and Cressie 2007; Moran 1950). Moran's I value ranges from −1 to 1, where a value greater than 0 indicates a positive spatial correlation, less than 0 indicates a negative spatial correlation and a value of 0 signifies a random spatial distribution without correlation. The calculation formula for Moran's I is as follows:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{1}$$

where n is the number of factors, x_i and x_j represent the attribute values of the i-th and j-th factors, $\bar{x}$ is the mean value of attribute values, and w_{ij} denotes the spatial weight between the i-th and j-th factors.

LISA is used to analyze the clustering and associational relationships of spatial data at the local level, it is necessary to introduce the local Moran's I, known as LISA (Anselin 1995). This method assesses the spatial clustering or dispersion phenomena of individual features by calculating the ratio of the covariance to the variance between each spatial element and its neighboring elements (Anselin 2005). The computational formula is as follows:

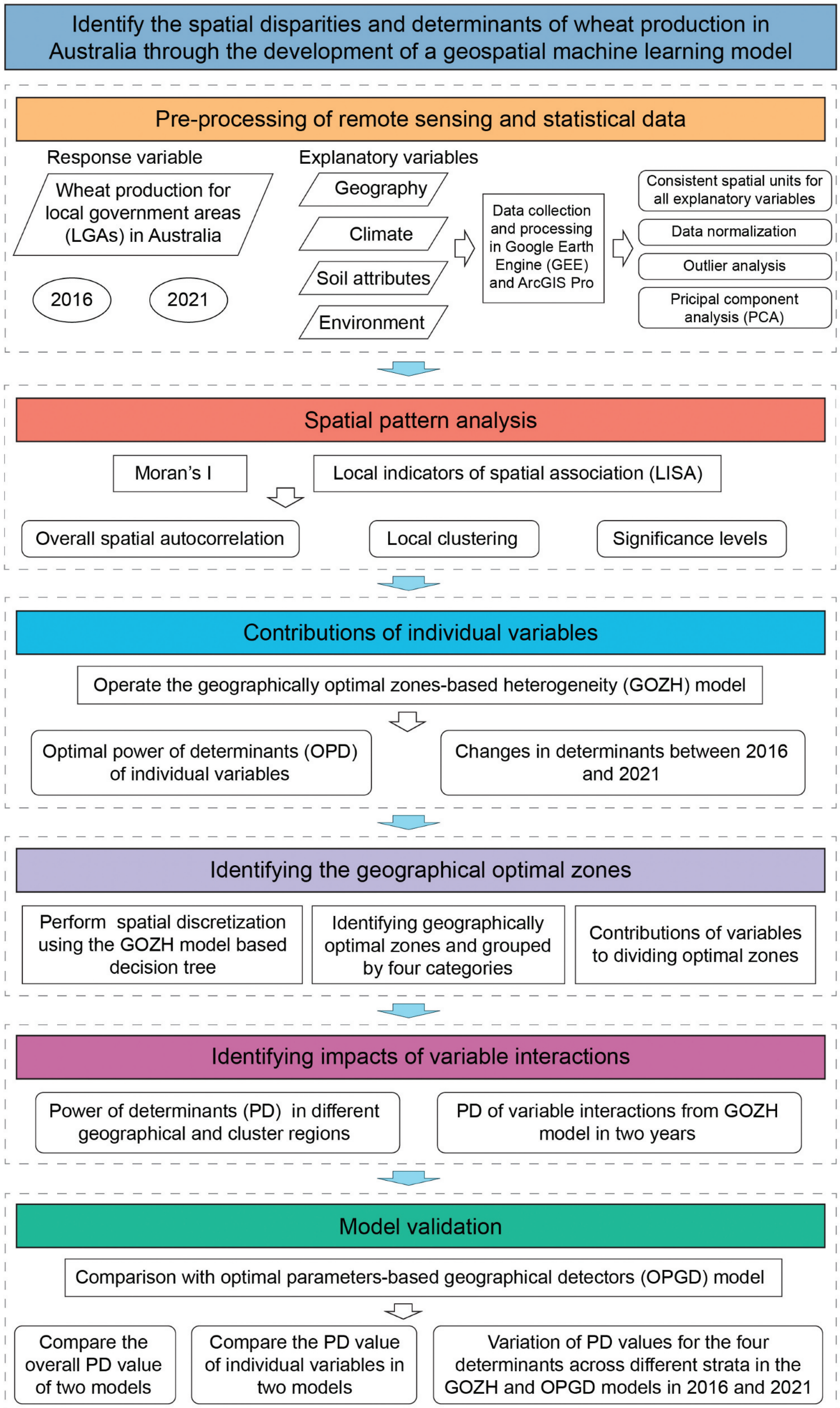

Figure 3. Schematic overview of identifying the spatial disparities and determinants of wheat production in Australia based on the geospatial machine learning model.

Table 3. Methods and procedures for pre-processing explanatory variables.

Pre-processing	Variable
Compute LGA mean	DEM, SLP, ASP
Compute annual LGA mean	NPP
Compute growing-season mean per LGA and convert to °C	AT
Outlier removal and PCA, then compute LGA mean	BDW, SOC
PCA applied, then compute LGA mean	AWC, CLY, NTO, PTO, SLT, SND, pHc
Compute growing-season mean per LGA	TP, WS, ETa, NDVI, EVI

$$I_i = \frac{x_i - \bar{x}}{s^2} \sum_{j=1}^{n} w_{ij}(x_j - \bar{x}) \quad (2)$$

where I_i is the value of local Moran's I, x_i and x_j represent the attribute values of the *i*-th and *j*-th factors, $\bar{x}$ is the mean value of attribute values. The w_{ij} is the spatial weight matrix, and s^2 is the variance of variable x.

In this study, the distance weight method selected is the Gaussian Kernel function, wherein diagonal weights are set to 1.

3.3. Contributions of individual variables

The geographical detector (GD) model consists of four main components: factor detection, interaction detection, risk detection, and ecological detection (Jinfeng and X 2017). The factor detection is a core component of the GD model, and it allows for the comparison of the relative importance of various explanatory variables within the study area by assessing the magnitude of their power of determinants (PD) values. The calculation formula for PD is as follows (J. F. Wang et al. 2010):

$$PD = 1 - \frac{\sum_{z=1}^{h} N_z \sigma_z^2}{N\sigma^2} = 1 - \frac{SSW}{SST} \quad (3)$$

where *SSW* and *SST* represent the sum of squares within the region strata and the total sum of squares determined by the explanatory variables, respectively. N_z and σ_z represent the count and standard deviation within geographical zone z ($z = 1, \ldots, h$), respectively. Similarly, N and σ denote the count and standard deviation throughout the entire study area. The value of PD, which falls within the range [0,1], indicates the explanatory power of the variables on the response variable, with a higher PD value signifying a stronger explanatory power of the explanatory variables concerning the response variable and, consequently, a higher degree of association between them.

The geographically optimal zones – based heterogeneity (GOZH) model is designed to delineate regions that maximize inter-zone dissimilarity while maintaining intra-zone similarity (P. Luo et al. 2022). It achieves this through a decision-tree-driven, optimized spatial stratification that simultaneously captures the interactions among multiple explanatory variables. Unlike classical GD, which depends on user-defined discretization, or OPGD, which carries out an exhaustive grid search for PD-maximizing cut-points, GOZH integrates the tree algorithm with an iterative search for the optimal power of determinants (Ω). At each split, the procedure maximizes between-zone variance, tracks the marginal gain in Ω, and terminates when the incremental improvement falls below 1% (P. Luo et al. 2022; Song et al. 2020). GOZH preserves GD's interpretability while automating zone delineation and safeguarding against both under- and over-segmentation.

In addition, GOZH enhances the reliability of estimating interactions among multiple explanatory variables in the study of spatial determinants. In this study, the GOZH model defines the PD value to explain the relationship of explanatory variables and geographical zones and also defines an optimal power of determinants (OPD) value to indicate the maximum explanatory power of variables on the geographically optimal zones (P. Luo et al. 2022):

$$\Omega = \max(\mathrm{PD}) = \gamma(\mathrm{X}, \mathrm{D}) = 1 - \frac{\min(\mathrm{SSW}_{\mathrm{X,D}})}{\mathrm{SST}} \quad (4)$$

where X is one or multiple explanatory variables, and D represents the stratified variable describing geographical regions, and $SSW_{X,D}$ is the sum of squares within geographical regions defined by D which are determined based on explanatory variable X. The Ω value signifies the OPD of explanatory variables and is the maximum value of PD.

The GOZH model quantifies the spatial heterogeneity contribution of each explanatory variable to wheat production. To prevent excessive regional partitioning and ensure a rational delineation of optimal geographic areas, decision trees were constructed for LGAs with a minimum observation size of 10 per node, denoted as minsplit = 10 (Breiman 2017). In the stepwise spatial discretization process, groups with fewer than 10 LGAs were not further discretized. The PD for each individual variable within each category was computed throughout the spatial discretization, which is the Ω value and elucidates the influence and contribution of each variable toward the spatial distribution of wheat production.

3.4. Identifying the geographically optimal zones

Optimal geographical zones require minimal within-area variance and maximum between-area variance. The minimum value of $SSW_{X,D}$ is calculated as follows for deriving Ω.

$$min\,(SSW_{X,D}) = min\left\{\sum_{z=1}^{h}\sum_{j=1}^{N_z}(y_{zj} - \bar{c}_z)^2\right\} \tag{5}$$

where y_{zj} and $\bar{c}_z$ represent the j-th ($j = 1, 2, \ldots, N_z$) observed value and the mean value of wheat production in zone z, respectively. To solve the above equation, it is necessary to discretize the spatial wheat production with explanatory variables step by step, selecting the optimal discrete breakpoints and cutoff variables. The detailed derivation of Ω is presented in the paper published by Peng Luo (P. Luo et al. 2022). This spatial discretization process is similar to the classification and regression tree (CART) algorithm (Breiman 2017). The GOZH model mainly uses the R software packages "rpart" (Therneau and Atkinson 2023) and "GD" (Song et al. 2020).

This study employed the GOZH model to discretize LGAs with wheat production into a binary tree structure using 20 explanatory variables. Throughout the iterative discretization process, the LGAs were segmented into several geographically optimal zones. Within these zones, regions of the same type displayed significant homogeneity, whereas regions of differing types exhibited pronounced heterogeneity.

3.5. Identifying impacts of variable interactions

The identification of geographically optimal zones using the geographically optimal zones-based heterogeneity (GOZH) model is accomplished through stepwise spatial discretization and the interactions among multiple variables. In this study, the GOZH model examines determinants of spatial disparities in wheat production from four categories of input explanatory variables. This examination considers how these factors influence the spatial disparities of wheat production and their relationships with each other. Through comparing the contribution of these factors to delineating geographically optimal zones, a systematic analysis is conducted to determine which group of factors primarily influences these spatial disparities.

Following that, the GOZH model is applied at multiple spatial scales using consistent input and output variables, encompassing the entire Australian continent, hotspot and non-hotspot regions, as well as individual states. The resulting Ω values generated for each region are compared to evaluate the model's robustness and effectiveness across different geographic contexts, thereby providing insights into the scale-dependent performance of the GOZH model.

3.6. Model evaluation

The developed geospatial machine learning model is evaluated through another geospatial heterogeneity model, optimal parameters-based geographically detectors (OPGD). OPGD model has been developed building upon the foundation of the GD model (Song et al. 2020). This model optimizes the processes of spatial data discretization and spatial scaling, determining the best parameter combination for the

Geographic Detector model. This optimization assists in extracting essential geographic features embedded within spatial explanatory variables.

The effectiveness and reliability of the GOZH model in analyzing spatial heterogeneity of wheat production will be validated by comparing the results of the OPGD model, which was executed in the "GD" package in the R language. The GOZH model's applicability and superiority in analyzing spatial disparities in wheat production were assessed by comparing the PD values of explanatory variables and their variations in individual PD values between the two models. This was further complemented by examining the trends in PD values differences of the four principal variables across different strata within each model.

4. Results

4.1. Data pre-processing

After a series of operations including spatial unit consistency processing, data standardization, outlier processing, and principal component analysis, the wheat production and explanatory variable data meet the model input requirements. Table 3 shows the pre-processing operations for each explanatory variable. As a result, we obtained 186 data sets in 2016 and 179 sets in 2021.

4.2. Spatial pattern analysis

Figure 4 illustrates the spatial distribution patterns of average wheat production in Australia for 2 years. From Figure 4(a,d), the Moran's I values for wheat production in 2016 and 2021 are 0.532 and 0.669, respectively, indicating a significant spatial clustering of wheat production in Australia. Figure 4(b,e) identify the hotspots (high-production areas) and cold spots (low-production areas) for wheat production, providing a clearer understanding of the clustering characteristics. In 2016, there were 41 hotspot regions, mainly distributed in south-central New South Wales (NSW), southern South Australia (SA), and the mid-western and southern regions of Western Australia (WA), accounting for 22% of the LGAs in the study area. In 2021, the number of hotspot regions increased to 44, primarily concentrated in NSW and Victoria (VIC), representing 24.6% of the

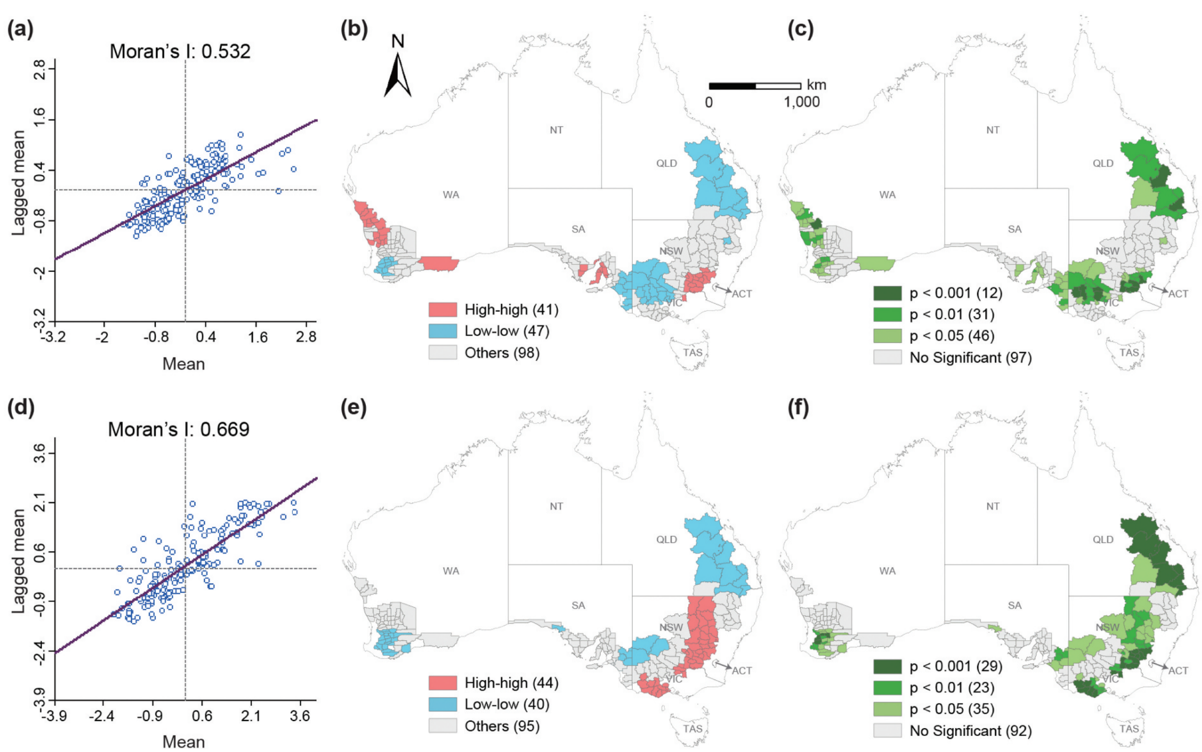

Figure 4. Results of Moran's I and LISA analysis. Value of Moran's I in 2016 (a) and 2021 (d), hotspot analysis based on LISA in 2016 (b) and 2021 (e), significance analysis based on LISA in 2016 (c) and 2021 (f).

study area. The changes in clustering areas over the 2 years may be related to variations in climate and environmental conditions, such as changes in rainfall, temperature fluctuations, and the El Niño-Southern Oscillation.

Figure 4(c,f) display the significance levels of different regions (e.g. $p < 0.001$, $p < 0.01$, $p < 0.05$) for both years, showing significant results in both hotspot and cold spot regions. This indicates that the clustering of wheat production in hotspot areas is not random, and the spatial autocorrelation is statistically significant. The cold spot regions in both 2016 and 2021 also show strong significance, suggesting that the low production in these areas is not due to random variation but is caused by persistent, significant factors such as poor soil quality and unfavorable climatic conditions.

Figure 5 presents the statistical characteristics of wheat production in hotspots, cold spots, and other regions within the study area. The boxplots for hotspots indicate a relatively high median wheat production, with values of 1.05 $t \cdot ha^{-1}$ and 1.72 $t \cdot ha^{-1}$ for 2016 and 2021, respectively, demonstrating that wheat production in hotspot regions consistently outperforms that in other areas. The median wheat production in 2021 is 63.8% higher than in 2016. The interquartile range (IQR) for hotspots in 2016 is narrower compared to 2021, suggesting less variability in wheat production in 2016. In addition, some outliers in 2016 indicate exceptionally high production in certain LGAs. In contrast, the median of cold spots is significantly lower, reflecting the generally lower production in these areas. Other areas are between hot and cold spots, with moderate IQR, indicating that the wheat production changes in these areas are more balanced and the production level is more stable.

Table 4 provides statistical details on the number of LGAs, total average production, and average production characteristics across hotspots, cold spots, and other regions. From Table 4, NSW had a majority of hotspots in 2021, totaling 30, which accounted for 68.2% of the total hotspot regions. The average production in these NSW hotspots was 1.62 $t \cdot ha^{-1}$, significantly higher than the 1.18 $t \cdot ha^{-1}$ recorded in the NSW hotspots in 2016. Additionally, the proportion of cold spots in WA increased to 50% in 2021, likely due to the region's drought conditions. Furthermore, the proportion of cold spots in WA increased from 25.5% in 2016 to 50% in 2021, which is likely due to drought conditions in this region. The number of hotspot regions in VIC increased from 1 to 14, with the proportion rising from 2.4% to 31.8%. Meanwhile, the number of cold spot regions decreased from 12 to 1, indicating a significant upward trend in wheat production in VIC.

In summary, wheat production within the Australian wheat belt exhibits significant spatial clustering and autocorrelation patterns, reflecting the uneven distribution of favorable agroecological conditions across

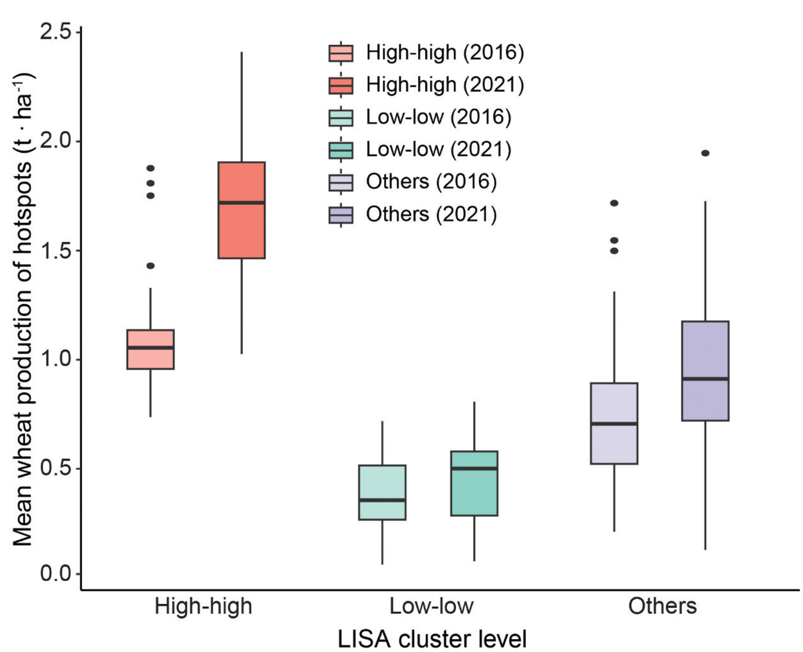

Figure 5. Mean wheat production of hotspot and non-hotspot areas based on the results of LISA analysis.

Table 4. Local Indicators of spatial association (LISA) clustering statistics for wheat production across different states in Australia.

LISA Statistics			NSW	VIC	WA	SA	QLD	Australia
2016	High-high	Count	13	1	18	9	–	41
		P_{sum}[1]	15.37	1.07	19.43	9.15	–	45.03
		P_{mean}[2]	1.18	1.07	1.08	1.02	–	1.08
	Low-low	Count	4	12	12	8	11	47
		P_{sum}	2.29	3.56	5.01	3.58	3.56	18.01
		P_{mean}	0.57	0.30	0.42	0.45	0.32	0.38
	Others	Count	30	13	35	19	1	98
		P_{sum}	19.47	9.52	27.31	15.12	0.69	72.11
		P_{mean}	0.65	0.73	0.78	0.80	0.69	0.74
2021	High-high	Count	30	14	–	–	–	44
		P_{sum}	48.68	25.01	–	–	–	73.69
		P_{mean}	1.62	1.79	–	–	–	1.67
	Low-low	Count	2	1	20	6	11	40
		P_{sum}	0.85	0.68	9.95	3.11	3.94	18.53
		P_{mean}	0.42	0.68	0.50	0.52	0.36	0.46
	Others	Count	15	11	38	30	1	95
		P_{sum}	14.34	9.84	36.35	26.36	1.11	88.00
		P_{mean}	0.96	0.89	0.96	0.88	1.11	0.93

[a]Sum of the mean production ($t \cdot ha^{-1}$).
[b]Average of the mean production ($t \cdot ha^{-1}$).

regions. The spatial pattern analysis using Moran's I and LISA not only reveals these regional disparities but also provides empirical justification for applying the GOZH model. By quantifying spatial autocorrelation, these methods help verify the existence of stratified heterogeneity, which is a prerequisite for the GOZH framework. Therefore, spatial pattern analysis serves as a foundation for the subsequent identification of geographically optimal zones and the assessment of explanatory variable contributions.

4.3. Contributions of individual variables

This study applied the geographically optimal zones-based heterogeneity (GOZH) model to simulate average wheat production data and explanatory variables affecting wheat growth in Australia's wheat belt. The goal was to determine the spatial distribution differences in wheat production and identify the decisive factors.

Figure 6 illustrates the contribution of individual variables to the heterogeneity of wheat production in 2016 and 2021 using the GOZH model. Despite the differences in the contributions of each determinant to the spatial pattern of wheat production, they all have significant impacts. The determining variables are

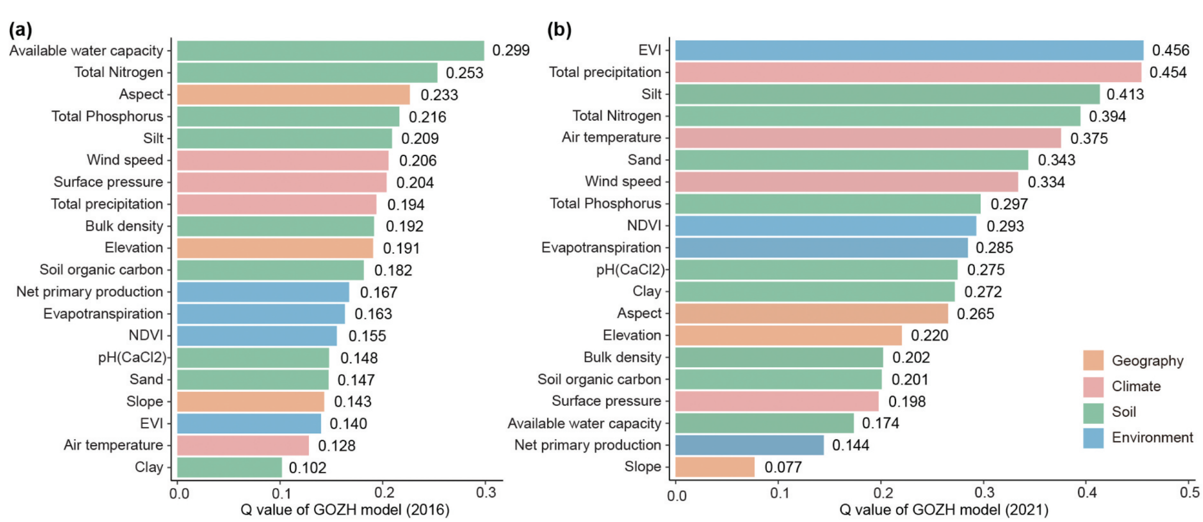

Figure 6. Results of the geographical optimal zones-based heterogeneity (GOZH) model for assessing power of determinants (PD) of wheat production.

categorized into four main types: geographical factors, climatic variables, soil attributes, and environmental conditions.

In 2016, soil attributes contributed the most, with nine soil attribute variables significantly affecting the distribution of wheat production. Among them, AWC (29.9%) and NTO (25.3%) had the highest contributions, closely related to the soil moisture and nutrient absorption required for wheat growth. The impact of topographical factors was also notable, especially ASP (23.3%), which influences light exposure, radiation, and water runoff, directly affecting wheat growth. Climatic variables are fundamental determinants of wheat production, with WS (20.6%), SP (20.4%), TP (19.4%), and AT (12.8%) collectively influencing pollination probability, photosynthetic efficiency, transpiration rate, and water absorption rate. Additionally, environmental variables related to wheat growth showed strong influence, particularly NPP (16.7%), ETa (16.3%), NDVI (15.5%), and EVI (14%). These vegetation-related variables are crucial for wheat production, as a better vegetative environment enhances photosynthesis and yield.

In 2021, there were notable differences from 2016. While the main influencing variables remained consistent, the top contributing variables were EVI (45.6%), TP (45.4%), and SLT (41.3%), each contributing over 40%, which indicates that these three variables significantly influenced the spatial distribution of wheat production in 2021. Soil attributes continued to show strong influence due to Australia's vast expanse, with most arable land located near coastal areas or hilly regions. Soil attributes significantly affect soil structure, water retention capacity, nutrient content, organic matter, and pH levels. Climatic variables also had substantial contributions in 2021, with TP (45.4%), AT (37.5%), WS (33.4%), and SP (19.8%) being the most impactful. The significant influence of TP and AT aligns with the general understanding of wheat's sensitivity to precipitation and temperature. Environmental variables, especially EVI, further underscored the importance of vegetation cover for wheat growth in 2021.

Figure 7 describes the changes in major determining factors between 2016 and 2021. Figure 7(a) shows a notable increase in EVI in Queensland (QLD), NSW, and VIC, a slight increase in the southwest of WA, and varying degrees of decrease in other areas. Figure 7(b) indicates a significant increase in total precipitation from 2016 to 2021 in NSW, VIC, and SA, with the largest increase being 37.27 mm, while WA and QLD experienced significant decreases. Figure 7(c) reveals an increase in temperature in QLD and western WA, with a maximum increase of 0.377 $^{\circ}C$, while other areas saw a decrease of less than 1 $^{\circ}C$. Figure 7(d,e) show these three variables' overall rate of change and coefficient of variation. The temperature had the highest rate of change and coefficient of variation at 9.13% and 6.17%, respectively, indicating a significant overall temperature increase in 2021 compared to 2016. EVI also showed a slight increase, suggesting an improvement in vegetation cover across Australia. Conversely, total precipitation decreased by 2.37% in 2021, with a coefficient of variation of 1.69%, indicating that even minor changes can significantly impact wheat growth.

4.4. Identifying the geographically optimal zones

The GOZH model, utilizing decision trees and stratified analysis, classified the study area into four groups. Although the classification variables for 2016 and 2021 differed, the grouped regions showed only slight differences. This method maximizes inter-group differences while minimizing intra-group regional disparities in wheat production.

Figure 8 illustrates the process of dividing the wheat production regions in 2016 and 2021 into four major groups using decision trees and stratified analysis methods. In 2016, the key variable was NTO, which is critical for nutrient absorption during the wheat growth process and directly affects wheat production. The secondary division was made using ASP and BDW. Figure 8(b) depicts the four geographically optimal groups using four different colors. The characteristics of Groups 1 to 4 are as follows: high total nitrogen and low aspect, high total nitrogen and high aspect, low total nitrogen and high bulk density, and low total nitrogen and low bulk density.

In 2021, the primary variable for defining optimal regions was EVI, followed by ASP and SLT (Figure 8(d,e)). Through two levels of stratification, these three variables divided the wheat production area into four major groups. Further detailed stratification was performed under other variables, ensuring significant geographical spatial differences between the four major groups while reducing significant differences within the LGA of each group. The characteristics of the four groups in 2021

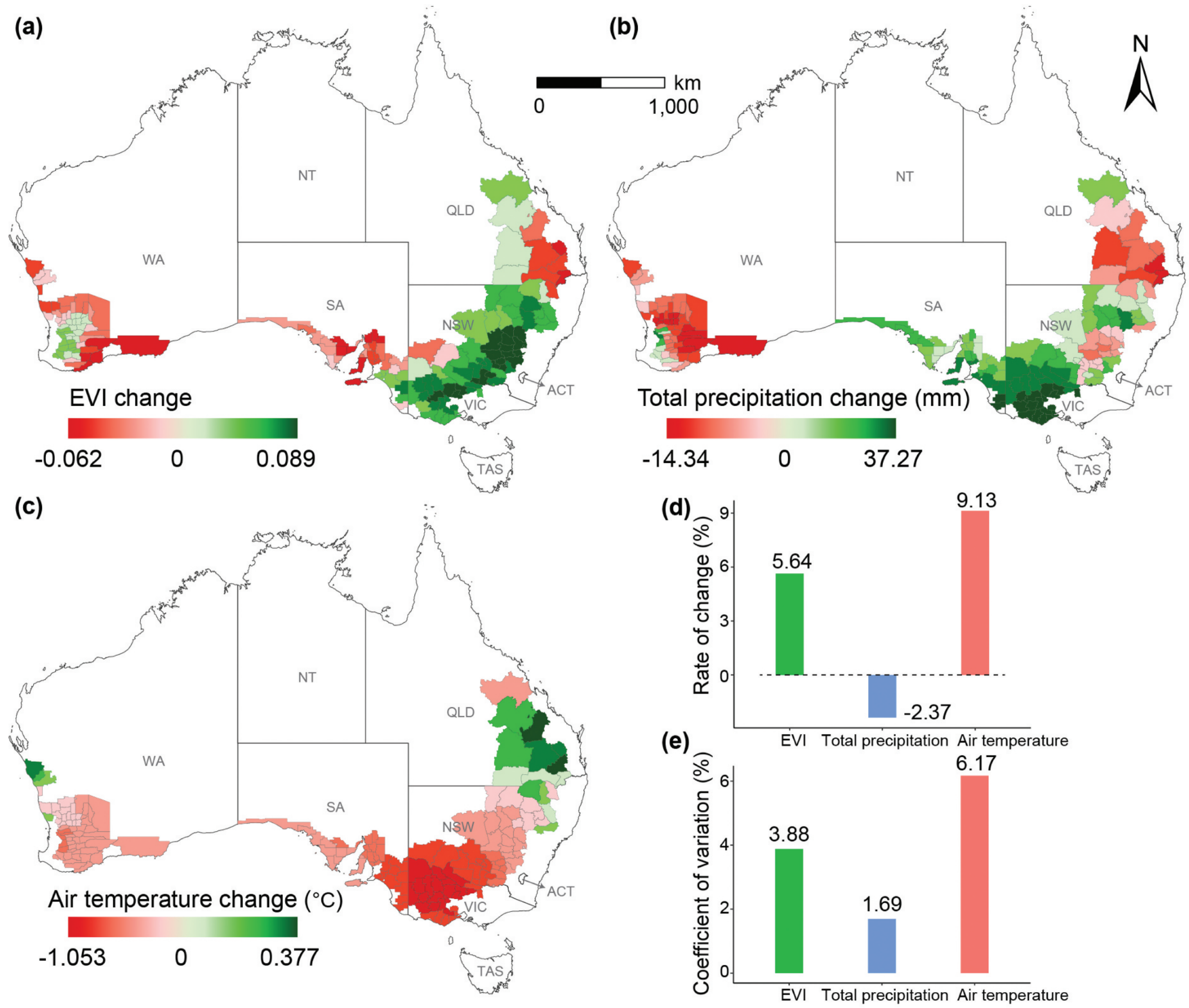

Figure 7. Changes in the top three determinants affecting spatial disparities in wheat production from 2016 to 2021, as well as their change rates and coefficients of variation.

are low EVI and low aspect, low EVI and high aspect, high EVI and low silt, and high EVI and high silt.

According to the box plots in Figure 8(c,f), it is evident that the three main variables for both years delineate the production regions based on different geographical spatial features. Identifying geographically optimal regions helps us better understand the spatial heterogeneity of wheat production under the influence of external determinants. This can assist decision-makers or farmers in managing wheat in different geographical regions adaptively to maximize yield and income.

Figure 9 and Table 5 depict the contribution percentages of each variable in defining the optimal geographical regions for wheat production. The figures illustrate that geography, soil, environment, and climate all influence wheat production distribution. In 2016, the primary factors affecting wheat production were geographical and soil conditions, suggesting minimal variability in climate and environmental conditions that year, resulting in a lesser impact than other variables. In contrast, in 2021, soil and environmental factors predominantly influenced the optimal regions for wheat production. This indicates that the impact of different factors on wheat production changes over time. This shift may be related to climate change, advancements in agricultural technology, and the implementation of environmental policies. For example, in 2021, enhanced soil improvement measures and stronger environmental protection policies significantly affected wheat production. In addition, fluctuations in climate variables such as precipitation and temperature could affect wheat production to varying degrees across different years.

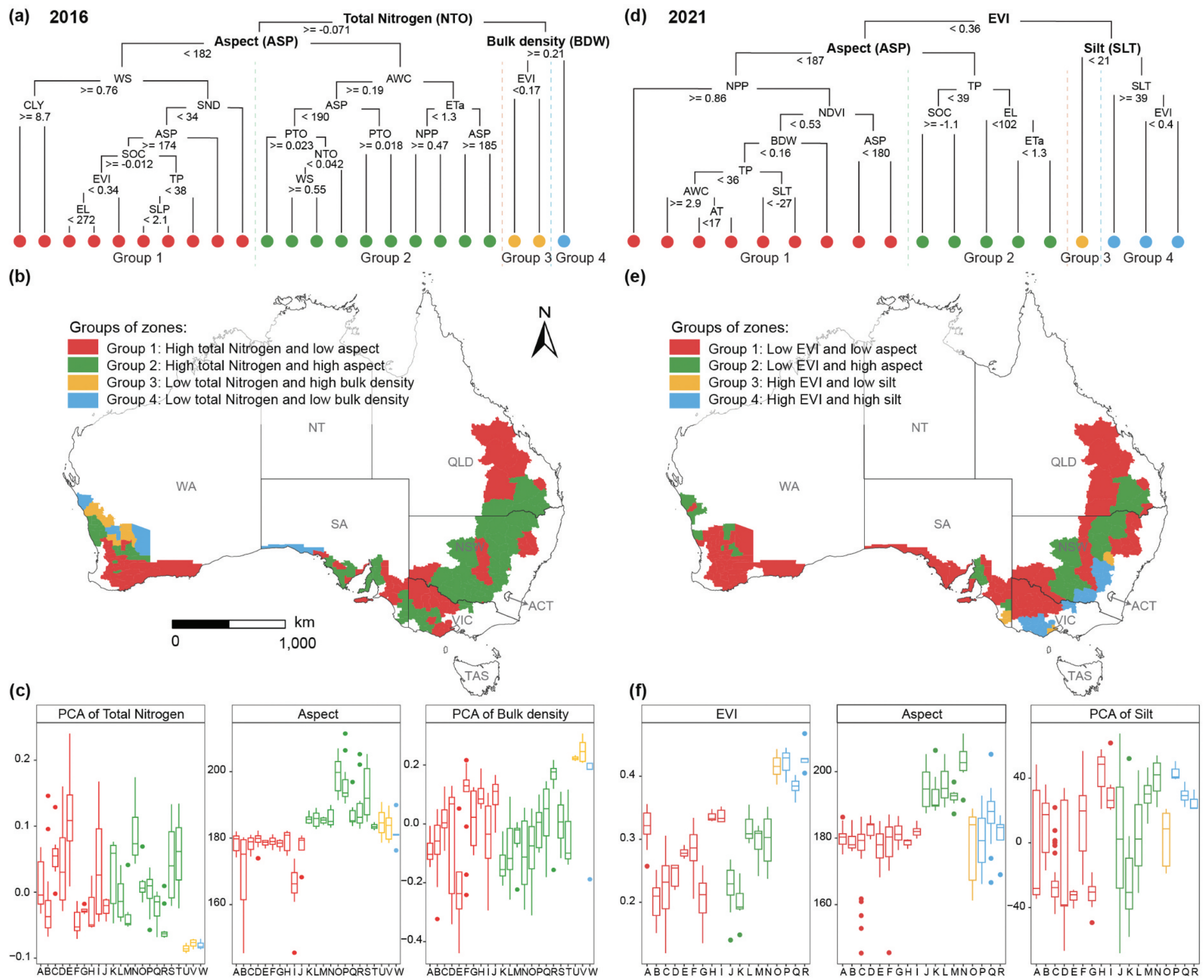

Figure 8. Decision tree of identifying optimal zones (a, d), geographically optimal zones of wheat production at the LGAs identified using the GOZH model (b, e), and statistical summaries of explanatory variables within zones for explaining characteristics of zones (c, f).

Table 5 details the contribution proportions of explanatory variables in defining the optimal geographical regions. In 2016, ASP, NTO, and AWC had higher contributions, at 42.29%, 24.51%, and 11.07%, respectively, indicating that terrain and soil attributes significantly influenced wheat production spatial disparities. In 2021, EVI and ASP had the largest contributions at 37.25% and 30.07%, followed by TP and NDVI. This suggests that vegetation status, terrain, and precipitation substantially influenced the spatial distribution of wheat production in 2021. ASP played a crucial role in both years, indirectly affecting light exposure, radiation, temperature, precipitation, and wind speed, thereby impacting photosynthesis, pollination, and nutrient absorption in wheat. These findings provide a foundational framework for understanding the spatial disparities of wheat production in Australia, aiding in the optimization of agricultural management and decision-making.

4.5. Identifying impacts of variable interactions

The model developed in this study examines the spatial disparities of wheat production across Australia and runs separately for different hotspot and non-hotspot areas and states. Figure 10 illustrates the model results at various spatial scales within Australia, with the Ω value representing the total impact of the interactions among geographic, climatic, soil, and environmental factors on the spatial distribution of wheat production. In 2016 and 2021, Ω values of the GOZH were 0.707 and 0.834, respectively, indicating that the combination of interacting determinants explained 70.7% and 83.4% of the wheat production patterns. For hotspot areas

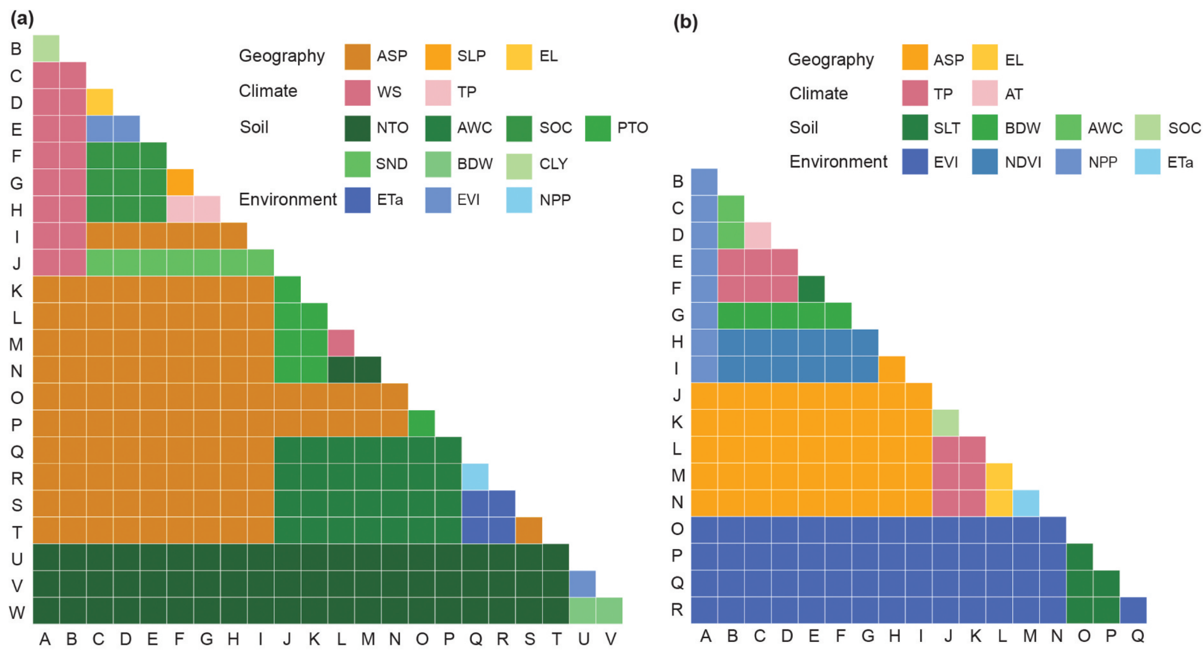

Figure 9. Contributions of explanatory variables for each pair of geographically optimal zones of mean wheat production in 2016 (a) and 2021 (b).

Table 5. Contributions of explanatory variables to dividing optimal zones.

2016		2021	
Variable	Contributes	Variable	Contributes
ASP	42.29%	EVI	37.25%
NTO	24.51%	ASP	30.07%
AWC	11.07%	NDVI	7.84%
WS	6.72%	TP	7.84%
SOC	3.56%	NPP	5.23%
PTO	3.16%	SLT	3.92%
SND	2.77%	BDW	3.27%
ETa	1.58%	AWC	1.31%
EVI	1.19%	EL	1.31%
TP	0.79%	ETa	0.65%
BDW	0.79%	AT	0.65%
NPP	0.40%	SOC	0.65%
SLP	0.40%		
EL	0.40%		
CLY	0.40%		

and state results, the combined Ω values were all greater than 0.5. Notably, VIC and WA in 2021 exhibited high Ω values, indicating that the interactions of determining factors substantially impacted wheat production distribution in these regions. Conversely, the non-hotspot areas in 2016 showed a lower Ω value of only 0.508, which is related to the more dispersed spatial distribution of wheat production in this region.

These findings significantly contribute to understanding the spatial disparities of wheat production in Australia and the complex interactions among different determining factors. The results indicate that the interactions of multiple variables in geographic space significantly influence the spatial patterns and variability of wheat production across Australia.

4.6. Model evaluation and comparison

This study evaluates the performance of the GOZH model in assessing the spatial disparities of wheat production in Australia by examining the contribution of individual variables, power of determinants (PD)

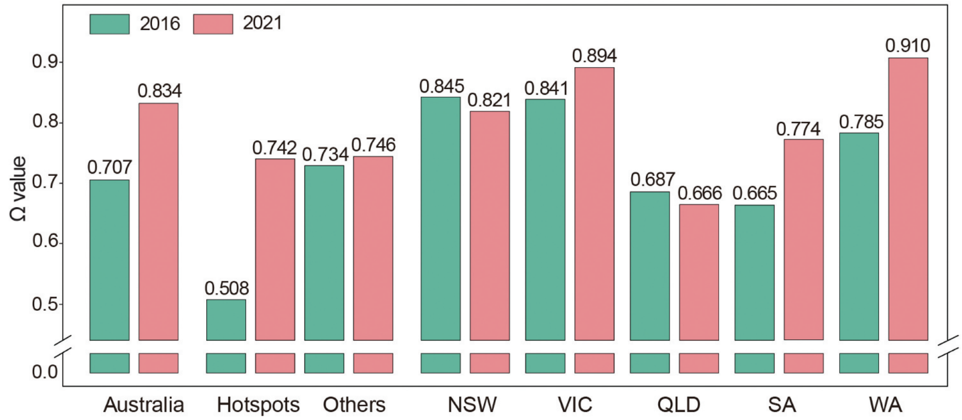

Figure 10. Ωvalues in Australia, hotspot and non-hotspot regions, and across individual states using the GOZH model.

values at different stratified levels, and the overall PD values of the model. The effectiveness of the GOZH model is further validated by comparing it with another geospatial heterogeneity model, optimal parameters-based geographical detector (OPGD), in the context of spatial heterogeneity research.

Figure 11(a, c) illustrate the contribution of individual variables to the OPGD model, using the same explanatory and response variables as inputs. It is evident that in 2016, the OPGD model showed significant contributions from variables such as BDW (14.2%), AWC (9.7%), ASP (8.8%), and NTO (8.2%). In 2021, the top four variables for the OPGD model were EVI, TP, AT, and SLT, contributing 36.7%, 35.5%, 29.5%, and 24.8%, respectively. Figure 11(b,d) depict the differences in individual variable contributions between the two models. Compared to the GOZH model, the OPGD model exhibited much lower contributions from individual variables. For instance, in 2016, the contribution of AWC in the GOZH model was 20.2% higher than in the OPGD model, with the lowest improvement observed for CLY in 2016. This indicates that individual variables in the GOZH model contribute more significantly to the spatial disparities of wheat production, demonstrating a more substantial impact.

Figure 12 shows the variation in PD values for the top four major variables under different hierarchical levels for both GOZH and OPGD models in 2016 and 2021. Figure 12(a)-(d) display the PD value changes for the top four major variables (AWC, NTO, ASP, and PTO) in 2016, while Figure 12(e)-(h) present the PD value changes for the top four variables (EVI, TP, SLT, and NTO) in 2021 across different spatial stratifications. It is evident from the figures that the overall PD values of the GOZH model are relatively higher than those of the OPGD model. The PD values of the GOZH model increase with the number of strata, quickly reaching the optimal PD value, whereas the PD values of the OPGD model fluctuate at lower levels. Specifically, the PD value for EVI in 2021 reaches its optimum at the seventh strata, while SLT and NTO reach their optimum at the eighth strata. Overall, the GOZH model consistently exhibits higher PD values relative to the OPGD model, signifying the GOZH model's robust performance in assessing spatial disparities in wheat production.

Finally, the overall PD values for the two models in 2016 and 2021 are as follows: 2016 OPGD: 0.42, 2016 GOZH: 0.71, 2021 OPGD: 0.65, and 2021 GOZH: 0.83. Clearly, the GOZH model demonstrates superior performance in explaining the spatial disparities of wheat production.

5. Discussion

5.1. Spatial disparities and determinants of wheat production in Australia

By integrating spatial autocorrelation, spatial stratified heterogeneity, and decision tree analysis, this study develops a geospatial machine learning model that identifies the determinants and analyzes the spatial disparities in wheat production. First, utilizing Moran's I and LISA methodology, this study analyzes the spatial autocorrelation of wheat production in Australia, incorporating hotspots and significance analyses. The study area was segmented into four geographically optimal zones based on the interactions of multiple variables, thereby identifying the key determinants influencing wheat production. Enhancing the geographic detector model, this model quantifies heterogeneity between regions through stepwise spatial

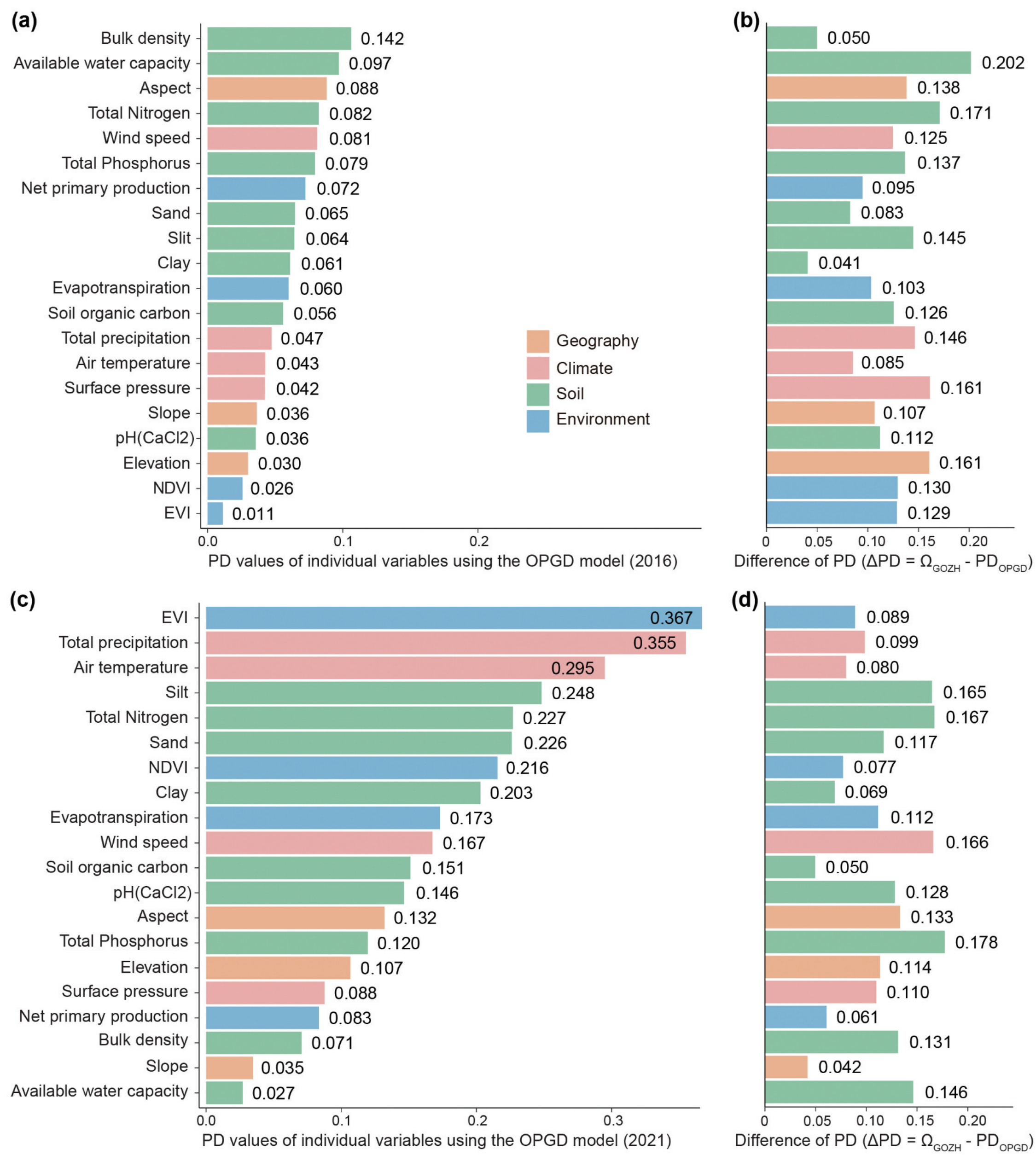

Figure 11. PD values of individual explanatory variables in the OPGD model in 2016 (a) and 2021 (c), differences in individual variable PD values between the OPGD and GOZH models in 2016 (b) and 2021 (d).

discretization and optimization of geographic areas. The GOZH model demonstrated commendable accuracy and precision in assessing the spatial variations of wheat production in Australia. This research provides decision support for understanding spatial patterns and disparities in wheat production and supports the implementation of measures to improve wheat production and ensure food security, such as precision agriculture, soil health management, water management, and the selection of optimal wheat varieties.

Australia is characterized by its vast, sparsely populated landscapes, particularly in the central and western regions, where extensive deserts and arid lands are prevalent. These areas typically feature soil types unsuitable for agricultural production. In contrast, the eastern, southeastern, and southwestern parts of Australia, with suitable rainfall and temperatures, have soils rich in organic matter and humus, making them more conducive to crop growth. Consequently, Australia's wheat-producing areas are concentrated in these regions. Climate elements, especially temperature, precipitation, atmospheric pressure, and wind speed, are vital for crop growth. Wheat is primarily cultivated in temperate climates and some subtropical and Mediterranean regions, providing favorable temperature and rainfall conditions for wheat growth. The

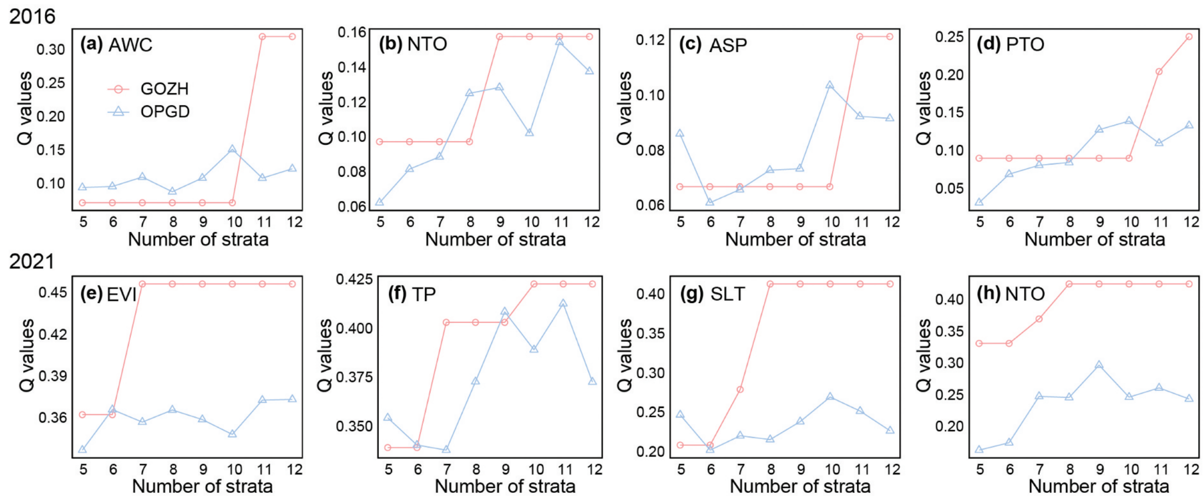

Figure 12. Variation of PD values for the four principal variables across different stratas in the GOZH and OPGD models in 2016 ((a) AWC, (b) Nto, (c) ASP, (d) PTO) and 2021 ((e) EVI, (f) TP, (g) SLT, (h) NTO).

temperate climate, mainly found in NSW, VIC, SA, and the southern part of WA, is crucial for wheat growth, particularly the winter rainfall in these regions. The subtropical climate, predominantly in eastern QLD and northern NSW, offers relatively mild winters. The Mediterranean climate in southwestern WA and parts of SA provides beneficial winter rainfall and moderate temperatures for wheat cultivation. Given the significant regional environmental differences among Australian states, this study employs spatial autocorrelation methods to analyze wheat production's clustering and hotspot distribution characteristics. This approach allows for a more comprehensive understanding of the geographical and spatial variations and correlations of wheat production within the study area.

The factors influencing wheat production are complex, with terrain, climate, soil fertility, and environmental conditions being widely recognized as primary influences (Ajami et al. 2020; Fan et al. 2018; Hodson and White 2007). This research incorporates four categories of variables directly affecting wheat production: geographic, climatic, soil properties, and environmental variables. Through spatial clustering analysis, delineating geographically optimal regions, exploring univariate and multivariate interactions, and model comparison and validation, this study comprehensively examines the spatial disparities in wheat production across Australia. It identifies the determinative factors contributing to these disparities based on model outcomes. Terrain and soil properties directly impact the growth environment and nutrient uptake of wheat, while climate and environmental conditions influence the crop's entire growth cycle and developmental processes. These factors interact, collectively shaping the ultimate production and spatial distribution of wheat. For example, the wetter La Niña conditions during the 2020–21 season intensified canopy greenness, which explains why the Enhanced Vegetation Index (EVI) emerged as the leading determinant that year (Bureau of Meteorology 2022); irrigation-efficiency upgrades in New South Wales and Victoria further amplified this vegetation signal (Hughes, Donoghoe, and Whittle 2020).

The geospatial machine learning model developed in this study integrates machine learning, geographical detector, and stratified heterogeneity models, applying it to analyze the spatial heterogeneity of wheat production in Australia. The model combines geographic spatial elements with external environmental factors, considering the spatial distribution differences in wheat production and the interrelationships among various variables influencing these variations. This approach addresses the limitations of existing heterogeneity models that fail to account for geographic spatial characteristics and typically consider only single-variable impacts (P. Luo et al. 2022). By adopting a comprehensive perspective, this study explores the determinants affecting wheat production differences from a geographical standpoint and within multi-variable interactions. This method represents a novel exploration in the agricultural domain, advancing the study of spatial correlations and disparities in crop production.

In summary, studying the spatial heterogeneity of wheat production contributes significantly to enhancing the efficiency of grain production, maintaining food security, improving resource management, formulating more effective policies, and advancing agricultural science. These aspects are crucial for the sustainability of the global agricultural system and the future of food supply, underscoring the importance of such research efforts.

5.2. Limitations and future recommendations

This study still has limitations. Because Local Government Areas are the finest resolution at which the Australian Bureau of Statistics publishes a fully harmonized, nation-wide dataset pairing wheat yield with all environmental covariates, our analysis focused on the two Agricultural Census seasons 2015–16 and 2020–21. Although a single season is normally sufficient for GD-family models, using two contrasting years provides an additional robustness check of the GOZH framework. Local Government Areas remain relatively coarse, and future work should downscale wheat-growing regions to 1 km or even 100 m grids to capture variability within each LGA more precisely. Once longer, consistently compiled time-series data become available, we will apply GOZH in a spatio-temporal framework to examine interannual dynamics and confirm the persistence of the spatial determinants identified here. Recent advances such as Robust GOZH introduce a Q function to optimize the complexity parameter (CP) within the GOZH partitioning process (H. Luo, Luo, and Meng 2025), offering an objective, data-driven means to select CP and enhance the stability of geographic zone delineation. We will evaluate this approach in future studies to further improve the performance and interpretability of our zoning results.

6. Conclusion

This study develops a geospatial machine learning model, integrating spatial autocorrelation, spatial stratified heterogeneity, and decision tree analysis to analyze spatial disparities and identify determinants in Australian wheat production. This research explores the spatial disparities of wheat production under the influence of multiple factors, identifying the determinants behind these variations. The development of the geospatial machine learning model has opened new paths for exploring spatial disparities and heterogeneity in wheat production. The findings indicate that the interaction of terrain, climate, soil attributes, and environmental conditions are the most significant factors affecting wheat production in Australia. Based on two-year case studies, the model identifies the determinants causing spatial disparities. The wheat production areas are divided into four geographically optimal zones, each exhibiting different determinant distribution characteristics, providing strong evidence for adaptive planting, cultivation, and wheat management. A comparison of the GOZH and OPGD models showed that the GOZH model consistently had higher PD values for individual variables and reached optimal PD values faster than the OPGD model. This indicates that the GOZH model is more effective in explaining the spatial disparities in crop production. This research contributes to an enhanced understanding of the spatial disparities and determinants of wheat production across Australia on a large scale. The developed model can be applied to other crops and regions to analyze the spatial disparities of different crops in various areas, thereby improving global grain production and ensuring food security more effectively.

Acknowledgment

We acknowledge Professor Mark Gibberd, Dr Julia Easton and Dr Arnold Salvacion from the Centre for Crop and Disease Management (CCDM), Curtin University, for their valuable comments on the data and results of our research.

Author contributions

CRediT: **Kai Ren:** Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft; **Yongze Song:** Methodology, Supervision, Visualization, Writing – review & editing; **Linchao Li:** Supervision, Writing – review & editing; **Francesco Mancini:** Writing – review & editing; **Zhuoyao Xiao:** Formal analysis, Writing – review & editing; **Xueyuan Zhang:** Formal analysis, Writing – review & editing; **Rui Qu:** Formal analysis, Writing – review & editing; **Qiang Yu:** Supervision, Writing – review & editing.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the National Key Research and Development Program of China [2023YFB3907403] and supported by the China Scholarship Council Grant No. [202206300058].

ORCID

Yongze Song http://orcid.org/0000-0003-3420-9622
Rui Qu http://orcid.org/0000-0003-0575-9537

Data availability statement

Data and codes supporting the findings of this study are available at https://doi.org/10.6084/m9.figshare.28644923.

References

Ajami, M., A. Heidari, F. Khormali, M. Zeraatpisheh, M. Gorji, and S. Ayoubi. 2020. "Spatial Variability of Rainfed Wheat Production Under the Influence of Topography and Soil Properties in Loess-Derived Soils, Northern Iran." *International Journal of Plant Production* 14 (4): 597–608. https://doi.org/10.1007/s42106-020-00106-4.

Anselin, L. 1995. "Local Indicators of Spatial Association—Lisa." *Geographical Analysis* 27 (2): 93–115. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x.

Anselin, L. 2005. "Exploring Spatial Data with GeoDaTM: A Workbook." In *Center for Spatially Integrated Social Science* 157. Vol. 1963.

Australian Bureau of Statistics. 2017. "Agricultural Commodities, Australia, 2015-16 Financial Year." https://www.abs.gov.au/AUSSTATS/abs%40.nsf/DetailsPage/7121.02015-16?OpenDocument.

Australian Bureau of Statistics. 2020. "Australian Statistical Geography Standard (Asgs): Volume 3 - Non Abs Structures." https://www.abs.gov.au/AUSSTATS/abs%40.nsf/DetailsPage/1270.0.55.003June%202020?OpenDocument.

Australian Bureau of Statistics. 2021. "Australian Statistical Geography Standard (Asgs) Edition 3, Digital Boundary Files." https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/access-and-downloads/digital-boundary-files.

Australian Bureau of Statistics. 2022. "Agricultural Commodities, Australia, 2020-21 Financial Year." https://www.abs.gov.au/statistics/industry/agriculture/agricultural-commodities-australia/2020-21.

Australian Export Grains Innovation Centre. 2022. "Australian Wheat Information Brochure." https://www.aegic.org.au/australian-grains/wheat/.

Breiman, L. 2017. *Classification and Regression Trees*. Boca Raton, FL: Routledge.

Bureau of Meteorology. 2022. Annual Climate Statement 2021. *Government Report*. Australian Government Bureau of Meteorology. Accessed June 12, 2025. https://www.bom.gov.au/climate/current/annual/aus/.

Cammarano, D., J. Payero, B. Basso, P. Wilkens, and P. Grace. 2012. "Agronomic and Economic Evaluation of Irrigation Strategies on Cotton Lint Yield in Australia." *Crop and Pasture Science* 63 (7): 647–655. https://doi.org/10.1071/CP12024.

Chakraborty, A., A. Biswal, V. Pandey, S. Shadab, K. Kalyandeep, C. Murthy, M. Seshasai, et al. 2022. "Developing a Spatial Information System of Biomass Potential from Crop Residues Over India: A Decision Support for Planning and Establishment of Biofuel/Biomass Power Plant." *Renewable and Sustainable Energy Reviews* 165:112575. https://doi.org/10.1016/j.rser.2022.112575.

Chu, L., C. Huang, Q. Liu, C. Cai, and G. Liu. 2019. "Spatial Heterogeneity of Winter Wheat Yield and Its Determinants in the Yellow River Delta, China." *Sustainability* 12 (1): 135. https://doi.org/10.3390/su12010135.

Diacono, M., A. Castrignanò, A. Troccoli, D. De Benedetto, B. Basso, and P. Rubino. 2012. "Spatial and Temporal Variability of Wheat Grain Yield and Quality in a Mediterranean Environment: A Multivariate Geostatistical Approach." *Field Crops Research* 131:49–62. https://doi.org/10.1016/j.fcr.2012.03.004.

Didan, K. 2021. "Modis/Terra Vegetation Indices 16-Day L3 Global 1km Sin Grid v061. Distributed by NASA EOSDIS Land Processes Distributed Active Archive Center." https://doi.org/10.5067/MODIS/MOD13A2.061.

Dong, J., H. Lu, Y. Wang, T. Ye, and W. Yuan. 2020. "Estimating Winter Wheat Yield Based on a Light Use Efficiency Model and Wheat Variety Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 160:18–32. https://doi.org/10.1016/j.isprsjprs.2019.12.005.

Duncan, E. G., C. A. O'Sullivan, M. M. Roper, M. B. Peoples, K. Treble, and K. Whisson. 2017. "Crop and Microbial Responses to the Nitrification Inhibitor 3, 4-Dimethylpyrazole Phosphate (Dmpp) in Mediterranean Wheat-Cropping Systems." *Soil Research* 55 (6): 553–566. https://doi.org/10.1071/SR16327.

Fan, D., F. Yang, Z. Pan, X. Su, Y. Pan, G. Han, J. Wang, D. Wu, and Z. Dong. 2018. "Development of an Improved Model to Evaluate Vulnerability in Spring Wheat Under Climate Change in Inner Mongolia." *Sustainability* 10 (12): 4581. https://doi.org/10.3390/su10124581.

FAO (Food and Agriculture Organization of the United Nations). 2024. "Production/Crops and Livestock Products." https://www.fao.org/faostat/en/.

Feng, P., B. Wang, I. Macadam, A. S. Taschetto, N. J. Abram, J.-J. Luo, A. D. King, et al. 2022. "Increasing Dominance of Indian Ocean Variability Impacts Australian Wheat Yields." *Nature Food* 3 (10): 862–870. https://doi.org/10.1038/s43016-022-00613-9.

Fletcher, A. L., C. Chen, N. Ota, R. A. Lawes, and Y. M. Oliver. 2020. "Has Historic Climate Change Affected the Spatial Distribution of Water-Limited Wheat Yield Across Western Australia?" *Climatic Change* 159 (3): 347–364. https://doi.org/10.1007/s10584-020-02666-w.

Fu, Y., J. Huang, Y. Shen, S. Liu, Y. Huang, J. Dong, W. Han, T. Ye, W. Zhao, and W. Yuan. 2021. "A Satellite-Based Method for National Winter Wheat Yield Estimating in China." *Remote Sensing* 13 (22): 4680. https://doi.org/10.3390/rs13224680.

Geoscience Australia. 2015. "Digital elevation model (dem) of australia derived from lidar 5 metre grid." https://developers.google.com/earthengine/datasets/catalog/AU_GA_DEM_1SEC_v10_DEM-S.

Gomez-Macpherson, H., and R. Richards. 1995. "Effect of Sowing Time on Yield and Agronomic Characteristics of Wheat in South-Eastern Australia." *Australian Journal of Agricultural Research* 46 (7): 1381–1399. https://doi.org/10.1071/AR9951381.

Guarin, J. R., P. Martre, F. Ewert, H. Webber, S. Dueri, D. Calderini, M. Reynolds, et al. 2022. "Evidence for Increasing Global Wheat Yield Potential." *Environmental Research Letters* 17 (12): 124045. https://doi.org/10.1088/1748-9326/aca77c.

Guerschman, J. P., T. R. McVicar, J. Vleeshower, T. G. Van Niel, J. L. Peña-Arancibia, and Y. Chen. 2022. "Estimating Actual Evapotranspiration at Field-To-Continent Scales by Calibrating the cmrset algorithm with modis, Viirs, Landsat and Sentinel-2 Data." *Journal of Hydrology* 605:127318. https://doi.org/10.1016/j.jhydrol.2021.127318.

Gutierrez, L., and L. A. Braunstein. 2017. "Impacts of el niño-Southern Oscillation on the Wheat Market: A Global Dynamic Analysis." *PLoS One* 12 (6): e0179086. https://doi.org/10.1371/journal.pone.0179086.

Han, J., Y. Liu, C. Hou, J. Li, J. Wang, Q. Zhang, Q. Yang, X. Chen, and J. Wu. 2020. "A 1ns Disomic Addition from Psathyrostachys Huashanica Keng Confers Resistance to Powdery Mildew in Wheat." *Agronomy* 10 (2): 312. https://doi.org/10.3390/agronomy10020312.

He, Y., J. Fang, W. Xu, and P. Shi. 2022. "Substantial Increase of Compound Droughts and Heatwaves in Wheat Growing Seasons Worldwide." *International Journal of Climatology* 42 (10): 5038–5054. https://doi.org/10.1002/joc.7518.

Hernandez-Ochoa, I. M., S. Asseng, B. T. Kassie, W. Xiong, R. Robertson, D. N. L. Pequeno, K. Sonder, et al. 2018. "Climate Change Impact on Mexico Wheat Production." *Agricultural and Forest Meteorology* 263:373–387. https://doi.org/10.1016/j.agrformet.2018.09.008.

Hersbach, H., B. Bell, P. Berrisford, G. Biavati, A. Horányi, J. Muñoz Sabater, J. Nicolas, et al. 2023. "Era5 Monthly Averaged Data on Single Levels from 1979 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS)."

Hodson, D., and J. White. 2007. "Use of Spatial Analyses for Global Characterization of Wheat-Based Production Systems." *The Journal of Agricultural Science* 145 (2): 115. https://doi.org/10.1017/S0021859607006855.

Hou, Q., Z. Hu, Y. Yang, X. Fan, J. Zhou, L. Zhang, and Y. Duan. 2023. "The Influence Mechanism of Farmer Behavior on the Spatial Pattern Evolution of Agricultural Production in the Yanhe River Basin." *The Land* 12 (7): 1281. https://doi.org/10.3390/land12071281.

Hu, J., Y. Song, and T. Zhang. 2025. "A Local Indicator of Stratified Power." *International Journal of Geographical Information Science* 39 (4): 925–943. https://doi.org/10.1080/13658816.2024.2437811.

Hughes, N., M. Donoghoe, and L. Whittle. 2020. "Farm Level Effects of On-Farm Irrigation Infrastructure Programs in the Southern Murray–Darling Basin." *Australian Economic Review* 53 (4): 494–516. https://doi.org/10.1111/1467-8462.12396.

Jin, N., B. Tao, W. Ren, L. He, D. Zhang, D. Wang, and Q. Yu. 2022. "Assimilating Remote Sensing Data into a Crop Model Improves Winter Wheat Yield Estimation Based on Regional Irrigation Data." *Agricultural Water Management* 266:107583. https://doi.org/10.1016/j.agwat.2022.107583.

Jinfeng, W. A. N. G., and C. X. 2017. "Geodetector: Principle and Prospective." *Acta Geographica Sinica* 72 (1): 116. https://doi.org/10.3724/SP.J.1041.2017.00116.

Jiu-Jiang, W., W. Nan, S. Hong-Zheng, and M. Xiao-Yi. 2022. "Spatial–Temporal Variation of Climate and Its Impact on Winter Wheat Production in Guanzhong plain, China." *Computers and Electronics in Agriculture* 195:106820. https://doi.org/10.1016/j.compag.2022.106820.

Kingwell, R. 2020. "The Changing Trade Landscape in Asian Grain Markets: An Australian Perspective." *Cereal Foods World* 65 (5): 51–53. https://doi.org/10.1094/CFW-65-5-0051.

Kitchen, N., S. Drummond, E. Lund, K. Sudduth, and G. Buchleiter. 2003. "Soil Electrical Conductivity and Topography Related to Yield for Three Contrasting Soil–Crop Systems." *Agronomy Journal* 95 (3): 483–495. https://doi.org/10.2134/agronj2003.4830.

Langridge, P., M. Alaux, N. F. Almeida, K. Ammar, M. Baum, F. Bekkaoui, A. R. Bentley, et al. 2022. "Meeting the Challenges Facing Wheat Production: The Strategic Research Agenda of the Global Wheat Initiative." *Agronomy* 12 (11): 2767. https://doi.org/10.3390/agronomy12112767.

Li, H., C. A. Calder, and N. Cressie. 2007. "Beyond moran's I: Testing for Spatial Dependence Based on the Spatial Autoregressive Model." *Geographical Analysis* 39 (4): 357–375. https://doi.org/10.1111/j.1538-4632.2007.00708.x.

Liu, Z., H. Lei, H. Sheng, and Y. Wang. 2023. "Analysis of Soil Organic Matter Influencing Factors in the Huangshui River Basin by Using the Optimal Parameter-Based Geographical Detector Model." *Geocarto International* 38 (1): 2246935. https://doi.org/10.1080/10106049.2023.2246935.

Luo, H., P. Luo, and L. Meng. 2025. "A Robust Geographically Optimal Zones-Based Heterogeneity Model for Analyzing the Spatial Determinants of National Traffic Accidents." *GIScience & Remote Sensing* 62 (1): 2448283. https://doi.org/10.1080/15481603.2024.2448283.

Luo, P., Y. Song, X. Huang, H. Ma, J. Liu, Y. Yao, and L. Meng. 2022. "Identifying Determinants of Spatio-Temporal Disparities in Soil Moisture of the Northern Hemisphere Using a Geographically Optimal Zones-Based Heterogeneity Model." *Isprs Journal of Photogrammetry & Remote Sensing* 185:111–128. https://doi.org/10.1016/j.isprsjprs.2022.01.009.

Luo, Q., B. Bryan, W. Bellotti, and M. Williams. 2005. "Spatial Analysis of Environmental Change Impacts on Wheat Production in Mid-Lower North, South Australia." *Climatic Change* 72 (1–2): 213–228. https://doi.org/10.1007/s10584-005-5361-1.

Lv, X., H. Zhang, Y. Zhang, and J. Li. 2023. "Spatial Correlation Evolution and Driving Factors of Wheat Production in China Based on Social Network Analysis." *Agronomy* 13 (2): 515. https://doi.org/10.3390/agronomy13020515.

Lv, Z., X. Liu, W. Cao, and Y. Zhu. 2017. "A Model-Based Estimate of Regional Wheat Yield Gaps and Water Use Efficiency in Main Winter Wheat Production Regions of China." *Scientific Reports* 7 (1): 6081. https://doi.org/10.1038/s41598-017-06312-x.

Mao, Y., Y. Liu, L. Zhuo, W. Wang, M. Li, B. Feng, and P. Wu. 2021. "Quantitative Evaluation of Spatial Scale Effects on Regional Water Footprint in Crop Production." *Resources, Conservation and Recycling* 173:105709. https://doi.org/10.1016/j.resconrec.2021.105709.

Millar, N., A. Urrea, K. Kahmark, I. Shcherbak, G. P. Robertson, and I. Ortiz-Monasterio. 2018. "Nitrous Oxide (N2o) Flux Responds Exponentially to Nitrogen Fertilizer in Irrigated Wheat in the Yaqui Valley, Mexico." *Agriculture, Ecosystems & Environment* 261:125–132. https://doi.org/10.1016/j.agee.2018.04.003.

Miller, M. P., M. J. Singer, and D. R. Nielsen. 1988. "Spatial Variability of Wheat Yield and Soil Properties on Complex Hills." *Soil Science Society of America Journal* 52 (4): 1133–1141. https://doi.org/10.2136/sssaj1988.03615995005200040045x.

Min, M., W. Zhao, T. Hu, J. Chen, and X. Nie. 2014. "Influential Factors of Spatial Distribution of Wheat Yield in China During 1978–2007: A Spatial Econometric Analysis." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7 (11): 4453–4460. https://doi.org/10.1109/JSTARS.2014.2325898.

Moran, P. A. 1950. "Notes on Continuous Stochastic Phenomena." *Biometrika* 37 (1–2): 17–23. https://doi.org/10.1093/biomet/37.1-2.17.

Nabiollahi, K., E. Heshmat, A. Mosavi, R. Kerry, M. Zeraatpisheh, and R. Taghizadeh-Mehrjardi. 2020. "Assessing the Influence of Soil Quality on Rainfed Wheat Yield." *Agriculture* 10 (10): 469. https://doi.org/10.3390/agriculture10100469.

Obanor, F., S. Neate, S. Simpfendorfer, R. Sabburg, P. Wilson, and S. Chakraborty. 2013. "Fusarium Graminearum and Fusarium Pseudograminearum Caused the 2010 Head Blight Epidemics in Australia." *Plant Pathology* 62 (1): 79–91. https://doi.org/10.1111/j.1365-3059.2012.02615.x.

Orton, T. G., T. Mallawaarachchi, M. J. Pringle, N. W. Menzies, R. C. Dalal, P. M. Kopittke, R. Searle, Z. Hochman, and Y. P. Dang. 2018. "Quantifying the Economic Impact of Soil Constraints on Australian Agriculture: A Case-Study of Wheat." *Land Degradation & Development* 29 (11): 3866–3875. https://doi.org/10.1002/ldr.3130.

Pang, A., M. W. L. Chang, and Y. Chen. 2022. "Evaluation of Random Forests (Rf) for Regional and Local-Scale Wheat Yield Prediction in Southeast Australia." *Sensors (Switzerland)* 22 (3): 717. https://doi.org/10.3390/s22030717.

Qader, S. H., J. Dash, and P. M. Atkinson. 2018. "Forecasting Wheat and Barley Crop Production in Arid and Semi-Arid Regions Using Remotely Sensed Primary Productivity and Crop Phenology: A Case Study in Iraq." *Science of the Total Environment* 613:250–262. https://doi.org/10.1016/j.scitotenv.2017.09.057.

Richter, G., A. Beblik, K. Schmalstieg, and O. Richter. 1998. "N-Dynamics and Nitrate Leaching Under Rotational and Continuous Set-aside—A Case Study at the Field and Catchment Scale." *Agriculture, Ecosystems & Environment* 68 (1–2): 125–138. https://doi.org/10.1016/S0167-8809(97)00142-4.

Rossel, R. A., C. Chen, M. Grundy, R. Searle, D. Clifford, and P. Campbell. 2015. "The Australian Three-Dimensional Soil Grid: Australia's Contribution to the Globalsoilmap Project." *Soil Research* 53 (8): 845–864. https://doi.org/10.1071/SR14366.

Running, S., and M. Zhao. 2021. "Modis/Terra Net Primary Production Gap-Filled Yearly L4 Global 500m Sin Grid v061. Distributed by NASA EOSDIS Land Processes Distributed Active Archive Center." https://doi.org/10.5067/MODIS/MOD17A3HGF.061.

Sbahi, M., A.-R. Ziboon, and K. Hassoon. 2018. "Sentinel-2 Imageries with Meteorological Data Utilize Remote Sensing and Gis for Wheat Crop Prove Production." In *MATEC Web of Conferences*, 03019. Vol. 162. Les Ulis, France: EDP Sciences.

Sharps, K., M. Vieno, R. Beck, F. Hayes, and H. Harmens. 2021. "Quantifying the Impact of Ozone on Crops in Sub-Saharan Africa Demonstrates Regional and Local Hotspots of Production Loss." *Environmental Science and Pollution Research* 28 (44): 62338–62352. https://doi.org/10.1007/s11356-021-14967-3.

Shen, J., and F. H. Evans. 2021. “The Potential of Landsat NDVI Sequences to Explain Wheat Yield Variation in Fields in Western Australia.” *Remote Sensing* 13 (11): 2202. https://doi.org/10.3390/rs13112202.

Song, Y., H. W. Linderholm, C. Wang, J. Tian, Z. Huo, P. Gao, Y. Song, and A. Guo. 2019. “The Influence of Excess Precipitation on Winter Wheat Under Climate Change in China from 1961 to 2017.” *Science of the Total Environment* 690:189–196. https://doi.org/10.1016/j.scitotenv.2019.06.367.

Song, Y., J. Wang, Y. Ge, and C. Xu. 2020. “An Optimal Parameters-Based Geographical Detector Model Enhances Geographic Characteristics of Explanatory Variables for Spatial Heterogeneity Analysis: Cases with Different Types of Spatial Data.” *GIScience and Remote Sensing* 57 (5): 593–610. https://doi.org/10.1080/15481603.2020.1760434.

Therneau, T., and B. Atkinson. 2023. “Rpart: Recursive Partitioning and Regression Trees. R Package Version 4.1.23.” https://CRAN.R-project.org/package=rpart.

Wang, B., C. Chen, D. Li Liu, S. Asseng, Q. Yu, and X. Yang. 2015. “Effects of Climate Trends and Variability on Wheat Yield Variability in Eastern Australia.” *Climate Research* 64 (2): 173–186. https://doi.org/10.3354/cr01307.

Wang, B., D. L. Liu, G. J. O’Leary, S. Asseng, I. Macadam, R. Lines-Kelly, X. Yang, et al. 2018. “Australian Wheat Production Expected to Decrease by the Late 21st Century.” *Global Change Biology* 24 (6): 2403–2415. https://doi.org/10.1111/gcb.14034.

Wang, E., J. Xu, Q. Jiang, and J. Austin. 2009. “Assessing the Spatial Impact of Climate on Wheat Productivity and the Potential Value of Climate Forecasts at a Regional Level.” *Theoretical and Applied Climatology* 95 (3–4): 311–330. https://doi.org/10.1007/s00704-008-0009-5.

Wang, J. F., Y. Ge, L. F. Li, B. Meng, J. Wu, Y. Bo, S. Du, Y. Liao, M. Hu, and C. Xu. 2014. “Spatiotemporal Data Analysis in Geography.” *Acta Geographica Sinica* 69 (9): 1326–1345.

Wang, J. F., X. H. Li, G. Christakos, Y. L. Liao, T. Zhang, X. Gu, and X. Y. Zheng. 2010. “Geographical Detectors-Based Health Risk Assessment and Its Application in the Neural Tube Defects Study of the Heshun region, China.” *International Journal of Geographical Information Science* 24 (1): 107–127. https://doi.org/10.1080/13658810802443457.

Wang, Y., X. Xu, L. Huang, G. Yang, L. Fan, P. Wei, and G. Chen. 2019. “An Improved Casa Model for Estimating Winter Wheat Yield from Remote Sensing Images.” *Remote Sensing* 11 (9): 1088. https://doi.org/10.3390/rs11091088.

Xiong, W., D. Conway, I. Holman, and E. Lin. 2008. “Evaluation of Ceres-Wheat Simulation of Wheat Production in China.” *Agronomy journal* 100 (6): 1720–1728. https://doi.org/10.2134/agronj2008.0081.

Xu, E., R. Wang, H. Zhang, and Z. Yu. 2019. “Coupling Index of Water Consumption and Soil Fertility Correlated with Winter Wheat Production in North China Region.” *Ecological Indicators* 102:154–165. https://doi.org/10.1016/j.ecolind.2019.02.045.

Zeng, X., Z. Li, F. Zeng, F. Caputo, and T. Chin. 2023. “Spatiotemporal Evolution and Antecedents of Rice Production Efficiency: From a Geospatial Approach.” *Systems* 11 (3): 131. https://doi.org/10.3390/systems11030131.

Zhang, H., Y. Tang, A. A. Chandio, G. R. Sargani, and M. Ankrah Twumasi. 2022. “Measuring the Effects of Climate Change on Wheat Production: Evidence from Northern China.” *International Journal of Environmental Research and Public Health* 19 (19): 12341. https://doi.org/10.3390/ijerph191912341.

Zhang, W., M. Xi, H. Liu, and H. Zheng. 2023. “Low Sensitivity of Net Primary Productivity to Climatic Factors in Three Karst Provinces in Southwest China from 1981 to 2019.” *Ecological Indicators* 153:110465. https://doi.org/10.1016/j.ecolind.2023.110465.

Zhang, Y., and B. Li. 2022. “Detection of the Spatio-Temporal Differentiation Patterns and Influencing Factors of Wheat Production in Huang-Huai-Hai Region.” *Foods* 11 (11): 1617. https://doi.org/10.3390/foods11111617.

Zhang, Z., Z. Li, and Y. Song. 2024. “On Ignoring the Heterogeneity in Spatial Autocorrelation: Consequences and Solutions.” *International Journal of Geographical Information Science* 38 (12): 2545–2571. https://doi.org/10.1080/13658816.2024.2391981.

Zhang, Z., Y. Song, L. Karunaratne, and P. Wu. 2024. “Robust Interaction Detector: A Case of Road Life Expectancy Analysis.” *Spatial Statistics* 59:100814. https://doi.org/10.1016/j.spasta.2024.100814.