Front. Earth Sci. 2022, 16(2): 401–410
https://doi.org/10.1007/s11707-021-0939-0

RESEARCH ARTICLE

Determination of the effective utilization coefficient of irrigation water based on geographically weighted regression

Rui SHI[1,2,3], Gaoxu WANG (✉)[2,3], Xuan ZHANG[2,3], Yi XU[2,3], Yongxiang WU[2,3], Wei WU[2,3]

1 College of Hydrology and Water Resource, Hohai University, Nanjing 210098, China
2 State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Nanjing 210029, China
3 Hydrology and Water Resources Department, Nanjing Hydraulic Research Institute, Nanjing 210029, China

Abstract This study uses geographically weighted regression to determine the spatial distribution of the effective utilization coefficient of irrigation water in Zhejiang Province, China, owing to the influences of spatial attributes on the irrigation efficiency. The sample set of this study comprised 165 agricultural test sites. A multivariate linear regression model and a geographically weighted regression model were established using the effective utilization coefficient of agricultural irrigation water as the dependent variable in addition to a suite of independent variables, including the actual irrigation area, the percentage of farmland using water-saving irrigation, the type of irrigation area, the net water consumption per mu, the water intake method, the terrain slope, and the soil field capacity. Results revealed a positive spatial correlation and noticeable agglomeration features in the effective utilization coefficient of irrigation water in Zhejiang Province. The geographically weighted regression model performed better in terms of fit and prediction accuracy than the multivariate linear regression model. The obtained findings confirm the suitability of the geographically weighted regression model for determining the spatial distribution of the effective utilization coefficient of irrigation water in Zhejiang, and offer a new approach on a regional scale.

Keywords effective utilization coefficient of irrigation water, spatial autocorrelation, multivariate linear regression, geographically weighted regression

1 Introduction

The agricultural sector accounts for more than 60% of China's total water use, and more than 90% of which is used for irrigation (Ministry of Water Resources of the People's Republic of China, 2020). There is significant potential for improving the low efficiency of the country's agricultural water consumption via the adoption of water-saving measures. With an increasing conflict between water supply and demand, it has become particularly important to improve the efficiency of agricultural irrigation and to exploit the potential for water saving in this sector. The effective utilization coefficient of irrigation water is the ratio of water diverted from its source that actually reaches the crops and contributes to their growth. It is an important indicator for measuring the efficiency of irrigation water use (Omezzine and Zaibet, 1998; Yilmaz et al., 2009) and for implementing the most rigorous water resources management system in China. Moreover, it is of great scientific importance for the objective evaluation and improvement of efficiency in irrigation water use in agriculture (Tang et al., 2015).

The effective utilization coefficient of irrigation water in an irrigated area is typically measured using the head and end measurement (Feng et al., 2020). The regional value of the variable is obtained by determining the weighted average of the coefficient and the water consumption at irrigated sites of different scales and types. Although the head and end measurement method is improved compared with traditional methods, the collected data on the crop water demand and irrigation quotas exhibit significant variations based on the differences in site irrigation, natural environments, and engineering conditions; these issues result in problems for the continuous and accurate measurement of data. Traditional measurement methods do not consider the geospatial attribute of effective utilization coefficient of irrigation water. In recent years, some studies have adopted regression model, remote sensing monitoring and other methods to calculate effective utilization coefficient of irrigation water. A process using a combination of remote sensing and ground level measurements was developed to evaluate on-farm

Received March 9, 2021; accepted July 26, 2021

E-mail: gxwang@nhri.cn

irrigation efficiency across the watershed (Ahadi et al., 2013). A spatially distributed model to evaluate efficiencies of irrigation scenarios was developed (Tromboni et al., 2014). Deterministic and stochastic models could be used to measure irrigation efficiency (Pereira and Marques, 2017). The spatial attributes of the irrigation water utilization coefficient typically result in spatial non-stationarity, e.g, an area with good irrigation conditions and flat terrain usually has an above average effective utilization coefficient. The head and end measurement can only obtain the effective utilization coefficient of irrigation water at typical irrigation sample sites and not in areas wherein large-scale irrigation is performed, and traditional regression models typically perform analyses from a global perspective without considering spatial attributes, reducing the scientific credibility and practicality (Kaneko et al., 2004). It was found that the GWR model considers the non-stationariness of spatial processes and is more suitable for spatial modeling than multiple linear regression (MLR) (Szymanowski and Kryza, 2011).

Geographically weighted regression (GWR) was derived from the general linear regression model and aims to address the spatial nonstationarity and heterogeneity of elements in a statistical regression analysis. This technique has been widely applied, and its advantages in terms of spatial data analysis and digital mapping have drawn attention of many researchers (Harris et al., 2010; Li et al., 2019). The GWR model considers the influence of local spatial samples and performs distance-weighted analyses based on the measurements of adjacent sample points. It more accurately reveals the quantitative relationship between elements and independent variables, improving the fit of the model. It has been proved that GWR model is more suitable for spatial modeling than multiple linear regression (MLR) account for considering the non-stationariness of spatial processes (Szymanowski and Kryza, 2011). In the calculation of crop water demand, the regression coefficient of GWR shows strong variability in space, which can better explain the spatial differences of the influences of various influencing factors on crop water demand (Wang et al., 2013). In the analysis of spatial interaction between regional industrialization level and various factors, the GWR method is obviously superior to the traditional regression analysis method, and can better reflect spatial variability (Huang and Leung, 2002).

Based on the existing research results and the multivariate linear regression method, this study selected the factors influencing the irrigation water-use efficiency and performed a spatial autocorrelation analysis to identify the spatial clustering of these independent variables. Then, a regression model representing the relationship between the efficiency of irrigation and its factors of influence was established using the GWR method. The fit of the GWR model and the multivariate linear regression model were analyzed and compared along with the spatial relationship between the independent variables and irrigation water use efficiency. By combining these two methods, this study explored a new approach to determining the large-scale spatial distribution of the effective utilization coefficient of irrigation water using both satellite and statistical data when actual measurements were limited.

2 Research area and data source

The research area of this study is Zhejiang Province in eastern China. It comprised 165 sample sites of irrigation area, which were used in the determination of the effective utilization coefficient of irrigation water. Figure 1 presents a distribution map of the sample sites in the irrigation area. Many scholars have made abundant research achievements on the influencing factors affecting the effective utilization coefficient of irrigation water (Song et al., 2018). The role of large-scale drivers, such as climate, population and adaption of efficient irrigation practices, in the water application rate, an indicator of irrigation efficiency in the United States (Das Bhowmik et al., 2020). Two new indices, the fraction of irrigation water (FIW), and the irrigation intensity of agriculture (IIA), were developed to depict the contribution of irrigation water (Chen et al., 2014). The effects on the project irrigation efficiency of some parameters, including irrigation rate, density of irrigation facilities, average farm size, density of distribution scheme and density of personnel per unit area, were studied (Koç, 2013). Five aspects — natural conditions, irrigation types and size, irrigation project status, irrigation technology, and management level — which could affect the irrigation water use efficiency factor in China were analyzed (Feng, 2013). The influential factors of effective utilization coefficient of irrigation water in south China were analyzed from the aspects of water-saving reconstruction project, irrigation patterns, planting structure, water resources management and land management of irrigation districts (Huang et al., 2018). The type of irrigation area, channel lining, water allocation management and field irrigation technology were considered to have significant impact on the effective utilization coefficient of irrigation water in Zhejiang Province (Cao, 2010). The existing research results, especially those related to Zhejiang Province, the research area of this study, as well as the availability of data acquisition were considered in this paper. Six quantifiable indexes were selected as indicators for analyzing the effective utilization coefficient of irrigation water, as the actual irrigation area and the net irrigation water consumption per mu represented for the scale of irrigation area, the percentage of farmland using water-saving irrigation represented for the situation of water-saving engineering measures, the irrigation types represented for the irrigation method, the terrain slope and the soil field capacity of each sample site represented for topographic features of the area. The respective definitions and calculation methods are provided in Table 1. The data for the coefficient, the actual irrigation area, the percentage of farmland using water-saving irrigation, the type of irrigation area, and the net irrigation water consumption per mu were derived from the analysis of the coefficient of the effective utilization of irrigation water in the irrigated areas of Zhejiang Province in 2017 and 2018. The study used ArcGIS tools to process data from the global 30-m SRTM DEM (Jarvis et al., 2008) to obtain the terrain slope data. Data for soil field capacity were obtained from the global soil data set from the Institute of Land–Atmosphere Interaction at Sun Yat-sen University. The hydraulic properties of the soil were estimated from soil-transfer functions using readily available physical and chemical soil parameters. The types of irrigation at the sample sites in Zhejiang included gravity and lift irrigation. Dummy variables were used to quantify the irrigation type, i.e., 0 and 1 were used to represent gravity irrigation and lift irrigation, respectively.

3 Methods

3.1 Spatial autocorrelation

Spatial autocorrelation evaluates the degree of interdependence of various elements in adjacent environments within a spatial context and provides the foundation for the construction of a GWR model. If the spatial correlation is not substantial, it indicates that the distance between the elements has little effect on the value of the variable, i.e., the spatial attribute of the element does not significantly influence the variable. In this case, the GWR loses its practical value.

A spatial autocorrelation analysis involves both global and local spatial autocorrelations. Global spatial autocorrelation only determines if agglomeration or outliers exist and reveals no specific spatial information regarding data agglomeration. Local spatial autocorrelation can identify the specific geographic locations wherein outliers or agglomeration occur; furthermore, it can reveal spatial heterogeneity to determine spatial dependences between neighboring areas with changes in geographic locations. Moran’s I is the primary measurement index for spatial autocorrelation analyses and is expressed as follows:

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

where I is the value of the Moran index, x_i and x_j are the values of the influential factor x in the adjacent spaces, $\overline{x}$ is the average of the influential factor x, n is the total number of sample sites in the irrigation area, and w_{ij} is the spatial weight of elements i and j.

When Moran’s $I > 0$, there is a positive spatial correlation; the higher the value the more substantial the spatial correlation. However, when Moran’s $I < 0$, a negative spatial correlation exists, as a lower value indicates greater spatial heterogeneity. Finally, when Moran’s $I = 0$, the space exhibits complete randomness.

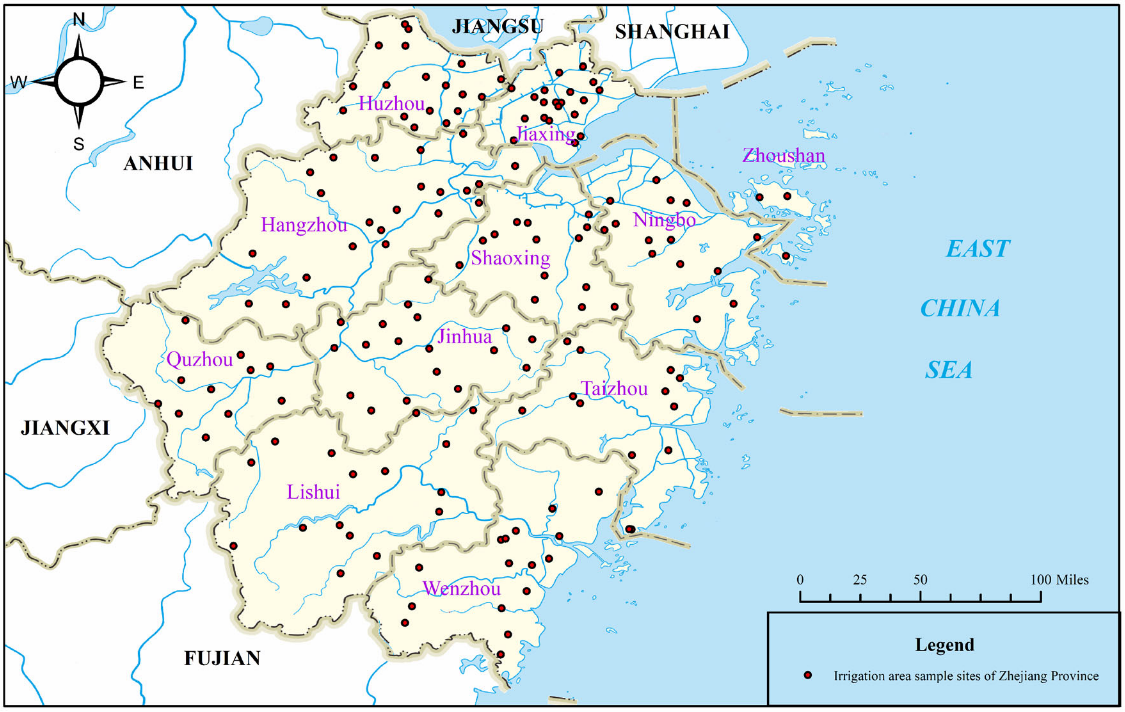

Fig. 1 Distribution of irrigation area sample sites in Zhejiang Province.

3.2 Multivariate linear regression model

Multivariate linear regression refers to a linear correlation model that illustrates the quantitative relationship between two or more independent variables and a dependent variable in a regression analysis. The ordinary least squares method is a commonly used multivariate linear regression approach that predicts the response variable via a series of predictor variables. The equation is as follows:

$$Y_t = \alpha + \beta x_t + \mu_t, \tag{2}$$

where $t = (1,2,\ldots,n)$ represents the number of observations, Y_t is the dependent variable, x_t is the independent variable, α and β are the regression coefficients, and μ_t is the term of random error.

In the model for predicting the effective utilization coefficient of irrigation water, Y_t is the matrix of the coefficient, x_t is the matrix of each influential factor, β is the regression coefficient matrix, and μ_t is the term of random error.

3.3 Geographically weighted regression model

The GWR uses regression to investigate the quantitative relationship between multiple variables with spatial (or regional) distribution characteristics. The essence of GWR is the local weighting of the least square method, where the weight is described with the distance function of the spatial position of the point to be estimated and the spatial position of the remaining observation points. The mathematical form of the GWR model is expressed as follows:

$$y_i = a_0(u_i,v_i) + \sum\nolimits_k a_k(u_i,v_i)x_{ik} + \varepsilon_i, \tag{3}$$

where y_i is the dependent variable at point i, x_{ik} is the value of the kth independent variable at the ith point, k is the number of independent variables; i is the number of sample points, ε_i is the residual, (u_i,v_i) are the geographic coordinates of the ith point, and $a_k(u_i,v_i)$ is the value of the function $a_k(u,v)$ at point i.

Since the GWR model cannot solve the problem of the variation of regression coefficients over time, Huang et al. (2010) merged temporal distance and spatial distance to construct a measurement representation of temporal and spatial distance, and proposed a temporal and spatial geographic weighted regression (GTWR) model. The mathematical form of the GWR model is expressed as follows:

$$y_i = \beta_k(u_i,v_i) + \sum\nolimits_{k=1}^{p} \ (u_i,v_i)x_{ik} + \varepsilon_i, \tag{4}$$

where (u_i,v_i) are the geographic coordinates of the ith point, $\beta_k(u_i,v_i)$ is the kth independent variable coefficient of the sample point i, which is related to the spatial position of the point, and ε_i is the error term of the ith point.

According to the mathematical expression of GTWR model, if it is determined that the time of the model does not change, the model degenerates into a GWR model.

The regression coefficient $a(u_i,v_i)$ in GWR model was predicted through local regression using the subsample data information from nearby observations. The equation is as follows:

$$a(u_i,v_i) = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}(u_i,v_i)\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{W}(u_i,v_i)y, \tag{5}$$

where $\boldsymbol{W}(u_i,v_i)$ is the distance weight matrix (a diagonal matrix with entries $[W_{i1}, W_{i2}, \ldots, W_{in}]$) and n is the number of samples. The nondiagonal elements are 0. W_{ij} shows the influence of the jth point on the ith point, which is generally defined as $W_{ij} = \exp(-d_{ij}^2/h^2)$, where d_{ij} is the distance between points i and j, and h is the customized bandwidth.

The spatial weight matrix uses a monotonically decreasing function (typically a Gaussian function) to reveal the relationship between distance and weight (W_{ij}). To avoid extensive calculations from the discrete data, this study selected Bi-square kernel function as a substitute for the Gaussian calculation. To perform the calculations faster, points with little or no influence were excluded.

$$w_{ij} = \begin{cases} \left[1-\left(\dfrac{d_{ij}}{b}\right)^2\right]^2 & d_{ij} < b \\ 0 \ d_{ij} \geqslant b \end{cases}, \tag{6}$$

where d_{ij} is the distance from the data point to the regression point and b is the bandwidth, which is a nonnegative attenuation of the functional relationship between the distance and the weight.

The bandwidth determines the rate at which the weight

Table 1 Definitions of independent variables for irrigation water use efficiency

Independent variables	Definition
Actual irrigation area/(10^4 km^2)	The area of farmland area that was irrigated at least once during the year
Percentage of farmland using water-saving irrigation	The ratio of the irrigation area covered by the irrigation engineering facilities to the total area
Irrigation type	The intake type of the irrigation sources in the irrigation area
Net irrigation water consumption per mu	The ratio of the net irrigation water consumption to the actual irrigation area
Terrain slope/(°)	The steepness of the ground unit
Soil field capacity/mm	The maximum soil water holding capacity that can be steadily maintained in the soil

decays as the distance increases. The larger the bandwidth, the slower the decay; conversely, the smaller the bandwidth, the faster the decay. Bandwidth is an extremely sensitive parameter in the GWR model. An excessively large bandwidth will result in a significant deviation in the regression parameters. There are two methods to determine the bandwidth: *CV* (cross-validation) and AIC (Akaike information criterion). The core of *CV* method is to select the optimal bandwidth by minimizing *CV* (Farber and Páez, 2007). AIC is a model selection criterion based on the maximum likelihood principle (Akaike, 1973). AIC performs well in overcoming the phenomenon of overfitting. However, it is more suitable for models with fewer parameters and its calculations are more complicated. Herein, the bandwidth was determined based on the cross-validation (*CV*) criterion. The specific process of *CV* is described as follows:

For a given bandwidth, the *i*th set of observations (Y_i, X_i) is excluded, and the GWR is applied to estimate parameters under a given b using the remaining $(n-1)$ sets of data. The fitting value, $\widehat{Y_{\neq i}}(b)$, is acquired at X_i. If

$$CV(b) = \frac{1}{n}\sum\nolimits_{i=1}^{n} [Y_i - \widehat{Y_{\neq i}}(b)]^2, \tag{7}$$

then the optimal bandwidth is as follows:

$$b_0 = \mathrm{argmin} CV(b). \tag{8}$$

4 Results and discussions

4.1 Spatial autocorrelation analysis

4.1.1 Global spatial autocorrelation analysis

The ArcGIS global spatial autocorrelation analysis tool was used to analyze the global spatial autocorrelation of the effective utilization coefficient of irrigation water at 165 sample sites in Zhejiang. The results are presented in Fig. 2. The *P* value and *Z* score are critically important when analyzing the results of a global spatial autocorrelation. Generally, when *Z* is above the critical value of 1.65 and *P* is lower than 0.1, the confidence level exceeds 90%; i.e., the probability for significant clustering and positive spatial correlation is over 90%. The results of the global spatial autocorrelation analysis for the effective utilization coefficient of irrigation water produced a *Z* score of 18.008694, which was far above the critical value of 1.65. The *P* value was 0 indicated that less than a 1% chance that a variable was randomly created. Consequently, the abovementioned analysis theoretically confirmed the significant clustering characteristics and positive spatial correlation of the effective utilization coefficient of irrigation water. Herein, the GWR model was the most appropriate model for measuring this variable in the irrigated area of Zhejiang.

4.1.2 Local spatial autocorrelation

Local spatial autocorrelation identifies the level of similarity between the feature attribute values of a spatial unit and those of its neighboring units. Furthermore, it reveals the agreement of each local unit with the general global trend and elucidates how spatial dependence changes according to geographic location. The local Moran index is widely used in research. A high value indicates the spatial clustering of units with similar variables, whereas a low value indicates the spatial clustering of units with dissimilar variables. Therefore, the local Moran index can be subdivided according to its spatial association: high–high (HH), low–low (LL), high–low (HL), and low–high (LH). HH and LL are positive spatial associations, whereas HL and LH are negative. An HH association indicates that spatial units with attribute values above the mean are surrounded by units whose values are also above average; conversely, an LL association signifies that spatial units with attribute values lower than the mean are surrounded by units with below average values. HL and LH associations represent abnormal situations in which high values encompass low values and vice versa.

Based on the ArcGIS platform, this study analyzed the local spatial autocorrelation of the effective utilization coefficient of irrigation water at 165 sample sites in the irrigated region of Zhejiang. The results are displayed in Fig. 3. Red indicates a high terrain slope, whereas green represents the opposite. Table 2 shows the number of various units and reveals that the positive correlation type includes 41 sites with HH associations and 21 with LL associations, which account for 24.8% and 12.7% of the total sample sites, respectively; there are 6 HL-type negative correlations and 3 LH types, which comprise 3.6% and 1.8% of the total sample sites, respectively. The HH type is concentrated mainly in the north and northeast region of the province (Huzhou, Jiaxing, Ningbo, and Zhoushan). This terrain in this region is flat, and plains are the major landform. Irrigation water is abundant, and the farming conditions are good. The LL type is concentrated in the southwestern and southeast areas of the province (Quzhou, Jinhua, Wenzhou, and Lishui). This region is dominated by high slopes, large undulations, and shallow plough layers. In contrast to the north, the distribution of farmland in this region is more sporadic, and the irrigation water use efficiency is lower, which is consistent with an LL-type aggregation.

4.2 Development of the multivariate linear regression model

This study used the six variables defined in Table 1 as influential factors of the effective utilization coefficient of

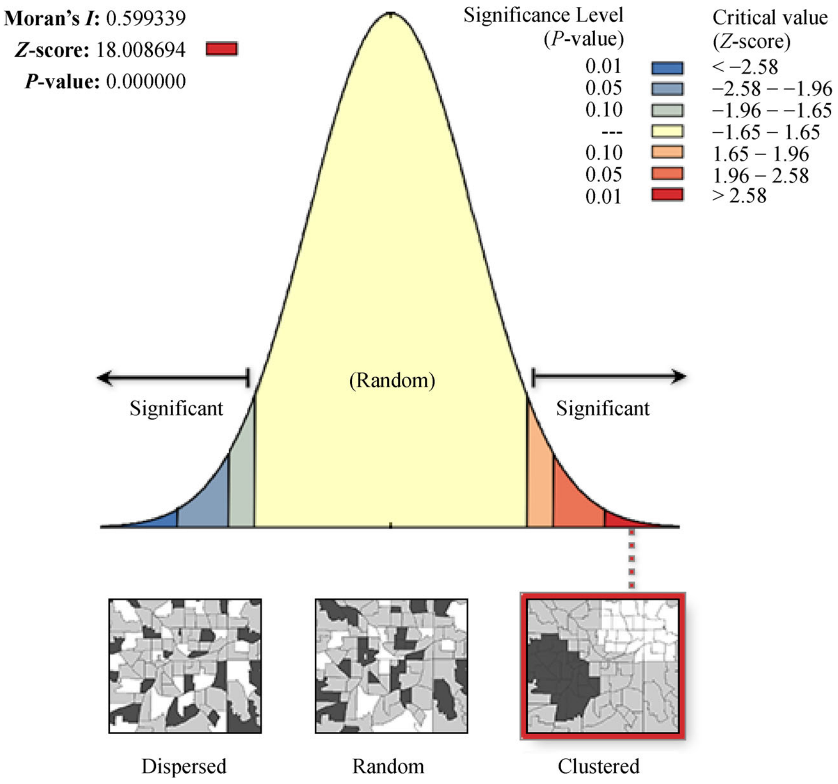

Fig. 2 Analysis results of the Moran index.

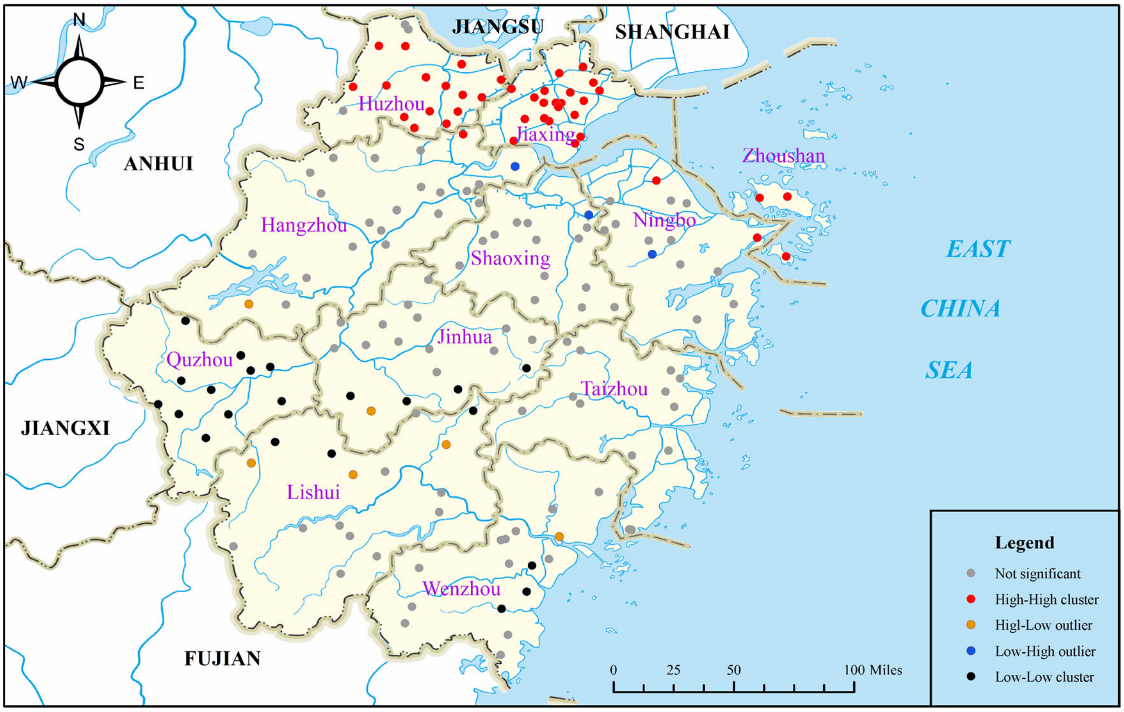

Fig. 3 Local spatial autocorrelation analysis results.

Table 2 Number of different types of units in the local spatial autocorrelation analysis

Type	Number of units	Percentage/%
Insignificant	94	57.0
HH type	41	24.8
LL type	21	12.7
HL type	6	3.6
LH type	3	1.8

irrigation water in Zhejiang. Then, the above coefficient was established as the dependent variable for the construction of a regression model. The results for the standardized regression coefficients for each independent variable are presented in Table 3. All the regression coefficients of the established models passed significance tests. The R^2 of the model was 0.2007, indicating a deficiency in the ability of the multivariate linear regression model to explain the relationship between the influential factors and the effective utilization coefficient of irrigation water.

Table 3 Multivariate linear regression analysis results

Influence factor	Standardized coefficient
Actual irrigation area/10^4 km^2	−0.350
Percentage of farmland using water-saving irrigation	0.176
Type of irrigation area	0.352
Net irrigation water consumption per mu	0.103
Terrain slope/(°)	0.103
Soil field capacity/mm	0.094

Table 3 reveals that the effective utilization coefficient of irrigation water is impacted mainly by the actual irrigation area and the type of irrigation area. The actual irrigation area variable produced a negative effect, suggesting that a larger irrigated area leads to a lower efficiency of use. There was a positive correlation in the type of irrigation area. This indicated that lift irrigation was more efficient than gravity irrigation, which is likely to be attributed to the poor leak-proof features of low-cost gravity irrigation systems. The percentage of farmland using water-saving irrigation, the terrain slope, and the soil field capacity had relatively small impacts, and all showed a positive correlation.

4.3 Development of the geographically weighted regression model

This study investigated the spatial heterogeneity and non-stationarity of the effective utilization coefficient of irrigation water in Zhejiang using a GWR model based on 165 sample sites. In our study of the sample points in the irrigation area, the focus is on the spatial attributes of the sample points rather than the temporal characteristics, and considering the data basis, a GWR model rather than a GTWR model was used here. The GWR model used several independent variables for its functions (please see Table 1 for details) and the effective utilization coefficient of irrigation water as the dependent variable.

Figure 4 provides the frequency distribution of the relative errors between the values predicted by the GWR model and the measured values of the effective utilization coefficient of irrigation water. The relative error is in the range of –11% to 10%, mainly between –5% and 5%. This indicates that the prediction of the GWR model is accurate, and it can be used to calculate the effective utilization coefficient of irrigation water in Zhejiang.

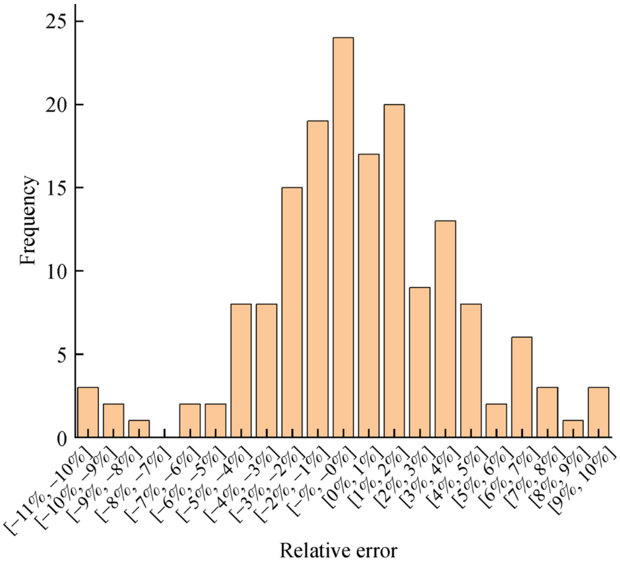

Fig. 4 Frequency distribution of relative errors.

4.4 Spatial distribution result analysis of the effective utilization coefficient of irrigation water

To compare the simulation results of the multivariate linear regression model and the GWR model, the 165 sample sites in Zhejiang Province (in 2017) were plotted on a graph using the predictions of the models as the y coordinates and the measured values as the x coordinates. Linear fitting was performed on the two sets of predictions, and the results are shown in Fig. 5. The GWR model's R^2 value of 0.6768 was significantly higher than that of the multivariate linear regression model, which confirms a better fit for the GWR model. Table 4 compares both sets of predictions. The accuracy of the models was evaluated using the R^2 value, the residual sum of squares, the mean absolute error, and the mean relative error. The table shows that in comparison with the multivariate linear regression approach, the GWR model achieved a

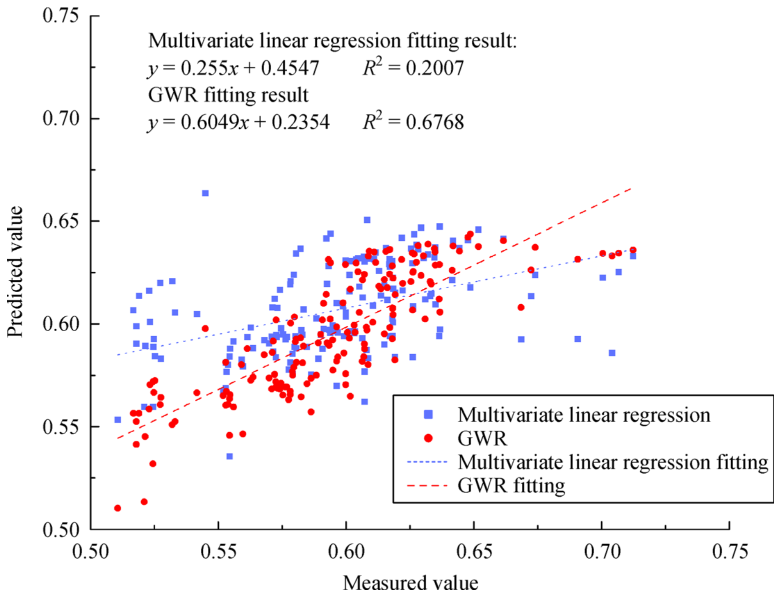

Fig. 5 Comparison of the predicted results between the two models in 2017.

Table 4 Comparison between the prediction accuracies of the two models in 2017

Model	R^2	Residual sum of squares	Mean absolute error	Mean relative error
Multivariate linear Regression model	0.2007	0.159	0.0276	4.73
GWR	0. 6768	0.086616	0.0170	2.86

significant improvement in R^2, which indicates a better fit. Meanwhile, decreases in the residual sum of squares, the mean absolute error, and the mean relative error indicate that the model effectively decreased errors. The above-mentioned data reveal that the GWR model performed better than the multivariate linear regression model in predicting the effective utilization coefficient of irrigation water in Zhejiang Province, and it provided a more adequate explanation for the impact of spatial attributes on the efficiency of irrigation water use.

This study confirmed the applicability of the GWR model in predicting the effective utilization coefficient of irrigation water in Zhejiang Province. For further validation, the two aforementioned regression models were used to predict the above coefficient in 165 irrigation districts in 2018. The inputs and outputs of the models remained the same. The results presented in Fig. 6 show the calculation of the variable in 2018. The GWR model had an R^2 value of 0.6294, which was slightly lower than that of 2017, whereas the R^2 of the multiple linear regression remained at a low level of 0.2061. Therefore, it is concluded that the GWR satisfactorily predicted the effective utilization coefficient of irrigation water in Zhejiang Province and outperformed the multiple linear regression model in terms of accuracy.

5 Conclusions

This study investigated the spatial distribution of the effective utilization coefficient of irrigation water in 165 irrigation districts in Zhejiang Province (in 2017 and 2018) with consideration of the impacts of spatial attributes. The conclusions are as follows.

1) Overall, the aforementioned coefficient exhibits a positive spatial correlation and noticeable agglomeration characteristics. Based on local spatial analyses, the HH type is concentrated mainly in the north and northeast of the province (Huzhou, Jiaxing, Ningbo, and Zhoushan), whereas the LL type is aggregated in the southwest and southeast of the region (Quzhou, Jinhua, Wenzhou, and Lishui).

2) A multivariate linear regression model was used to analyze the relationship between irrigation water use efficiency and different independent variables. The coefficient of irrigation water use is impacted mainly by the actual irrigation area and the type of irrigation area, whereas the influences of the percentage of farmland using water-saving irrigation, the terrain slope, and the soil field capacity are relatively low.

3) The GWR model, which considers the influence of spatial attributes, consistently outperformed the multi-

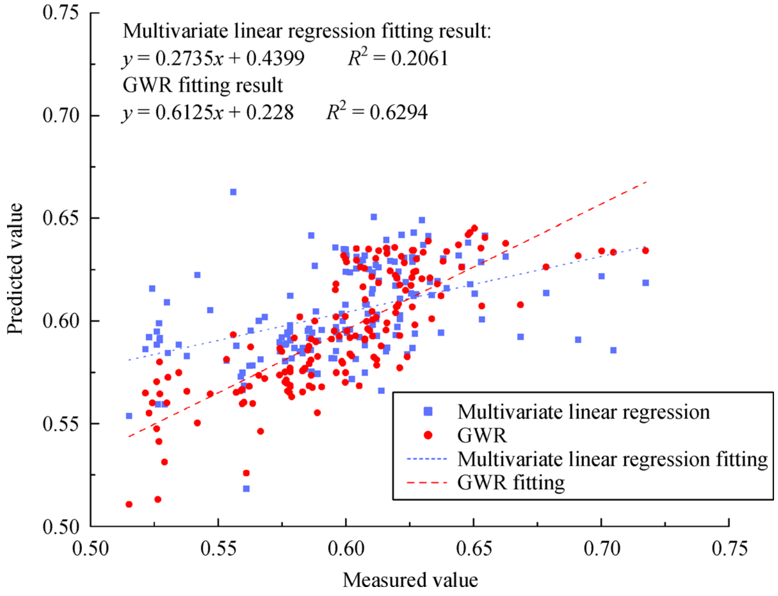

Fig. 6 Comparison of the predicted results between the two models in 2018.

variate linear regression model in terms of fitting performance and prediction accuracy. These results confirm that the GWR model is more suitable for calculating the spatial distribution of the effective utilization coefficient of irrigation water.

Owing to data limitations, the independent variables selected herein are insufficient by themselves to provide a comprehensive overview. However, additional variables can be selected in follow-up research to further improve the model's fitting performance and prediction accuracy.

Acknowledgements This study was supported by the National Key R&D Program of China (Nos. 2016YFC0401005 and 2016YFA0601703), the National Natural Science Foundation of China (Grant Nos. 42075191, 92047203 and 91847301), and Nanjing Hydraulic Research Institute Fund (No. Y520009). We thank Chinese Academy of Meteorological Sciences for providing monitoring data of the study area.

References

Ahadi R, Samani Z, Skaggs R (2013). Evaluating on-farm irrigation efficiency across the watershed: a case study of New Mexico's Lower Rio Grande Basin. Agric Water Manage, 124: 52–57

Akaike H (1973). Information theory and an extension of the maximum likelihood principle. In: Petrov B N, Csaki F, eds. International Symposium on Information Theory. Budapest: Akademiai Kiado: 267–281

Cao H (2010). Test-analysis and application of irrigation water use efficiency of Zhejiang Province. Dissertation for the Master's Degree. Hangzhou: Zhejiang University (in Chinese)

Chen D, Luo Z, Webber M, Chen J, Wang W (2014). Emergy evaluation of the contribution of irrigation water, and its utilization, in three agricultural systems in China. Front Earth Sci, 8(3): 325–337

Das Bhowmik R, Seo S B, Das P, Sankarasubramanian A (2020). Synthesis of irrigation water use in the united states spatiotemporal patterns. J Water Resour Plan Manage, 146(7): 04020050

Farber S, Páez A (2007). A systematic investigation of cross-validation in GWR model estimation: empirical analysis and Monte Carlo simulations. J Geogr Syst, 9(4): 371–396

Feng B (2013). Study on the evaluation and management of irrigation water use efficiency for different scales in countrywide. Dissertation for the Doctoral Degree. Beijing: China Institute of Water Resources & Hydropower Research (in Chinese)

Feng B, Cui J, Wu Q, Han Z (2020). Application of a monitoring and evaluation method in irrigation water efficiency. Irrig Drain, 69(S2): 161–170

Harris P, Fotheringham A S, Crespo R, Charlton M (2010). The use of geographically weighted regression for partial prediction: an evaluation of models using simulated data sets. Math Geosci, 42(6): 657–680

Huang B, Wu B, Barry M (2010). Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. Int J Geogr Inf Sci, 24(3): 383–401

Huang S, Li J, He X (2018). Analysis on influencing factors of effective utilization coefficient of irrigation water in typical southern irrigation district. Pearl River, 39(5): 41–44 (in Chinese)

Huang Y, Leung Y(2002). Analysing regional industrialisation in Jiangsu Province using geographically weighted regression. J Geograph Systems 4: 233–249

Jarvis A, Reuter H I, Nelson A, Guevara E (2008). Hole-filled seamless SRTM data V4. International Centre for Tropical Agriculture (CIAT). Available at CGIAR-CSI SRTM website

Kaneko S, Tanaka K, Toyota T, Managi S (2004). Water efficiency of agricultural production in China: regional comparison from 1999 to 2002. Int J Agric Resour Gov Ecol, 3(3/4): 231–251

Koç C (2013). A study on some parameters which can affect project irrigation efficiency in irrigation networks. Irrig Drain, 65(5): 586–

591

Li Y, Dong S, Wang C, Deng X, Gao Q, Li X, Luo Y, Yang X (2010). Impact of irrigation and other inputs on farm productivity in the upper and middle parts of the Yellow River basin, China. Front Earth Sci China, 4(3): 345–356

Li Y, Zhang Y, He D, Luo X, Ji X (2019). Spatial downscaling of the tropical rainfall measuring mission precipitation using geographically weighted regression Kriging over the Lancang River Basin, China. Chin Geogr Sci, 29(3): 446–462

Ministry of Water Resources of the People's Republic of China (2020). China Water Resources Bulletin, 2019. Beijing: China Water & Power Press (in Chinese)

Omezzine A, Zaibet L (1998). Management of modern irrigation systems in Oman: allocative vs. irrigation efficiency. Agric Water Manage, 37(2): 99–107

Pereira H, Marques R C (2017). An analytical review of irrigation efficiency measured using deterministic and stochastic models. Agric Water Manage, 184: 28–35

Song C, Oxley L, Ma H (2018). What determines irrigation efficiency when farmers face extreme weather events? A field survey of the major wheat producing regions in China. J Integr Agric, 17(8): 1888–1899

Szymanowski M, Kryza M (2011). Application of geographically weighted regression for modelling the spatial structure of urban heat island in the city of Wroclaw (SW Poland). Procedia Environ Sci, 3: 87–92

Tang J, Folmer H, Xue J (2015). Technical and allocative efficiency of irrigation water use in the Guanzhong Plain, China. Food Policy, 50: 43–52

Tromboni F, Bortolini L, Martello M (2014). The use of water in the agricultural sector: a procedure for the assessment of large-scale irrigation efficiency with GIS. Irrig Drain, 63(4): 440–450

Wang J, Kang S, Sun J, Chen Z F (2013). Estimation of crop water requirement based on principal component analysis and geographically weighted regression. Chin Sci Bull, 58(27): 3371–3379

Yilmaz B, Yurdusev M A, Harmancioglu N B (2009). The assessment of irrigation efficiency in Buyuk Menderes Basin. Water Resour Manage, 23(6): 1081–1095