Science of the Total Environment 572 (2016) 450–466

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Impacts of land use and population density on seasonal surface water quality using a modified geographically weighted regression

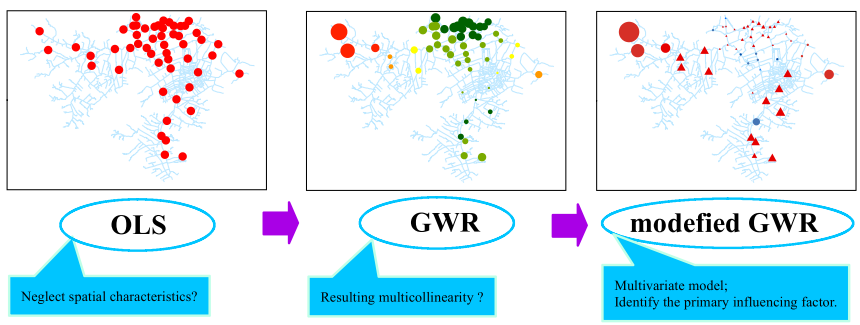
Qiang Chen [a], Kun Mei [a], Randy A. Dahlgren [a,b], Ting Wang [a], Jian Gong [a], Minghua Zhang [a,b,*]

[a] Zhejiang Provincial Key Laboratory of Watershed Sciences and Health, Southern Zhejiang Water Research Institute (iWATER), Wenzhou Medical University, China
[b] Department of Land, Air and Water Resources, University of California, Davis, USA

HIGHLIGHTS

- The modified GWR models had higher R^2 and reflected the actual spatial features.
- A manual variable excluding-selecting method is explored to avoid multicollinearity.
- Influences of the dominant indicator on water quality varied with space and seasons.
- Protection policies need consider site-specific conditions and seasonal variations.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 12 April 2016
Received in revised form 5 August 2016
Accepted 7 August 2016
Available online 18 August 2016

Editor: D. Barcelo

Keywords:
Water pollution
GWR
OLS
Multicollinearity
Wen-Rui Tang River

ABSTRACT

As an important regulator of pollutants in overland flow and interflow, land use has become an essential research component for determining the relationships between surface water quality and pollution sources. This study investigated the use of ordinary least squares (OLS) and geographically weighted regression (GWR) models to identify the impact of land use and population density on surface water quality in the Wen-Rui Tang River watershed of eastern China. A manual variable excluding-selecting method was explored to resolve multicollinearity issues. Standard regression coefficient analysis coupled with cluster analysis was introduced to determine which variable had the greatest influence on water quality. Results showed that: (1) Impact of land use on water quality varied with spatial and seasonal scales. Both positive and negative effects for certain land-use indicators were found in different subcatchments. (2) Urban land was the dominant factor influencing N, P and chemical oxygen demand (COD) in highly urbanized regions, but the relationship was weak as the pollutants were mainly from point sources. Agricultural land was the primary factor influencing N and P in suburban and rural areas; the relationship was strong as the pollutants were mainly from agricultural surface runoff. Subcatchments located in suburban areas were identified with urban land as the primary influencing factor during the wet season while agricultural land was identified as a more prevalent influencing factor during the dry season. (3) Adjusted R^2 values in OLS models using the manual variable excluding-selecting method averaged 14.3% higher than using stepwise multiple linear regressions. However, the corresponding GWR models had adjusted R^2 ~59.2% higher than the optimal OLS models, confirming that GWR models demonstrated better prediction accuracy. Based on our findings, water resource protection policies should consider site-specific land-use

* Corresponding author at: Zhejiang Provincial Key Laboratory of Watershed Sciences and Health, Southern Zhejiang Water Research Institute (iWATER), Wenzhou Medical University, China.
E-mail address: mhzhang@ucdavis.edu (M. Zhang).

http://dx.doi.org/10.1016/j.scitotenv.2016.08.052

conditions within each watershed to optimize mitigation strategies for contrasting land-use characteristics and seasonal variations.

1. Introduction

Degradation of surface water quality is a crucial global environmental issue (Zielinski et al., 2016; Roebeling et al., 2015). It's particularly apparent in China due to the rapid processes of urbanization and economic development (Chen et al., 2016a, 2016b; Mei et al., 2014). Non-point source pollution is of great importance due to its prevalence on surface water impairment and the difficulty of identifying specific sources for remediation (Dowd et al., 2008; Zhang et al., 2011; Sun et al., 2013; Huisman et al., 2013). As an important regulator of pollutants in overland flow and interflow, land use/land cover has become a critical research topic for elucidating the relationship between surface water quality and non-point source pollutants. Several statistical methods are widely used in these studies, such as correlation analysis, cluster analysis, principal component analysis, linear regression models, linear mixed effects models and exponential models (Wang et al., 2014; Wan et al., 2014; Chen and Lu, 2014; Seeboonruang, 2012; Ahearn et al., 2005; Madrinan et al., 2012). Previous results often show that land-use types closely related to human activities, such as agriculture and urban, were positively correlated with river pollution indicators (e.g., nitrogen, phosphorus, ammonia), while woodlands and grasslands that were less affected by human activities had negative correlations. These previous analyses and approaches were all based on the assumption that relationships between water quality indicators and land-use patterns were constant over the entire study area.

These global statistical methods express the average of existing relationships, which may neglect some significant spatial characteristics and hide local variations (Tu and Xia, 2008; Tu, 2013). Geographically weighted regression (GWR) models (Fotheringham et al., 1996; Boots, 2003) were first applied to assess the relationship between land use and water quality by Tu and Xia (2008). They demonstrated better results for GWR models than from traditional linear regression models, such as ordinary least squares (OLS) regression. Through embedding location data into regression parameters to explore local variations between independent and dependent variables, GWR effectively addresses the spatial non-stationarity issue (Boots, 2003). In addition, GWR considers spatial autocorrelation, which is difficult to deal with in traditional statistical models (Brown et al., 2012). As an emerging technique, GWR has recently been applied in several disciplines, such as identification of high crime areas (Cahill and Mulligan, 2007), forest damage evaluations (Pineda et al., 2010), human health and disease analysis (Carrel et al., 2011), and atmospheric pollutant assessment (Song et al., 2014).

In studies applying the GWR technique to evaluation of land use on surface water quality (e.g., Tu and Xia, 2008; Tu, 2011; Brown et al., 2012; Tu, 2013; Sun et al., 2014), it is common that only a single land-use indicator was selected as the independent variable because of the high potential for multicollinearity among different land-use variables (Griffith, 2008). This would result in an invalid GWR model when variables experiencing collinearity are selected. It is common for several land-use variables to be correlated since different land-use indicators will be interrelated when calculated as a percentage of total land area. Using this modeling approach with a single explanatory variable avoids multicollinearity and logically reflects the simple correlation. However, the univariate models may miss one or more important explanatory variables. For example, it would be misleading to explain variations in riverine nitrogen levels based only on agricultural land percentage as an explanatory variable when water quality impairments resulted primarily from runoff associated with impervious surfaces in subcatchments dominated by urban area with low agricultural land use.

As early as 2005, Wheeler investigated multicollinearity problems in GWR models and pioneered the ridge regression method to introduce multiple variables and eliminate the collinearity problem in GWR models (Wheeler and Tiefelsdorf, 2005; Wheeler, 2007). Additionally, Wheeler (2009) developed the use of LASSO (least absolute shrinkage and selection operator) algorithms in GWR models to limit the effects of explanatory variable correlation. Other multiple variable selection approaches such as Principal Component Analysis (PCA) (Wang et al.,

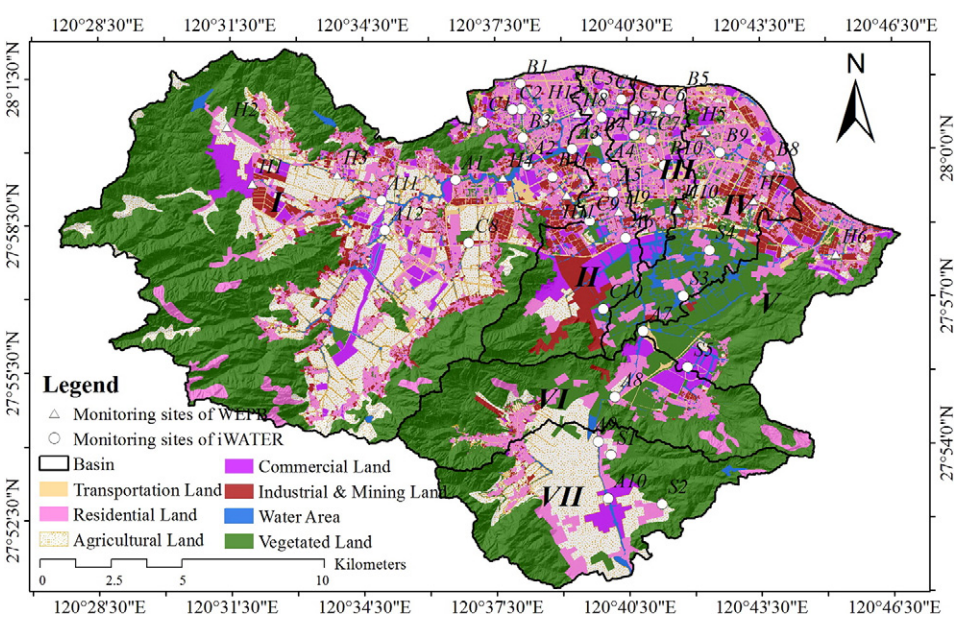


Fig. 1. Location of monitoring sites, drainage subcatchments, and major land-use units in the Wen-Rui Tang River watershed.

2013) and Geographically Weighted Principal Component Analysis (GWPCA) (Kumar et al., 2012) have been explored with GWR models. Specific to land-use impacts on water quality, the application of multivariate GWR models is rarely utilized. Pratt and Chang (2012) used stepwise multiple linear regression (SMLR) to eliminate collinear land-use variables, with the remaining significant independent variables ($p < 0.05$) used to analyze the effects of land cover and topography on water quality. An advantage of this approach is that the regression excludes redundant variables automatically, but sometimes variables eliminated by SMLR are not without statistical significance (Yang et al., 2005).

In practice, model variables should be determined based on site-specific information and professional knowledge of the study site. On this basis, a manual variable excluding-selecting method was explored in this study to modify the multivariable modeling process. The objectives of this paper were to examine the impact of land use/population density (LU/PD) on water quality in both dry and wet seasons in the Wen-Rui Tang River watershed in eastern China. The investigation attempts to answer the following questions: (1) Does the influence of LU/PD on specific water quality parameters change spatially and with seasons? (2) Is there a dominant LU/PD indicator affecting water quality impairment in individual subcatchments? and (3) Does the dominant factor have a spatial distribution within the greater watershed? In order to identify the LU/PD indicator that has the strongest correlation to water quality indicators, we introduced the concept of standard regression coefficients (Mayer and Younger, 1976) into GWR models. To avoid the influence of external factors in selection of independent variables (Yuan and Chan, 2011; Nimon and Oswald, 2013), cluster analysis was combined with the calculation of standard regression coefficients to assure that only those subcatchments with similar land-use structure were compared. The manual variable excluding-selecting method and the re-processing method for GWR output used in this study are not only suitable for the Wen-Rui Tang River watershed, but also can be used in other study areas having watershed-scale, water quality data and a small number of explanatory variables. In addition, the results of this study provide important information to inform water resource management and remediation in watersheds with mixed land use and strong seasonal climate variations.

2. Study area

The Wen-Rui Tang River watershed located in Wenzhou, Zhejiang Province on the east coast of China has a drainage area of 740 km^2 with a population of ~9.2 million (Fig. 1). The watershed originates from Lishui mountain streams and flows eastward through an urban district. From the urban area, ~70% of the water flows south to join the Fei-Yun River which flows into the East China Sea and the remaining ~30% of the water flows eastward to the Oujiang estuary. The region has a subtropical monsoon climate with an average annual rainfall of 1818 mm, approximately 70% falls in April to September. The Wen-Rui Tang River played an important role in irrigation, aquaculture, industrial water and transportation in the past.

Based on the distribution of the river network and topographic features, the study area was divided into 7 drainage basins (Fig. 1). Basin I is a low population area surrounded by mountains and has a primary land cover of trees/shrubs, fruitwood and agriculture, which is considered as suburban and rural areas (Mei et al., 2014). Basin II, III, IV, and V are more densely populated and urbanized areas with the northern area containing Wenzhou city center and the southern area comprising an important aquatic habitat called Sanyang wetland. Sanyang wetland has been used for agriculture for centuries with 47% of its land area currently used for citrus groves. Land use in the lower reaches of basin VI and VII are dominated by agriculture, including rice, soybean, and waxy corn. Basin VII is considered a rural area.

The Wen-Rui Tang River is severely polluted and has experienced reoccurring hypoxia since rapid urbanization and economic development began in the 1980s (Li et al., 2013). In 2010, 92% of the river segments were classified in the inferior Type V national water quality category, the lowest water quality classification that supports aquatic ecosystem health. Only ~60% of the sewage is collected for centralized processing at wastewater treatment facilities. Nitrogen pollution, especially from ammonium, contributing to low dissolved oxygen is considered the most serious pollution problem in the watershed (SEPBC, 2002a). Water quality has been significantly improved in the past few years in some regions due to improved sewage collection, removal of river sediments, ecological water diversion, relocation of animal husbandry, and riparian green belt construction (providing buffer strips)

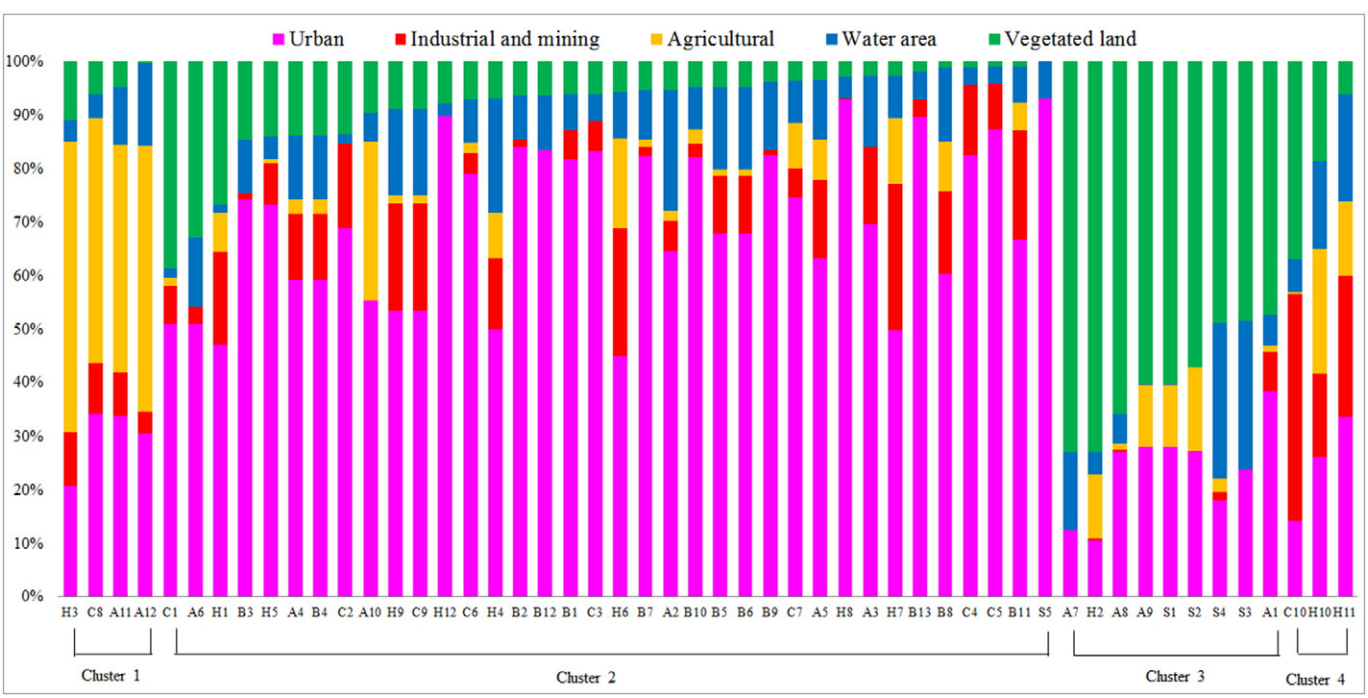


Fig. 2. Comparison of land use structure in the 52 sub-catchments where monitoring sites are located.

(Mei et al., 2014). However, additional water quality remediation is required and studies providing insights to the spatial and temporal controls on water pollution are critically needed to develop better water pollution management strategies.

3. Data collection and analysis methods

3.1. Water quality data

Monthly water samples were collected from 40 sites established by iWATER (Institute of Wenzhou Applied Technology for Environmental Research) from May 2008 to December 2010. In this study, five water quality indicators were selected: total nitrogen (TN), ammonia nitrogen (NH_4^+-N), total phosphorus (TP), dissolved oxygen (DO) and chemical oxygen demand (COD). Additionally, water quality data for 12 monitoring sites from the Wenzhou Environmental Protection Bureau (WEPB) were included for the same water quality parameters and sampling dates. All water quality samples were measured using standard analytical methods (SEPBC, 2002b).

To consider the influence of seasonal factors within each land-use category, mean values for each water quality indicator in the wet (April–September) and dry (October–next March) seasons were calculated. Data for the five water quality variables were not normally distributed (as determined by the Kolmogorov-Smirnov test) and therefore we used a natural logarithmic transformation to provide a normal distribution for subsequent statistical analyses.

3.2. Land-use data and GIS analysis

Land-use data were interpreted from aerial imagery by the Land Use Survey Project of Wenzhou (2005) having a resolution of 0.5 m. The original 96 land-use categories were aggregated by merging similar

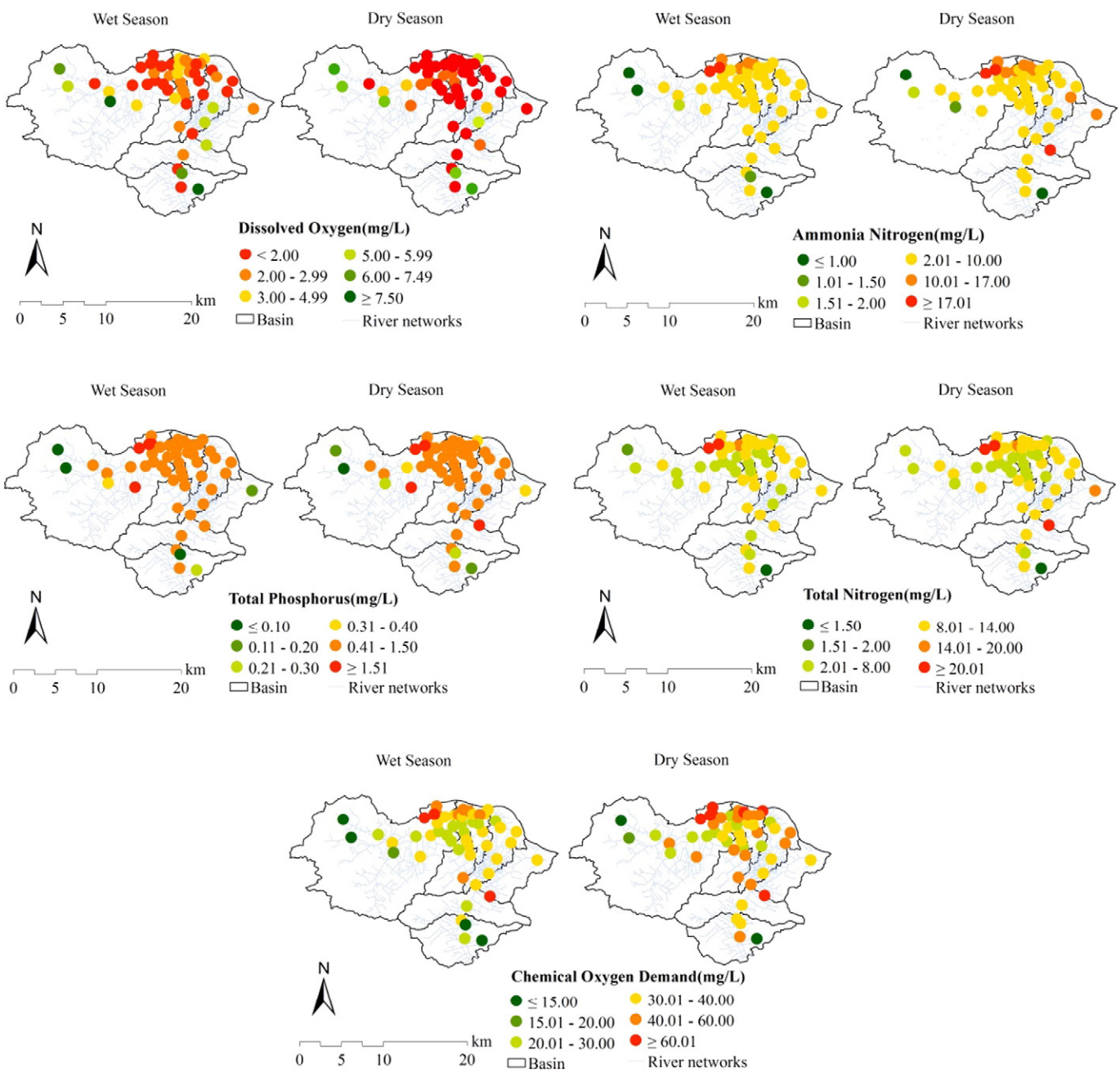


Fig. 3. Spatial-temporal distribution of water quality in the Wen-Rui Tang River watershed.

Table 1
Mean concentration of water quality indicators in different clusters during the wet and dry season (mg/L).

	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
	Wet	Dry	Wet	Dry	Wet	Dry	Wet	Dry
DO	4.21[a]	3.45[b]	2.23[a]	1.84[b]	4.47[a]	4.77[a]	1.67[a]	1.11[b]
NH_4^+-N	5.43[a]	6.04[b]	8.57[a]	9.95[b]	4.27[a]	5.89[b]	6.65[a]	8.37[b]
TP	0.84[a]	1.09[b]	0.98[a]	0.97[a]	0.51[a]	0.52[a]	0.82[a]	0.91[a]
TN	7.59[a]	8.34[b]	10.44[a]	11.51[b]	5.86[a]	7.06[b]	9.75[a]	10.54[b]
COD	29.1[a]	34.6[b]	37.7[a]	51.0[b]	25.1[a]	30.4[b]	33.6[a]	37.3[b]

Values with different lower case letters in a certain cluster are statistically different ($p \le 0.05$).

land-use types into 7 broader categories: agriculture, vegetated land (forest, grassland and urban green belts), commercial (commercial, administrative, cultural entertainment, municipal utility lands), transportation, industrial and mining, residential, and water (Fig. 1). The few unused land categories (0.03% of total land area), such as bare rock, gravel and waste lands, were aggregated with the residential land category due to their similar characteristics with respect runoff and erosion.

The water flow direction was derived from DEM data with a spatial resolution of 5 m obtained from the Wenzhou Urban Planning Bureau (WUPB), and 7 drainage basins (Fig. 1) were created by locating the pour points at the edges of the study area using a flow direction raster. We also calculated flow accumulation, a dimensionless number in ArcGIS defined as the pixel quantity that flows from upstream cells. The grids with flow accumulation values >10,000 were subsequently extracted as the river networks and the corresponding subcatchment polygons were generated using Arc Hydro Tools. Then we overlapped the subcatchment land-use polygons to the 7 basins and eliminated the ensuing sliver polygons. Catchments with areas <100,000 m^2 were merged with adjacent catchments according to the benchmark drainage basin boundary. As a result, there were a total of 201 subcatchments in the study area. Land-use percentages based on the 7 land-use categories were calculated for each subcatchment using ArcGIS statistical tools.

Table 2
Optimized multivariate OLS models using the manual variable excluding-selecting method.

	Variable selection method	Regression model	Adjusted R^2
Wet Season			
DO	SMLR	−0.322 [PD] + 2.342	0.192
	Manual	−0.340 [PD] + 0.015 [WA] + 2.288	0.203
NH_4^+-N	SMLR	0.416 [PD] + 0.005 [Urban] − 0.438	0.294
	Manual	0.457 [PD] + 0.140 [Urban] − 0.017 [AG] + 0.016	0.322
TP	SMLR	0.014 [Urban] − 1.122	0.226
	Manual	0.483 [Urban] − 0.017 [AG] −0.102	0.242
TN	SMLR	0.01 [Urban] + 1.569	0.163
	Manual	0.409 [Urban] − 0.031 [AG] + 0.598	0.180
COD	SMLR	0.204 [PD] + 2.504	0.186
	Manual		\
Dry season			
DO	SMLR	−0.364 [PD] + 2.384	0.236
	Manual	−0.530 [PD] + 0.200 [WA] + 2.309	0.290
NH_4^+-N	SMLR	0.012 [Urban] + 1.318	0.127
	Manual	0.329 [Urban] − 0.107 [AG] + 1.451	0.153
TP	SMLR	0.012 [Urban] − 1.014	0.170
	Manual	0.426 [Urban] − 0.013 [AG] −0.999	0.186
TN	SMLR	0.009 [Urban] + 1.656	0.114
	Manual	0.315 [Urban] − 0.105 [AG] + 0.761	0.139
COD	SMLR	0.008 [Urban] + 3.265	0.261
	Manual	\	\

Urban = transportation + residential + commercial lands (%); IN = industrial and mining land (%); AG = agricultural land (%); WA = water area (%); GR = grassland & forest (vegetated land) (%); PD = population density (people·ha^{-1}).

3.3. Population density data

The 2010 population census for administrative subdistricts was obtained from the Wenzhou Statistical Yearbook published by Wenzhou Municipal Bureau of Statistics (WSB). Population data were input to ArcGIS to calculate the population density for each monitoring site: (1) the centroid for each administrative subdistrict was determined and the corresponding population assigned to these centroids; (2) the centroid point was interpolated into a population density raster using the Kernel algorithm (ArcGIS Help Library, 2014); and (3) the cells of the density raster based on the set of coordinate monitoring points were extracted as grid values.

3.4. Modeling methods

Multivariate OLS regression and GWR models were developed using SPSS 21 and ArcGIS 10.2, respectively, to explore the relationship between land use and water quality indicators. To avoid multicollinearity, a manual variable excluding-selecting method was utilized: (1) all of the explanatory variables were entered to develop OLS functions for each water quality indicator, and Spearman coefficients for both dependent and independent variables were calculated; (2) pairwise land-use variables in the OLS model with higher correlations were selected and variables with weaker correlations in the pair were removed; (3) the independent variable with the highest variance inflation factor (VIF) was removed (if the maximum VIF value was <2, then jump to the fourth step); and (4) the OLS model was reformulated with the independent variables having the highest impact on selected water quality indicators according to our previous findings (Lu et al., 2011; Li et al., 2013; Mei et al., 2014). If there was a significant correlation among variables, step 2 was repeated until all argument VIF values were <2. In addition, a stepwise multiple linear regression (SMLR) with the same water quality parameters was run using the independent variables identified as significant ($p \le 0.05$). The results of the two methods were contrasted and the one with higher R^2 was chosen as the optimized multivariate OLS model. These variables were then used to run a corresponding GWR model. With five water quality parameters (TN, NH_4^+-N, TP, DO and COD), two seasons (wet and dry seasons), and two modeling methods (SMLR and the manual variable excluding-selecting method), a total of 20 models were evaluated.

As an extension of global statistical models such as OLS, GWR embeds location data into the regression parameters to assess the local relationships between independent and dependent variables (Boots, 2003). The GWR model can be defined as:

$$y_j = \beta_0(u_j, v_j) + \sum_{i=1}^{p} \beta_i(u_j, v_j)\chi_{ij} + \varepsilon_j \quad (1)$$

where (u_j,v_j) represents the coordinates for location j, $\beta_i(u_j,v_j)$ represents the local regression coefficient for independent variables χ_i at location j, and $\beta_0(u_j,v_j)$ and ε_j represent the intercept and error term, respectively. $\beta_i(u_j,v_j)$ was estimated by Eq. (2) to determine a minimum:

$$\beta_0(u_j, v_j) = \sum_{k=1}^{n} w_{jk}\left(y_k - \beta_0(u_j, v_j) - \sum_{i=1}^{p} \beta_i(u_j, v_j)\chi_{ij}\right)^2 \quad (2)$$

where w_{jk} represents the distance decay function for location j and k, with the basic assumption that observations closer to sample point j have a higher impact on local regression parameters. As the core of the GWR model, w_{jk} can be calculated using the distance threshold method, Gauss function method, bi-square function method, etc.

(Brunsdon et al., 2002). The Gauss function was used in this analysis due to its greater efficiency:

$$w_{jk} = \exp\left(-d_{jk}^2/b^2\right) \quad (3)$$

where $d_{jk}d_{jk}$ represents the distance between location j and k, and b represents the kernel bandwidth. With ArcGIS 10.2, both fixed and adaptive bandwidths were provided. GWR calculates the optimal distance for fixed kernel or optimal number of neighbors for the adaptive kernel. Unlike the literature using adaptive bandwidth (Tu and Xia, 2008; Pratt and Chang, 2012; Tu, 2013), due to the distribution of monitoring sites, we found that the fixed bandwidth method had a significant advantage in developing GWR models for the Wen-Rui Tang River watershed.

3.5. Statistical analysis and model assessment

To elucidate the principal factor(s) affecting water quality in an individual subcatchment, cluster analysis was utilized and standardized regression coefficients (Beta-coefficients) were calculated. We assume that predictors with larger standardized coefficients are more important than predictors with smaller coefficients, particularly when the variables are uncorrelated (Yuan and Chan, 2011). We previously minimized the collinearity among variables using the manual variable

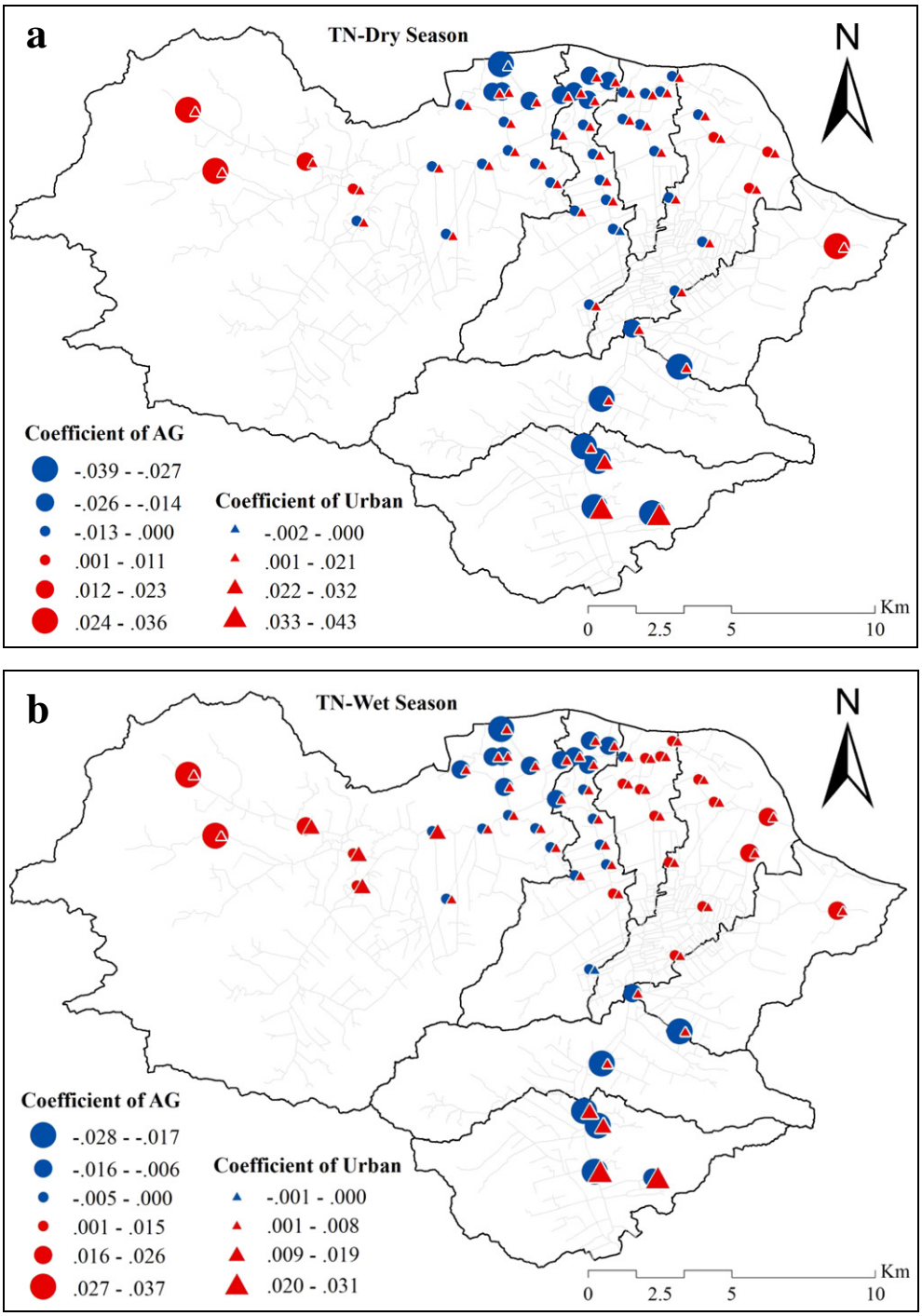


Fig. 4. Local coefficients in a dry season and b wet season in the GWR model for TN.

excluding-selecting method. Moreover, the contrast process demands that all other predictors in the regression of $\boldsymbol{X}$ (i.e., $(X_1, X_2, \ldots X_{j-1}, X_{j+1}, \ldots X_m)^T$) are controlled or have little fluctuation when the contribution of X_j to Y is estimated (Nimon and Oswald, 2013). We applied cluster analysis to assure that only monitoring sites with similar land-use structure were compared, which largely stabilizes the selected variables. The standardized regression coefficients b'_j can be calculated as follows:

$$b'_j = b_j(S_j/S_Y) \tag{4}$$

where b_j are the raw regression coefficients, S_j and S_Y are the standard deviation of independent variable X_j and dependent variable Y in a particular cluster, respectively.

The 52 water quality monitoring sites generated 4 clusters based on land-use characteristics (Fig. 2). Cluster 1 corresponded to the suburban area that was dominated by agriculture (48.1%) but also had mixed residential areas (Mei et al., 2014). Cluster 2 was composed of 36 sites characterized by a relatively developed region with an average proportion of 70.0% urban land. Two special subcatchments that included monitoring sites H1 and H2, were considered rural area (Mei et al., 2014) as their surrounding subcatchments had an average proportion of impervious surface area <13.2%. Cluster 3 included 9 monitoring sites possessing an average of 59.3% non-agricultural, vegetated land. Cluster 4 consisted of three sites (C10, H10 and H11) having an average of 28.1% industry land, 24.7% urban land and 12.4% agricultural land.

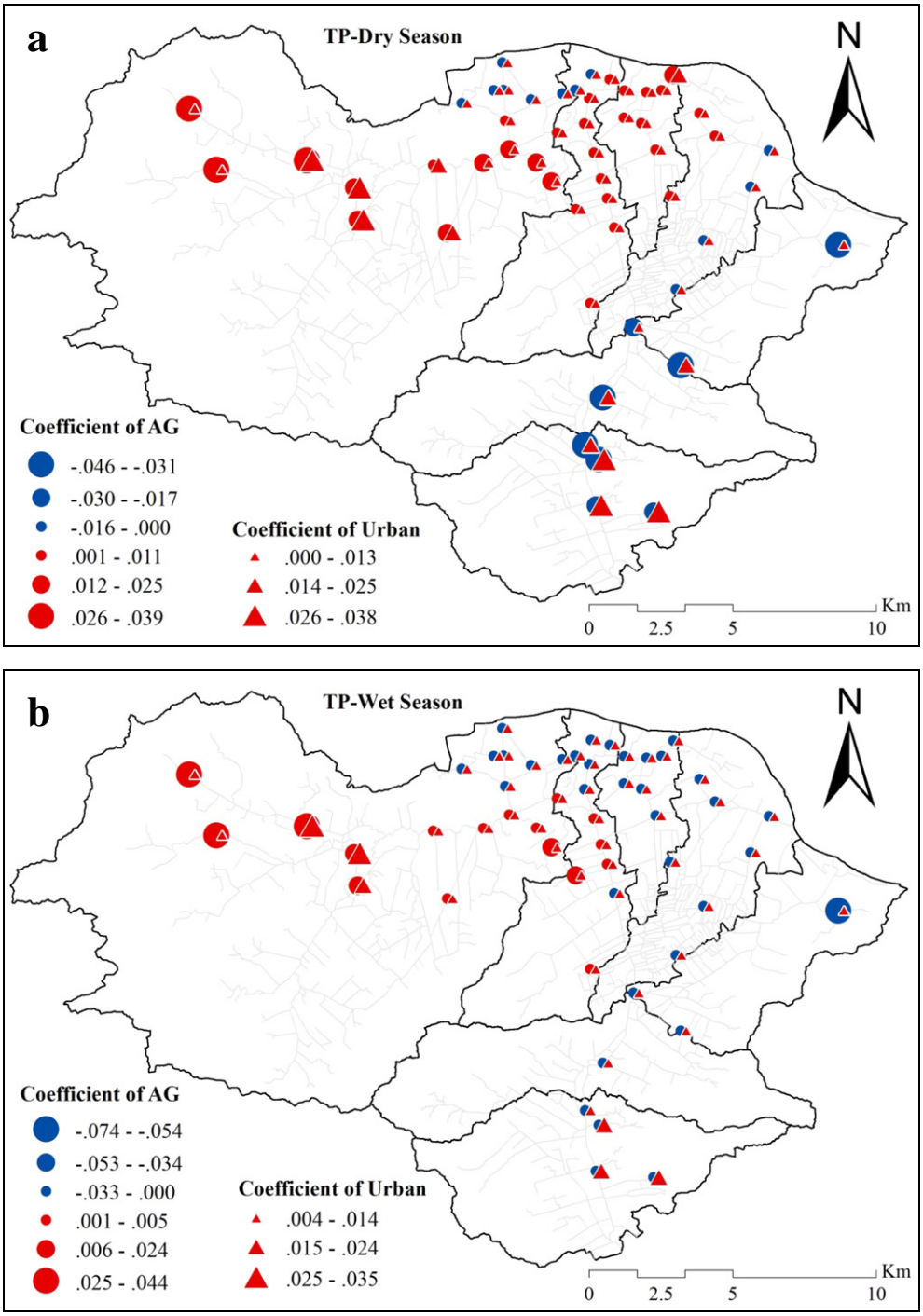


Fig. 5. Local coefficients in a dry season and b wet season in the GWR model for TP.

4. Results

4.1. Spatial-temporal patterns of water quality

Higher concentrations of NH_4^+-N, TP, TN, and COD were found in the north of basins II, III and IV, with lower values in basins VI, VII and the western portion of basin I (Fig. 3). Specific to the different cluster types (Table 1), mean NH_4^+-N, TN and COD concentrations were highest in cluster 2. In contrast, the mean concentrations of NH_4^+-N, TP, TN and COD were lowest in cluster 3. In general, the water quality indicators within the different clusters reflect the pattern of urbanization intensity within the watershed.

The water quality variables showed significant seasonal variation in the Wen-Rui Tang River. For example, mean NH_4^+-N, TN and COD concentrations in clusters 1–4 during the wet season were significantly lower compared with the dry season ($p < 0.05$), while DO concentrations showed an inverse trend with higher DO during the wet season. In contrast, cluster 2 with the highest degree of urbanization did not show a significant difference for TP concentrations between the wet and dry seasons. This same pattern for TP was shown in clusters 3 and 4. These seasonal variations are consistent with precipitation in less urbanized subcatchments causing dilution of pollutants (Cunningham et al., 2010) while in subcatchments with greater impervious area, rainfall-runoff leads to increasing pollutant transport to rivers. Additionally,

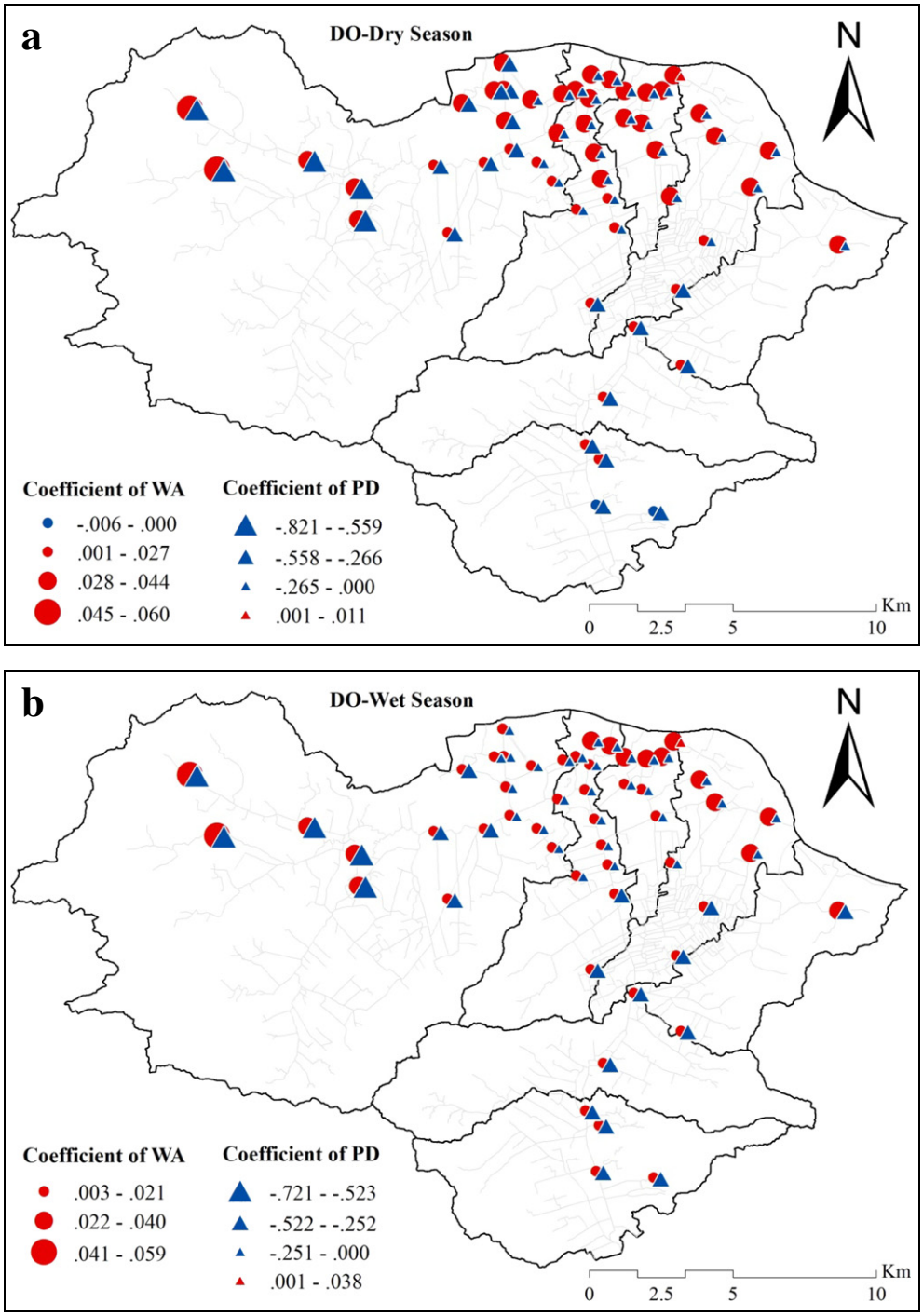


Fig. 6. Local coefficients in a dry season and b wet season in the GWR model for DO.

microorganisms might consume and transform more DO and nutrients during the dry season due to longer water residence times in the river network (Tsegaye et al., 2006).

4.2. Relationships between land use and water quality in wet versus dry season

Significant improvements were found in 8 OLS models using the manual variable excluding-selecting method compared to the SMLR method (Table 2). TN, TP and NH_4^+-N were best modeled by a combination of AG and Urban during the two seasons. The optimized explanatory variables for COD and NH_4^+-N varied with season. For instance, PD was not a significant factor for COD in the dry season, but became a positive explanatory variable in the wet season. Additionally, all optimized multivariate OLS models, with the exception of DO, had higher adjusted R^2 values during the wet season compared to the dry season (Table 2). However, none of the optimized models explained more than 50% of the variance for any of the water quality indicators.

In contrast to the optimized OLS models, relationships among water quality parameters and independent variables in the GWR models had complex local characteristics at both seasonal and spatial scales (Figs. 4–8). In the case of TN (Fig. 4), the sign of the coefficients for AG switched from positive in the western and northeastern subcatchments to negative in the north-central and southern subcatchments. Urban

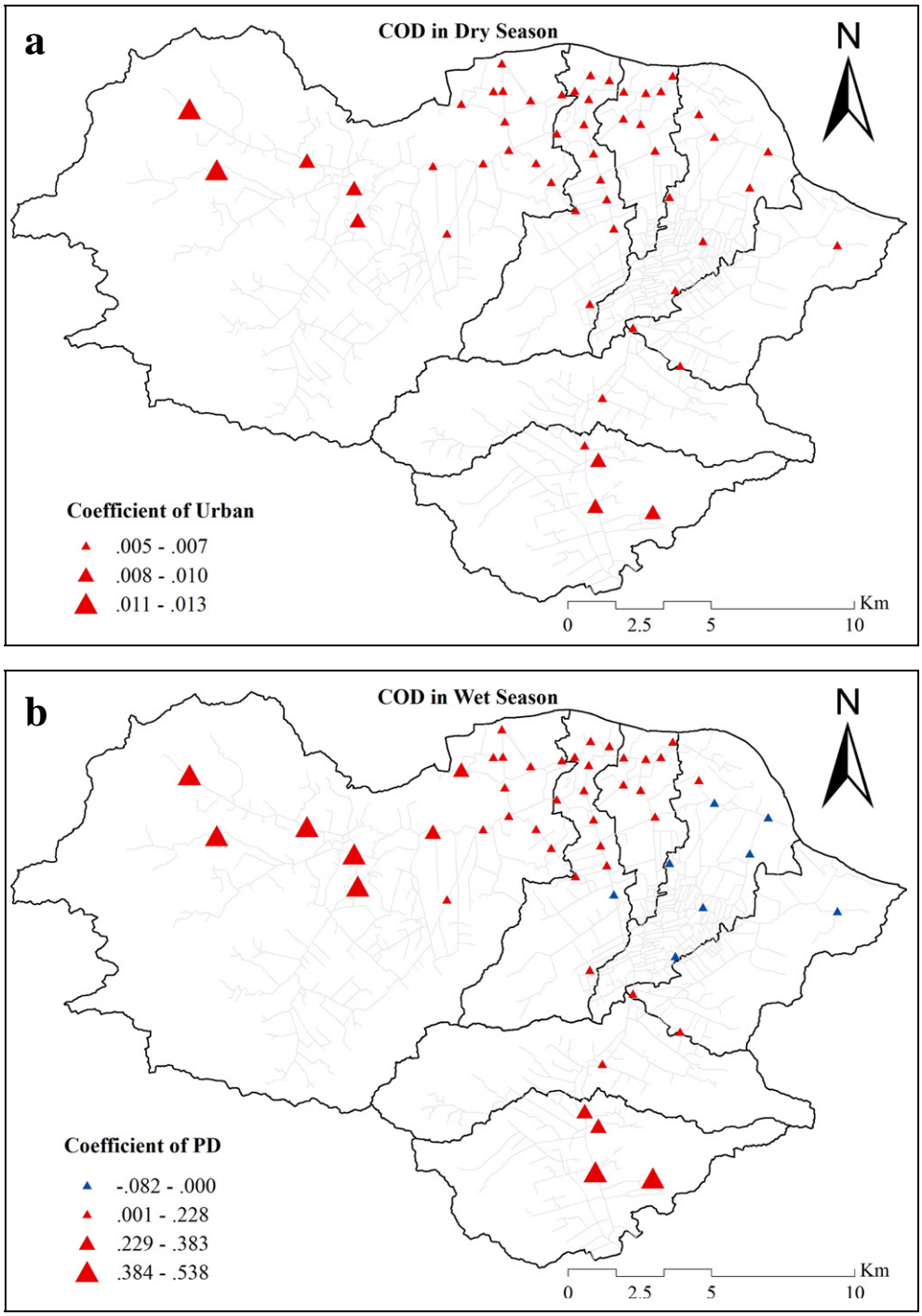


Fig. 7. Local coefficients in a dry season and b wet season in the GWR model for COD.

showed positive coefficients in the majority of regions except for sites A6 and B1 during the dry season and C10 during the wet season, however, all of these negative coefficients were close to zero. In contrast, the local coefficient for AG during the dry season in subcatchments in western basin I and southern basins VI and VII had higher absolute values. Wet season results (Fig. 4b) share similar features with the dry season except that more sites in northern basin III demonstrated a positive correlation between AG and TN.

The GWR model for TP during dry and wet seasons had positive coefficients for Urban throughout the entire study area (Fig. 5), similar to the OLS model (Table 2). Positive AG coefficients were found to be higher during both seasons at sites H1 and H2, which were close to the river source and had GR = 50.0% and AG = 9.6%. During the dry season, a stronger negative relationship appeared at different AG levels from 0 to 29.5% in the southern watershed. More monitoring sites in the urban area showed a stronger positive influence from AG during the dry season than during the wet season.

More than 95% of the local coefficients for both WA and PD in the DO model for the wet versus dry seasons showed positive and negative

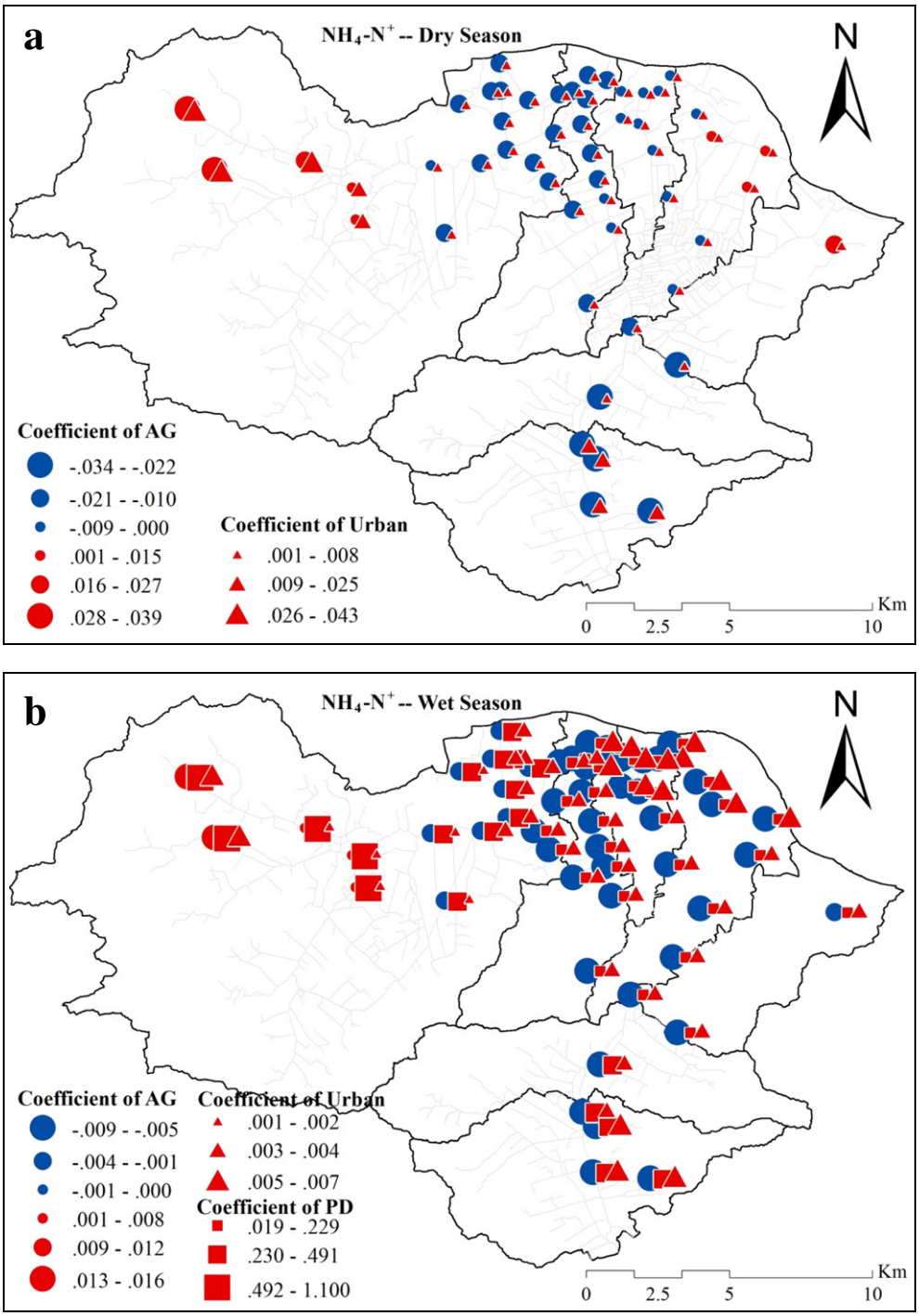


Fig. 8. Local coefficients in a dry season and b wet season in the GWR model for NH_4^+-N.

correlations with DO, respectively (Fig. 6). Higher positive WA coefficients were found in western suburban and rural areas. Specific to the dry season, these higher coefficient values extended to the urban area and maintained a stable level (Fig. 6a).

The dominant factor for COD in the wet season was Urban, while in the dry season COD was best determined by PD (Fig. 7). In contrast, the subcatchments located in western basin I and southern basin VII have higher PD/Urban coefficient values. A consistent positive correlation between Urban and COD with lower valued regression parameters was found in the urban area. During the wet season, PD had a weak negative effect on COD in basins IV and V, which were the watersheds closest to Sanyang wetland.

4.3. Analysis of the primary factors influencing water quality

The primary factor having the largest contribution to water quality parameters varied with the spatial and seasonal analysis scale. PD played a dominant role in DO variations with higher local R^2 values in western basin I, which was identified as rural area with a lower population density, while in the wetland area DO was best explained by water area during both seasons (Fig. 9). In contrast, during the dry season both WA and PD had a better predictability in the urban area with higher local R^2 values. Specific to the suburban area in cluster 1, WA played a more important role compared to PD in the dry season, while in the wet season the negative effects from PD on DO were more obvious.

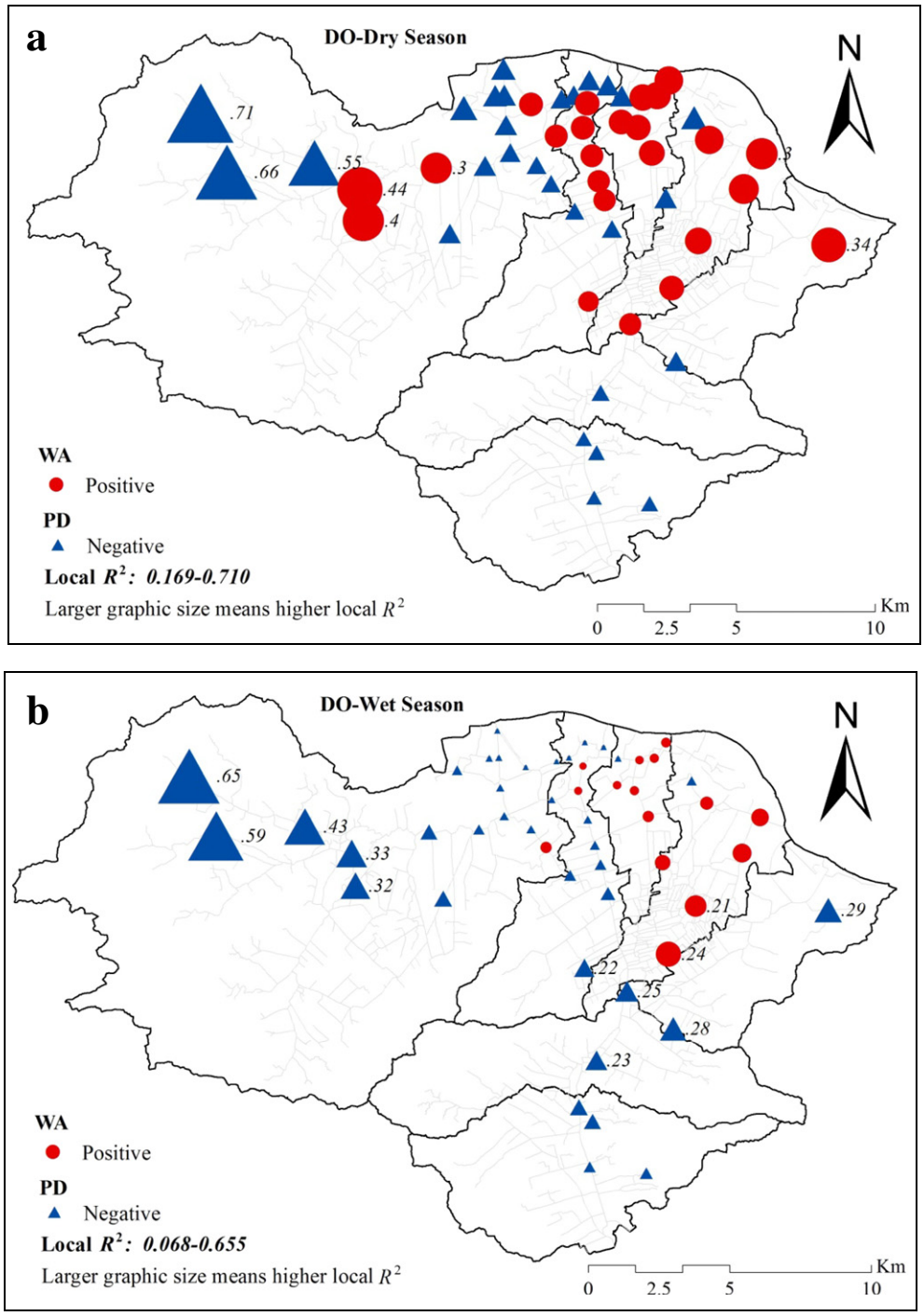


Fig. 9. Primary influencing factor results for DO in a dry season and b wet season.

The primary factors influencing TN exhibited similar spatial variations as TP (Figs. 10–11). AG had a stronger impact on both TN and TP than Urban at sites H1 and H2 during both seasons. The subcatchments located near the central city with higher impervious surface areas and higher PD suffer a greater effect from Urban than AG, however, the corresponding local R^2 was at a low level (<0.2). In contrast, in the northeastern portion of basin V, especially at site H6, AG showed a negative influence on TP and a positive influence on TN during both seasons. Specific to the subcatchments in cluster 1 where AG dominated, Urban played a more important role than AG, but the percentage of urban land explained only ~25% of the variance in TP and TN. In Sanyang wetland, there was a negative correlation between AG and TP (Fig. 5) and a positive influence from Urban (Fig. 10). Differences among the primary factors influencing TN between the wet and dry season (Fig. 11) indicate that more subcatchments are affected by Urban during the wet season, while more subcatchments are influenced by AG in the dry season. These seasonally sensitive subcatchments are mostly located along the perimeter of the city (northeastern basin I and upstream of Sanyang wetland).

The spatial distribution of the primary factors influencing NH_4^+-N during the dry season (Fig. 12a) suggested that AG was a major pollutant source with high predictability at sites H1 and H2 (local $R^2 = 0.66$ & 0.53) and with similar results for TP (Fig.10a) and TN (Fig. 11a). During the wet season (Fig. 12b), PD appears to be most important in

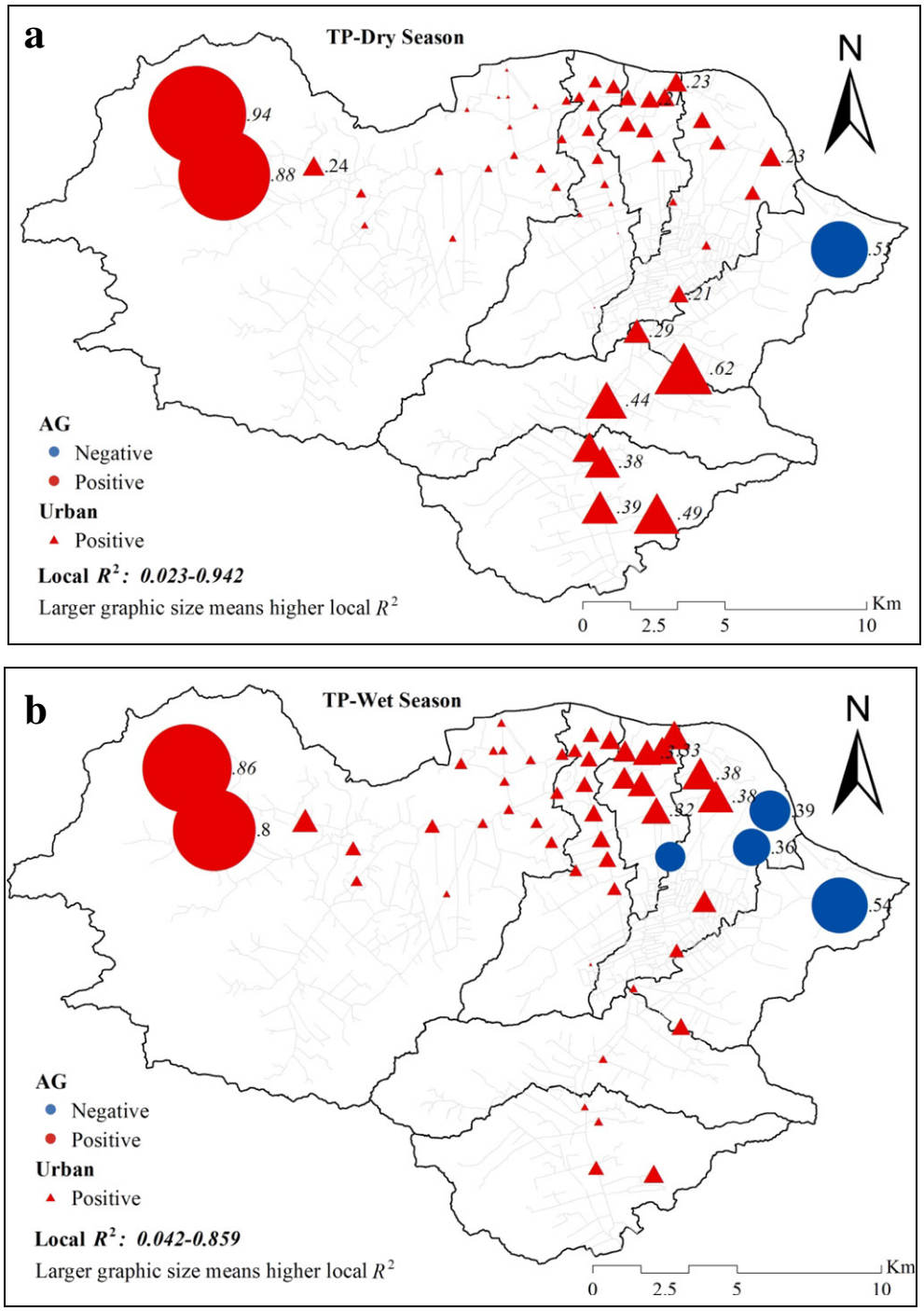


Fig. 10. Primary influencing factor results for TP in a dry season and b wet season.

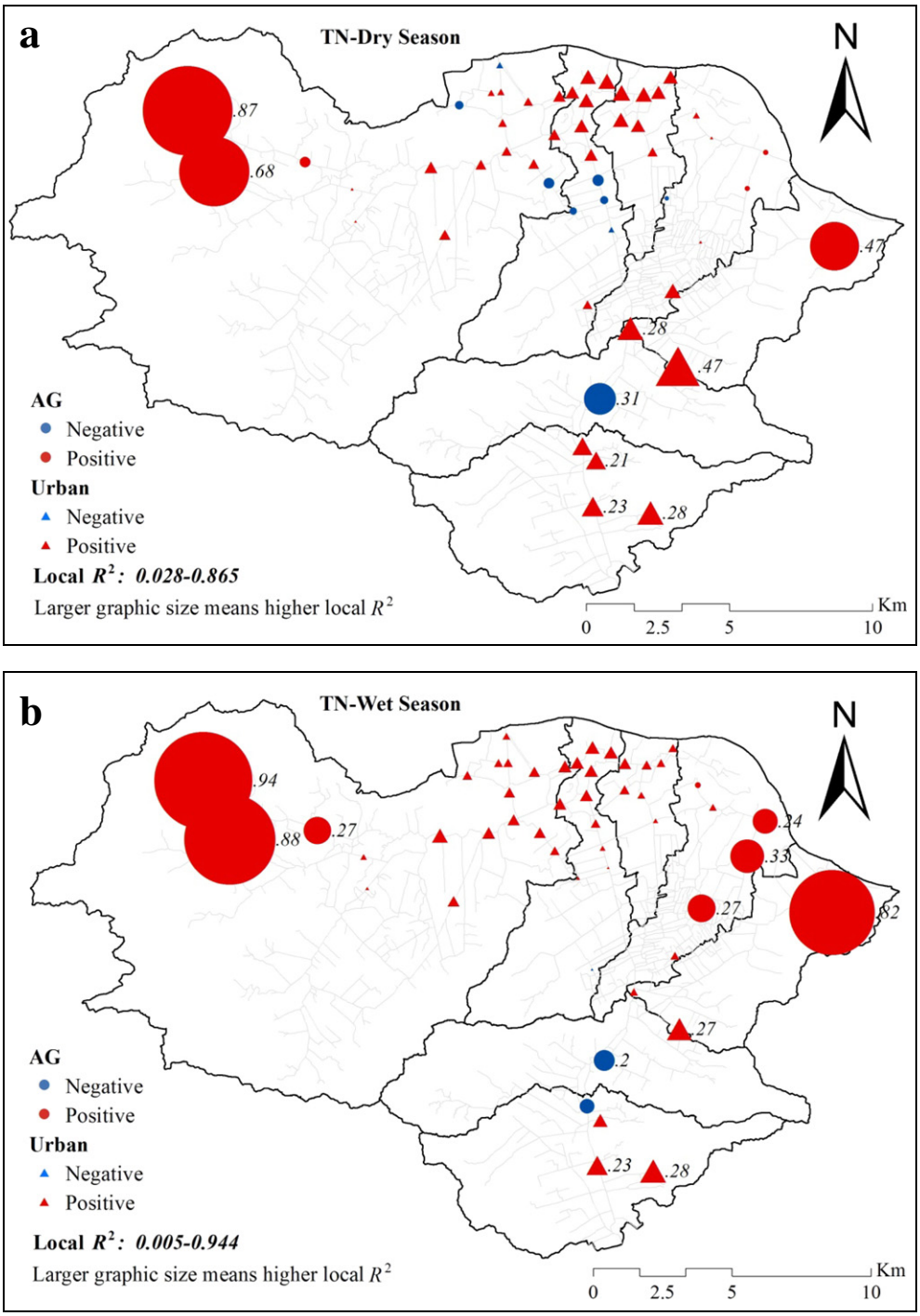


Fig. 11. Primary influencing factor results for TN in a dry season and b wet season.

determining NH_4^+-N for subcatchments in the central city, similar to the subcatchments in cluster 1 and cluster 3. The local R^2 was relatively stable ~0.26 across clusters 1 and 3.

5. Discussion

5.1. Interpretation of water quality and predicting variables

Positive correlations between urban land use with TN, TP, NH_4^+-N and COD in the majority of subcatchments with few exceptions were expected (Figs. 4–8), as urban lands are associated with various anthropogenic and economic activities generating pollutants, such as discharge of residential and industrial sewage. In addition, a higher percentage of impervious surfaces in urban areas prevents rainfall from infiltrating into soil resulting in transport of soluble and particulate forms of pollutants to nearby streams through surface runoff. These results are consistent with several previous studies examining pollutant inputs from urban areas (Walsh and Wepener, 2009; Peters, 2009; Tu, 2011; Liu et al., 2013; Mei et al., 2014; Wang et al., 2014).

The positive and negative relationships between AG and TN, TP and NH_4^+-N in GWR analysis (Figs.4, 5 and 8) showed divergence from our OLS results (Table 2) and other studies using traditional global statistical methods (e.g., Seeboonruang, 2012; Wang et al., 2014; Wan et al., 2014; Chen and Lu, 2014). In these studies, AG had a strong positive effect on

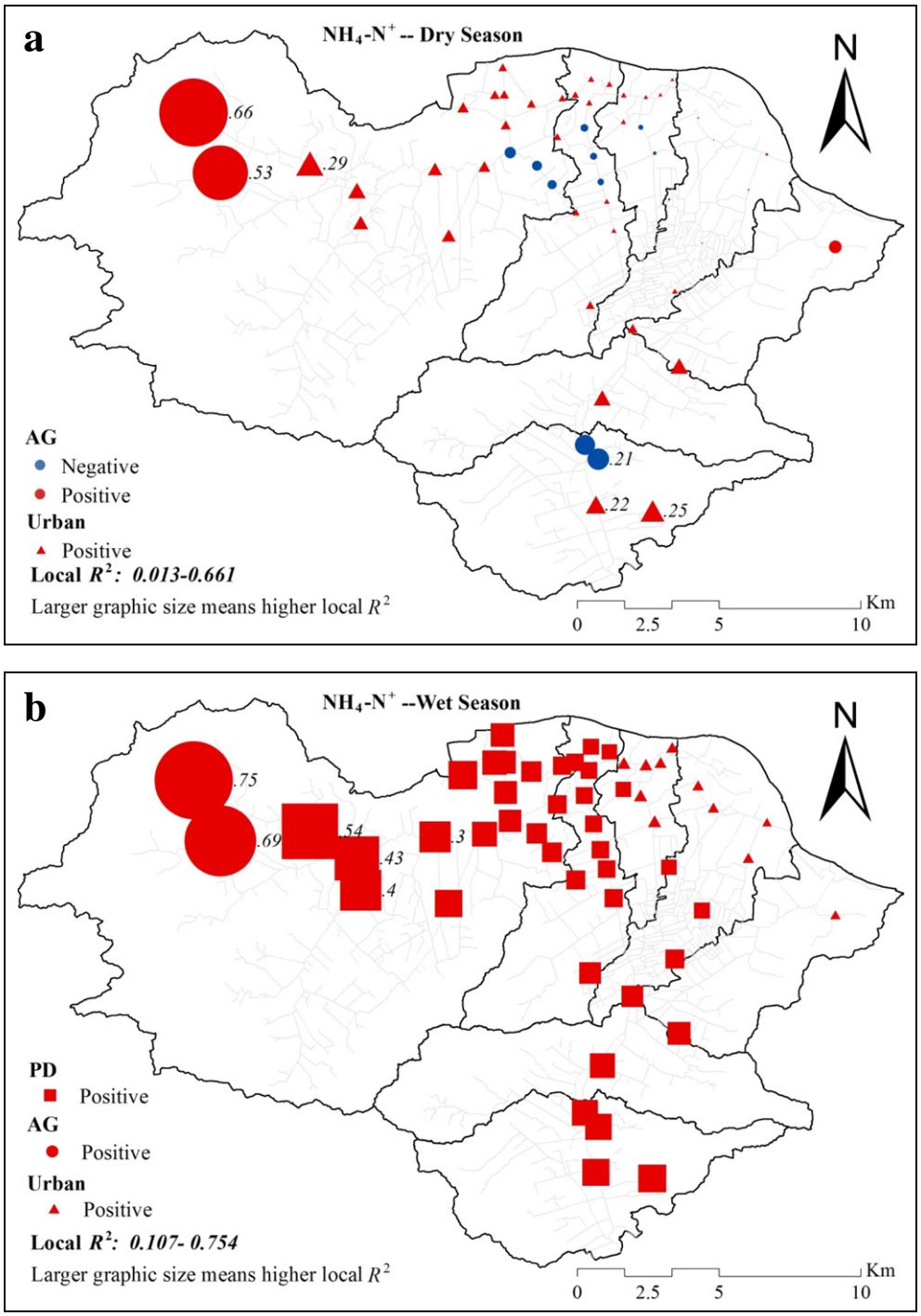


Fig. 12. Primary influencing factor results for NH_4^+-N in a dry season and b wet season.

water pollution indicators because agricultural activities, such as fertilizer and pesticide application and livestock farming, are often the main non-point pollution source. However, the strong relationship with AG was also present in less urbanized areas, especially in agriculture-dominated subcatchments. This is demonstrated by the subcatchments in cluster 1 and sites H1 and H2 that have an average AG of ~40%. A weak negative influence from AG in our study was mainly apparent in the central city, where the percentage of agricultural land was low relative to urban land. These GWR results are consistent with findings from northern Georgia (Tu, 2013) and eastern Massachusetts (Tu, 2011) indicating that agricultural lands are an important pollution source in less-urbanized areas, but in highly urbanized areas its contribution to pollution is negligible and usually dominated by urban sources.

Population density in most subcatchments except Sanyang wetland was found to be positively correlated with COD during the wet season (Fig. 7b) along with a corresponding negative effect on DO (Fig. 6) during both seasons. The higher population density results in higher organic contaminant discharge, such as food waste, human sewage and industrial wastes, which contribute to increased COD concentrations (Xu et al., 2005; Campos et al., 2012; Mohseni-Bandpei and Yousefi, 2013; Sivri et al., 2014).

5.2. Spatial distribution of primary influencing factors

In the Wen-Rui Tang River watershed, the impact from AG was found to be the primary factor influencing TP and TN concentrations at sites H1 and H2 (Figs. 10–11). However, the dominant land use in these subcatchments was natural vegetation (GR) rather than AG (Figs. 1–2). In contrast, urban land use appeared to be a dominant factor determining TP in subcatchments comprising cluster 1 that had an average AG of 45% and average impervious surface area (including urban and industry land) of 40% (Fig. 10). This suggests that urban land might play a more important role in the suburban area where AG and Urban share a similar proportion. This might be due to urban lands transporting more pollutants by overland flow, while in less impervious areas the infiltration process mitigates pollutants by soil retention processes while other pollutants are transferred to groundwater by infiltration and interflow (Cunningham et al., 2010; Seo and Schmidt, 2012).

Our analysis also showed that more subcatchments had urban as the primary influencing factor during the wet season while more subcatchments had AG as the primary influencing factor in the dry season (Figs. 10–11). The subcatchments experiencing seasonal changes in the primary influencing factor were mostly located in the suburban areas. The Wen-Rui Tang River watershed experiences a subtropical monsoon climate with high intensity rainfall (including typhoons) in the wet season and half of the rainfall events lasting for more than 1 h. Such prolonged precipitation events can result in stable or even decreased pollutant runoff from impervious surfaces due to a dilution effect (pollutant supply limited; Seo and Schmidt, 2012). In addition to the seasonal climate effect, landscape characteristics also played a role in pollutant runoff. The intensity, volume and duration of rainfall during the dry season are much smaller and the interval between rainfall events is much larger. Thus, the influence from runoff mobilization of pollutants is much stronger than rainfall dilution during the dry period. As for subcatchments located in the northeastern portion of basin I and the Sanyang wetland, they tend to be more strongly affected by upland drainage from the eastern portions of basins I and VI, which delivers higher quality waters to the lowlands.

Both positive and negative correlations with water quality indicators were observed in the GWR models (Figs. 4–8), but only one variable can be considered the primary factor affecting a given water quality parameter. Other variables may be important as well; however, they will be masked by the effects of the stronger predictive variable. Thus, if the predicting factor-pollutant indicator relationship does not match our understanding of water quality dynamics for a particular water quality indicator, it is important to reevaluate the modeling process to verify if a significant variable was lost during the variable selection process. Based on the optimized model results, environmental protection agencies may devise water quality remediation plans specific to local areas and seasonal considerations advised by the model.

5.3. Variable selection

Of all the explanatory variables considered, Urban was the most common variable selected for OLS models irrespective of SMLR or manual variable excluding-selecting methods (Table 2). The IN and GR variables fell out of the OLS models while PD maintained a significant positive correlation with COD and NH_4^+-N, similar to the findings of Ahearn et al. (2005). These results can be considered as a reflection of the fact that urban land percentage and PD are different metrics of human habitation (Baker, 2003). In addition, several other explanatory variables, such as topographic and landscape indexes, are commonly applied as potential explanatory variables (Chen et al., 2002; Pratt and Chang, 2012; Sun et al., 2014). However, our previous research did not find strong relationships between these subcatchment characteristics and water quality in the Wen-Rui Tang River watershed (Mei et al., 2014; Lu et al., 2011).

5.4. Performance and uncertainty of the modeling system

The distribution of local R^2 results (Figs. 9–12) confirmed our assertion that the influence of LU/PD on specific water quality parameters change spatially and with seasons. The higher local R^2 values appearing in rural areas suggest that the GWR model is more applicable in less-urbanized areas. This relationship reflects that non-point pollution makes a greater contribution to water quality in rural areas.

The majority of studies have found that the influence of land use on water quality varies with different scales of analysis. The watershed scale has been verified to be more effective for spatial analysis than buffer-strip analysis (Lee et al., 2010; Pratt and Chang, 2012; Nielsen et al., 2012; Sun et al., 2014). Moreover, the method of delineating subcatchments may strongly affect results as different methods can lead to different land-use classifications for the same monitoring site. Similar to Kang et al. (2010) and Tu (2013), we designed the subcatchments with the sampling sites located on the subcatchment boundary rather than in the interior. In other words, the sampling sites were considered the outlet of the subcatchments. In highly fragmented subcatchments, it is very difficult to delineate subcatchments that accurately reflect the characteristics of the basin.

Additionally, subcatchment division using direct DEM processing methods is inaccurate in plain river networks, such as the Wen—Rui Tang River watershed. The relatively flat landscapes are usually beyond the required vertical resolution required to accurately identify realistic drainage flow paths (Callow et al., 2007; Getirana et al., 2009). Thus, further research is needed to explore more precise methods for subcatchment delineation to assist in the identification of the primary factors influencing water pollution at the different monitoring sites.

6. Conclusion

This study examined the influence of six water quality predictors including land-use types and population density on five water quality parameters in the Wen-Rui Tang River watershed during both wet and dry seasons. A manual variable excluding-selecting method was explored to resolve multicollinearity among independent variables in the GWR models. To determine which subcatchment characteristics had the greatest control on water quality indicators, we introduced the concept of standard regression coefficient analysis along with cluster analysis. The main results are as follows:

(1) Impact of land use on water quality changes along with the spatial and seasonal scales. For instance, the relationships between TN and LU/LC percent in GWR results showed a positive correlation in rural areas and negative correlation in urban areas. The absolute values of local regression coefficients were lower in urban areas and higher in suburban and rural areas ($p < 0.05$). More subcatchments in suburban areas showed a negative correlation between agricultural land and TN during the dry season, while more subcatchments showed a positive correlation during the wet season. This may reflect surface runoff being a more important pathway of pollutant transport to rivers in suburban areas having greater impervious area.

(2) The factor having the largest contribution to water quality parameters varied with space and time. Urban land was found to be the dominant influencing factor on N, P and COD in the highly urbanized regions, but the relationship was weak as the pollutants mainly derive from point sources. Agricultural land was found to be the primary influencing factor on N and P in suburban and rural areas; the relationship here was strong as the pollutants are mainly deriving from agricultural surface runoff. Urban land played a more important role in the suburban area where AG and Urban share a similar proportion of land area. This may result from urban land having pollutants transported through overland flow, while in less impervious areas infiltration mitigates pollutants by soil retention and some pollutants percolate to the groundwater system. Subcatchments located in suburban areas were identified with urban land as the primary influencing factor during the

wet season while agricultural land was identified as a more prevalent influencing factor during the dry season.

(3) The manual variable excluding-selecting method was found to be effective in solving the multicollinearity issue. Adjusted R^2 values in OLS models using the manual variable excluding-selecting method were 14.3% higher than using stepwise multiple linear regression, indicating that the optimized multivariate model was efficacious and highly interpretability. However, the corresponding GWR models had adjusted R^2 values an average of 59.2% higher than the optimal OLS models, indicating that GWR models had better prediction accuracy than OLS models and better reflect the actual watershed spatial characteristics.

The manual variable excluding-selecting method and the re-processing method for GWR output used in this study have wide spread suitable in other study areas having watershed-scale, water quality data and a small number of explanatory variables. This watershed analysis provides an important tool for local water resource agencies to establish more scientific-based water quality management and remediation plans.

Acknowledgements

This research was supported by the Science and Technology Department of Zhejiang Province (No. 2008C03009) and Wenzhou City (No. 20140037, No. 20140039). The authors would like to express appreciation to the Wenzhou Environmental Protection Bureau (WEPB) for the data they provided.

References

Ahearn, D.S., Sheibley, R.W., Dahlgren, R.A., Anderson, M., Johnson, J., Tate, K.W., 2005. Land use and land cover influence on water quality in the last free-flowing river draining the western Sierra Nevada, California. J. Hydrol. 313 (3/4), 234–247.

ArcGIS Help Library, 2014. How Kernel Density Works [EB/OL]. (http://resources.arcgis.com/en/help/main/10.2/index.html#/How_Kernel_Density_works/009z00000011000000/).

Baker, A., 2003. Land use and water quality. Hydrol. Process. 17 (12), 2499–2502.

Boots, B., 2003. Geographically weighted regression: the analysis of spatially varying relationships. Int. J. Geogr. Inf. Sci. 17 (7), 717–719.

Brown, S., Versace, V.L., Laurenson, L., Lerodiaconou, D., Fawcett, J., Salzman, S., 2012. Assessment of spatiotemporal varying relationships between rainfall, land cover and surface water area using geographically weighted regression. Environ. Model. Assess. 17 (3), 241–254.

Brunsdon, C., Fotheringham, A.S., Charlton, M., 2002. Geographically weighted summary statistics - a framework for localised exploratory data analysis. Comput. Environ. Urban. Syst. 26 (6), 501–524.

Cahill, M., Mulligan, G., 2007. Using geographically weighted regression to explore local crime patterns. Soc. Sci. Comput. Rev. 25 (2), 174–193.

Callow, J.N., Van Niel, K.P., Boggs, G.S., 2007. How does modifying a DEM to reflect known hydrology affect subsequent terrain analysis? J. Hydrol. 332 (1–2), 30–39.

Campos, I., Alcaniz, M., Aguado, D., Barat, R., Ferrer, J., Gil, L., Marrakchi, M., Martinez, M.R., Soto, J., Vivancos, J.L., 2012. A voltammetric electronic tongue as tool for water quality monitoring in wastewater treatment plants. Water Res. 46 (8), 2605–2614.

Carrel, M., Escamilla, V., Messina, J., Giebultowicz, S., Winston, J., Yunus, M., Streatfield, P.K., Emch, M., 2011. Diarrheal disease risk in rural Bangladesh decreases as tubewell density increases: a zero-inflated and geographically weighted analysis. Int. J. Health Geogr. 10 (14), 1761–1776.

Chen, J.B., Lu, J., 2014. Effects of land use, topography and socio-economic factors on river water quality in a mountainous watershed with intensive agricultural production in East China. PLoS One 9 (8), e102714. http://dx.doi.org/10.1371/journal.pone.0102714.

Chen, L.D., Fu, B.J., Zhang, S.R., Qiu, J., Guo, X.D., Yang, F.L., 2002. A comparative study on nitrogen-concentration dynamics in surface water in a heterogeneous landscape. Environ. Geol. 42 (4), 424–432.

Chen, D.J., Hu, M.P., Guo, Y., Dahlgren, R.A., 2016a. Modeling forest/agricultural and residential nitrogen budgets and riverine export dynamics in catchments with contrasting anthropogenic impacts in eastern China between 1980–2010. Agric. Ecosyst. Environ. 221, 145–155.

Chen, D.J., Hu, M.P., Wang, J.H., Guo, Y., Dahlgren, R.A., 2016b. Factors controlling phosphorus export from agricultural/forest and residential systems to rivers in eastern China, 1980–2011. J. Hydrol. 533, 53–61.

Cunningham, M.A., Menking, K.M., Gillikin, D.P., Smith, K.C., Freimuth, C.P., Belli, S.L., Pregnall, A.M., Schlessman, M.A., Batur, P., 2010. Influence of open space on water quality in an urban stream. Phys. Geogr. 31 (4), 336–356.

Dowd, B.M., Press, D., Los, H.M., 2008. Agricultural nonpoint source water pollution policy: the case of California's Central Coast. Agric. Ecosyst. Environ. 128 (3), 151–161.

Fotheringham, A.S., Charlton, M., Brunsdon, C., 1996. The geography of parameter space: an investigation of spatial non-stationarity. Int. J. Geogr. Inf. Syst. 10 (5), 605–627.

Getirana, A.C.V., Bonnet, M.P., Rotunno, O.C., Mansur, W.J., 2009. Improving hydrological information acquisition from DEM processing in floodplains. Hydrol. Process. 23 (3), 502–514.

Griffith, D.A., 2008. Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR). Environ. Plan. A 40 (11), 2751–2769.

Huisman, N.L.H., Karthikeyan, K.G., Lamba, J., Thompson, A.M., Peaslee, G., 2013. Quantification of seasonal sediment and phosphorus transport dynamics in an agricultural watershed using radiometric fingerprinting techniques. J. Soils Sediments 13 (10), 1724–1734.

Kang, J.H., Lee, S.W., Cho, K.H., Ki, S.J., Cha, S.M., Kim, J.H., 2010. Linking land-use type and stream water quality using spatial data of fecal indicator bacteria and heavy metals in the Yeongsan river basin. Water Res. 44, 4143–4157.

Kumar, S., Lal, R., Lloyd, C.D., 2012. Assessing spatial variability in soil characteristics with geographically weighted principal components analysis. Comput. Geosci. 16 (3), 827–835.

Lee, Yang, J.S., Kim, D.K., Han, M.Y., 2010. Relationship between land use and water quality in a small watershed in South Korea. Water Sci. Technol. 62 (11), 2607–2615.

Li, J., Liu, H.X., Li, Y.C., Mei, K., Dahlgren, R.A., Zhang, M.H., 2013. Monitoring and modeling dissolved oxygen dynamics through continuous longitudinal sampling: a case study in Wen-Rui Tang River, Wenzhou, China. Hydrol. Process. 27 (24), 3502–3510.

Liu, Z.H., Wang, Y.L., Li, Z.Z.G., Peng, J., 2013. Impervious surface impact on water quality in the process of rapid urbanization in Shenzhen, China. Environ. Earth Sci. 68 (8), 2365–2373.

Lu, P., Mei, K., Zhang, Y.J., Liao, L.L., Long, B.B., Dahlgren, R.A., Zhang, M.H., 2011. Spatial and temporal variations of nitrogen pollution in Wen-Rui Tang River watershed, Zhejiang, China. Environ. Monit. Assess. 180 (1–4), 501–520.

Madrinan, M.J.M., Al-Hamdan, M.Z., Rickman, D.L., Ye, J., 2012. Relationship between watershed land-cover/land-use change and water turbidity status of Tampa Bay major tributaries, Florida, USA. Water Air Soil Pollut. 223 (5), 2093–2109.

Mayer, L.S., Younger, M.S., 1976. Estimation of standardized regression coefficients. J. Am. Stat. Assoc. 71 (353), 154–157.

Mei, K., Liao, L.L., Zhu, Y.L., Lu, P., Wang, Z.F., Dahlgren, R.A., Zhang, M.H., 2014. Evaluation of spatial-temporal variations and trends in surface water quality across a rural-suburban-urban interface. Environ. Sci. Pollut. Res. 21 (13), 8036–8051.

Mohseni-Bandpei, A., Yousefi, Z., 2013. Status of water quality parameters along Haraz River. Int. J. Environ. Res. 7 (4), 1029–1038.

Nielsen, A., Trolle, D., Sondergaard, M., Lauridsen, T.L., Bjerring, R., Olesen, J.E., Jeppesen, E., 2012. Watershed land use effects on lake water quality in Denmark. Ecol. Appl. 22 (4), 1187–1200.

Nimon, K.F., Oswald, F.L., 2013. Understanding the results of multiple linear regression: beyond standardized regression coefficients. Organ. Res. Methods 16 (4), 650–674.

Peters, N.E., 2009. Effects of urbanization on stream water quality in the city of Atlanta, Georgia, USA. Hydrol. Process. 23 (20), 2860–2878.

Pineda, J.N., Bosque, S.J., Gomez, D.M., Franco, P.R., 2010. Exploring the driving forces behind deforestation in the state of Mexico (Mexico) using geographically weighted regression. Appl. Geogr. 30 (4), 576–591.

Pratt, B., Chang, H.J., 2012. Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. J. Hazard. Mater. 209, 48–58.

Roebeling, P.C., Cunha, M.C., Arroja, L., van Grieken, M.E., 2015. Abatement vs. treatment for efficient diffuse source water pollution management in terrestrial-marine systems. Water Sci. Technol. 72 (5), 730–737.

Seeboonruang, U., 2012. A statistical assessment of the impact of land uses on surface water quality indexes. J. Environ. Manag. 101, 134–142.

Seo, Y.W., Schmidt, A.R., 2012. The effect of rainstorm movement on urban drainage network runoff hydrographs. Hydrol. Process. 26 (25), 3830–3841.

SEPBC, 2002a. Environmental Quality Standards for Surface Water; Prepared and Published Jointly by State Environment Protection Bureau of China (SEPBC), State Quality Supervision, Inspection and Quarantine Bureau of China (SQSIQBC), Beijing.

SEPBC, 2002b. Water and wastewater analysis methods. Prepared and Published by State Environment Protection Bureau of China (SEPBC). China Environmental Science Press, Beijing.

Sivri, N., Ongen, A., Aydin, S., Gungor, Y., Azaz, D., 2014. Water quality assessment and monitoring pollution in an unsanitary dumpsite: case study on Narman (Erzurum). Fresenius Environ. Bull. 23 (12B), 3374–3383.

Song, W.Z., Jia, H.F., Huang, J.F., Zhang, Y.Y., 2014. A satellite-based geographically weighted regression model for regional $PM_{2.5}$ estimations over the Pearl River Delta region in China. Remote Sens. Environ. 154, 1–7.

Sun, H.J., Zhang, H.L., Yu, Z.M., Wu, J.S., Jiang, P.K., Yuan, X.Y., Shi, W.M., 2013. Combination system of full-scale constructed wetlands and wetland paddy fields to remove nitrogen and phosphorus from rural unregulated non-point sources. Environ. Geochem. Health 35 (6), 801–809.

Sun, Y.W., Guo, Q.H., Liu, J., Wang, R., 2014. Scale effects on spatially varying relationships between urban landscape patterns and water quality. Environ. Manag. 54 (2), 272–287.

Tsegaye, T., Sheppard, D., Islam, K.R., Johnson, A., Tadesse, W., Atalay, A., Marzen, L., 2006. Development of chemical index as a measure of in-stream water quality in response to land-use and land cover changes. Water Air Soil Pollut. 174 (1–4), 161–179.

Tu, J., 2011. Spatially varying relationships between land use and water quality across an urbanization gradient explored by geographically weighted regression. Appl. Geogr. 31, 376–392.

Tu, J., 2013. Spatial variations in the relationships between land use and water quality across an urbanization gradient in the watersheds of northern Georgia, USA. Environ. Manag. 51 (1), 1–17.

Tu, J., Xia, Z.G., 2008. Examining spatially varying relationships between land use and water quality using geographically weighted regression I. Model design and evaluation. Sci. Total Environ. 407 (1), 358–378.

Walsh, G., Wepener, V., 2009. The influence of land use on water quality and diatom community structures in urban and agriculturally stressed rivers. Water SA 35 (5), 579–594.
Wan, Y.S., Qian, Y., Migliaccio, K.W., Li, Y.C., Conrad, C., 2014. Linking spatial variations in water quality with water and land management using multivariate techniques. J. Environ. Qual. 43 (2), 599–610.
Wang, R.Z., Xu, T.L., Yu, L.Z., Zhu, J.J., Li, X.Y., 2013. Effects of land use types on surface water quality across an anthropogenic disturbance gradient in the upper reach of the Hun River, Northeast China. Environ. Monit. Assess. 185 (5), 4141–4151.
Wang, G.Q., A, Y.L., Xu, Z.X., Zhang, S.R., 2014. The influence of land use patterns on water quality at multiple spatial scales in a river system. Hydrol. Process. 28 (20), 5259–5272.
Wheeler, D.C., 2007. Diagnostic tools and a remedial method for collinearity in geographically weighted regression. Environ. Plan. A 39 (10), 2464–2481.
Wheeler, D.C., 2009. Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted LASSO. Environ. Plan. A 41 (3), 722–742.
Wheeler, D., Tiefelsdorf, M., 2005. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. J. Geogr. Syst. 7 (2), 161–187.
Xu, M.Q., Cao, H., Xie, P., Deng, D.Q., Feng, W.S., Xu, J., 2005. Use of PFU protozoan community structural and functional characteristics in assessment of water quality in a large, highly polluted freshwater lake in China. J. Environ. Monit. 7 (7), 670–674.
Yang, X.W., Belin, T.R., Boscardin, W.J., 2005. Imputation and variable selection in linear regression models with missing covariates. Biometrics 61 (2), 498–506.
Yuan, K.H., Chan, W., 2011. Biases and standard errors of standardized regression coefficients. Psychometrika 76 (4), 670–690.
Zhang, X.D., Huang, G.H., Nie, X.H., 2011. Possibilistic stochastic water management model for agricultural nonpoint source pollution. J. Water Resour. Plan. Manag. 137 (1), 101–112.
Zielinski, M., Dopieralska, J., Belka, Z., Walczak, A., Siepak, M., Jakubowicz, M., 2016. Sr isotope tracing of multiple water sources in a complex river system, Notec River, central Poland. Sci. Total Environ. 548, 307–316.