An analysis of the relationship between spatial patterns of water quality and urban development in Shanghai, China

By: Zhi-Yong Yin, Susan Walcott, Brian Kaplan, Jian Cao, Weiqing Lin, Minjian Chen, Dongsheng Liu, and Yuemin Ning

Yin, Z-Y, Walcott, S., Kaplan, K., Cao, J., Lin, W., Chen, M., Liu, D., and Ning, Y. (2005) "An Analysis of the Relationship between Spatial Patterns of Water Quality and Urban Development in Shanghai, China" *Computers, Environment and Urban Systems* 29(2): 197-221.

Abstract:
Recent urban development in Shanghai, the largest city in China, and its impact on the water environment are examined in this study. The area of built-up surface was obtained from the classification of the Landsat 7 ETM+ images of the year 2000 for Shanghai. The proportion of built-up surface and population density were extracted from buffer zones with radii ranging from 100 to 2000 m, and used in regression analysis against various water quality parameters at 44 water quality monitoring stations across metropolitan Shanghai. Results suggest that in most cases, the pattern of urban land use as represented by the built-up surface was a stronger predictor than population density in explaining spatial patterns of water quality parameters in Shanghai. The best models of most water quality parameters were found for buffer zones of 2000 m radius rather than for smaller buffers, indicating the regional nature of the factors that influence water quality in the study area. Evidence suggests that strong associations between land use, population density, and water quality result from the contribution of untreated domestic wastewater and non-point pollution sources to waterways in Shanghai. Such relationships should remain strong in the near future until measures to increase the capacity of wastewater treatment and control of non-point pollution sources are fully implemented.
Keywords: Water quality; Shanghai; Urbanization; GIS; Remote Sensing

Article:

1. Introduction

Shanghai is the largest city in China with a population of over 13 million in the year 2000 (Shanghai Statistical Bureau, 2002). The municipality occupies an area of approximately 6400 km^2 on the Yangtze River Delta (Fig. 1). Shanghai is a major industrial center of China, with many of its over 10,000 factories and industrial enterprises concentrated in the city's urban core. The industries range from petrochemical, textiles and metallurgy to paper production and food processing. With approximately 1.2% of the total population in China, Shanghai generates approximately 4.9% of the nation's industrial output and 4.2% of the national gross domestic product (GDP) (Shanghai Statistical Bureau, 2002). With Shanghai and other surrounding smaller cities, the Yangtze River Delta is becoming the largest mega-urban region in Asia (Douglass, 2000).

Major waterways in Shanghai's vicinity are overburdened with sewage demands. From the mid-1880s to 1949, a few wastewater treatment plants served only the then relatively new foreign concessions areas, leaving the rest of the city without any sewage collection and delivery system. Little discernable improvements occurred after 1949 (Wu & Shi, 1999). Recent rapid urban development in and around Shanghai puts an even heavier burden on the environment. Industries in the Shanghai metropolitan area generate 70% of the 5.4 million tons of liquid effluents daily (Wu & Shi, 1999). Similar to the situation in many developing countries, economic development and industrialization often have a higher priority than environmental protection in China (Shen, Cheng, Gu, & Lu, 2002). First appearing in the late 1970s, the legal framework regarding environmental protection evolved gradually and matured during the next 25 years. However, even though now there are numerous regulations and policies regarding pollution control, the political, financial and legal systems often do not offer sufficient support or provide enforcement mechanisms and incentives (Shen et al., 2002; Ward & Liang, 1995). In a recent survey (Shanghai EPB, Shanghai Water Affairs Bureau, & Shanghai Urban Planning & Management Bureau, 2001), over 55,000 point sources were identified throughout the metropolitan area, with the highest density in the central city districts. Among them, over 33,000 are pollution sources discharging directly into waterways. Although the municipal government claims that over 98% of total industrial wastewater discharge is treated (Shanghai Statistical Bureau, 2002), it is the cumulative effect of large amount of industrial effluents and the density of pollution sources along the waterways that pose great threats to Shanghai's water environment. Another problem that is getting more attention now is non-point pollution, which is estimated to contribute up to 50–60% of water pollution in Shanghai (SAES, 1998).

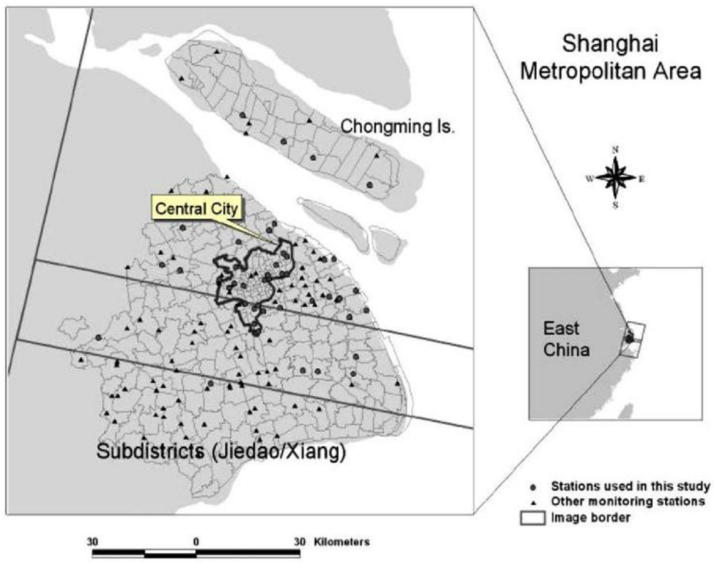


Fig. 1. Shanghai metropolitan area, subdistricts (*Jiedao* and *Xiang*), and water quality monitoring stations.

Untreated domestic wastewater, frequent flooding and street runoff all contribute to water quality deterioration for waterways around Shanghai, which also severely affects Shanghai's drinking water sources. The major domestic water supply intake had to be relocated to further upstream of the Huangpu River for less polluted water (Yusuf & Wu, 2002). Another example of severe water quality problems faced by Shanghai is the Suzhou Creek, a tributary of the Huangpu River. Suzhou Creek is approximately 125 km long, with 24 km flowing through the highly urbanized part of Shanghai before it joins the Huangpu River. Since the 1920s, Suzhou Creek has been known for its black-colored water and foul odor because it receives a large quantity of untreated domestic and industrial wastewater. It became a major focus of environmental rehabilitation because the government used Suzhou Creek as a symbol of its determination to curb environmental pollution. A massive 12-year rehabilitation program financed by the World Bank, the Asian Development Bank (ADB) and the Shanghai Municipal Government is near completion (ADB, 1999). Although water quality problems in Shanghai have been recognized for a long time, very few studies have addressed this issue in the past (e.g., Kung & Ying, 1991; Ward & Liang, 1995), partially because many environmental data are considered sensitive information and unavailable to general public and scholars.

The purpose of this study is to investigate the relationship between spatial patterns of urbanization processes and water quality in Shanghai using water quality data of 1995. Such understanding may help future planning and efforts to alleviate water quality problems in the region. The period of the late 1980s and early 1990s represents an important milestone of Shanghai's recent history, as manifested by new areas opened for development, reform of the housing system, and significant changes in the city's infrastructure and physical environment (Binns, 1991; He, 1993; Shi & Hamnett, 2002; Ward & Liang, 1995; Yusuf & Wu, 2002). Results from this study can serve as the baseline condition for future comparisons to identify the trend of improvement or deterioration of Shanghai's water environment. The municipal government has set very ambitious goals for Shanghai (Shi & Hamnett, 2002). With rising living standards and increasing public awareness of the status of Shanghai's water environment, there is now a great sense of urgency for dealing with the water quality issues in order for Shanghai to accomplish its aspiration as a “global city” (Leman, 2002; Shi & Hamnett, 2002; Yusuf & Wu, 2002).

2. Methods and data

The intensity of urbanization can be measured by different parameters, such as population density and industrial output. In this study, population density and the amount of land surface used for buildings and infrastructure are considered. For the purpose of this study, all types of non-vegetated artificial land surfaces are defined as built-up surface. It has been found that the percentage of built-up surface or urban impervious surface is a key indicator of the effects of non-point pollution on water quality (Slonecker, Jennings, & Garofalo, 2001). Shanghai is located in a humid subtropical region suitable for intense farming activities. Unless the land is used for industrial/commercial or residential developments, it is usually cultivated for farming year around. Therefore, the intensity of urbanization can be represented by the proportion of built-up surface in a given area, which can then be related to water quality parameters.

Remote sensing has become an effective tool in urban studies. One of the widely used sources of images is the Landsat TM imagery. Since its first deployment in 1982, with a ground resolution of 30 × 30 m (for visible, NIR and MIR bands) and seven spectral bands, it soon became an excellent tool in regional studies and has been used in many studies in China (e.g., Ji et al., 2001; Li & Yeh, 1998; Weng, 2001) as well as in other countries. Remote sensing technology is especially useful in areas where maps of high accuracy cannot keep up with rapidly changing surface conditions (Ward, Phinn, & Murray, 2000). A review of applications of remote sensing of built-up surfaces or urban impervious surfaces can be found in Slonecker et al. (2001). In this study, the Landsat 7 ETM+ images (launched in 1999) are used as the main source to estimate built-up surface spatial distribution in Shanghai. The images (Path 118/Rows 38 and 39), obtained from the US Geological Survey (USGS), were acquired on March 26, 2000 (Fig. 1).

Supplementary imagery sources used in this study include the panchromatic 2000 Indian Remote Sensing (IRS) image, 1993 multispectral SPOT XS image, and the 2000 IKONOS image (covering the central city only). These higher spatial resolution images, such as IRS (5 m) and IKONOS (1 m), were mostly used in validation and accuracy assessment of image classifications. Because the two 2000 Landsat ETM+ images display significant spectral differences, separate classifications were conducted and then reconciliation was performed based on the overlapped area. In classification, masks were first used to remove most of the water body pixels from the two images to reduce spectral complexity of the images. Then unsupervised classification was conducted to arrive at 200 clusters. Each cluster was then examined and assigned one of the following land use/cover types: water, built-up surface, vegetation and farmland, coastal land, and mix of vegetation and water (wetlands and rice paddies). This approach is similar to the multiresolution land characterization (MRLC) program of the US EPA (USGS, 2002; Volgelman, Sohl, Howard, & Shaw, 1998). All image processing was conducted using ERDAS Imagine™ (ERDAS, Atlanta, GA).

Population density has been used as a measure of urbanization in many studies based on census data and geographic information systems (GIS) (e.g., Chen, Zeng, & Xie, 2000; Mesev, 1998; Zhou & Ma, 2000). It is calculated as the number of people per unit area in a given region, such as a census tract or a county. In this study, population density is calculated at the level of subdistricts (Jiedao and Xiang) based on population data from 1997, obtained from East China Normal University (Fig. 2). This is the census level with a higher spatial resolution than the county—district level. Population density, together with built-up surface, is then used to explain the spatial variation pattern of water quality parameters.

Water quality data for 1995 at 44 stations are available for this study (Fig. 1), provided by the Shanghai Academy of Environmental Science (SAES). These stations are a subset of the monitoring system managed by the Shanghai Environmental Monitoring Center. Since domestic wastewater and non-point sources contribute a significant proportion of water pollution, this study examines water quality parameters that represent domestic (organic) pollutants as well as industrial pollutants, including dissolved oxygen (DO), chemical oxygen demand (COD Index as the permanganate index, $KMnO_4$), 5-day biological oxygen demand (BOD5), nitrogen as ammonia or NH_3 (NH_3-N), oil, phenol, total phosphorous (TP), total nitrogen (TN), and total mercury (THG). Table 1 contains descriptive statistics of the water quality parameters examined. At each monitoring station, at least two samples were taken under different hydrological conditions during the year (normally 4–6 times), e.g., high water, medium water, and low water conditions (Cao, 2002). The annual average values are used in this study. While it is true that the water quality data used in this study are not concurrent with the time of image acquisition and therefore, the relationships obtained this way may not represent the current condition, we believe that the urban land use pattern in Shanghai had maintained its general characteristics in the late 1990s and that we can still capture the general nature of such relationships at the regional scale.

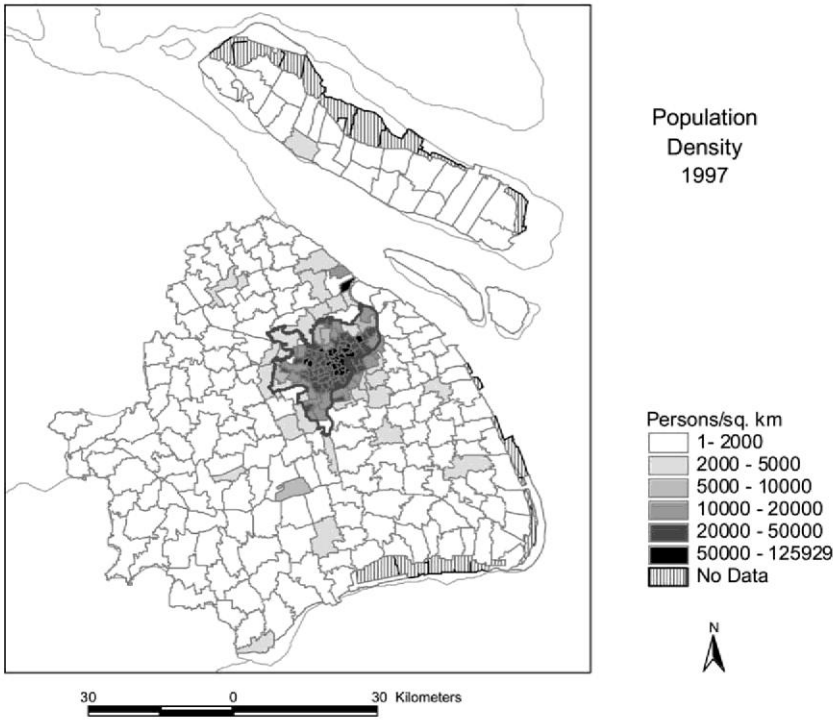


Fig. 2. Population density at the subdistrict (*Jiedao* and *Xiang*) level (persons/km^2).

Shanghai is located on the delta of the Yangtze River, characterized by thick unconsolidated fluvial sediments and a dense network of various types of waterways of natural streams and artificial canals, over 21,000 km in total length (Fig. 3) (Shanghai EPB et al., 2001). Flow velocity and flow direction of many waterways are affected by tidal fluctuation and controlled by sills and gates for flood management purposes. For most waterways, especially small ones, water flow is sluggish and bidirectional due to low channel gradient, tidal action and the interference of the gates. These factors significantly hamper the flushing of pollutants out of their source reaches. Obviously hydrological condition has a great impact on water quality. However, no stream flow data are available for those monitored cross-sections. Instead, a ranking system is used in this study as a proxy of long-term flow condition (Table 2). The ranking data are part of the attributes of the stream GIS coverage obtained from the SAES. Each monitoring station is assigned the rank of the waterway (river rank). When a station is located at the intersection of two waterways with different ranks, the higher rank is used. In general, larger waterways (lower ranks) have faster flows and stronger mixing mechanisms as compared to smaller waterways (higher ranks).

Table 1
Descriptive statistics of water quality parameters (mg/l)

Water quality parameters	N	Mean	Std. dev.
DO	44	3.763	2.823
COD Index	44	14.026	15.254
BOD5	44	20.220	31.619
NH_3-N	44	8.020	8.250
Oil	44	0.854	1.386
Phenol	44	0.037	0.066
Total P	44	0.758	0.741
Total N	44	10.816	9.949
Total HG	44	0.00085	0.00200

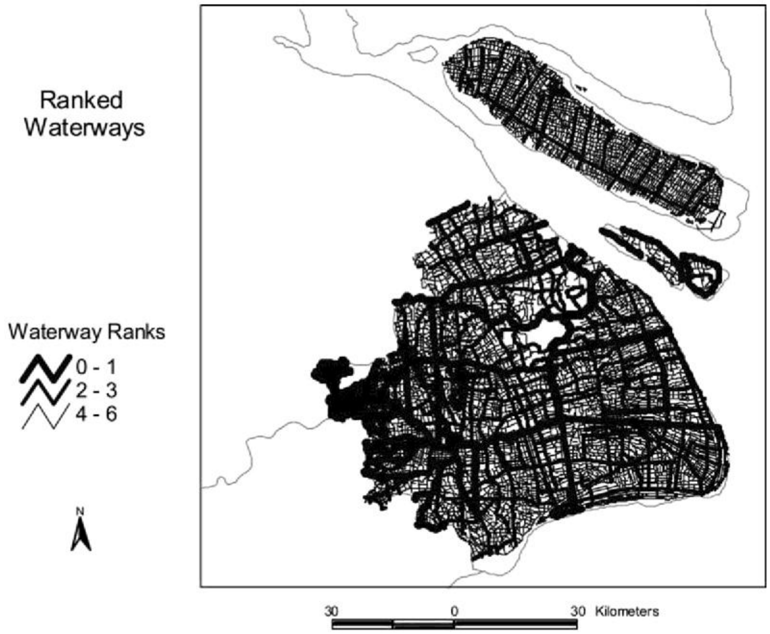


Fig. 3. Ranked waterways in Shanghai.

To simplify the analysis, it is assumed that the interaction between nearby streams/waterways through surface and subsurface water exchange facilitates the dispersion and mixing of pollutants. Therefore, each water quality monitoring station not only represents the reaches immediately above and below its location, but also the nearby waterways. The water quality parameters measured at a given station should then also be related to the urban land use intensity and the distribution of pollution sources in the surrounding area. Due to the aforementioned hydrological characteristics of the waterway network in Shanghai, it is all but impossible to determine the drainage area for each monitored cross-section. Therefore, the conventional approach of relating water quality monitored at the pour point of a drainage basin to land use conditions within the basin is not applicable. To investigate how strongly the land use pattern within a given area is related to water quality, for each of the 44 stations, buffers with radii from 100 to 2000 m are developed (Fig. 4). The buffer zones are then overlaid on the classified images and the land use/cover pattern (built-up surface proportion) is extracted for the buffer zones. Similarly, the population density at the subdistrict level is first converted to a grid of 30 m × 30 m resolution, and then the mean population density is extracted for the same set of buffer zones.

Table 2
Ranking of waterways and general hydrologic condition (Cao, 2002)

Rank	Description	Hydrologic condition
0	Large lakes and waterways that are presented as double lines on the map	Faster flow and stronger mixing
1	Waterways that are open for navigation of ships larger than 100 tons	–
2	Waterways that are open for navigation of ships 60–100 tons	–
3	Waterways that are open for navigation of ships smaller than 60 tons	–
4	Waterways that are not suitable for navigation	–
5	Insignificant small waterways	–
6	Small lakes	Slower flow and weaker mixing

Spatial analyses were conducted using ArcView, a GIS software package from ESRI (Redlands, CA). In searching for the relationship between urbanization and water quality, built-up surface proportion, population density, and river rank are used in correlation and regression analyses against water quality parameters. Step-wise regression is used to find the best models for the water quality parameters, in which only independent variables with statistical significance of 0.05 are included. Two types of relationships are considered in analysis: linear and exponential. All statistical analyses were performed using SAS (SAS Institute, Carey, NC).

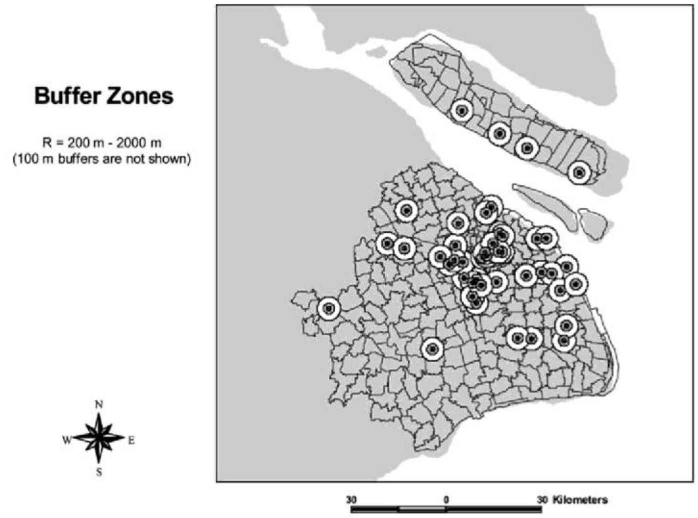


Fig. 4. Buffers of radii from 200 to 2000 m surrounding the water quality monitoring stations.

3. Results and analysis

3.1. Land use/cover classification

In unsupervised classification, the ETM+ images were first classified into 200 clusters using the routine of ISODATA of ERDAS Imagine. Even with such fine divisions, it was found that some clusters still represented pixels of multiple land use/ cover types, mostly found in the southern image (Path 118/Row 39). To solve this problem, these pixels of mixed classes were extracted and then further classified into 60 clusters. Then each cluster was examined against the original false color composite image and the high-resolution images (IRS and IKONOS) to determine its land use/cover type. A printed image atlas at a scale of 1:8000 and a declassified general use map published by CIA were also used in this process (CIA, 1982; Shanghai Academy of Surveying & Mapping, 2001). A more detailed account of the imagery classification process is presented in a separate article (Walcott et al., forthcoming).

To ensure that classifications on the northern and southern images are consistent, the overlapped area between the two images was compared quantitatively. Table 3 shows the result of comparison of the pixels in the overlapped area. The differences between the two images are minor for the major classes (built-up surface and vegetation/farmland). The impact on the following analysis should be insignificant since the area represented by the southern image is small (Fig. 1). After the classification, accuracy assessment was conducted by randomly selecting points on the images (Table 4). With an accuracy of 87%, the result is comparable to or better than other land use/cover studies using similar imagery sources (Allen & Kupfer, 2000; De Bruin & Gorte, 2000; Phinn & Stanford, 2001). The major source of error seems to be the barren coastal land being confused with built-up surface. Fig. 5 shows the classified image with the simplified classes of built-up surface and non-built-up surfaces.

Table 3
Comparison between pixels on the northern image (Path 118/Row 38) and southern image (Path 118/Row 39) in the overlapped area

Class name	Class	Pixels in north image	Pixels in south image	Difference (%)
Water/border[a]	0	55,170,474	55,181,931	0.02
Unclassified (cloud)	1	9977	4478	−55.12
Built-up surface	2	387,249	390,955	0.96
Vegetation/farmland	3	1,259,445	1,272,391	1.03
Remaining water	4	202,699	273,200	34.78
Coastal	5	3981	752	−81.11
Water/vegetation mix	6	124,959	114,997	−7.97
Total pixels		57,158,784	57,238,704	

[a] Including pixels outside the overlapped area.

3.2. Relationship between land use and population density

The area of the built-up surface based on the classification of the Landsat imagery was calculated and then divided by the total area for each subdistrict to obtain the proportion of built-up surface (Fig. 6). The proportion of built-up surface has a spatial pattern similar to that of population density (Fig. 2), with values ranging from 0.8 to 1.0 in the central core of the city; then dropping quickly into the surrounding rural areas. The population density and the proportion of built-up surface at the subdistrict level are statistically correlated with a correlation coefficient of 0.8640 (exponential relationship).

Table 4
Error matrix of classification accuracy assessment

Classified data	Reference data							
	Background	Unclassified (clouds)	Built-up surface	Vegetation/ farmland	Remaining water	Coastal	Water/ vegetation	Row total
Water/border	0	0	0	0	0	0	0	0
Unclassified (clouds)	0	50	0	0	0	0	0	50
Built-up surface	0	0	134	4	1	29	11	179
Vegetation/farmland	0	0	3	166	0	2	7	178
Remaining water	0	0	0	1	56	1	3	61
Coastal	0	0	0	0	3	58	0	61
Water/vegetation	0	0	2	4	2	1	46	55
Column total	0	50	139	175	62	91	67	584

Class name	Reference totals	Classified totals	Number correct	Producers accuracy	Users accuracy
Accuracy totals					
Background	0	0	0	–	–
Water/border	50	50	50	100.00%	100.00%
Built-up surface	139	179	134	96.40%	74.86%
Vegetation/farmland	175	178	166	94.86%	93.26%
Remaining water	62	61	56	90.32%	91.80%
Coastal	91	61	58	63.74%	95.08%
Water/vegetation	67	55	46	68.66%	83.64%
Totals	584	584	510		

Overall classification accuracy = 87.33%

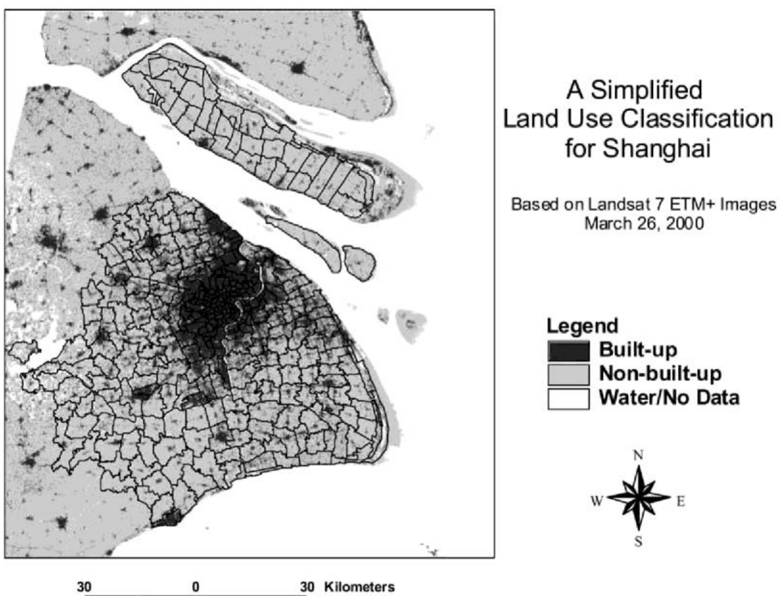


Fig. 5. Land use classification based on the Landsat 7 ETM+ imagery.

We also examined the relationship between the proportion of built-up surface and population density at different scales as represented by the buffers of different sizes surrounding the water monitoring stations. Fig. 7 suggests that the relationships between built-up surface proportion and population density summarized for the buffers become stronger as the buffer size increases. Table 5 contains the results of bivariate regression of built-up surface proportion against population density. The R2 values become higher for larger buffers, partially because population density is calculated from the subdistrict averages, which has a lower spatial resolution than the proportion of built-up surface derived from the satellite images. For smaller buffers, local land use patterns may or may not entirely conform to the pattern of population distribution. For example, non-residential land uses may be a significant portion of a buffer of 100 m along the stream of interest. As the buffer increases, the land use pattern within the buffer becomes more and more similar to the regional land use pattern that is strongly associated with the spatial pattern of population distribution.

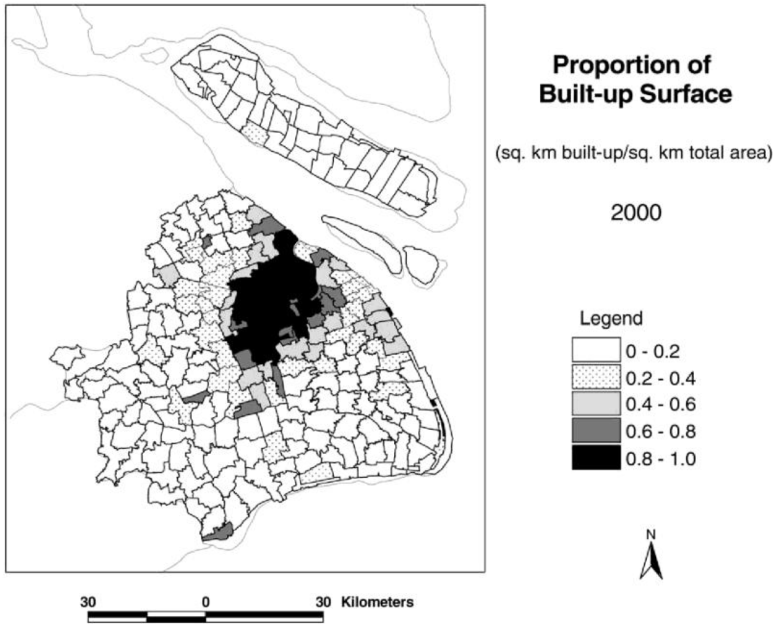


Fig. 6. Proportion of built-up surface at the subdistrict (*Jiedao* and *Xiang*) level.

Shanghai's urban infrastructure development strongly mirrors its population distribution. Before China reformed the housing system in the mid-1990s, state-owned enterprises were responsible for providing housing for their employees. With bicycles and public transit systems as the major mode of transportation, housing developments were located as close to the locations of employment as possible. Even today, the employers are still heavily involved in arranging housing for their employees. This explains the close association between urban land use and population density for larger buffers as well as at the subdistrict level.

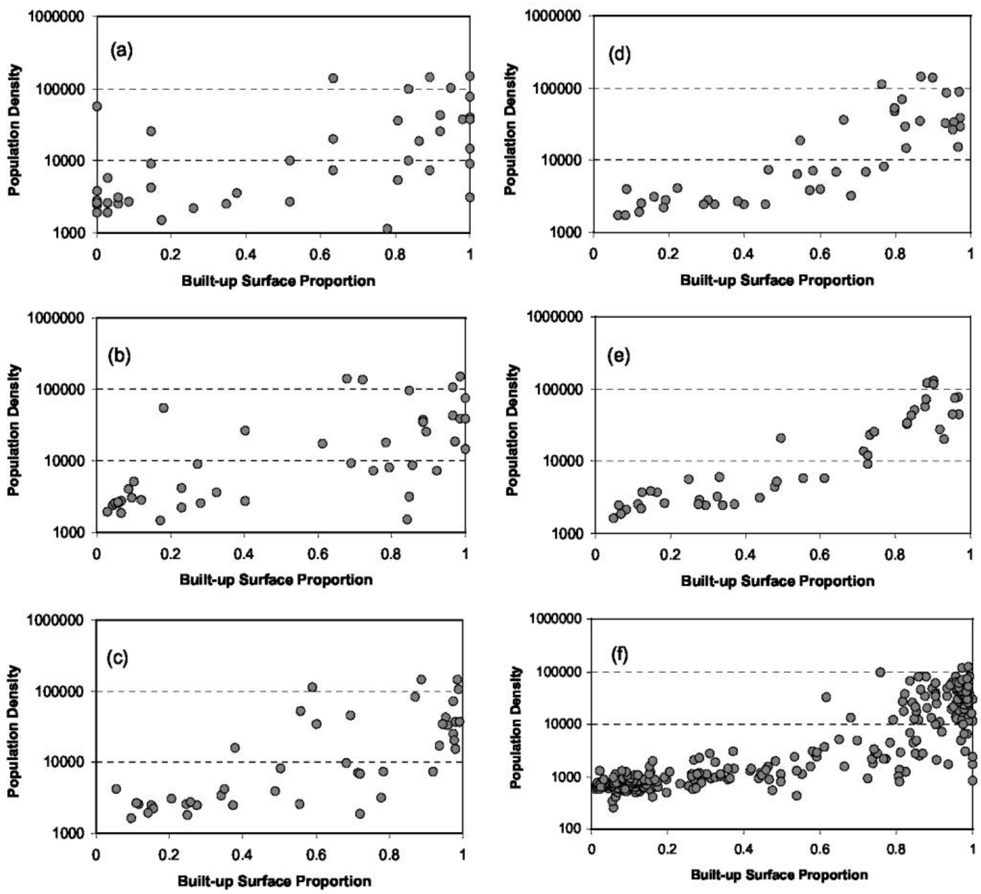


Fig. 7. Relationship between proportion of built-up surface and population density at different scales, represented by buffers of different sizes. (a) 100 m buffers, (b) 200 m buffers (c) 500 m buffers, (d) 1000 m buffers, (e) 2000 m buffers and (f) subdistricts.

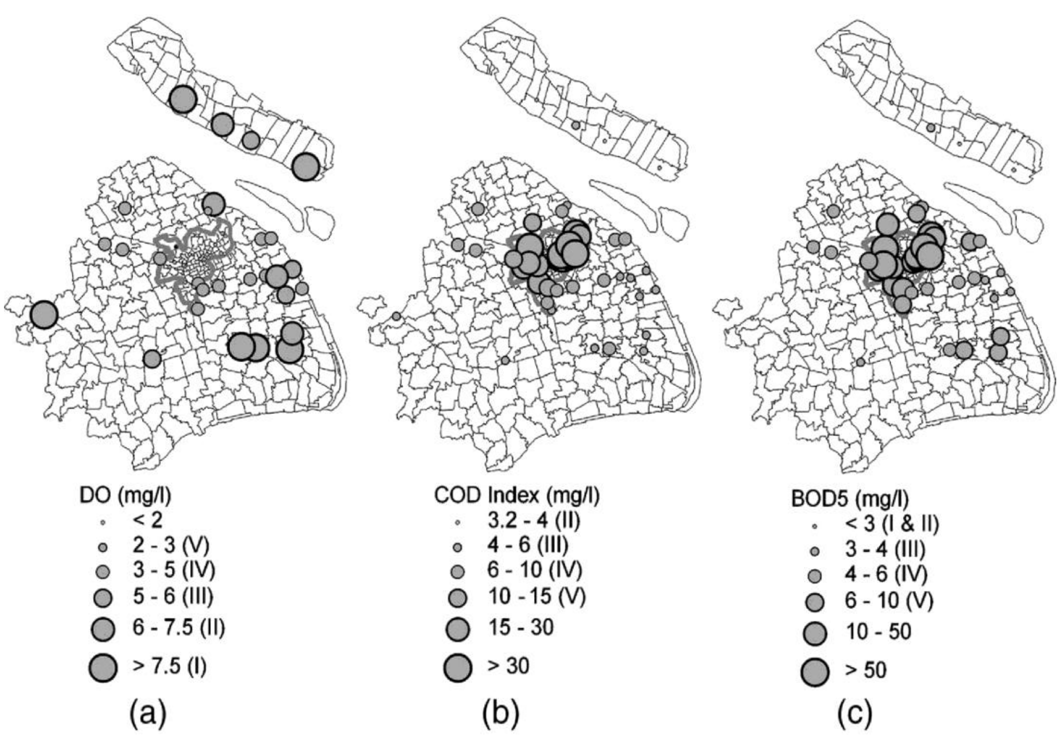
Table 5
Results of bivariate regression of built-up surface proportion against population density for buffers of different sizes

Buffer size	Model	R^2
100	$P = 3212.3e^{2.2002B}$	0.3756
200	$P = 2491.7e^{2.6202B}$	0.4646
500	$P = 1478.5e^{3.2656B}$	0.5603
1000	$P = 1128.4e^{3.8323B}$	0.7127
2000	$P = 1229.2e^{3.9695B}$	0.8457
Subdistrict	$P = 459.26e^{3.7641B}$	0.7465

P = population density; B = built-up surface proportion.

3.3. Spatial pattern of water quality parameters

Fig. 8 shows spatial variation patterns of the water quality parameters. DO has low values for stream segments within the city and high values in the rural area, while all other parameters have higher values in the city and lower values in the rural area. To help interpreting the data, the surface water quality standards (State Environmental Protection Administration, 2002) regarding those parameters are presented in Table 5. Type III water is considered moderately polluted, while Type IV and Type V represent heavily polluted water. The graduated symbol classes in Fig. 8 conform to the water quality standards, except for levels exceeding the limits. Most monitoring stations in 1995 registered levels corresponding to Type IV or V water or worse for almost all water quality parameters examined in this study, indicating heavily polluted water across the study area with the poorest water quality found in the central city area (Fig. 8). It can be argued that the intensity of urban contribution of pollutants much outweighed the agricultural non-point source pollution, as indicated by some of the common measures of non-point pollution (total N, total P, and NH3-N). In fact, the effect of urban pollution obscures the effect of agricultural non-point source pollution (Table 6).



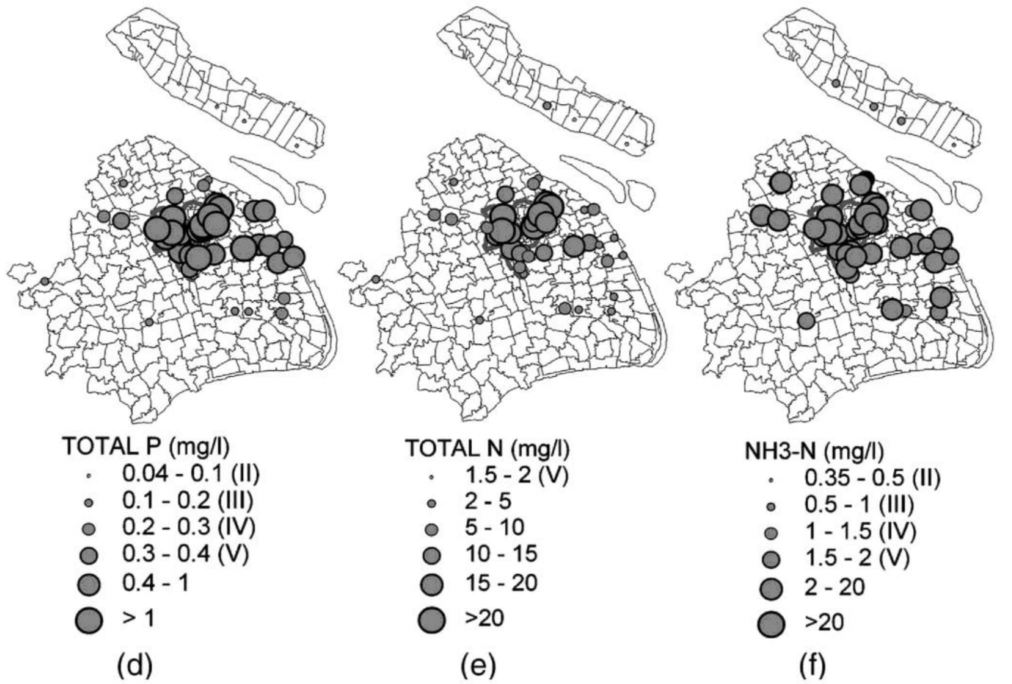


Fig. 8. Spatial patterns of water quality parameters. The unit of total Hg is μg/l on the map, although the original unit is mg/l in the dataset. (a) DO, (b) BOD5, (c) COD Index, (d) NH_3-N, (e) oil, (f) phenol, (g) total P, (h) total N, (i) total Hg.

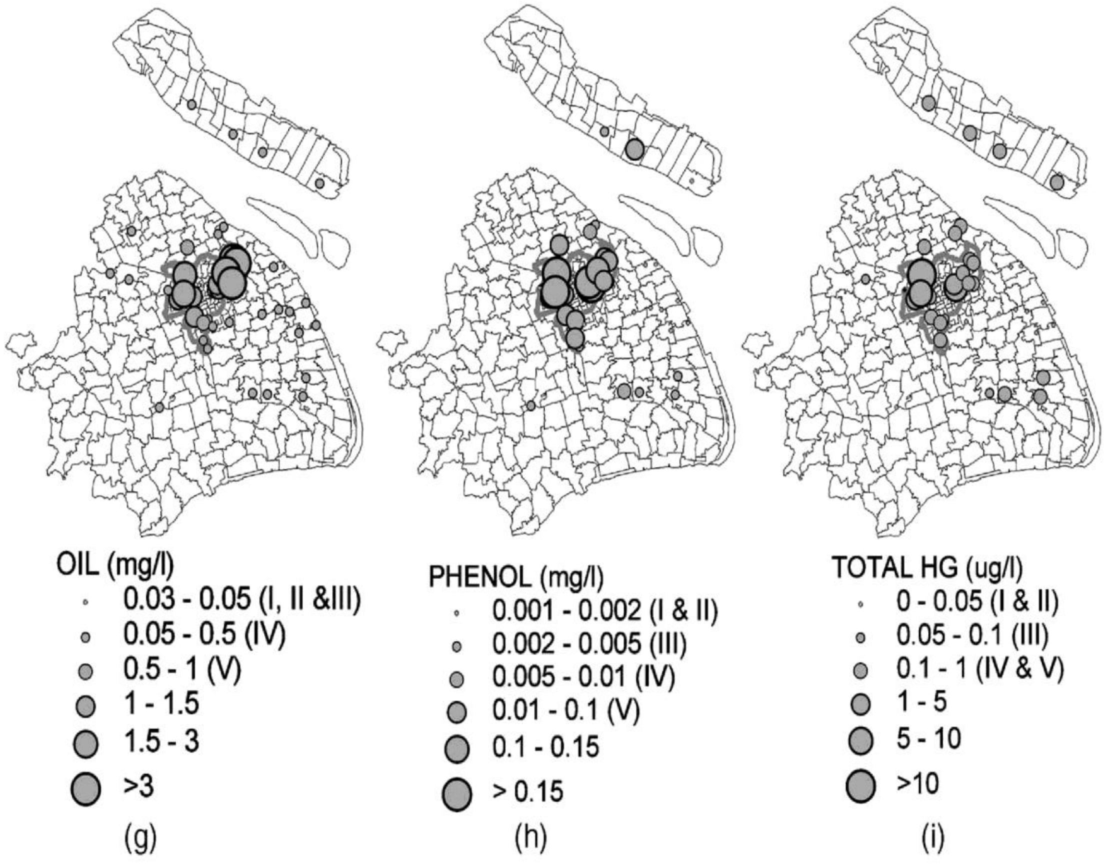


Fig. 8 (*continued*)

By examining Figs. 2, 5 and 8, one would assume that both population density and proportion of built-up surface should be similarly correlated with water quality parameters. This is further illustrated by the distribution of point pollution sources in Shanghai (1995 data available only for the area west and north of the Huangpu River). Fig. 9 shows that the concentration of point pollution sources is very high in the central city, decreasing outwards with a few clusters.

Table 6
Water quality standards (State Environmental Protection Administration, 2002)

Water quality parameter	Standards (for waterways)				
	Type I	Type II	Type III	Type IV	Type V
DO (mg/l)⩾	7.5[a]	6.0	5.0	3.0	2.0
COD Index (mg/l)⩽	2.0	4.0	6.0	10.0	15.0
BOD5 (mg/l)⩽	3.0	3.0	4.0	6.0	10.0
NH_3-N (mg/l)⩽	0.15	0.5	1.0	1.5	2.0
Oil (mg/l)⩽	0.05	0.05	0.05	0.5	1.0
Phenol (mg/l)⩽	0.002	0.002	0.005	0.01	0.1
Total P (mg/l)⩽	0.02	0.1	0.2	0.3	0.4
Total N (mg/l)⩽	0.2	0.5	1.0	1.5	2.0
Total Hg (μg/l)⩽	0.05	0.05	0.1	1.0	1.0

[a] Or 90% saturation.

The Shanghai Municipal Government claims that almost all industrial wastewater is treated before discharging into waterways (Shanghai Statistical Bureau, 2002). Since the late 1970s, major industrial projects are required to include pollution control measures in the design stage, so that the effluents will meet the standards for discharging into waterways at the time when the factory begins its operation—the "Three Synchronicities Policy" (Shen et al., 2002; Ward & Liang, 1995). There are also laws and policies that require Environmental Impact Assessment (EIA) for major projects and the polluters to pay fines if the effluents do not meet the standards—the "Polluter-Pays Principal'' (Shen et al., 2002). However, enforcement of these laws and policies has been difficult and inefficient. For example, over time, as the factory expands, the designed capacity of wastewater treatment may not be updated in many cases. When effluents fail to meet the standards, the fines assessed tend to be so low that the enterprises simply pay the fines and continue the pollution process, rather than upgrade the pollution control measures.

In recent years, there are more and more privately owned small manufacturing enterprises that do not follow the government guidance in pollution control. Above all, the monitoring of point sources takes the form of self-reporting since the local environmental agencies simply do not have the personnel and financial means to monitor all pollution sources. This situation is especially serious in the rural outskirts and along the fringe of the central city (Leman, 2002; Shen et al., 2002). According to Ward and Liang (1995), only one-third of domestic wastewater generated by the city is processed by wastewater treatment plants. The balance is used in agriculture (a diminishing practice), collected by canals and discharged to the Yangtze River or the ocean, or discharged to local waterways. It should also be pointed out that the capacities of wastewater collection and treatment are much lower in the surrounding rural areas. Additionally, over 2000 small shallow landfills together with three major landfills used by the municipal government allow leachate to seep or flow laterally into nearby waterways. All of the above factors explain the spatial patterns of water quality parameters in Shanghai in the mid-1990s and the close associations between water quality, population density, and urban built-up surface.

Results of stepwise regression clearly demonstrate the impact of spatial scale on such relationships. In general, R^2 values increase with buffer size, from a mean of 0.549 for models of 100 m buffers to a mean of 0.642 for models of 2000 m buffers (Table 7). Built-up surface appears to be a better predictor of water quality than population density. For all 45 models (9 water quality parameters and 5 buffer sizes), built-up surface is included in 42 models, while population density and river rank are included in 26 and 21 models, respectively. Additionally, in stepwise selection built-up surface is the first variable to enter 33 models, explaining most variance in water quality parameters. In comparison, population density is the first variable to enter only 12 models. With increasing buffer size, built-up surface becomes more dominant, especially for buffers 1000 and 2000 m in size (Table 7). For smaller buffers, there is a greater tendency for both built-up surface and population density to be included in the models (17 out of 27 models for buffers 100–500 m). For larger buffers (1000–2000 m), two thirds of the models (12 out of 18) contain either variable, but not both, resulting from increased colinearity between the two variables (Fig. 7). Most models are based on exponential relationships, except for DO and TP that have linear models for larger buffers (500–2000 m). The exponential relationships suggest that the impact of urbanization on water quality tends to amplify with an accelerated pattern as urbanization intensifies.

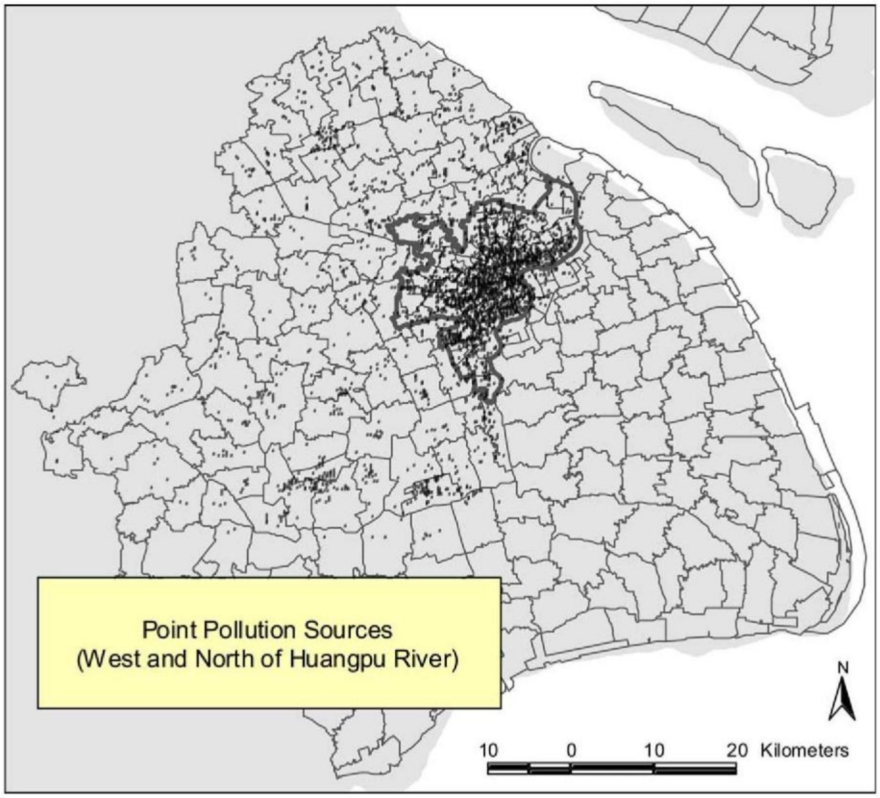


Fig. 9. Distribution of point pollution sources north and west of Huangpu River, 1995 (Source: SAES).

Table 7
Summary of regression analysis for different buffer sizes, with water quality parameters as the dependent variables

Buffer size (m)	Mean R^2	Models that include			First variable to enter	
		B	*R*	*P*	*B*	*P*
100	0.549	7	3	7	6	3
200	0.568	9	4	6	6	3
500	0.564	9	4	6	7	2
1000	0.616	9	5	4	7	2
2000	0.642	8	5	3	7	2
All models	0.588	42	21	26	33	12

The independent variables include proportion of built-up surface (*B*), river rank (*R*), and population density (*P*).

As to individual water quality variables, DO is among the best-explained with high R^2 values for all buffer sizes (Table 8). Models of smaller buffers (100–500 m) contain both built-up surface and population density. Population density is the first variable to enter models of 100–200 m buffers, but built-up surface first enters the rest of the models. River rank is not significant for any of the DO models. COD's models consistently include all three independent variables, but built-up surface is the dominant variable entering first into all models. BOD5's models include all three variables for buffers 100–500 m in size; then population density drops out for the larger buffers. Built-up surface enters first into all models. Of the three water quality parameters that represent nutrient levels, all NH_3-N's models include only built-up surface. The same is true for TN's models of the buffers 200–2000 m in size. For TP, built-up surface and river rank enter the models for buffers 200–2000 m in size. Oil, phenol and THG represent mostly industrial pollution. For smaller buffers (100–500 m), both built-up surface and population density are included in oil's models, while built-up surface and river rank are included in the models for larger buffers. For all phenol's models, all three independent variables enter the models. However, unlike the COD models that also include all three independent variables, population density is the dominant variable for phenol's models. Models for THG have low R^2 values overall, which is partially caused by many observations below the detection limits being registered as half of the detection limits in the data set. Population density is the single predictor for buffers 100–200 m in size. Although built-up surface is included in models of larger buffer sizes, population density remains the first variable to enter all the models of THG. Additionally, THG is the only exception that the best model was found for the 200 m buffers (R^2 = 0.436). All other water quality parameters have their best models found for the largest buffer.

Table 8

Summary of stepwise regression for individual water quality parameters

Water quality parameter	100 m buffers		200 m buffers		500 m buffers		1000 m buffers		2000 m buffers	
	R^2	Variables included	R^2	Variables included	R^2	Variables included	R^2	Variables included	R^2	Variables included
DO	0.610	*P*, *B*	0.605	*P*, *B*	0.628*	*B*, *P*	0.709*	*B*	0.776*	*B*
COD Index	0.685	*B*, *P*, *R*	0.692	*B*, *P*, *R*	0.669	*B*, *P*, *R*	0.684	*B*, *P*, *R*	0.723	*B*, *R*, *P*
BOD5	0.659	*B*, *P*, *R*	0.682	*B*, *P*, *R*	0.666	*B*, *P*, *R*	0.691	*B*, *R*, *P*	0.691	*B*, *R*
NH_3-N	0.465	*B*	0.495	*B*	0.517	*B*	0.631	*B*	0.658	*B*
Oil	0.522	*B*, *P*	0.561	*B*, *P*	0.545	*B*, *P*	0.579	*B*, *R*	0.611	*B*, *R*
Phenol	0.592	*P*, *R*	0.633	*P*, *B*, *R*	0.609	*P*, *R*, *B*	0.629	*P*, *R*, *B*	0.653	*P*, *R*, *B*
TP	0.523	*B*	0.520*	*B*, *R*	0.538*	*B*, *R*	0.607*	*B*, *R*	0.637*	*B*, *R*
TN	0.511	*B*, *P*	0.490	*B*	0.507	*B*	0.611	*B*	0.659	*B*
THG	0.371	*P*	0.436	*P*, *B*	0.397	*P*, *B*	0.404	*P*, *B*	0.374	*P*

R^2 values marked with "*" indicate linear relationships, while all others are exponential relationships. Variables included in the models are represented by *B* for built-up surface, *P* for population density, and *R* for river rank, listed according to the order these variables entered the models in the stepwise selection process.

The results found above suggest that the major factors of the spatial pattern of water quality have a regional or diffuse nature. This also supports the earlier assumption that water quality monitored at a given cross-section of the waterways actually represents the effect of all pollution sources in a large area rather than the area immediately surrounding the monitoring site. Although it would have been optimum if the size of buffers could be further increased, the radius of 2000 m seems to have reached the limit for data independence since at this scale, many buffer zones begin to overlap significantly. Table 9 contains the best regression models of the water quality parameters, most of them for the buffers of 2000 m radius. Built-up surface proportion is included in all models. Only three models contain both built-up surface proportion and population density due to high levels of colinearity for large buffers. River rank is included in 5 out of the 9 models.

Table 9

Best regression models for the water quality parameters

Parameters	Models	R^2	Buffer size (m)
DO	$7.925–7.708 * B$	0.776	2000
COD Index	$3.071e^{(1.248*B+0.171*R+0.00000709*P)}$	0.723	2000
BOD5	$1.660e^{(2.633*B+0.254*R)}$	0.691	2000[a]
NH_3-N	$0.876e^{(3.019*B)}$	0.658	2000
Oil	$0.0646e^{(2.718*B+0.224*R)}$	0.611	2000
Phenol	$0.000812e^{(0.0000255*P+0.383*R+1.904*B)}$	0.653	2000
TP	$-0.269 + 1.431 * B + 0.211 * R$	0.637	2000
TN	$2.321e^{(2.173*B)}$	0.659	2000
THG	$0.0000703e^{(0.0000170*P+1.249*B)}$	0.436	200

All independent variables included are statistically significant at 0.05.

B—built-up surface proportion; *P*—population density; *R*—river rank.

[a] For BOD5, the R^2 values of the models for buffers 1000 and 2000 m are the same.

Tables 8 and 9 show that built-up surface proportion entered more models than population density. Both variables are used as the measure of the intensity of urbanization. Since built-up surface is obtained from satellite images of 30 m resolution, it should represent mostly local land use patterns, especially for the smaller buffers. Population density, on the other hand, is derived from the census data at the subdistrict level of a much coarser spatial resolution, and therefore, represents the intensity of urbanization from a regional perspective or the background condition for water quality at a specific cross-section. For smaller buffers, there is a better chance that both built-up surface proportion and population density are included in the model because these two variables tend to be more different in nature. At these scales, water quality is influenced by both regional land use patterns as well as local conditions. This signifies the need to manage water environment through local planning measures, such as maintaining greenspace along riverbanks as buffer zones and removing pollution sources within vicinity of waterways. As stated earlier, with increasing buffer size, the land use pattern within the buffers more and more approaches that of the regional scale. As the home of 13 million people, residential land uses dominate large portions of the total built-up surface across the city. Therefore, the built-up surface proportion within the larger buffers will acquire the characteristics of regional pattern of population distribution and replace the effect of population density in regression analysis. Regional planning measures that can be effective in water quality management may include redistribution/relocation of population and polluting industries, increasing regional wastewater treatment capacity, and implementing land use policies that limit further expansion of the city.

A recent study pointed out that in the foreseeable future, the conditions for housing, social services, public health, greenspace, and drinking water will be substantially improved for Shanghai, while the air and water quality will continue to deteriorate into the following decades (Ying & Kung, 2000). As Shanghai's population continues to grow, with intensive residential housing construction continuing into the foreseeable future and infrastructure lagging behind demands, the quantitative relationships between built-up surface, population density, and water quality as defined in this study should remain strong. It is noted that the mismatch of the dates of image acquisition and water quality data may have introduced errors to the empirical models in this study. However, results from this study indicate strong linkages between water quality and land use patterns and the nature of these linkages should not change substantially in the near future. It also calls the need for further studies to investigate possible changes in such relationships between water quality and land use patterns due to implementation of pollution control measures and changes in functionality of land uses at regional scales.

4. Conclusion

In this study, land use/cover classification was conducted using Landsat 7 ETM+ images for Shanghai, the largest city and industrial center of China. The focus was the land used for urban development, categorically defined as built-up surface. It was found that the proportion of built-up surface and population density were closely associated spatially, and therefore, satellite images can provide an excellent measure of the intensity of urbanization. In regression analysis, the proportion of built-up surface appeared to be a better predictor of water quality than population density. Regression analysis indicated that built-up surface and/or population density can explain 60–70% or more of the variances for most water quality parameters examined. As the buffer size increased, the model performance improved for most water quality parameters.

Shanghai has gone through a similar path as many developing or newly developed countries, focusing on economic development while ignoring environmental degradation. Lessons learned from the past indicate that this can be a costly process in the long run (Chakrabarti, 2001; Chi, 1994; Kabala, 1985; Reich, 1983; Takahasi, 2001). The Shanghai Municipal Government has realized that to achieve the goal of transforming Shanghai into a "global city", Shanghai must improve its image as a heavily polluted city on the list of the World Health Organization (Yusuf & Wu, 2002). Investment in Shanghai's urban infrastructure skyrocketed since the early 1990s (Ning, 1999). Environment-related investment amounted to 3% of the total GDP by 2001 (Shanghai Municipal Government, 2002), much higher than the national goal of 1.0– 1.5% of the total GDP (Shen et al., 2002). The government has defined the future for the city as a global financial center with preferential development of high value-added manufacturing, such as hi-tech and automotive industries. Many old, low value-added and polluting industries in the central city core are either relocated or closed (Leman, 2002). Between 1991 and 1998, 12,000 enterprises along with 400,000 households were relocated from the central city to the suburbs and the rural outskirts as the result of industrial restructuring, new developments for housing, and the reconstruction of the central core (Yusuf & Wu, 2002). By 2020, more than a dozen new wastewater treatment plants will be completed. With expanded capacity of the current 31 wastewater treatment plants, the total capacity will reach 5.74 million tons/day (Shanghai EPB et al., 2001). This represents more than 5 times of the current capacity and will almost meet the demand of the projected domestic wastewater discharge increases. These measures will certainly change the spatial pattern of water quality in and around Shanghai. By then the relationships between built-up surface, population density and water quality will be much different from those examined in this study, with new challenges of pollution control mostly in the suburban and exurban towns and the rural outskirts of the city. However, under the current condition, it is possible that the relationships derived in this study can be used to predict regional pattern of water quality in the near future for areas in Shanghai with no or few water quality monitoring data, since it is relatively easy to obtain updated satellite images and determine the spatial distribution of built-up surface using the methods outlined in this study. These relationships also provide insight into planning and treatment measures to alleviate water pollution problems and offer the basis for future comparisons to determine the trend of improvement or deterioration of the water environment in Shanghai.

References

Allen, T., & Kupfer, J. (2000). Application of spherical statistics to change vector analysis of Landsat data: Southern Appalachian spruce-fir forests. Remote Sensing of Environment, 74, 482–493.

Asian Development Bank (ADB) (1999). Project Profiles by County—People's Republic of China. Available: http://www.adb.org/Documents/Profiles/LOAN/32121013.ASP.

Binns, T. (1991). Shanghai's Pudong development project. Geography, 76, 362–365.

Cao, J. (2002). Personal communication. Mr. Cao is a senior engineer of SAES and a co-author of this article.

Central Intelligence Agency (CIA) (1982). Shanghai. Paper map, original scale 1:30,000. Washington, DC: CIA.

Chakrabarti, P. G. D. (2001). Urban crisis in India: new initiatives for sustainable cities. Development in Practice, 11, 260–272.

Chen, S., Zeng, S., & Xie, C. (2000). Remote sensing and GIS for urban growth analysis in China. Photogrammetric Engineering and Remote Sensing, 66, 593–598.

Chi, C.-C. (1994). Growth with pollution: unsustainable development in Taiwan and its consequences. Studies in Comparative International Development, 29, 23–47.

De Bruin, S., & Gorte, B. G. H. (2000). Probabilistic image classification using geological map units applied to land-cover change detection. International Journal of Remote Sensing, 21, 2389–2402.

Douglass, M. (2000). Mega-urban regions and world city formation: globalization, the economic crisis and urban policy issues in Pacific Asia. Urban Studies, 37, 2315–2335.

He, X. (1993). Development of Pudong and optimization of urban area structure in Shanghai. Chinese Environment and Development, 4, 68–88.

Ji, C. Y., Lin, P., Li, X., Lui, Q., Sun, D., & Wang, S. (2001). Monitoring urban expansion with remote sensing in China. International Journal of Remote Sensing, 22, 1441–1455.

Kabala, S. J. (1985). Poland facing the hidden cost of development. Environment, 27, 6–13, and 37–42. Kung, H. T., & Ying, L. G. (1991). A study of lake eutrophication in Shanghai. China Geographical Journal, 157, 45–50.

Leman, E. (2002). Can Shanghai compete as a global city? The China Business Review, 2002/September–October, pp. 7–15.

Li, X., & Yeh, A. G. O. (1998). Principal component analysis of stacked multi-temporal images for the monitoring of rapid urban expansion in the Pearl River Delta. International Journal of Remote Sensing, 19, 1501–1518.

Mesev, V. (1998). The use of census data in urban image classification. Photogrammetric Engineering and Remote Sensing, 64, 431–438.

Ning, Y. M. (1999). City planning and urban construction in the Shanghai metropolitan area. In H. Foster, D. Lai, & N. Zhou (Eds.), The Dragon's Head: Shanghai, China's Emerging Megacity (pp. 229– 246). Victoria, Canada: Western Geographical Press.

Phinn, S., & Stanford, M. (2001). Monitoring land-cover and land-use change in a rapidly urbanizing coastal environment: the Maroochy and Mooloolah Rivers catchments, southeast Queensland, 1988– 1997. Australian Geographical Studies, 39, 217–232.

Reich, M. R. (1983). Environmental policy and Japanese society: Part II. Lessons about Japan and about Polity. International Journal of Environmental Studies, 20, 199–207.

Shanghai Academy of Environmental Sciences (SAES) (1998). Report on comprehensive treatment of Huangpu River Water Environment (Huangpujiang Shui Huanjing Zonghe Zhengzhi Baogao). Shanghai, China: SAES (in Chinese).

Shanghai Academy of Surveying and Mapping (2001). Image atlas of Shanghai (Shanghai Yingxiang Dituji). Shanghai, China: Shanghai Publishing House of Science and Technology (in Chinese).

Shanghai Environmental Protection Bureau (EPB), Shanghai Water Affairs Bureau, and Shanghai Urban Planning and Management Bureau (2001). Shanghai Water Environment Management and Protection Planning and the 10th 5-Year Plan. Shanghai, China: Shanghai EPB (in Chinese).
Shanghai Municipal Government (2002). Shanghai in the next 5 years. Available: www.sh.gov.cn. Shanghai Statistical Bureau (2002). The 2001 Year Book, online edition. Available: http://www.statssh.-gov.cn/shtj/tjnj/2001/tjnj2001.htm.
Shen, L., Cheng, S., Gu, S., & Lu, Y. (2002). Environmental policy and law for sustainable natural resources development—issues and challenges. Environmental Policy and Law, 32, 91–98.
Shi, Y., & Hamnett, C. (2002). The potential and prospect for global cities in China: in the context of the world system. Geoforum, 33, 121–135.
Slonecker, E. T., Jennings, D. B., & Garofalo, D. (2001). Remote sensing of impervious surfaces: a review. Remote Sensing Review, 20, 227–235.
State Environmental Protection Administration (2002). Environmental quality standards for surface water. Available: http://www.china-eia.com/LAW/biaozhun/gb3838-2002.htm (in Chinese).
Takahasi, Y. (2001). Towards a solution to the North–South problem. Water Resources Development, 17, 555–561.
US Geological Survey (USGS) (2002). Multi-resolution land characteristics (MRLC). Available: http://edc.usgs.gov/glis/hyper/guide/mrlc.
Volgelman, J. E., Sohl, S. M., Howard, S. M., & Shaw, D. M. (1998). Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. Environmental Monitoring and Assessment, 51, 415–428.
Walcott, S., Yin, Z.-Y., Kaplan, B., Lin, W., Cao, J., & Chen, M. (forthcoming). Assessing environmental impacts of urbanization using multi-sensor data: Shanghai, 1970–2000 (Land use change in Shanghai). Asian Geographer.
Ward, D., Phinn, S. R., & Murray, A. T. (2000). Monitoring growth in rapidly urbanizing areas using remotely sensed data. Professional Geographer, 52, 371–386.
Ward, R. M., & Liang, W. (1995). Shanghai water supply and wastewater disposal. Geographical Review, 85, 141–156.
Weng, Q. (2001). A remote sensing-GIS evaluation of urban expansion and its impact on surface temperature in the Zhujiang Delta, China. International Journal of Remote Sensing, 22, 1999– 2014.
Wu, G., & Shi, C. (1999). Shanghai's water environment. In H. Foster, D. Lai, & N. Zhou (Eds.), The Dragon's Head: Shanghai, China's Emerging Megacity (pp. 93–104). Victoria, Canada: Western Geographical Press.
Ying, L. G., & Kung, H. T. (2000). Forecasting up to year 2000 on Shanghai's environmental quality. Environmental Monitoring and Assessment, 63, 297–312.
Yusuf, S., & Wu, W. (2002). Pathways to a world city: Shanghai rising in an era of globalisation. Urban Studies, 39, 1213–1240.
Zhou, Y., & Ma, L. (2000). Economic restructuring and suburbanization in China. Urban Geography, 21, 205–236.