VISUALIZATION OF SPATIAL DISTRIBUTION AND TEMPORAL CHANGE OF LAND PRICES FOR RESIDENTIAL USE IN TOKYO 23 WARDS USING SPATIO-TEMPORAL KRIGING

Ryo INOUE
Assistant Professor
Department of Civil Engineering
School of Engineering
University of Tokyo
7-3-1, Hongo, Bunkyo
Tokyo 113-8656
Japan
Tel: +81 3 5841 6129
Fax: +81 3 5841 7453
E-mail: rinoue@civil.t.u-tokyo.ac.jp

Naoyuki KIGOSHI
Chief Official
Ship Machinery Industries Division
Maritime Bureau
Ministry of Land, Infrastructure and Transport
2-1-3, Kasumigaseki, Chiyoda
Tokyo 100-8918
Japan
Tel: +81 3 5253 8111
E-mail: kigoshi-n22aa@hkd.mlit.go.jp

Eihan SHIMIZU
Professor
Department of Civil Engineering
School of Engineering
University of Tokyo
7-3-1, Hongo, Bunkyo
Tokyo 113-8656
Japan
Tel: +81 3 5841 6126
Fax: +81 3 5841 7453
E-mail: shimizu@civil.t.u-tokyo.ac.jp

Abstract: It has been pointed out that the disclosure of the detailed land price data is significant in improving transparency in the real estate market. Although the officially published land prices had played a major role in price making in Japan, and transaction prices will do the same in the future, they do not provide enough information for all market participants. Thus, it is necessary to estimate the land prices of interested land lots using the observed land price data, which is distributed spatially and temporally. In this study, we focus on the recently proposed spatio-temporal stationary covariance function, which models the correlation through space and time. We empirically analyze the applicability of the spatio-temporal covariance function with universal kriging to land price estimation by using the officially published residential land prices in Tokyo 23 wards, and visually illustrate the land price distribution and its change in Tokyo in 30 years.

Keywords: spatio-temporal stationary covariance function, universal kriging, land price, visualization

1. INTRODUCTION

There is an increasing awareness that the disclosure of the detailed land price data is significant to establish land markets with a high level of transparency and fairness, since the national government of Japan promotes the realization of “effective land use” through market mechanisms.

In Japan, the officially assessed land prices by the Ministry of Land, Infrastructure and Transport had played a major role in land price making. In the near future, transaction prices, which have been open to the public since 2005, will play an equally important role. However, they do not provide enough information for all market participants, since the assessed land price data is limited in number and is subject to the budget constraint; in addition, the transaction land price data would be unevenly distributed in space and time. Thus, it is necessary to estimate the land prices of interested land lots using the land price data, which is distributed spatially and temporally.

Recently, spatio-temporal stationary covariance functions, which model the correlation through space and time, have been proposed (Cressie and Huang, 1999; Gneiting, 2002). They are the extension of spatial covariance functions that have been discussed in geostatistics and used for spatial interpolation (or prediction). These spatio-temporal stationary covariance functions have been applied to environment datasets such as wind data and ozone density by the ordinary kriging, and their applicability has been evaluated.

However, when applying them to land price estimation, the ordinary kriging, which does not use explanatory variables, would not function adequately. The reason is that socioeconomic data such as land price data is highly dependent on other socioeconomic variables. For instance, the land prices would be largely influenced by economic conditions, accessibility, and zone restrictions. Thus, it is difficult to estimate land prices at unobserved land lots from the stationary covariance function on only spatial distance and time.

In this study, we apply the universal kriging, which uses the explanatory variables such as accessibility, lot area, and floor space index, and assume the spatio-temporal stationarity of covariance on the error terms of the land price model. We empirically analyze the applicability of the spatio-temporal stationary covariance functions for land price estimation using the residential land price data in Tokyo 23 wards in 30 years, and visualize the land price change in Tokyo using the estimated result.

2. SPATIO-TEMPORAL KRIGING

Kriging is the spatial interpolation or prediction method that has been discussed in geostatistics (e.g., Cressie (1993)). It assumes that the spatial data is a result of observations on the stochastic process, and the stochastic process has the second-order stationarity over space. Then, kriging provides the best linear estimators that depend on the stochastic process at any location. The process is usually expressed by the covariogram, or the spatial stationary covariance function.

An attempt is made to extend the spatial stationary covariance function to the spatio-temporal covariance function. The difficulty in expansion is that the covariance function must be positive-definite. Then, separable spatio-temporal covariance functions, which are a product of spatial covariance function and temporal covariance function, have been proposed; however, only the simple covariance structure was able to be modeled, not the space-time interaction.

Then, Cressie and Huang (1999) and Gneiting (2002) proposed non-separable stationary covariance functions, which were positive-definite and could consider the interaction between spatial and temporal covariance. They also proposed several examples of covariance functions; there are several examples of the application of these spatio-temporal stationary covariance functions to environmental dataset. Cressie and Huang (1999) applied the covariance functions the wind speed data, and Huang and Hsu (2004) applied them to the ground-level ozone concentration data.

However, to our knowledge, spatio-temporal stationary covariance functions have never been applied to socioeconomic data, which cannot be explained only by spatial distances and time interval.

In this study, we apply the spatio-temporal covariance function to the land price estimation, and evaluate its applicability to the socioeconomic dataset in the next section.

3. APPLICATION OF SPATIO-TEMPORAL STATIONARY COVARIANCE FUNCTION TO RESIDENTIAL LAND PRICE DATA IN TOKYO

In this section, we apply a spatio-temporal stationary covariance function to residential land price data in Tokyo. First, we describe the dataset and the land price model in this study, and then present the result of parameter estimation and the validation of land price estimation.

3.1 Dataset and land price model

The explained variable in this application is the residential land price in Tokyo 23 wards for 30 years, from 1975 to 2004. The data source is “the officially assessed residential land prices” by the Ministry of Land, Infrastructure and Transport of Japan, and is distributed as “Time-series dataset of officially assessed land prices” from the Land Information Center in Japan.

In this study, the definition of “residential land” is that the land lot is located inside residential zones that are designated by city-planning law. Note, however, that “residential land” does not mean that the building on the land lot is a residence. The number of land price observation (or assessed) points differ by year: 1,003 points in 2004 and 25,525 in 30 years in total. Figure 1 illustrates the spatial distribution of observation points in 2004. It can be noted that there are not too many observation points in the center of the study area and waterfront areas. These areas are designated as business and industrial zones; note that the result of the residential land price estimation shown later need not necessarily be in agreement with the actual land price in these areas.

Several attributes of officially assessed land price dataset are utilized as explanatory variables of the residential land price model; they are the names of the nearest train stations, distance to the nearest station (in meters), floor area ratio (in percentages), area of land lot (in square meters), and width of the front road (in meters).

There are two additional explanatory variables for land price estimation. One is an accessibility index. We define the accessibility index for each train station according to the travel time taken from the major stations near CBD. The distance to the nearest station and the accessibility index of the nearest train station express the accessibility of each land lot in the land price model.

The accessibility index of each train station is set by the weighted average travel time data to the major five stations. Figure 2 shows the alignment of train network and the location of train stations. The five major stations near CBD are selected by the number of passengers according to the 1995 Transportation Census of Urban Cities, by the Ministry of Land, Infrastructure and Transport. The major stations are Shinjuku, Ikebukuro, Tokyo, Shibuya, and Ueno, and are identified as red points in Figure 2. The travel time to the five major stations are weighted by the number of passengers.

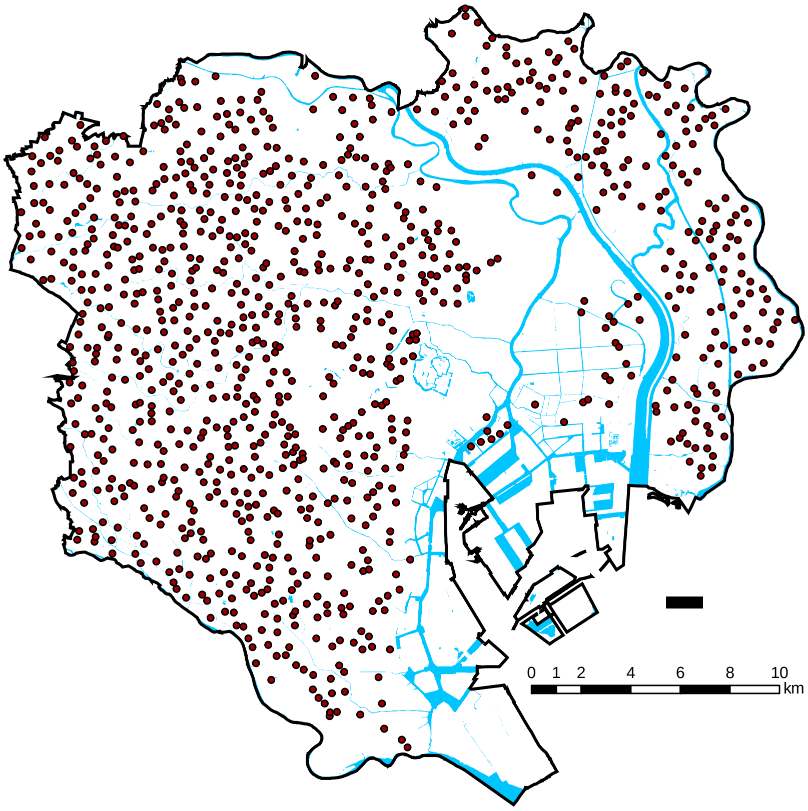
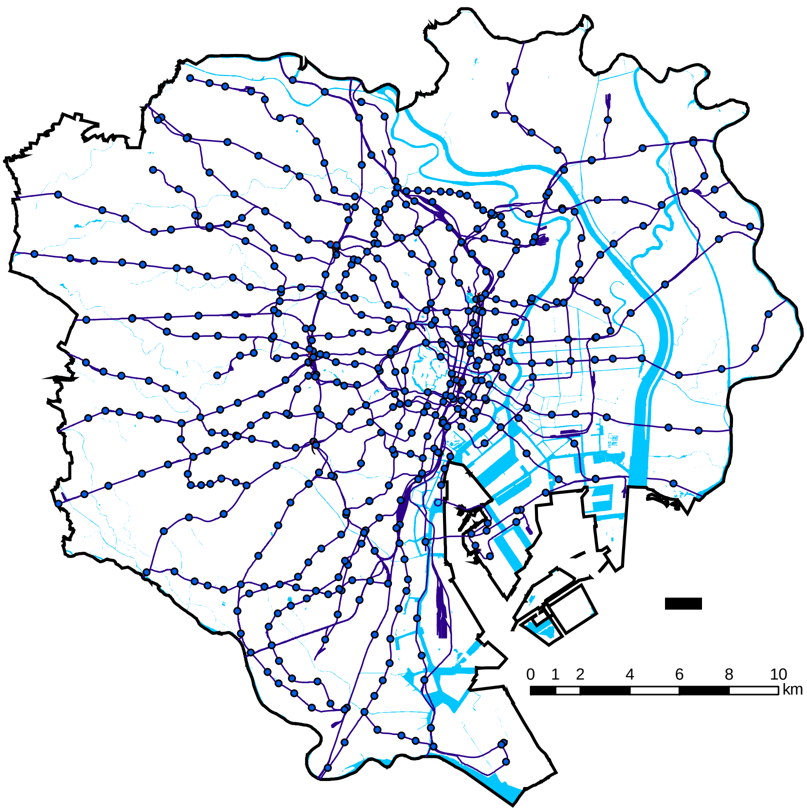


Figure 1 Residential Land Price Observation Point in 2004



Blue points: Train stations; Red points: Major train stations

Figure 2 Railway Network and Stations in the Study Area

The travel time data is obtained from the train route navigation software “EKI-spert (expert)” by VAL Laboratory Corporation, by inputting the railway network data of each year from 1975 to 2004.

Another explanatory variable is the yearly average of the Nikkei 225 stock average. This data is used to explain the economic environment in Japan.

The residential land price model in this study is the following:

$$\ln\left(p_i\right)=\beta_0+\sum_{k=1}^{5}\beta_k\ln\left(x_{ik}\right)+\beta_6 x_{i6}+\beta_7\ln\left(x_7\right)+\varepsilon_i \tag{1}$$

where p_i denotes the land price per square meters at land lot i (JPY/square meters), x_{i1} is the accessibility index of the nearest train station from land lot i (minutes), x_{i2} is the distance to the nearest station from land lot i (meters), x_{i3} is the floor area ratio (%) at land lot i, x_{i4} is the area of land lot i (square meters), x_{i5} is the width of the front road of land lot i (meters), x_{i6} is the dummy variable for sewerage at land lot i, x_7 is the yearly average of the Nikkei 225 stock average (JPY), β_i is the parameters, and ε_i is the error term.

The spatio-temporal covariance function is written as follows:

$$C(\mathbf{h};u)=\sigma^2\left(\frac{2}{(a^2u^2+1)^{3/2}}\right)\exp\left\{-b\|\mathbf{h}\|\right\} \tag{2}$$

where **h** is the spatial vector between two observation points, u is the time between observations, and a, b, and σ^2 are the parameters (Cressie and Huang, 1999). We set the range as 15 kilometers in space and 15 years in time.

The estimation of parameters by spatio-temporal kriging in this study is as follows. First, βs, the parameters of land price model, are estimated by ordinary least squares (OLS). Then, a, b, and σ^2, the parameters of spatio-temporal stationary covariance function, are estimated using the residual of land price model by the weighted least squares criterion (Cressie, 1993). After estimating the covariance function, we estimate the parameters of land price model using the generalized least squares method and inputting the estimated covariance structure. We iterate the estimation of the land price model and the spatio-temporal stationary covariance function and stop the iteration when the parameters converge.

Since the estimation process requires inverting the large matrix, which size is the number of observations by the number of observations, large computation power is required. For this, we used a super computer at the University of Tokyo for the estimation.

3.2 Result of Parameter Estimation

We present the parameter estimation result in Table 1. The right-hand side of the table shows the result of parameter estimation of land price model by OLS for comparison. Note that the t-value of the spatio-temporal kriging is calculated on the assumption that the variance covariance matrix of error terms is given. However, the variance covariance matrix is estimated using the spatio-temporal stationary covariance function; in fact, the real t-value should be smaller than that shown in Table 1.

Table 1 Result of Parameter Estimation

Variables	Spatio-temporal Kriging		OLS	
	Coefficient	t-value	Coefficient	t-value
Constant	11.42	17.0	5.88	78.2
Accessibility	–0.0399	–4.09	–0.673	–70.5
Distance to nearest station	–0.0714	–25.3	–0.190	–48.0
Floor area ratio	0.0340	8.91	0.0247	4.62
Area of land lot	0.0514	19.2	0.244	50.5
Width of front road	0.149	37.8	0.204	26.7
Dummy variable for sewerage	0.00914	3.11	0.0968	12.9
Nikkei 225 stock average	0.0310	5.65	0.931	211
σ^2	0.120	-	-	-
a	0.0842	-	-	-
b	0.120	-	-	-

Table 1 reveals that the coefficients of “accessibility” and “distance to the nearest station” in the spatio-temporal kriging model are smaller than those in the OLS estimation. These explanatory variables have strong spatial autocorrelation and would explain some of the parts of spatial autocorrelation of land price data. However, since the spatial correlation between the error terms of the land price model could be explained by the spatio-temporal covariance function, coefficients of these explanatory variables would be smaller. The same thing occurs with the coefficient of the Nikkei 225 stock average. This variable explains the time-series correlation of land price in the OLS estimation. However, once the time-series correlation is structured by the covariance function, the coefficient of the Nikkei 225 drops.

3.3 Validation of Estimated Land Prices

Here, we show the validation result of the estimated land price data.

For the purpose of validation, first, we randomly choose 80% of the observations from each year’s land price data and estimate the parameters of the model. Then, we estimate the land prices at the remaining 20% of the observation points and compare the estimated value with the observation data. The process from random sampling to land price estimation is repeated five times.

Figure 3 illustrates the validation result of land price estimation by spatio-temporal kriging by five times of random samplings. The vertical axis represents the exponential of root mean squared errors (RMSE) between the natural logarithm of estimated land prices and the natural logarithm of observed land prices. When the value of the vertical axis—exponential (RMSE)—is 1.1, it indicates that the estimation error is 10% on average. The validation result of the land price estimation by OLS is also shown in Figure 3 for comparison.

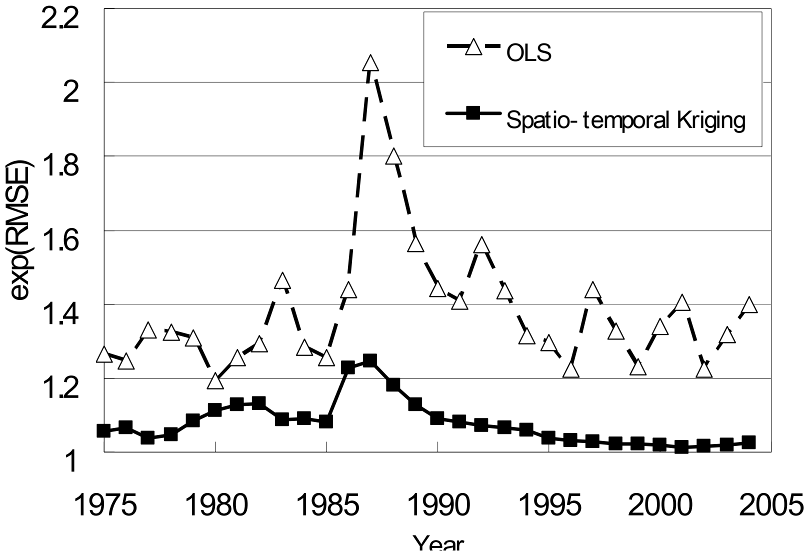


Figure 3 Validation of Estimated Land Prices

It is apparent that land prices can be estimated in high precision with the spatio-temporal kriging. Even in the late 1980s, when land prices were raised sharply, the estimation error of spatio-temporal kriging is approximately 20%. The estimation error is around 10%, except during the bubble economy period between 1986 and 1991, and the estimation error is almost 2% in later years when the land price move calms.

As mentioned above, it is possible to state that the spatio-temporal kriging enables a highly accurate estimation of land prices.

4. VISUALIZATION OF SPATIAL DISTRIBUTION AND CHANGE OF RESIDENTIAL LAND PRICE DATA IN TOKYO

We estimated the residential land price of every city block in the Tokyo 23 wards from 1975 to 2004. The centroid of city blocks is calculated from the city block data of “the Digital Map 2500 (Spatial Data Framework)” distributed by the Geographical Survey Institute of Japan. Then, we estimated the land price at the centroid of each city block by the spatio-temporal kriging using the estimation result shown in Table 1. The Tokyo 23 wards have a total number of 107,164 city blocks.

Since there is no available data for most of the explanatory variables in every city block, it is required to set values. The values of “floor area ratio,” “area of land lot,” and “width of front road” are calculated from the officially assessed land price dataset using the spatial ordinary kriging by each year.

The “distance to the nearest station” is calculated by the software package SANET (Okabe (2003)) using the road line network data and railway station data of “the Digital Map 2500 (Spatial Data Framework).” SANET outputs the network distance between the centroids of city blocks and railway stations.

Figure 4 is an example of land price estimation result. It is possible for the spatio-temporal kriging to output graphs that represent land price change for every city block. Figure 5 represents the spatial distribution of residential land prices in the Bunkyo ward. It is clear that the land price estimation can visualize the spatial distribution of land price in a small area.

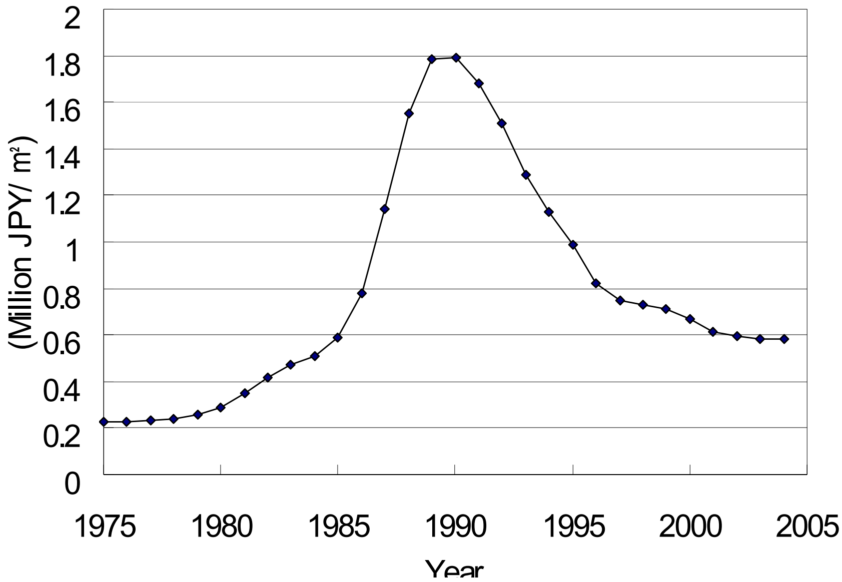


Figure 4 An Example of Land Price Estimation
Address: 2-4-47, Hakusan, Bunkyo, Tokyo

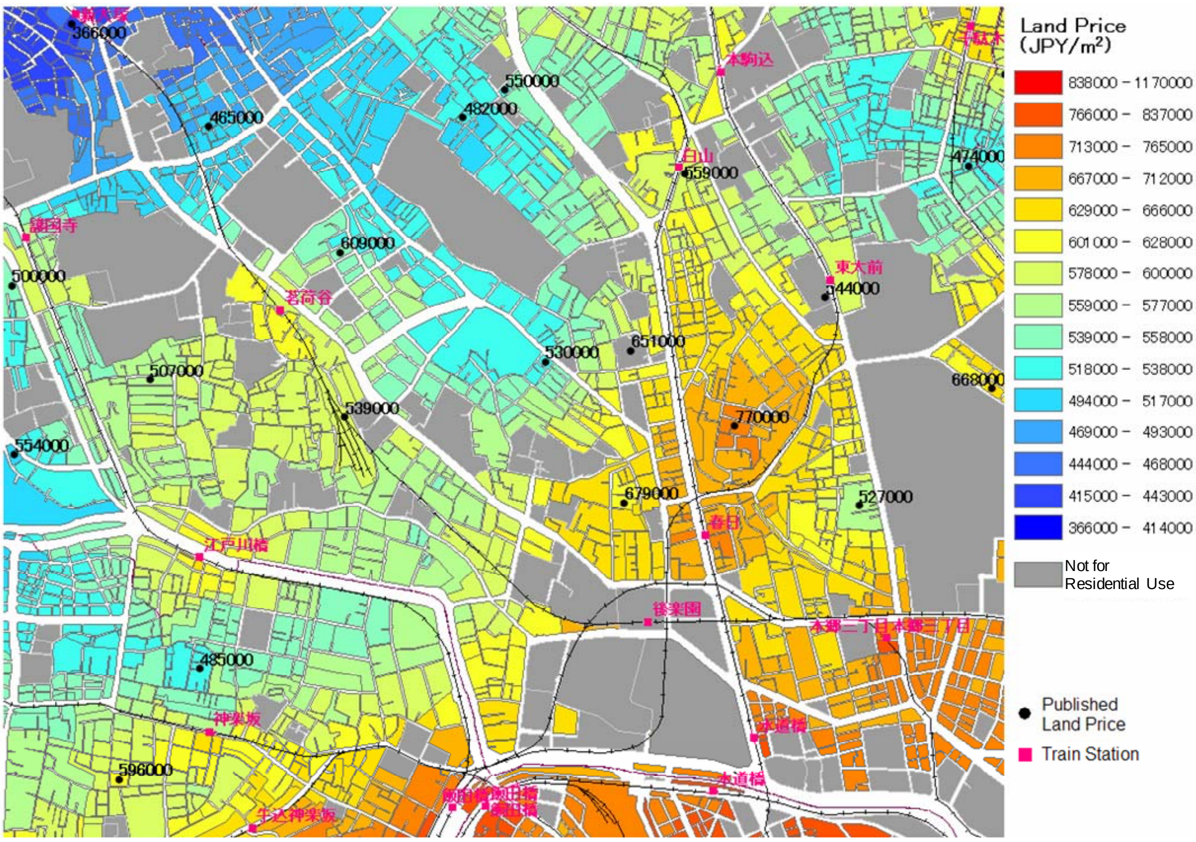


Figure 5 Visualization of Land Price Distribution in Bunkyo Ward in 2004

Needless to say, it is also possible that the estimation can visualize the land price changes in the whole study area. Figure 6 visually illustrates the spatial distribution and change of land price in Tokyo. These figures would be effective in sharing the information about the trend of the real estate market, and the animated presentation would be more useful.

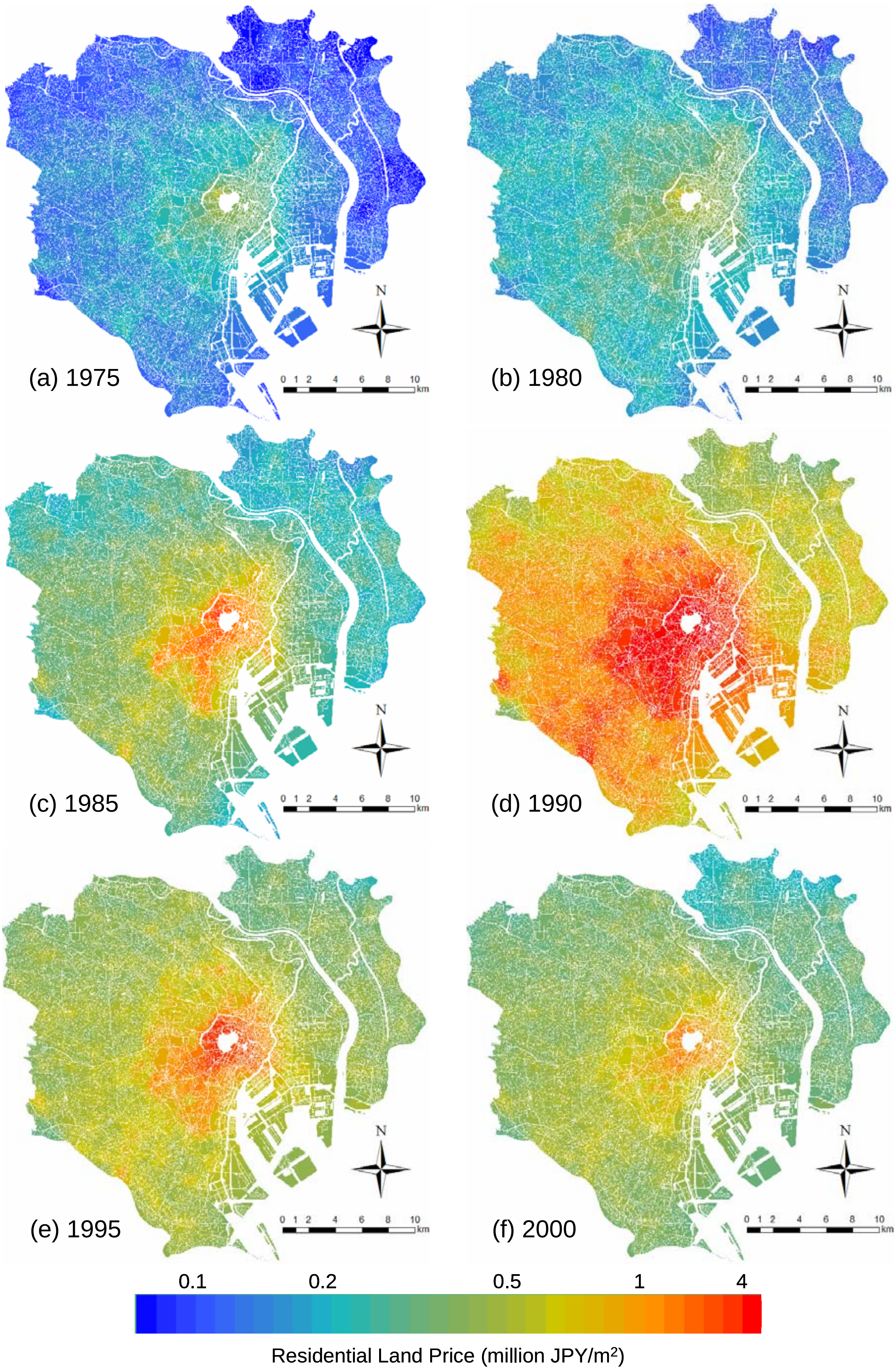


Figure 6 Visualization of Land Price Change in Tokyo

5. CONCLUDING REMARKS

This study focused on the spatio-temporal covariance function, which models the correlation structure in space and time. We applied it to the residential land price estimation in Tokyo 23 wards in 30 years and evaluated its applicability.

The validation result reveals that the estimation has high accuracy with less than 10% error, except during the period of the bubble economy when the land price boomed. It is confirmed that the spatio-temporal stationary covariance function could model the spatio-temporal correlation of land prices, and the spatio-temporal kriging using the estimated covariance function would output land prices in the study area at a high precision.

We estimated land price at every land lot in Tokyo using the estimation result of spatio-temporal stationary covariance function. Since the spatio-temporal kriging is costless compared to the real estate appraisement, it is certain that the land price estimation by the spatio-temporal kriging is one of the tools to create land price information. Moreover, the visualization of land price distribution and change would be effective in sharing the information on the real estate market trend.

The following are remaining topics for future works.

This study used the officially assessed land price data. The observation points for this data are intentionally placed separately, and the pairs of points that have a short distance between them are scarce. Thus, the coefficients of spatio-temporal covariance function tend to be unstable. It is necessary to validate that these coefficients become stable if input is observation data at random locations such as transaction land prices.

Due to limited computing power, this study used data only from inside the 23 wards of Tokyo, which is not a very wide area. This precluded setting the spatial range at more than 15 kilometers; however, the propriety of this value was not discussed sufficiently. It requires further discussion by using the data of a wider area.

We visually illustrated the spatial distribution and change of residential land price data. However, as mentioned above, there are many commercial and industrial zones in Tokyo. We are planning to estimate the land price that includes not only residential but also commercial and industrial zones together, and visualize the change in the Tokyo metropolitan area.

ACKNOWLEDGEMENTS

We would like to thank Supercomputing Division of the Information Technology Center at the University of Tokyo for providing super computers for the estimation of the land price model. We would also like to thank Professor Atsuyuki Okabe for providing SANET, the spatial analysis software package.

REFERENCES

Cressie, N. (1993) **Statistics for Spatial Data.** New York: Wiley.

Cressie, N., and Huang, H-C (1999) Classes of Nonseparable, Spatio-Temporal Stationary Covariance Functions, **Journal of the American Statistical Association,**

Vol. 94, No. 448, 1330–1340.

Gneiting, T. (2002) Nonseparable Stationary Covariance Functions for Space-Time Data, **Journal of the American Statistical Association, Vol. 97, No. 458**, 590–600.

Huang, H-C, and Hsu, N-J (2004) Modeling Transport Effects on Ground-Level Ozone Using a Non-Stationary Space-Time Model, **Environmetrics, Vol. 15**, 251–268.Okabe, A. (2003) A New Approach to Spatial Analysis: Spatial Analysis on a Network with SANET, **The 8th International Conference on Computers in Urban Planning and Urban Management CUPUM '03 Sendai**, Sendai, Japan, 27–29, May 2003.