

PAPER • OPEN ACCESS

Spatial Dynamics Model for Sustainability Analysis of Urban Settlements, Case Study of Surabaya Urban Wetlands, Indonesia

To cite this article: Habib Subagio *et al* 2019 *IOP Conf. Ser.: Earth Environ. Sci.* **280** 012016

View the [article online](#) for updates and enhancements.

Spatial Dynamics Model for Sustainability Analysis of Urban Settlements, Case Study of Surabaya Urban Wetlands, Indonesia

Habib Subagio^{1,3}, Raldi H Koestoer², Yuri M Zagloel³, Linda Darmajanti⁴

¹ Geospatial Information Agency, Jl. Raya Jakarta Bogor Km. 46 Cibinong, habib.subagio@big.go.id

² School of Environment Science, University of Indonesia, Jl. Salemba Raya 4 Jakarta Pusat 10430

³ Department of Industrial Engineering, Faculty of Engineering, University of Indonesia, Depok 16424.

⁴ Department of Sociology, Faculty of Social and Political Science, University of Indonesia, Depok 16424.

ABSTRACT The city of Surabaya is the second largest city in Indonesia which still has active wetlands in the form of mangrove and ponds cultivation areas. This wetland has an important contribution to the environmental balance, protection of biodiversity, as well as maintaining the ecological function of the urban area to continue properly. Specifically this region contributes to surface water control which is an absolute prerequisite for the sustainability of settlements in urban zones. Urbanization and economic dynamics continue to drive increasingly highly demands of land, so that wetland areas continue to experience pressure and turn into built-up land use. The spatial dynamics model of land use change is one of the important instruments that can be used to analysis of urban dynamics complexes. This model using 2 methods in stages. First, the spatial data mining (SDM) method used to determine indicators of agents that contribute on the fragmentation of wetlands. Second, the integration method of Markov Change and Cellular Automata (CA) models for spatial land use projection. The results showed that the integration of SDM methods and Markov-CA contributed significantly to improving the accuracy of the results of land use projection. This model is then referred to as guided spatial projection where the process of urban sprawl can be shown to be more rational and logical. The results of the research are projections of the dynamics of land use that are used as input in calculating the surface runoff as one indicator of sustainability for urban settlements.

Keywords: urban wetland, spatial dynamic, fragmentation, land use projection

1 Introduction

Wetlands are biologically diverse and occupy highly productive transitional areas ('ecotones') between soil and water, characterized by shallow water submerged in surface water and interspersed with submerged or growing vegetation [1]. Wetlands are an important part that is integrated with a global ecosystem that has an important function in maintaining environmental balance such as preventing or reducing the impact of flooding, accommodating surface water and providing groundwater, and providing a unique habitat both flora and fauna [2]. Therefore, wetland ecosystems become an important area that is designated as an area that must be protected and monitored by various world institutions recognized through the international agreement scheme on the Ramsar Convention on wetlands [3].

Urbanization is a topic of great concern for urban researchers [4, 5, 6]. Delays in anticipating high population growth are one of the important factors causing various urban problems in Indonesia. Land conversion in urban areas in Indonesia is largely uncontrolled due to economic pressure on land [6], on the other hand uncontrolled land conversion in urban development will reduce the ecological function of the city and create new problems for the physical, social and economic environment [7, 8, 9].



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Analysis of land use change is currently an important research cluster to understand the dynamics of the environment with complex and constantly evolving indicators [10,11,12,13]. The land use change analysis provides two key interrelated questions, namely "what drives/causes land use changes" and "what (social and economic environment) impacts of land use change". The driving factor or driving force which is often referred to as the driver of land use change does not always depend on the condition of the study area, although it is generally accepted and understood logically by researchers analyzing land use change [14].

The driving factors for land use change that have been accepted in this research cluster consist of two main categories: bio-physical and socio-economic drivers. The combination of driving factors is often referred to as the driving force of the dynamics of land cover / use changes. Bio-physical drivers include characteristics and processes of the natural environment such as: weather and climate variations, land forms, topography, geomorphological processes, natural disasters, vegetation succession, soil types, drainage patterns, and availability of natural resources. Socio-economic drivers include several factors and processes such as demography, industrial structure and dynamics, technological change, market economy, public sector institutions and government policies, economic and political institutions, community values and norms, community organizations, and policies that regulate property and property ownership. Research into the dynamics of land use and land cover changes (LULC: Land Use and Land Cover) agrees that bio-physical drivers usually do not cause directed on land use change. These drivers cause land cover changes which in turn can affect land use decisions from land owners. The driving force forms complex systems and forms interdependent, interacting, and influencing relationships at various temporal and spatial levels [15]. In addition, changes in land use can result in land cover changes which then provide feedback for land use decisions, so that the dynamics of land use change is a cycle that will continuously take place [14, 16].

The remote sensing system and Geographic Information System approach have added a new dimension to understanding the dynamics of this change including for urban landscapes [17]. A more detailed study of urban landscape changes requires monitoring instruments that have high temporal spatial values to determine the effect of climate variability on the dynamics of land use change [18]. The method of integrating sensing systems and Geographic Information Systems is very important to be used in the field of monitoring environmental changes [19, 20]. Urban growth is a process of spatial and social evolution related to changes in urban areas and transformation of people's lifestyles at different scales [21, 22, 23, 24]. To modelling the historical urban landscape and to project future scenarios, experts have developed a large number of approaches to simulate changes in urban land use.

Land use modeling and prediction has been developed with various approaches and methods which are Markov Change model, logistic regression, Celular Automata and others [25, 26]. Meanwhile, modeling the dynamics of land change for urban areas requires more complex variables including socio-cultural elements to produce more rational and logical projections of future land use [17, 24, 26]. This complex variable is expected to be able to analyze not only at the level of the surface landscape but can reach the ecology of the city. In the last decade, researchers have continued to develop new methods that are able to answer the dynamics of the development and projection phenomena of a more rational city landscape such as the phenomenon of urban sprawl. Adaptation of spatial dynamics modeling continues to evolve and adapts methods from other fields of science such as data mining and more focused spatial data mining in the process of gathering urban socio-cultural variables that influence the dynamics of urban land use [28, 29]. In its development, the development of agent-based land use dynamics modeling methods in Geographic Information Systems has provided new breakthroughs that are more scientifically accepted as some of the studies conducted by Zhang *et al.* (2010), Filatova *et al.* (2013), Tan *et al.* (2015).

Technological advances in computerized data acquisition, storage and processing have resulted in rapid and broad database growth. This causes the phenomenon of data production that accumulates continuously to exceed the ability of humans to fully interpret and use it, in the field of geo-spatial science this phenomenon becomes more serious in the near and future [30]. The process of extracting data and finding knowledge of database data almost leads to the same technique, then these two things

are often referred to together as data mining and knowledge discovery (DMKD). At present most of the existing data has geo-reference properties, the need has been directed to consider the spatial characteristics in the DMKD and develop new branches of geo-spatial science, namely Spatial Data Mining and Knowledge Discovery / HRD [31, 32]. This study aims to produce driving spatial information that has a strong relationship with the dynamics of land use change in wetland areas and to develop projections of future land use through dynamic spatial modeling of several scalable and logical scenarios. The results of the study can be used to improve and improve the accuracy of spatial modeling of future land use projections.

2 Methodology

1.1. Study Area

The research location is East Surabaya Area which consists of 7 sub-districts namely; Gubeng, Gununganyar, Rungkut, Sukolilo, Tambaksari, Tenggilis Mejoyo, and Mulyorejo. Geographically located at coordinates 7 ° 13'59.10 "LS-7 ° 20'45.78" LS and 112 ° 44'31.42 "BT-112 ° 50'48.96" LS. The area of East Surabaya is 98.18 km² (30.14%) of the total area of Surabaya City which has an area of 326.81 km² (BPS, 2017). This area is a very fertile area that is formed from the alluvial plains and the active process of the formation of the Brantas watershed delta that continues to this day. Topographically this area is a wetland landscape with heights between 0-8 mdpal. Wetland ecosystems have important value in maintaining the function of urban ecological sustainability in coastal areas where mangrove expansions in Surabaya City are currently only lagging behind in this region with a very high level of threat. The important value of the region's ecosystem is to become one of the points of stop-over of the world's coastal bird migration on the East-Asian Australian Flyway on precisely the mangrove wetland area in Wonorejo East Surabaya.

1.2. Availability of spatial data

The preliminary data available is a urban resource dynamics dataset sourced from the results of the Surabaya City Land Use Dynamics Mapping conducted by the Geospatial Information Agency (BIG) in 2014 consisting of;

Raster data	Year of acquisition	Spatial data and year of mapping	Scale
Ikonos images	2002	Series of Land Use of Kota Surabaya 2002, 2008, 2012, and 2017	1:10.000
Quick Bird image	2008, 2012		
Aerial Photograph, akuisisi 2016	2016		1:1.000
		Digital base map of Kota Surabaya	1:1.000

The results of this activity are dynamics analysis (spatial balances) of land resources that provide information on the transition of changes in land cover during the 2002-2008, 2008-2012, and 2002-2012 periods at 1: 10,000 scale. Geometric references throughout the spatial process use of RBI Maps 1: 10,000 scale data. In ongoing process, the City of Surabaya has acquired aerial photography using 2 sensors together, namely Large Format Aerial Photographs and LIDAR at the beginning of 2016. This data also used as a research data catalog by taking a subset of Eastern Surabaya City.

2.3 Spatial Data Mining and Knowledge Discovery

Data mining and knowledge discovery are an iterative process that involves several steps, including data selection, data cleaning, pre-processing, and transformation; combining previous knowledge; analysis with computational algorithms and / or visual approaches, interpretation and analysis; formulation and modification of hypotheses and theories; data adjustment / uniformity and analytical

methods; evaluation of return results; and so on iteratively [33, 34]. Data mining and knowledge discovery are explorative, more inductive than traditional statistical methods. Naturally this is the initial stage of the deductive discovery process, where researchers develop and modify theories based on information found from observation data [35]. Knowledge discovery refers to a gradual process such as above, while mining data is defined more narrowly as the application of computational, statistical or visual methods. The application of the Spatial Data Mining (SDM) method in this study was carried out following the steps above to ensure meaningful and useful findings were represented statistically and visually. In addition, 'data mining' and 'geographical knowledge discovery' are carried out in an iterative process, both of which refer to the overall process of knowledge discovery. Spatial data extraction for spatial modeling of land use projections uses of sustainable environmental variables are; bio-physical, socio-cultural, and economic on spatial bases.

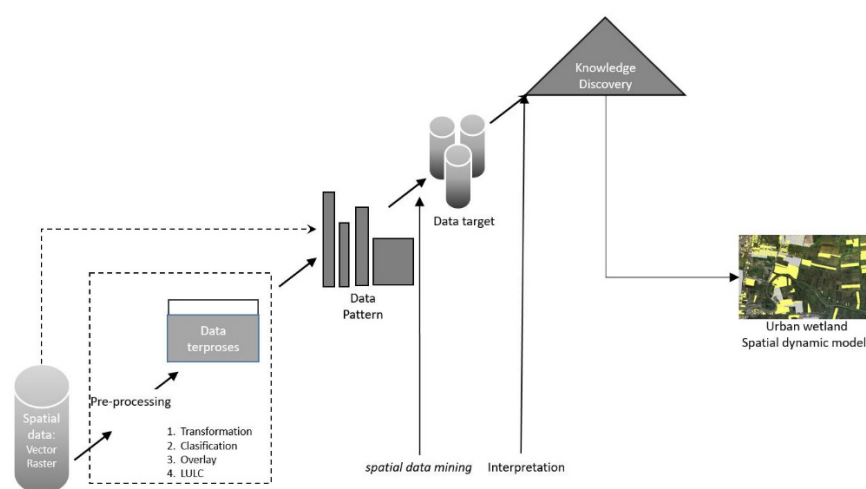


Figure 1. Process of Spatial Data Mining and Knowledge Discovery (SDMKD) on wetland spatial dynamics modelling (modification from Fayyad, Piatetsky-Saphiro, and Smyth, 1996)

Spatial data mining is the process of finding something that is interesting and not yet available, but is a potential spatial pattern that is useful from spatial datasets [36]. The sequential process in Spatial Data Mining (SDM) is the extraction of knowledge from spatial relationships or other interesting patterns that are not explicitly stored in the spatial data basics. The integration of HR with spatial database technology can be used to understand spatial data, find spatial relationships and the relationship between spatial and non-spatial data, compile spatial knowledge bases, reorganize spatial databases, and optimize spatial queries [35].

Currently, geospatial data with large volumes, types and thematic types have been, and continue to be collected with modern data acquisition techniques such as the use of very high-resolution remote sensing GPS applications, geolocation services, and internet-based geospatial information. There is an urgent need for the development of effective and efficient methods for extracting useful information and knowledge from massive and complex spatial databases [37].

2.4 Spatial Modelling

Generally, spatial modeling to calculate land use projections consists of 2 stages; calculate the weight of each driving factor to land conversion based on the results of time series land cover analysis and calculate the attractiveness of the land to find out which area has the most potential to change (high probability). The method used adopted from Constanza (1989) is 'Goodness of fit'. The weight of driving factor is calculated using the method of calculating the value of similarity or compatibility of a data (image / spatial) with other (image / spatial) data. This type of driving factor in modeling is

generated from the spatial data mining process above. Every driving factor in the form of spatial data (picture / raster) is seen in the similarity value to the expansion of land/increasing of landuse.

Weight calculation is the average similarity value of various sizes according to the Goodness of fit model using the following formula;

$$F_t = \frac{\sum_{s=1}^{t_w} f_w e^{-k(w-1)}}{\sum_{w=1}^n e^{-k(w-1)}} \quad \dots\dots\dots(1)$$

Where f_t = a weighted average of the fits over all window sizes, f_w = the fit for sampling window of linear dimension w , k = a constant and w = linear dimension of a sampling window. This equation gives exponentially less weight to the fit at lower resolution. The value of k determines how much weight is to be given to small vs. large sampling windows. If $k = 0$, all window sizes are given the same weight. At $k = 1$, only the first few window sizes will be important.

Land attractiveness calculated in this dynamic spatial modelling have three steps of calculation, there are: the nearest neighbour factor (called cellular automata), historical factors (markov models), and driving factors. Calculation of land attractiveness from nearest neighbour factor using this formula:

$$p_{ij} = \frac{p_{ij}^0 - p_{ij}^r}{r_{ij}^n + \max r_{ij}^n} \quad \dots\dots\dots(2)$$

where p_{ij} = land attractiveness from nearest *nearest neighborhood* at coordinate (i,j) , p^0 = determinant cost of the type of land to be converted, p^r = determinant cost of the type of land that will convert the land, r^n = distance at coordinates (i,j) against land classes that will convert (n), where the data has been distributed normally after *euclidean distance*.

Land attractiveness from the historical factors of land use change must have at least two land use data, where the data is seen as over land function and percentage of change, either converted or converted to other land use.

$$h_{ij} = \frac{a_1^1 - a_n^0 / a_1^1}{r_{ij}^1 + \max r_{ij}^1} \quad \dots\dots\dots(3)$$

where h_{ij} = land attractiveness from opportunity value (*probability*) at coordinates (i,j) , a^1 = the extent of a land type that converts the previous land or which has land expansion, a^0 = the area of the initial type of land converted into new land use, n = the type or class of land converted, and 1 = the type or class of land that is expansion, r^1 = distance to land expansion that converts other land use (n), where the distance data is already in the normal distribution after euclidean distance.

Driving factor is a factor that can make an attraction to increase the function over the surrounding land. Every driving factor has a different role in determining the attractiveness of land and each class or type of land has its own land attractiveness.

$$d_{ij} = \frac{(1 + \sum_m w_m \cdot b_{ij}^m) + \max(1 + \sum_m w_m \cdot b_{ij}^m)}{(1 + \sum_n w_n \cdot c_{ij}^n) + \max(1 + \sum_n w_n \cdot c_{ij}^n)} \quad \dots\dots\dots(4)$$

where d_{ij} = land attractiveness at coordinate (i,j) , b_{ij} = determinants of the benefits of driving factors that have been normally distributed at coordinates (i,j) , c_{ij} = cost determinants of the driving factor that has been normally distributed at coordinates (i,j) , dan w_m = weighted of (F_t) from the determinants of costs m , dan w_n = weighted (F_t) from the determinant of cost n .

The three determinants of the attractiveness of the land are calculated in the ratio of 1: 1: 1 or it is assumed that the three have the same power effect in determining the location of the most attractive land to the unattractive.

$$a_{ij}^n = d_{ij}^n * p_{ij}^n * h_{ij}^n \quad \dots\dots\dots(5)$$

a_{ij} = attractiveness land for classes that grow or convert to other land use, and n = class or type of land inputted in the spatial model. So that all land use classes that will grow or expand will have a total value of land attraction.

3 Result and Discussion

The earliest spatial analysis process was carried out by overlaying the time series land use layers, LU 2002, LU 2008, and LU 2012 after ascertaining the fix topology on these three data features. This research applied 17 classes of land use which are the reclassification of 23 land use classes from the interpretation of the previous land use data. Determination of 17 classes of land use is seen as sufficient to represent the representation of the land use classification of the research area and consider computational processing in subsequent spatial modeling. This process results in land use in the history of the research area which is then carried out by Markov analysis as presented in **Table 1**.

Tabel 1. Result of GIS subset analysis of land use history of East Surabaya Timur

No	Class of LU	Area 2002	Area 2008	Area 2012	Rate (ha/yr)
1	High level education area	319,769	322,655	322,619	+0,285
2	Lake dwelling	17,379	25,338	37,101	+1,972
3	Industrial area	293,868	302,979	302,571	+0,870
4	Roads	553,140	601,742	630,112	+7,697
5	Cemetery	3,710	6,095	9,578	+0,587
6	Mangrove area	316,600	330,750	405,800	+8,920
7	Median of roads	53,163	55,424	60,360	+0,720
8	Sport area	29,146	31,634	31,634	+0,249
9	Settlements	3.705,106	3.858,146	4.021,423	+31,632
10	Commecial area	251,016	276,608	291,598	+4,058
11	Flat mud	171,253	32,865	6,355	-16,490
12	Railroads	6,557	6,557	6,557	+0,000
13	Paddy fields	366,600	283,048	238,505	-12,809
14	Train station	13,522	13,522	13,522	+0,000
15	River	155,070	157,943	156,004	+0,093
16	Ponds	3.174,479	3.214,676	3.012,398	-16,208
17	Open area	844,396	754,484	797,148	-4,725

Source: analysis result, note: '+' increase '-' decrease

Refer from Table 1. above can be seen the dynamics of land use change in the 10 years period starting from 2002-2012, some land use classes have increased in area such as settlements with the rate of addition of 31.6 ha / year and trade services at the rate of addition of 4.05 ha / year , others experienced a broad decline such as rice fields and ponds with an average reduction rate of 12.8 ha / year and 16.2 ha / year respectively. Markov analysis of the spatial dynamics of land use change shows the probability value of each land class in all land use classes.

Related to the mangrove area, an interesting factor in this study is the area of land use in the mangrove area which was initially assumed to be reduced but the results of Markov analysis on spatial data indicate that the mangrove area has an increase in area with a positive rate of 8.9 ha / year. The addition of mangrove areas occurs with a pattern away from the initial coastline along the coast of the study area, new mangroves continue to grow in the area of the mud flats or sediment deposition zones.

In the land area, mangrove areas are transformed into pond areas, open land, and road construction on the other side of the mangrove area also continues to grow in the sediment deposition zone, namely the mud flat area. The history of land use change history then becomes the basis for compiling the land use change transition matrix as input in modeling simulations of future land use change projections.

Other land uses that continue to increase are roads identified as new roads or road widening. Interpretation results show that additional roads occur as new infrastructure built as a means of connectivity between old residential areas and new residential area built by large developers. This new residential area occupies on ponds (recorded in the 2002 image), becomes open land with land drying and maturation (recorded in the 2008 image), and then becomes a new residential area (recorded in the 2012 image). The use of time series data acquisition is able to provide information on the transition process over better land use change analysis. The transition path of land use change is important information that can ensure the accuracy of the projected land use model.

3.1 Spatial Driving Factor Analysis

Determination of driving factors that affect dynamics is based on previous research which is then referred to as the category of Business as Usual (BaU) and the results of data mining exploration through following all procedures that have been established in spatial data mining theory. Recent research shows that there are agents that influence the occurrence of urban sprawl in urban areas and urban peripherals. Land use decisions at the macro scale (analysis of urban details) on their reality are more determined by the interdependence of socio-cultural variables, on the other hand the interaction between humans and biophysical variables simultaneously is indirectly or continues to be improved by local government policies. Interaction between these agents is then more dominant in changing urban landscapes. Spatial information openness is currently a new opportunity for spatial data mining implementation for more precise and accurate application of spatial modeling. Table 1. below shows the results of spatial driving factor produced and used in spatial land use modeling, as well as the calculation of the weight (similarity) of each feature driving factor to the expansion of residential land and commercial areas in the study area. Some of these new spatial features are then categorized as the results of Spatial Data Mining (SDM).

Table 2 List of spatial driving factor and the weight of similarities to the expansion of settlements and commercial land

No	Spatial driving factor	Land expansion of settlements	Land expansion of commercial area	Variabel	Category
1	Dist from apartemen dan rusun	0,6623	0,7229	Socio Cultural	BaU
2	Dist from high level education area	0,6855	0,7591	Socio Cultural	BaU
3	Dist from industrial area	0,5889	0,6290	Socio Cultural	BaU
4	Dist from artery/collector roads	0,7270	0,6927	Biophysic	BaU
5	Dist from local roads	0,5782	0,4775	Biophysic	BaU
6	Dist from economic center	0,6112	0,7057	Biophysic	BaU
7	Dist from bus station	0,4897	0,5665	Biophysic	BaU
8	Dist from commercial	0,7010	0,8645	Biophysic	BaU
9	Dist from train station	0,4593	0,5277	Biophysic	BaU
10	Dist from river	0,5746	0,6012	Biophysic	BaU
11	Dist from mangrove recreation	0,4742	0,5089	Socio Cultural	BaU
12	Slope	0,6344	0,5805	Biophysic	BaU
13	Site Plan (Izin Perumahan)	0,5323	0,5360	Socio Cultural	SDM
14	Land ownership rights	0,5329	0,5365	Socio Cultural	SDM
15	Land price zone	0,4912	0,4785	Economy	SDM
16	Allocation for settlements area (RTRW)	0,4570	0,4613	Policy	BaU
17	Allocation for commercial area (RTRW)	0,5542	0,5578	Policy	BaU

18	Build and use rights	0,5741	0,5506	Economy	SDM
----	----------------------	--------	--------	---------	-----

Note: BaU = Bussiness as Usual, SDM = Spatial Data Mining

Spatial data mining provides new knowledge while showing the heterogeneity of spatial information, heterogeneous spatial data processing and analysis iteratively provides feedback for interpreters (data analyst) in understanding the patterns, relationships and interrelationships between spatial and non-spatial data that can be used to answer complexity urban landscape phenomenon.

Based on **Table 2** it can be seen that the distance to the road factors of either arteries or collectors still have the highest spatial similarity weight (mean 0.7 from a scale of 0-1) to the pattern of settlement land expansion followed by a distance factor to the service and trade area. This shows that the existence of large universities has become a driving factor that has a great influence on the expansion and transfer of land functions. The distance factor to the objects of apartments and flats, the distance to the education area is high, and the distance to the economic center has a lower weight (average 0.6 from a scale of 0-1). Meanwhile, the spatial factors of the results of spatial data mining (HR) such as housing site plans, land rights, land value zones, and usufructuary rights have relatively similar weighting (average 0.49-0.53 of scale 0 -1) which is towards the pattern of settlement land expansion. The spatial factor of this human resource also has the same type of weight on the pattern of commercial expansion. Slightly different in the pattern of commercial area expansion where the distance factor to the location of apartments and flats, and the distance to the area of higher education gives the highest similarity weight effect (mean 0.7 from a scale of 0-1)).

The feature of spatial planning which focused on the designation zone for settlements and designation zones for commercial areas in the study area is used as a driving factor for the policy category, as well as spatial constraint on this model. These features become important spatial information that is used as an input in the spatial modeling process to analyze the stock of land allocation in the two function allotments. This variable of spatial policy is used in spatial modeling of future land use projections which are included as an inputs in scenarios and modeling interventions.

3.2 Spatial modelling, scenarios of land use projection, and validation

Spatial modeling of land use is the process of determining the projected land use in each cell unit analyzed. The process carried out is to integrate the results of Markov analysis as the basis for the transition rules and the rate of land use change in each land use class, then the CA (cellular automata) process based on an analysis of the conversion costs of each land use change to the nearest land use, and considering the driving factors in determining the most interesting location to convert. Processing of land use projections by utilizing the GIS operation function and coding support (script) for mathematical formula operations.

Baseline projections for land use were carried out starting in 2002 according to input data used. Land use projections are carried out until 2032 to see the future dynamics of land use. Projection scenario for land use consists of two, namely; Scenario-1, hereinafter referred to as BAU (Bussiness as Usual), only uses the driving factor BAU category and scenario-2 is referred to as Optimistic using the driving factor BAU + SDM. The results of the projection of land use in both scenarios can be seen in the **Figure 2**. below.

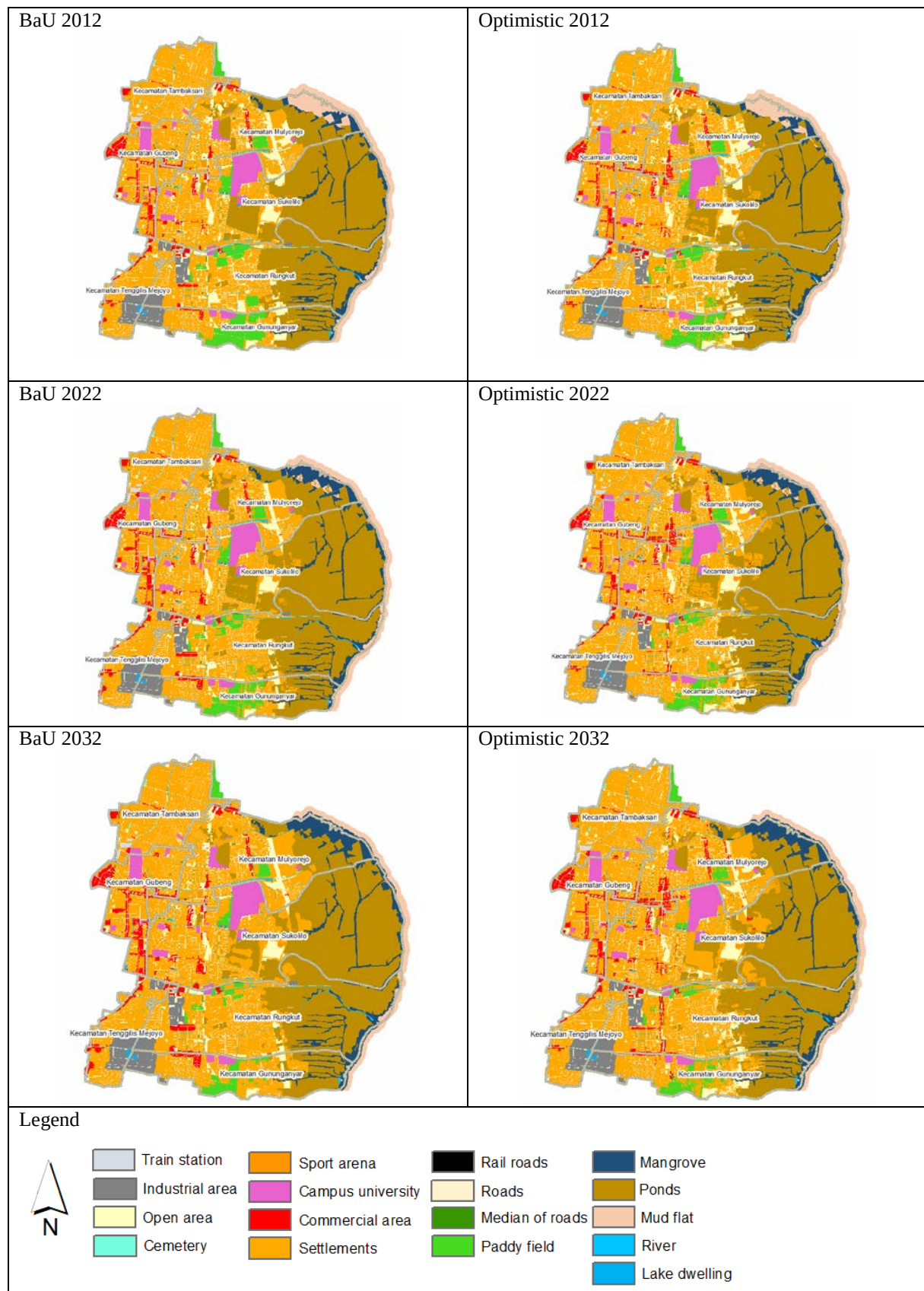


Figure 2. Comparison results from the land use projection model using BAU and Optimistic scenarios

Result of calculation accuracy from the land use projection model by applying the Kappa method to calculate overall accuracy by using land use data 2008, 2012 and 2016 as reference data. The results obtained for the BAU Scenario were 91.37%, 87.83% and 82.37% respectively. The results obtained for the Optimistic scenario are 90.75%, 88.18% and 83.18% respectively. This shows that the addition of the driving factor found based on the HR process such as the location of the housing permit and use rights and proven rights provides an increase in the accuracy of the resulting projection model. Although it has not shown a significant contribution to the validation results, this is relevant to the results of calculating the weight of the driving factor by 0.53 to the expansion of residential land and 0.57 to the expansion of the commercial area, this weight value can vary for different study areas as well.

One thing that needs to be addressed is the problem of the absence of metadata obtained from the spatial data mining process. Metadata becomes important information on a data, time information on site plan data will contribute to the results of spatial modeling scenarios of land use projection.

These results are important findings in the development of spatial modeling methods for projection of land use, especially in urban areas, where socio-cultural variables become driving factors that have influence and contribute to increasing the accuracy of a model based on spatial data.

4 Conclusion

The study of the dynamics of land use change is still an interesting phenomenon because of the characteristics of different regions. Development of methodology also continues to grow to get more ideal and logical spatial land use modeling. The dynamics of urban land use change have more complex spatial dimensions. The main environmental variables are biophysical-economy-socio-culture capable of providing deductive instructions in translating Spatial Data Mining and Knowledge Discovery in spatial dynamic research clusters.

Detailing the stages in the spatial data mining mechanism such as identification of complete metadata in each of spatial spatial feature gives positive implications for the results of the accuracy. The results of the research show that the SDM application in spatial modeling of land use projection influences the results of the accuracy of spatial modeling produced. SDM utilization also contributes to the discovery (knowledge discovery) that the spatial modeling of the phenomenon of "*urban sprawl*" can be done accurately. The results of the study can be used as consideration for the sustainability analysis of urban settlements in the future.

Acknowledgment

Gratefully thanks to BIG, Local Government of Surabaya City, and a joint survey team from ITS and Airlangga University for all the data and information provided in this study.

References

- [1] Lee *et al* 2006 Impact of urbanization on coastal wetland structure and function *Austral Ecology* **31** 92006 p 149–163 doi:10.1111/j.1442-9993.2006.01581.x.
- [2] Mitsch, W. J. and J. G. Gosselink 1993 *Wetlands, 2nd edition* (Van Nostrand Reinhold, New York, NY, USA).
- [3] Töyrä, J. and A. Pietroniro 2005 Towards operational monitoring of a northern wetland using geomatics- based techniques *Remote Sensing of Environment* **97** p 174–191.
- [4] Jantz, C. A., S. J. Goetz and M. K. Shelley 2004 Using the SLEUTH urban growth model to simulate the impacts of future policy scenarios on urban land use in the Baltimore-Washington metropolitan area. *Environment and Planning B-Planning & Design* **31(2)** p 251-271. Liu, Y 2012 Modelling sustainable urban growth in a rapidly urbanising region using a fuzzyconstrained cellular automata approach *International Journal of Geographical Information Science* **26(1)** p 151-167.
- [5] UN-DESA 2013 *Population facts*. New York: The United Nations.

- [6] Firman, T 2000 Rural to Urban Land Conversion in Indonesia During Boom and Bust Periods *Journal of Land Use Policy* **17** p 13-20
- [7] Pickett, S.T.A., Cadenasso, M.L., Grove, J.M., Nilon, C.H., Pouyat, R.V., Zipperer, W.C., Constanza, R 2001 Urban Ecological Systems: Linking Terrestrial Ecological, Physical, and Socioeconomic Components of Metropolitan Areas *Annu. Rev. Ecol. Syst* **32** p 127–57.
- [8] Mugerauer, R 2010 Toward a Theory of Integrated Urban Ecology: Complementing Pickett et al *Ecology and Society* **15**(4) p 31.
- [9] Ramsar 2013 Towards the wise use of urban and peri-urban wetlands. BN No. 6.
- [10] Constanza, R 1989 Model Goodness Of Fit: A Multiple Resolution Procedure *Ecological Modelling* **47** (Elsevier Science Publishers B.V., Amsterdam) p 199-215 doi:10.1016/j.gloplacha.2004.10.020, <http://dx.doi.org/10.1016/j.gloplacha.2004.10.020>.
- [11] Batty, M., Xie, Y., Sun, Z 1999 Modeling urban dynamics through GIS-based cellular automata *Computers, Environment and Urban Systems* **23** p 205-233.
- [12] Jensen, J.R. 2000 *Remote Sensing of the Environment: An Earth Resource Perspective* (Prentice Hall: New Jersey).
- [13] Odum, E.P., Barrett, G.W 2005 *Fundamentals of ecology* (5th ed.) (Thompson, Brooks Cole, 598).
- [14] Helen Briassoulis, H 2000 Factors Influencing Land-Use And Land-Cover Change *LAND USE, LAND COVER AND SOIL SCIENCES Vol. I* ©Encyclopedia of Life Support Systems (EOLSS).
- [15] Hersperger, A.M., Gennaio, M.P., Verburg, P.H., and Matthias Bürgi, M 2010 Linking Land Change with Driving Forces and Actors: Four Conceptual Models *Ecology and Society* **15**(4) 1 [online] URL: <http://www.ecologyandsociety.org/vol15/iss4/art1/>.
- [16] Tan, R., Liu, Y., Zhou, K., Jiao, L., Tang, W 2015 A game-theory based agent-cellular model for use in urban growth simulation: A case study of the rapidly urbanizing Wuhan area of central China. *Computers Environment and Urban Systems* **49** p 15–29.
- [17] Jat Mahesh Kumar., Garg, P. K. and Khare, Deepak 2008 Monitoring and modeling of urban sprawl using remote sensing and GIS techniques *International Journal of Applied Earth Observation and Geoinformation* **10** p 26-43 doi:10.1016/j.jag.2007.04.002 <http://dx.doi.org/10.1016/j.jag.2007.04.002>.
- [18] Milesi Cristina et al. 2005 Climate variability, vegetation productivity and people at risk *Global and Planetary Change* **47** p 221-231.
- [19] Turner, M. G 2005 Landscape ecology in North America: past, present and future *Ecology* **86** p 1967-1974. doi:10.1890/04-0890, <http://dx.doi.org/10.1890/04-0890>.
- [20] Lambin Eric, F., Turner, B. L., Geist, Helmut J., Agbola, Samuel B., Angelsen, Arild, Bruce, John W. et al 2001 The causes of land-use and land-cover change: moving beyond the myths *Global Environmental Change* **11**(4) p 261-169 doi:10.1016/S0959-3780(01)00007-3, [http://dx.doi.org/10.1016/S0959-3780\(01\)00007-3](http://dx.doi.org/10.1016/S0959-3780(01)00007-3).
- [21] Batty, M. and Y. Xie 1994 From cells to cities, *Environment and Planning B-Planning & Design* **21** (7) p 31-48.
- [22] Baynes, T. M. 2009 Complexity in Urban Development and Management *Journal of Industrial Ecology* **13** (2) p 14-227.
- [23] Wahyudi, A. & Liu, X. 2015 Spatial Dynamic Models for Inclusive Cities: A Brief Concept of Cellular Automata (CA) and Agent-based model (ABM) *Jurnal Perencanaan Wilayah dan Kota*. Vol **26** (1) p 54-70.
- [24] Fajar, Y, Taufik, M, Khomarudin , M.R 2018 Analysis of the dynamics of land use change and its prediction based on the integration of remotely sensed data and CA-Markov model, in the upstream Citarum Watershed, West Java, Indonesia *International Journal of Digital Earth*. <https://doi.org/10.1080/17538947.2018.1497098>.

- [25] Tatiana Filatova, T., Versburg, P.H., Parker, D.C., Stannard, C.A 2013 Spatial agent-based models for socio-ecological systems: Challenges and prospects *Environmental Modelling & Software* 45 p 1-7.
- [26] Torrens, P.M., O'Sullivan, D 2001 Cellular automata and urban simulation: where do we go from here? *Environment and Planning B-Planning & Design* 28 p 163-168.
- [27] Lau, K. H. and B. H. Kam 2005 A cellular automata model for urban land-use simulation *Environment and Planning B-Planning & Design* 32(2) p 247-263.
- [28] Malerba, D 2006 Mining Spatial Data: Opportunities and Challenges of a Relational Approach IASC 07 August 30th – September 1st, 2007, Aveiro, Portugal.
- [29] Spate, J.M., Gilbert, K., Sánchez-Marrè, M., Frank, E., Comas, J 2006 Data Mining as a Tool for Environmental Scientists *International Congress on Environmental Modelling and Software* 67.
- [30] Deren Li and Shuliang Wang. 2005 Concepts, Principles And Applications Of Spatial Data Mining And Knowledge Discovery *ISPRS XXXVI-2/W25*.
- [31] Li D.R., Cheng T., (1994), KDG: Knowledge Discovery from GIS - Propositions on the Use of KDD in an Intelligent GIS. In Proc. ACTES, The Canadian Conf. on GIS.
- [32] Ester, M. et al 2000 Spatial data mining: databases primitives, algorithms and efficient DBMS support. *Data Mining and Knowledge Discovery* 4 p 193-216.
- [33] Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth 1996 Advances in knowledge discovery and data mining *Data Mining to Knowledge Discovery: an Overview* (American Association for Artificial Intelligence, 1996a) p 1–34.
- [34] Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth 1996 From data mining to knowledge discovery in databases (a survey) *AI Magazine* 3(17) p 37–54.
- [35] Han, J and Kamber, M 2001 *Data Mining: Concepts and Techniques* (Academic Press, California).
- [36] Sumathi, N., Geetha, R. dan Bama, S. S 2008 Spatial Data Mining – Technique Trends and Its Applications *Journal of Computer Applications Vol-1* p 28-30.
- [37] Diansheng Guo & Jeremy Mennis 2009 Spatial data mining and geographic knowledge discovery - An introduction *Computers, Environment and Urban Systems*, 33 p 403–408, doi:10.1016/j.compenvurbsys.2009.11.001.