Desalination and Water Treatment
www.deswater.com
19 (2010) 184–190
July

1944-3994 / 1944-3986
doi: 10.5004/dwt.2010.1911

Calculation of diffuse pollution loads using geographic information

Ji-Heon Lee[a,*], Sung-Ryong Ha[b], Myong-Soon Bae[a]

[a]*Chungbuk Research Institute, 89 Munwha-dong, Cheongju, Chungbuk, Korea*
email: career03@hotmail.com
[b]*Department of Urban Engineering, Chungbuk National University, Cheongju, Korea*

Received 24 September 2009; Accepted 22 February 2010

ABSTRACT

The first grade Korean TMDL (K-TMDL) Master Plan for 2005–2010 has been established by the Ministry of Environment with statistical data from local governments to get the land-coverage information and calculate the discharged non-point source (NPS) loads from land coverage. The statistical information is easy to obtain but is not consistent with real land-coverage and can cause a considerable number of errors when calculating the watershed based pollution loads. As requests for GIS-based data processing are getting increasing, various approaches for calculating each land use have been considered. In this study three kinds of data resources such as SPOT5 remote sensing image, registration map and cadastral surveying result were used. This study mostly focuses on the comparison of area calculation results between existing cadastral surveying and registration map referencing of SPOT5 satellite and GIS digital map. Mostly, differences of NPS loads are caused by procedural changes of basin area calculation. It was converted into basin pollution loads after calculating loads based on the municipal administrative unit before, however, when using GIS-based method, conversion process from municipal based result to basin based result is not required. Basically, conventional text-based method apply the specific formula, building area + road/2, to get sewage collection area estimate while GIS-based method calculate the sewage collection area directly by the GIS overlay analysis. As a result of that, the larger sewage collection area has, the greater sewer inflow pollution loads are generated. And this causes less discharged pollution loads. On the contrary, the less sewage collection area calculated the greater individually discharged pollution loads are generated, which induce higher discharged pollution loads. Those generate the changes of basin and river purifying parameters of the water quality simulation model, which will affect the attainable pollution loads for development in the future.

Keywords: TMDL; NPS load; GIS; SPOT5 image; Land registration map; Cadastral surveying

1. Korean TMDL and accuracy improvements of NPS calculation

The Korean TMDL (K-TMDL) policy manages both point source and non-point source (NPS) pollutants. The precision of information on NPS is very important to allocate the appropriate NPS carrying capacity in TMDL policy. The latest information on land-coverage is especially important because it is changed every year depending on development plans. Currently, the first grade K-TMDL Master Plan for 2005–2010 has been established by the Ministry of Environment with statistical data from local governments to get the land-coverage information and calculate the NPS loads

*Corresponding author.

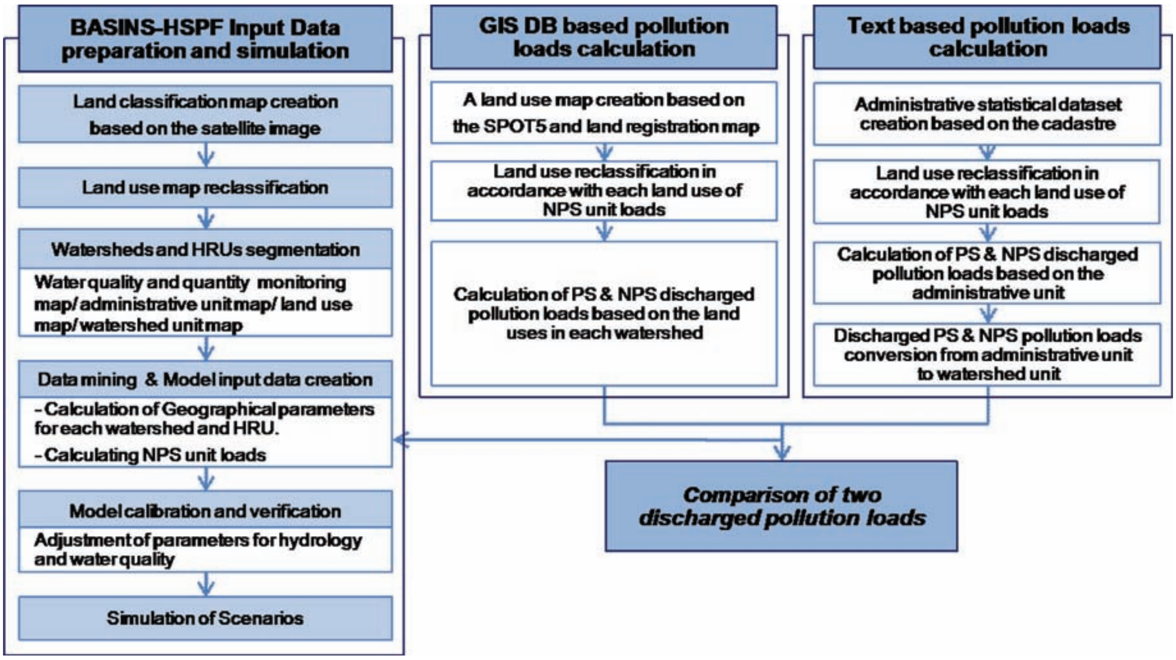

Fig. 1. A schematic structure of the study in whole process.

discharged from land coverage. The statistical information is easy to obtain but is not consistent with real land-coverage and can cause a considerable number of errors when calculating the watershed based pollution loads. In order to solve this problem, a number of studies are ongoing. [6]

In Korea currently, an efficient technical method of a location-based object data management have been studied because the National Geographic Information Systems (NGIS) did not update the latest information in a timely manner, which hinders the use of the first phase GIS framework data. This object-oriented database using a location based feature identifier (UFID; Unique Feature Identifier) makes it possible to access the newly updated management directly [1]. In an effort to take advantage of this dynamic data processing structure in the TMDL management, an area of land uses based on the GIS spatial DB has been calculated.

To improve modeling ability by using dynamic model instead of steady state model, studies are ongoing to apply the watershed hydrological and water quality models, such as Hydrological Simulation Program-Fortran (HSPF), SWAT, MIKE-SHE, and SWMM to simulation. This makes it possible to select the best assessment timing for water quality management contrary to limited simulation of conventional model. To extract input data for HSPF from the GIS DB, in the United States of America, BASINs have been implemented. To simulate the model more accurately, basin based pollution loads was calculated with GIS DB. This study focuses on the methodology which can improve the precision of calculation on the NPS loads using the geographic information such as land registration maps and artificial satellite images.

2. Data preparations and limitations

In this study, the land registration maps which were provided by the NGIS and the high-resolution artificial satellite (SPOT5) images were used to calculate the NPS loads discharged from land-coverage. The watershed based land-coverage information was acquired using watershed maps and land registration maps by overlay analysis with GIS tools. The SOPT-5 image was used to construct a standard for area comparison according to different data types and to reduce errors in land registration maps. Data acquirement constraints allow us to access only 2006 data for statistics based on cadastral surveying and SPOT5 image and 2007 data for land registration maps.

3. Land use acquirement limitations from the Korea Land Information System

To acquire more advanced geospatial management systems such as Parcel Based Land Information System (PBLIS) cadastral map digitalization has been carried

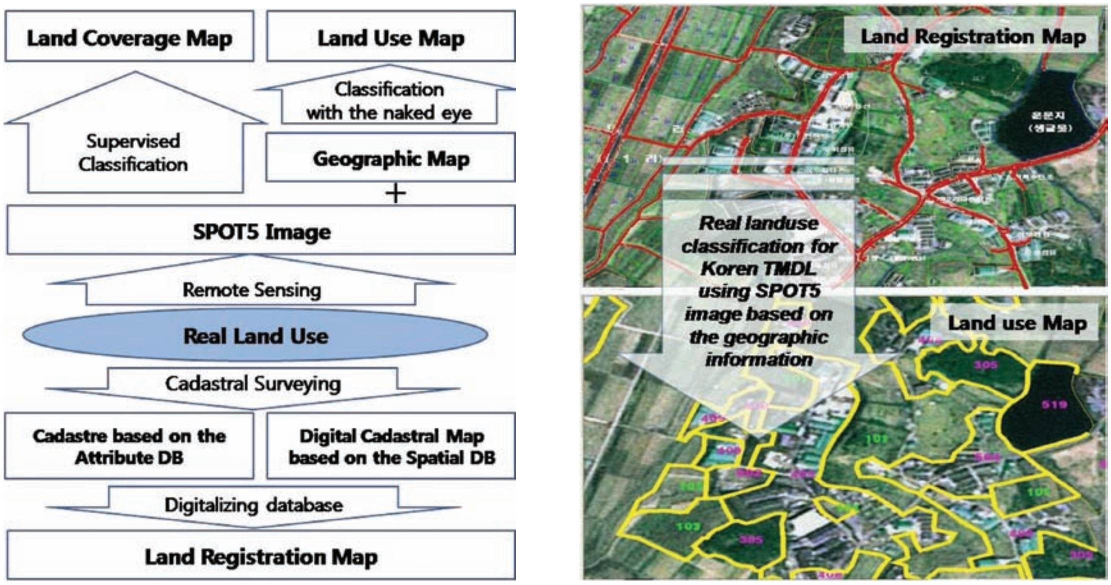

Fig. 2. (a) A process analysis for digital map creation. (b) Land use map creation using land registration map and SPOT5 image.

out by the Ministry of Government Administration and Home Affairs (MGAHA) using the latest equipment and survey technologies, such as the cadastral surveying using GPS satellite technology since 1994. Since 2006, Korea Land Information System (KLIS) has integrated Land Management Information System (LMIS) from the Ministry of Land, Transportation and Maritime Affaires (MLTMA) with PBLIS from the MGAHA to minimize redundant operating expenses [2]. With changes to the land management system, NPS loads calculated using land-coverage for K-TMDL are directly using areas calculated from the geospatial data instead of sheer statistical data. Even though these procedural changes offer convenient data processing with other GIS-based data, they have some limitations. Fundamentally, because a land registration map is created for administrative management purposes, results from the spatial data of a land registration map are different from the results from the SPOT5 images which represent real land use. Therefore modification of the land registration map referencing satellite image should be achieved to get more precise data.

Furthermore, though the KLIS infrastructure has various feedback processes to obtain the results reflecting real land use, it still includes several procedural and structural errors. Such errors include area or location errors caused by inaccurate data conversion into digital types, database inconsistencies due to update delay caused by separate management organizations and unsynchronized data due to individual management between attribute and spatial data.

4. Land use acquirement limitations from the SPOT5 image

The accuracy of land use classification is differentiated by the classification methods of SPOT5 image. Because applying the supervised classification to the high resolution satellite image cannot provide high accuracy [7], this study adopts land use modification referencing SPOT5 image with the naked eye based on the geographic information data. Meanwhile, it was revealed that setting up the classification criteria into five types of land use which are the principal components for the calculation of NPS loads should be the top priority.

5. A time and cost analysis for data preparation from the SPOT5 image

To improve accuracy of land registration map, naked eye classification result of SPOT5 which was referenced by GIS digital map was used as ancillary information [4]. Even though naked eye classification is more cost and time consuming task than pixel-based classification, more improved accuracy is expected. A result from the supervised classification of satellite image was too ambiguous to distinguish some land use from dry paddy field, Orchard and so on. Contrary to this, not only shows naked eye classification 90% accuracy but every land use is noticeable from the others. Table 1 shows that cost and time requirement for land use classification from the SPOT5 Image. GIS data acquirement and SPOT5 image preprocessing is the most demanding task because it needs 70~90% of total budget and 40~50% of total working time. Given that cost is consuming similarly, naked eye classification result which shows high accuracy is a better choice.

6. Discrepancy of classification categories between satellite image and discharged NPS unit-loads

In this study, though a land use map was created with 29 categories (28 land registration map categories and 1 golf courses category), it was reclassified into five categories that existing discharged NPS unit-loads have. Currently, a long-term investigation for NPS unit-loads is undergoing but there are some troubles to solve such as sampling methods and frequency of NPS monitoring. But criterions for each sampling area need to be consistent with the land use map (Table 2) [5]. A discharged

Table 1
Time and cost analysis for data preprocessing with the SPOT5 image.

Category	Work lists	Working cost (thousand won)	Working cost ratio (%)	Working time (day)	Working time ratio (%)
Naked eye classification	Data preparation for SPOT5	10,000	37.1	–	0.0
	Preprocessing of SPOT5	4,000	14.8	15	21.4
	Reference data preparation (GIS map and registration map and etc.)	4,950	18.4	15	21.4
	Land use map construction (9unit in 1/25,000)	6,000 (based on the staff salary)	22.3	30 (with three persons)	42.9
	Field work	2,000	7.4	10	14.3
	Subtotal	*26,950*	*100*	*70*	*100*
Pixel based classification	Data preparation for SPOT5	10,000	47.7	–	0.0
	Preprocessing of SPOT5	4,000	19.1	15	25.0
	Reference data preparation (GIS map and registration map and etc.)	4,950	23.6	15	25.0
	Supervised classification	2,000 (based on the staff salary)	9.5	30 (with one person)	50.0
	Subtotal	*20,950*	*100*	*60*	*100*

Table 2
Categorical discrepancy between land registration map and discharged BOD NPS unit loads.

	Land use classification				
Category name of land registration map	Dry paddy-fields/orchard	Paddy-field	Forestry	Building site/factory site / school site/parking lot/gas station site/warehouse site/ road/railroad site/gymnastics site/ recreation site/religion site/historic area	Pasture/mineral spring site/saltern/bank/ rivers/ditch/marsh/ fish farm/water supply site/park/burial/ miscellaneous area
Category name of discharged NPS unit loads	Paddy fields	Upland	Forest	Impervious area	The others
Discharged BOD NPS unit loads (kg/km^2·day)	1.59	2.30	0.93	85.90	0.960

NPS unit-load in impervious area shows the largest unit loads and this is composed of 12 different land uses in land registration map. Because land use unit for impervious area is very influential to the total NPS pollution loads, more divided and intensive monitoring for NPS loads in the impervious area should be established.

7. A procedural analysis of calculating pollution loads

When calculating discharged pollution loads, changes of NPS loads are caused by how much percentage of areas are involved in the sewage collection district. We compare the conventional text based method with the GIS DB based process. Because the text-based statistic data is offered on the basis of administrative unit, areas of sewage collection district in the watershed may not be calculated precisely. And specific formula has been developed to estimate approximate sewage collection area. But this formula can't get exact area but arithmetic estimate. However, GIS based area calculation can calculate precise area by GIS overlay analysis. This sewage collection area has more sensitive impact on the total NPS loads because most of those regions are classified as an impervious area representing high discharged NPS unit-loads.

With reference to Table 3, BOD and TP pollution loads are 5,658.4 kg/day, 345.6 kg/day respectively when using text based data. Meanwhile these results are decreasing to 5,567.2 kg/day, TP 341.1 kg/day respectively when reflecting spatial distribution of sewage collection districts. This was caused by not only the area gap but also methodological differences of sewage disposal area calculation. An existing method applies the specific formula, which is building area + road/2, as shown Table 4, to estimate sewage collection area while area calculation

Table 3
A comparisons of pollution loads between text based result and GIS DB based result.

Municipal category	Basin unit	A text based result				GIS DB based result			
		Discharged BOD pollution loads		Discharged TP pollution loads		Discharged BOD pollution loads		Discharged TP pollution loads	
Youngdong county	Geombon C	134.8		11.1		134.8		11.1	
	Geumbon E	1,322.8		86.6		1,323.0		86.5	
	Geumbon F	23.4		1.0		23.4		1.0	
	Chogang A	1,669.6	3,150.6	102.2	200.9	1,668.9	3,150.1	102.2	200.7
Okcheon county	Geumbon F	2,442.1		140.4		2,351.5		136.1	
	Chogang A	65.2		4.3		65.2		4.3	
	Bocheong A	0.4	2,507.8	0.0	144.7	0.4	2,417.1	0.0	140.3
Total			5,658.4		345.6		5,567.2		341.1

Table 4
A comparison of applying area when calculating pollution loads between text-based result and GIS DB based result.

Municipal category	Land use Class 1	Land use Class 2	Sewage collection area (km^2)				Applying area (km^2)		Area gap ($c = a - b$)
			Text based DB		GIS-based DB		Text based DB (a)	GIS-based DB (b)	
Okcheon county	Dry paddy-fields		9.88		3.18				
	Paddy-fields		8.48		2.73				
	Impervious area	Building area	3.08						
		Road	3.37	83.60	4.49	17.14	Building area + road/2 = 4.77	17.14	–12.37
		The others	1.18						
	Forestry		48.89		5.12				
	The others		8.74		1.63				
Youngdong county	Dry paddy-fields		4.38		1.16				
	Paddy-fields		3.34		1.34				
	Impervious area	Building area	1.09						
		Road	1.00	45.01	2.00	4.91	Building area + road/2 = 1.59	4.91	–3.32
		The others	0.04						
	Forestry		31.46		1.42				
	The others		2.99		0.93				

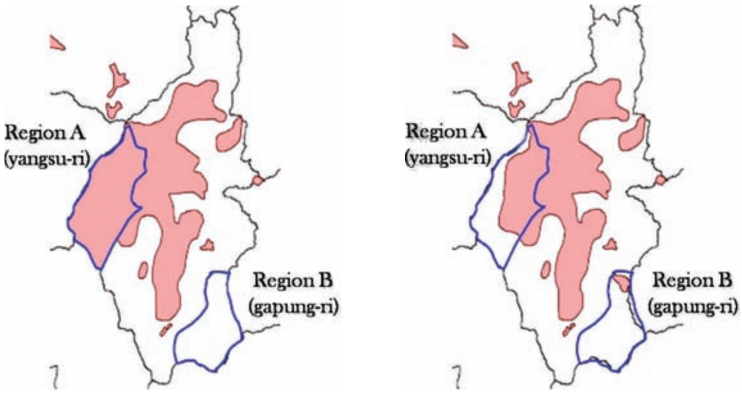

Fig. 3. (a) A text-DB based result. (b) A GIS DB based result.

Table 5
A comparison of separated pollution loads according to area calculation methods (kg/day).

Category	Region A (Yangsu-ri in Okcheon-county)				Region B (Gapung-ri in Okcheon-county)			
	Text-based results		GIS DB based results		Text-based results		GIS DB based result	
	BOD	TP	BOD	TP	BOD	TP	BOD	TP
Generated pollution loads	35.3	1.27	35.3	1.27	30.5	1.45	30.5	1.45
Sewer inflow pollution loads	25.4	0.62	33.6	1.07	0	0	2.8	0.09
Individually discharged pollution loads	9.9	0.65	1.7	0.2	30.5	1.45	27.7	1.36
Sewer overflow pollution loads	0.0	0.0	0.2	0.007	0	0	0.3	0.011
Discharged pollution loads	9.9	0.65	1.9	0.21	30.5	1.45	28.7	1.36

method using GIS overlay analysis calculate the real sewage collection area which reflects various land uses such as paddy field, upland, forest, urbanized area and etc.. In this context, the method using GIS overlay analysis represents real land use of sewage collection area precisely. Pollution loads resulted from the method using GIS overlay analysis shows a bit lower result than the text-based result, which may mean an increase in treated sewage.

8. A difference of spatial distribution of sewage collection district according to data sources

Generally, most regional or basin scale natural resource management problems must be addressed at a finer scale [3], we focused on the small region in Ok-cheon county. While area A and B are represented 100%, 0% sewage collection area according to the text-based methods (Fig. 3(a)), result based on the GIS overlay analysis (Fig. 3(b)) shows 53%, 4% sewage collection area respectively in the area A and B. The larger sewage collection area has, the greater sewer inflow pollution loads are generated. And this cause less discharged pollution loads. On the contrary, the less sewage collection area calculated, the greater individually discharged pollution loads is generated, which induce higher discharged pollution loads. When the method is changing from a text DB to GIS DB, BOD discharged pollution loads are decreasing from 9.9 to 1.9 in the region A and from 30.5 to 28.7 in the region B (Table 5).

9. A sensitivity analysis of discharged NPS loads at the different imperviousness ratio

As a result of this study, the amount of NPS pollution loads which were calculated using GIS in Youngdong-county and Okcheon-county decreased by about 1.6%, 3.6% respectively when compared to when calculations were done with existing methodology. To be more specific, NPS loads are largely affected by 5 kinds of land use types such as paddy fields, farm, forestry, urbanized areas and the others. Especially the urbanized area was identified as the most sensitive region to total NPS pollution loads. However, due to differences of discharged NPS unit-loads, changes of impermeable area rate had more significant impacts on the individual pollutant loads in forest regions than in urbanized areas.

Ideally, the best way to reflect real land use is creating the land use referencing SPOT5 based on the geographic

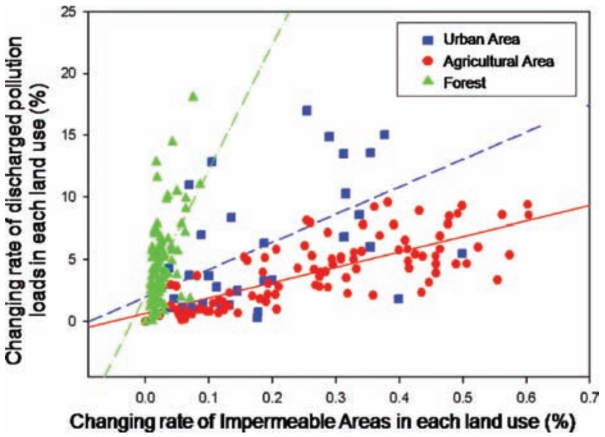

Fig. 4. Changes of pollution load ratio caused by impermeable area ratio in three kinds of land use types.

data but it is considered to be an uneconomic and inefficient method. To solve this problem, if it has more critical impact on NPS loads, we should focus on accurate calculations of the sensitive region. In this context, we should increase the accuracy of land use classification in urbanized area than the other area. Considering the high effects of urbanized land use on the NPS pollution loads, more delicate classifications of discharged NPS unit-loads for this area should be attained.

10. Discussions and further studies

This study applied to two counties which are Youngdong county and Okcheon country nearby the Dae-chung Dam which supplies over 1,000,000 ton drinking water per day. NPS pollution is considered to be a significant issue in those regions. As requests for GIS-based data processing are getting increasing, various approaches for calculating each land use have been considered. In this study three kinds of data resources such as SPOT5 remote sensing image, registration map and cadastral surveying result were used. And this study is mostly focused on the comparison of area calculation result between existing cadastral surveying and registration map referencing of SPOT5 satellite and GIS digital map. The studies show that a land use map referencing of SPOT5 image has more similarities with conventional method based on the cadastral surveying than land registration maps. However, this was mainly caused by similarities in the timing of the surveys and the geographic scope of comparison which was done on a large scale. Even though we used the SPOT5 image as reference data only in this study, it would be more useful when it apply to simulation of watershed hydrological and water quality models using demanding GIS-based data set.

Mostly, differences of NPS loads are caused by procedural changes of basin area calculation. It was converted into basin pollution loads after calculating loads based on the municipal administrative unit before, however, when using GIS-based method, conversion process from municipal based result to basin based result is not required. Furthermore, conventional method which is based on the statistic data apply the specific formula, building area + road/2, to get sewage collection area though it was not exact estimate. Therefore, it is essential to use GIS data to reflect the distributions of influential factors more accurately such as pollution, development plans and wastewater treatment facilities. Furthermore it is possible to have watershed based result directly by using overlay analysis with GIS DB without data conversion from administrative unit to basin unit.

Even though the area gap and pollution loads are relatively small between those data sources, discrepancy is significant enough when considering site is small and undeveloped area which has small sewage collection area. Therefore, applying GIS-based area calculation on larger and more urbanized basins will bring more influential and dominant impact on NPS loads.

When it comes to NPS unit loads, it decreases the efficiency of calculation of land use. In this study, even though we try to improve accuracy of each area calculation of 29 land use, we only have five NPS unit loads. Therefore we classified 29 categories (28 land registration map categories and 1 golf courses category) into five categories. Even though national long term investigation on the detailed NPS unit loads is undergoing, it is imperative to establish a standard for matching monitoring sites for NPS unit loads category with conventional registration map category.

References

[1] J-S. Lee, Y-S. Choi, C-W. Seo and C-D. Jeon. Development of an object-oriented framework data update system, Journal of the Korean Association of Geographic Information Studies, 11(1) (2008) 31–44
[2] S-M. Kim, J. Heo, Y-H. Lee, S-H. Kim and S-S. Kim, Introduction of ISO land administration domain model (LADM) to Korea Land Information System(KLIS), Journal of the Korean Society for Geo Spatial Information System, 17(1) (2009) 79–87.
[3] J.R. Jenson, Remote Sensing of the Environment: An Earth Resource Perspective, Prentice Hall, New Jersey, 2000.
[4] J.R. Anderson, , E.E. Hardy and J.T. Roach, , A land use and land cover classification system for use with remote sensor data, Geological Survey Professional Paper 964, US Government Printing Office, Washington, DC, 1976.
[5] D. Seok Shin, Basic monitoring concept for revised unit-load on NPS, Journal of Korean Society on Water Quality, 23(4) (2007) 429–433.
[6] Ministry of Environment, Construction of Land Cover Map Using Satellite Imagery Data, 3rd report, 2002.
[7] M-H. Park and M.K. Stenstorm, 2006, Using satellite imagery for stormwater pollution management with Bayesian networks, Water Research, 40 (2006) 3429–3438.