Application of the High Resolution Aerial Images to Estimate Nonpoint Pollutant Loads in the Unit Load Approach to Total Water Pollutant Load Management System in South Korea

Bum-Yeon Lee*, Chang-Hee Lee**

* Department of Environmental Engineering and Biotechnology, Myong-Ji University, San 38-2, Namdong Yong-In City, Gyeonggi-Do, 449-728 South Korea, E-mail : armdry83@ nate.com
** Department of Environmental Engineering and Biotechnology, Myong-Ji University, San 38-2, Namdong Yong-In City, Gyeonggi-Do, 449-728 South Korea, E-mail : changhee@ mju. ac.kr

Abstract

Unit load approach is currently used for the estimation of nonpoint source (NPS) pollutant load in total water pollutant load management system (TWPLMS) of Korea. It is based on land register data; however, it cannot always reflect the actual land surface coverage. This study presents the NPS pollutant loads estimation methods according to real land coverage using high resolution (0.4m) image in Kyeongan watershed, South Korea. Pearson's correlation and linear regression methodology was used to analyze the relationship between image data and land register data. Composition of land categories was analyzed using normal distribution methods and 95% probability value was selected for compositions of actual land use. For analysis of unit load categories pattern; the school, university, apartment complex and Golf course have 65%, 45%, 80% and 5% of ground characteristics respectively. From the results, 15 categories showed significant linear correlation at 0.01 levels. It was observed that application of linear relationships as conversion factor of land register to image data can be a possible option to improve the unit load method for the specific categories. The difference in unit load can also be applied using the portion of each unit load categories in land use. These improved methods can help to increase the accuracy in TWPLMS.

Keywords : Nonpoint source, Total water pollutant load management system, Unit load, High resolution image

1. Introduction

Non-point Source (NPS) pollution is affected by variety of factors like antecedent dry day, rain intensity, land use, topography and others, therefore, management of NPS is difficult and limited. NPS pollution flows directly into the public water systems such as stream, lake, and coastal without management (Choi and Shin, 2002). NPS pollutant load vary differently according to land use or land cover, therefore, it is difficult to estimate the load technically for NPS management (Go et al., 2009; Kim et al., 2003; Shin et al., 2001, WERF, 2003). Total Maximum daily load in USA is used for various methods to estimate NPS pollutant load such as pollutograph based on measured data, watershed model, statistical analysis, and unit load approach (USEPA, 2008; WERF, 2003).

In Total Water Pollutant Load Management System (TWPLMS) technical guidelines of Korea, estimation of NPS pollutant loads from the watershed is recommended using hydrograph and pollutograph that is derived by continuous measured flow and water quality data (NIER, 2008). For the reliability of hydrograph and pollutograph, long term monitoring data is required to derive the graph. However, monitoring data is not enough for making reliable graphs in Korea, therefore, unit load method determined by land use, area of land categories and rainfall discharge ratio is currently used for estimation of NPS pollutant load in TWPLMS.

The unit load method is easy to apply and to understand the calculation of NPS pollutant load. In TWPLMS, it is used to record the land register data for calculating the area of each land category. However, land register data does not always reflect the actual land use that determines runoff characteristics of NPS pollution. Also, limited number of unit loads (5 to 7 unit loads among 28 land register categories) was used to estimate the NPS pollutant loads. Therefore, unit loads inevitably have to represent several land use categories regardless of their actual land coverage characteristic (table 1).

Table 1. Applicable unit loads for the specific land-use categories in TWPLMS (MOE, 2006; NIER, 2008)

Land Use	BOD (kg/㎢/d)	TN (kg/㎢/d)	TP (kg/㎢/d)
Rice paddy	2.30 (2.30)*	6.56 (6.56)	0.61 (0.61)
Dry paddy	1.59 (1.60)	9.44 (9.44)	0.24 (0.24)
Forestry	0.93 (1.0)	2.20 (2.20)	0.14 (0.14)
Ground	85.90 (85.90)	13.69 (13.69)	2.10 (2.10)
Others	0.960 (1.0)	0.759 (0.06)	0.027 (0.03)
Pasture	- (35.1)	- (5.37)	- (1.72)
Golf course	- (1.0)	- (3.56)	- (2.76)

Notes 1) Dry paddy includes the dry paddy and orchard.

2) Ground includes the ground, factory, school, road(except the road slope), railroad (except railroad line), physical exercise(except golf course and skiing ground), recreation ground, religious site, historical site, parking lot, gas station site and warehouse site.

3) Others includes the mineral spring, salt marsh, embankment, ditch, pond, water supply site, park, burial ground, miscellaneous land, river, fish farm, railroad line and road slope.

4) Unit loads for golf course and skiing ground depend on the actual land use characteristics.

*( ) : Previous (1st step TWPLMS in Han river) unit loads (before 2009)

As a way to minimize this problem, Ministry of Environment in Korea tried to study two kinds of approaches, one is developing the subdivision unit load categories more than current 7 unit load categories to reflect the land use or land cover characteristics (Han River basin management committee, 2007), and the other is estimation of NPS pollutant loads by the subdivision of each land register categories to current unit load categories reflect real land cover situation using high resolution image. This study represents the estimation method of the NPS pollutant loads using second approach that derives to establish quantitative statistical relationship between land registration map and high resolution (0.4m) image which reflect actual land surface coverage, and analyze the unit load categories pattern in land use of school (elementary·middle·high school), university, apartment complex, and golf course using the high resolution

image and to suggest the improvement method for estimation of NPS pollutant loads in Kyeongan watershed, Korea.

2. Study Area

The Kyeongan watershed is located in Yongin and Gwangju cities, Gyeonggi-Do, South Korea (figure 1). The watershed is divided into two subunits; Kyeongan-A (198.5km^2) and Kyeongan-B (249km^2) unit watersheds for TWPLMS. Kyeongan stream is one of the three main tributaries of Paldang Reservoir, which supplies drinking water to the 23 million residents in metropolitan area of Seoul and Gyeonggi Province.

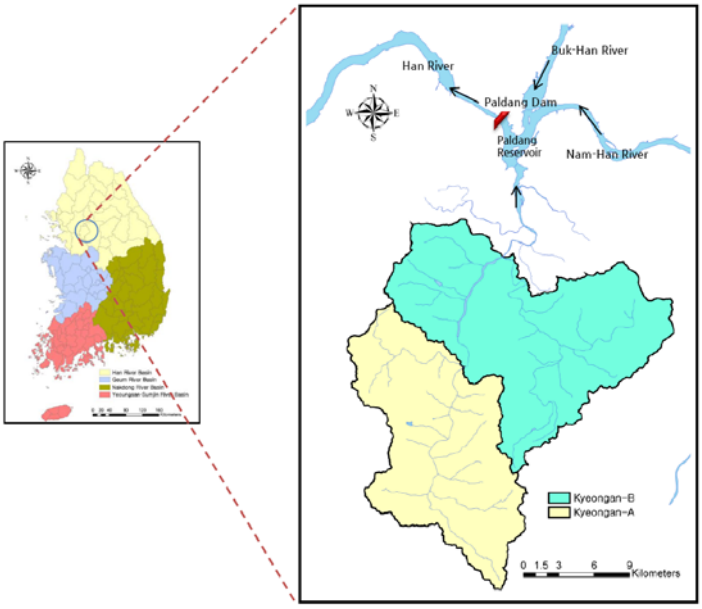


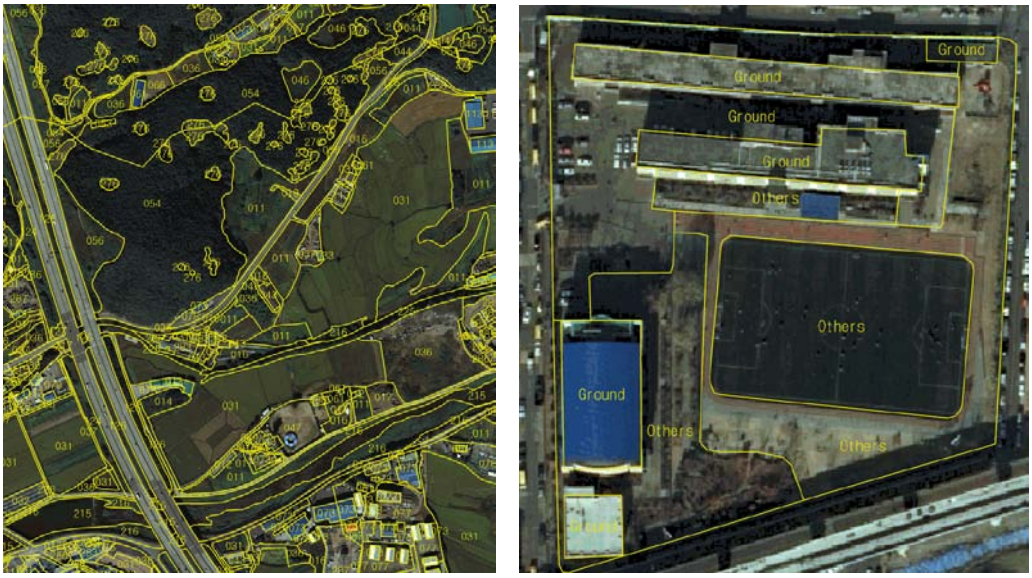
Figure 1. Study area

The Kyeongan watershed have 30 Eup·Myeon·Dong (Unit of administrative district in Korea), and this area has 26 land register categories among 28 except of mineral spring and salt marsh. However, 9 categories do not exist such as, railroad, recreation ground, historical site, embankment, fish farm, water supply site, and park. For the analysis of unit load categories pattern, 4 land use sites such as school (49), university (7), apartment complex (12), and golf course (15) have been selected considering the available data for statistical analysis, to analyze the land cover pattern using high resolution image in the study area.

3. Methods

In this study, two approaches were used to estimate the NPS pollutant loads to improve the unit load method, first was analyzing the relationship between image data (actual land use) and land register data in 19 land register categories, and second was statistical analysis of unit load categories patterns in 4 land use categories.
For these approaches, a high resolution (0.4m resolution) aerial image was used to represent the actual land use and was classified in 29 medium categories, and 218

subdivision categories. The medium categories were classified considering Korean land register and TWPLMS unit load categories and then subdivided according to the land cover patterns (e.g. facilities area, grass lands, bare grounds, forests, waters, roads, parking lots, farms, and paddy fields, etc). This classified actual land use map was reclassified into 28 land register categories considering actual land cover characteristics for correlation with land register data. For the analysis of unit load categories, four land use sites were subdivided according to unit load categories such as ground (includes facilities area, roads, and parking lots), forest, farms, paddy fields, golf course, and others (includes bare ground, grass lands, and waters) using land cover map classified by high resolution image.



(a) Classification of actual land use

(b) Classification of Unit load categories

Figure 2. Example of classification of image data

The Pearson's correlation method is generally used for computing the correlation among parameters in a numerical form. In this study, it was used to analyze the quantitative relationship between image data and land register data. Linear regression method was also used to show the mathematical functional formula about the relationship between image data and land register data. For the comparison between image data and land register data, percentage of each land register categories was used instead of area, because there was a difference in total area between image data (447.5 km^2) and land register data (455.8 km^2). This difference did not reflect to physical condition of current land, mismatch in the boundary between image data and land register data, error of starting point, and reduced scale during the making the land registration map.

In this research, normal distribution was used to analyze the composition of unit load categories. It is a continuous probability distribution that is often used as a first approximation to describe real-valued random variables that tend to cluster around a single mean value (Wikipedia). Kolmogorov-Smimov and Shapiro-Wilk methods (Kang, 2005) are used for verification of normal distribution. If the land use has normal distribution, 95% probability value was selected for compositions of each unit load category considering the error. Statistical Package for the Social Science (SPSS) version 17.0 program is used to statistical analysis.

4. Results and Discussion

4.1 Results of comparison between image data and land register data

From the estimation of area, targeted 19 land categories have 99.9% image data and 99.6% land register data (table 2). From the results, forestry, road, and river land categories showed similar area between image data and land register and they also showed identical relationship between image data and land register (figure 3). It means that these categories were reflected by actual land use.

Table 2. Area and portion of each land category

Land category	Image (km^2)	Land register (km^2)	Land category	Image (km^2)	Land register (km^2)
Dry paddy	27.264(6.09)[a)]	35.838(7.86)	Warehouse site	1.287(0.29)	3.063(0.67)
Orchard	0.145(0.03)	0.443(0.10)	Road	14.675(3.28)	12.279(2.69)
Rice paddy	22.675(5.07)	40.087(8.80)	River	9.201(2.06)	12.527(2.75)
Pasture	0.784(0.18)	1.656(0.36)	Ditch	1.773(0.40)	5.252(1.15)
Forestry	273.466(61.11)	286.987(62.97)	Pond	1.426(0.32)	0.546(0.12)
Ground	9.260(2.07)	19.015(4.17)	Physical exercise	0.157(0.04)	14.391(3.16)
Factory site	4.755(1.06)	9.476(2.08)	Religious site	0.145(0.03)	0.564(0.12)
School site	0.331(0.07)	1.689(0.37)	Burial ground	10.149(2.27)	2.664(0.58)
Parking lot	17.297(3.87)	0.069(0.02)	Miscellaneous land	52.138(11.65)	7.322(1.61)
Gas station site	0.035(0.01)	0.103(0.02)	Others [b)]	0.508(0.10)	1.813(0.40)
Total area				**447.471**	**455.785**

[a)] () : Area portion (%), [b)] Others : Excepted 9 land category's area from analysis

However, other 16 land categories showed difference in area between image data and land register data. Parking lot, pond, burial ground, and miscellaneous land has larger area in the image data than land register data. And other 12 land categories (dry paddy, orchard, rice paddy, pasture, ground, factory site, school site, gas station site, warehouse site, ditch, physical exercise, and religious site) showed that land register data has lareger area than image data (table2, figure 3).

The reason of difference between image data and land register is reclassification of land categories according to characteristics of actual land cover or land use from/to other land categories in image map. For the analysis of land use by image data, all land categories have difference between actual land use (image data) and adminstrative land use (land register data). It means that land register data which is used to estimate the NPS pollutant load does not reflect the real land use and it is possible to decrease the accuracy of calculated NPS pollutant bad.

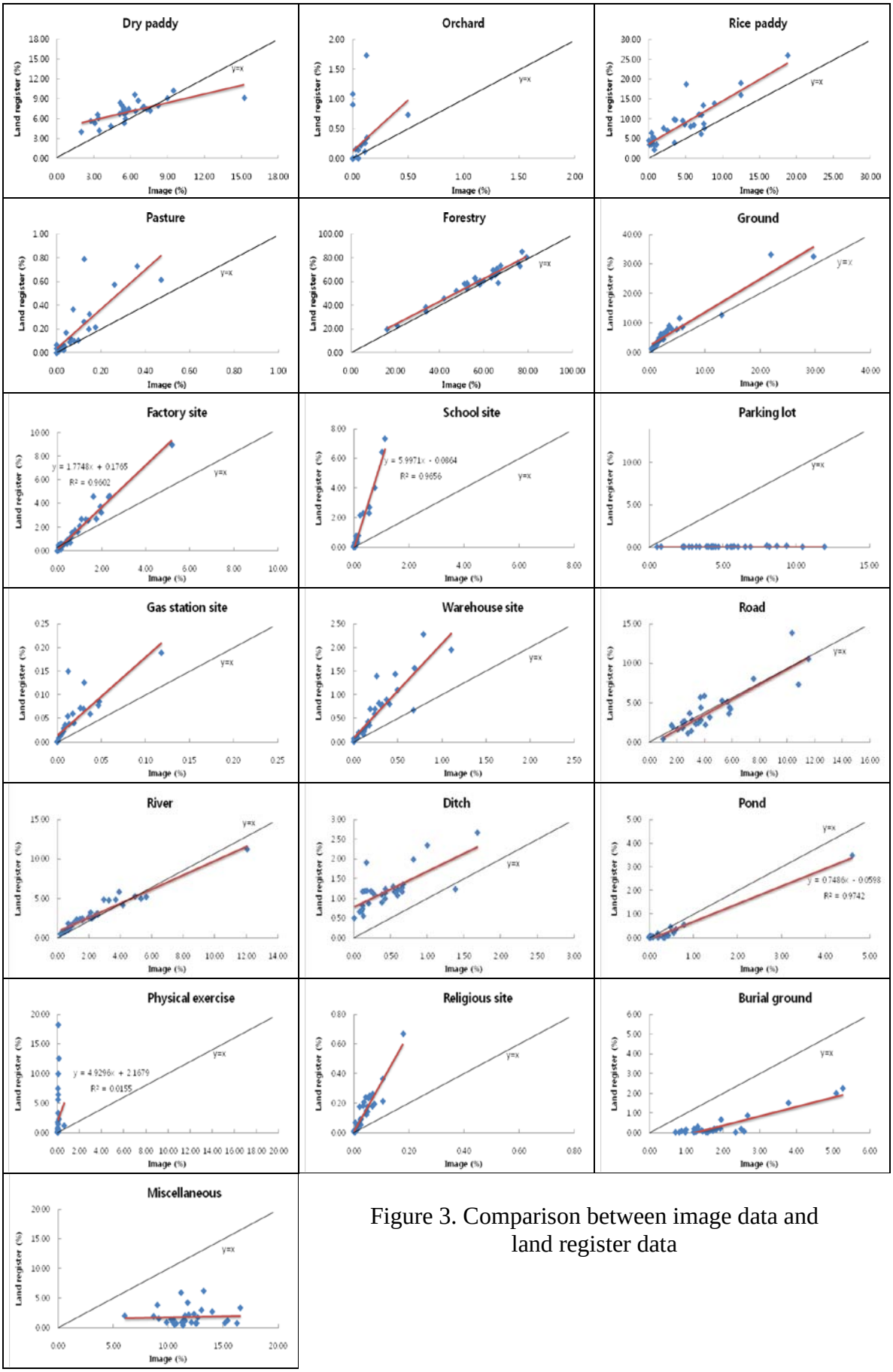


Figure 3. Comparison between image data and land register data

Table 3 shows the results of the Pearson's correlation between the image data and the land register data, and linear regression. Pond, school site, factory site, forestry, river, and ground land categories have high correlation coefficient (r). Among 19 land categories, 15 land categories have significant correlation at the 0.01 level that shows regularity relationship between image data and land register data, but 4 land categories (orchard, parking lot, physical exercise site, and miscellaneous land) have less correlation, that is significant over than 0.01 level.

Table 3. Results of statistical relationship between image data and land register data

Land category	r	y=Ax+B (%)	Land category	r	y=Ax+B(%)
Dry paddy	0.718**	y=0.434x+4.526	Warehouse site	0.890**	y=1.983x+0.108
Orchard	0.412*	y=1.729x+0.122	Road	0.881**	y=0.958x-0.284
Rice paddy	0.873**	y=1.078x+3.717	River	0.968**	y=0.898x+0.799
Pasture	0.828**	y=1.665x+0.036	Ditch	0.727**	y=0.915x+0.782
Forestry	0.976**	y=0.959x+4.995	Pond	0.987**	y=0.749x-0.060
Ground	0.966**	y=1.135x+2.293	Physical exercise	0.125	y=4.930x+2.168
Factory site	0.980**	y=1.775x+0.177	Religious site	0.929**	y=3.306x+0.011
School site	0.983**	y=5.997x-0.086	Burial ground	0.913**	y=0.477x-0.580
Parking lot	0.353	y=0.005x-0.001	Miscellaneous land	0.047	y=0.031x+1.510
Gas station stie	0.829**	y=1.656x+0.014			

** : Correlation is significant at the 0.01 level. * : Correlation is significant at the 0.05 level.

Therefore, it is analyzed that categories with identical relationship can be directly used to land register data, and categories with significant relationship can be used as a conversion factor from land register to image data to estimate the NPS pollutant loads in Kyeongan watershed.

4.2 Results of composition of unit load categories analysis

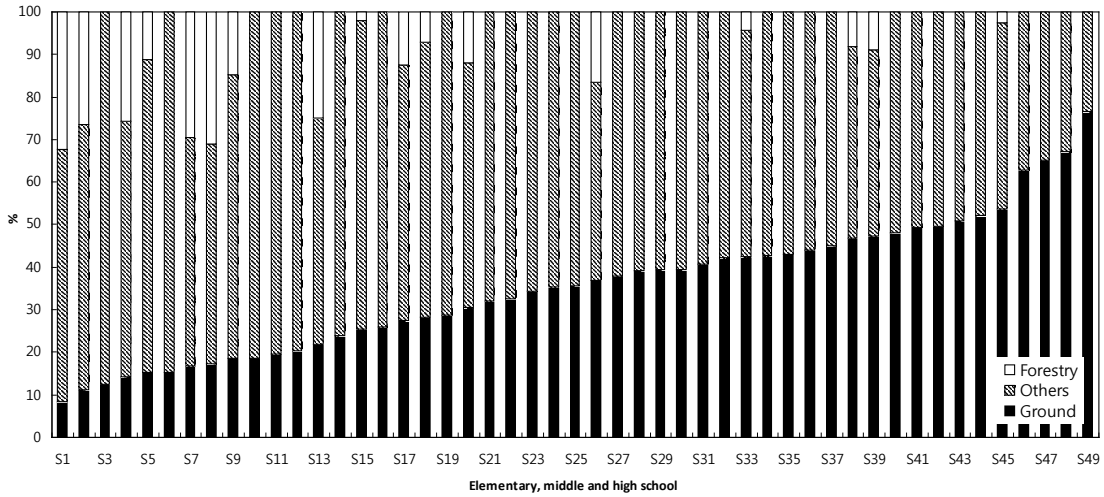


Figure 4. Relative portions of applicable unit load categories area in the Elementary·middle·high school sites

Figure 4 to figure 7 represents the results of estimated portions of current unit load categories in four land use based on high resolution image data. School sites currently

considered as 100% ground in Korea TWPLMS has approximately 10 to 80% of ground and 20 to 80% of others (figure 4) because school sites contains play ground and garden area.

For the university site, the ground unit, forestry, and other units showed 16.11 to 42.63%, 4.99 to 29.97%, and 32.89 to 67.77% (figure 5) instead of currently considered as 100% ground. Almost of all universities have approximately 40% ground, because university sites also contain the play ground, garden, open space, and forest.

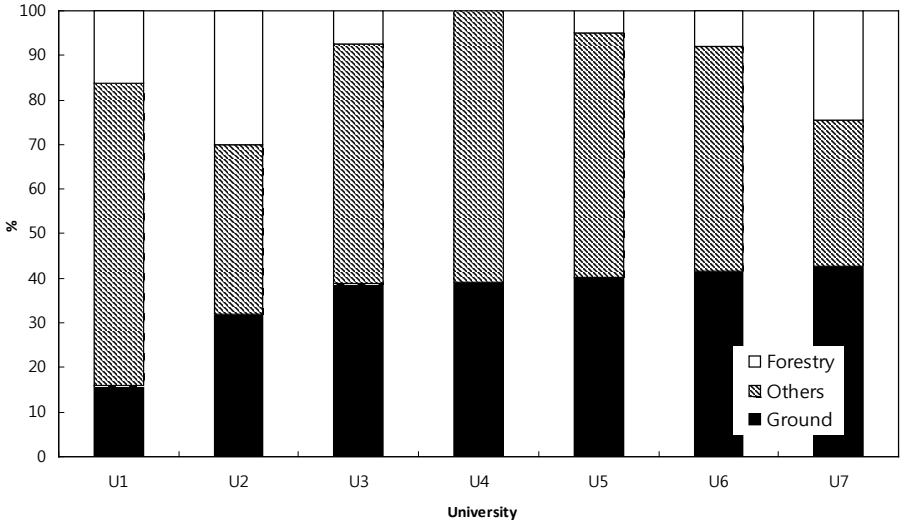


Figure 5. Relative portions of applicable unit load categories area in the university sites.

Apartment complex currently cosidered as 100% ground showed 51.20 to 80.50 % ground and 18.18 to 48.80% others (figure 6). This result is reflected because construction of multifamily housing should contain more than 30% green belt area by regulation of house construction standards (MLTM, 2010), therefore, apartment complex cannot exceed the 70% ground of land use. However, some of analyzed apartment complex exceeded 70% ground due to smaller area than classification size (10m x 10m) by image data, or change in the land use after construct permission.

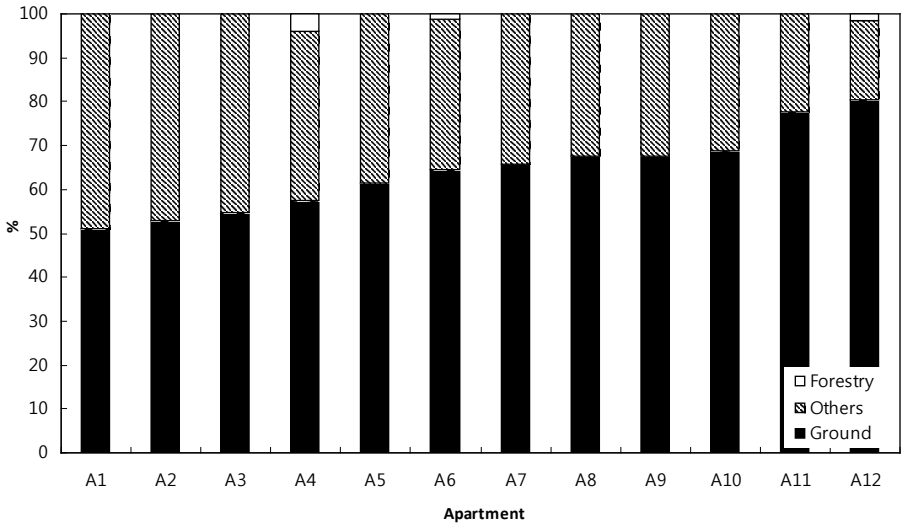


Figure 6. Relative portions of applicable unit load categories area in the apartment complex sites.

Golf course is currently considered as 100% golf course (before 2009 in Han River TWPLMS, Golf course include the corressponding with real land use characteristic after 2009.). It includes other land use sites such as golf cart road, club house, parking lot, and forestry, and these are classified as forest, ground, and others unit load categories. Results showed that 45.15 to 91.88 % of golf course land use was found instead of 100% golf course because it has 3.84 to 8.83% ground, 0.00 to 46.17% forest, and 0.24 to 7.31% other unit load categories (figure 7).

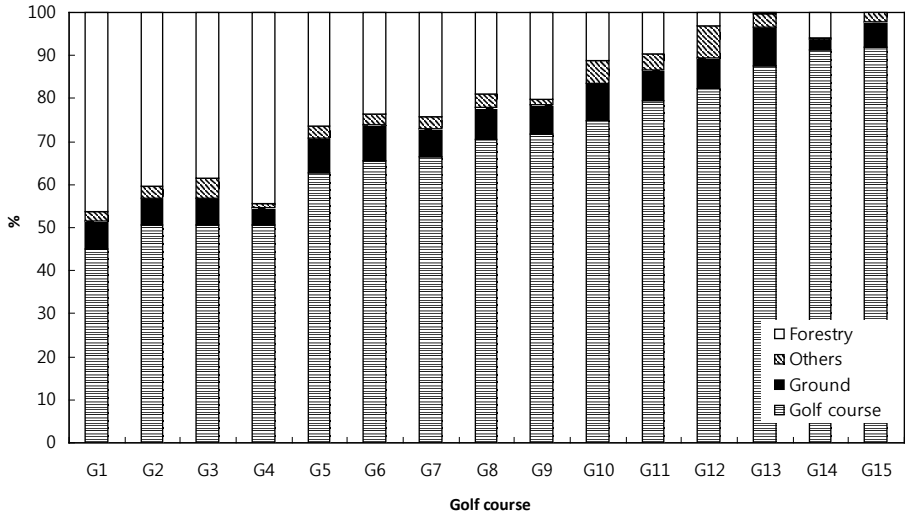


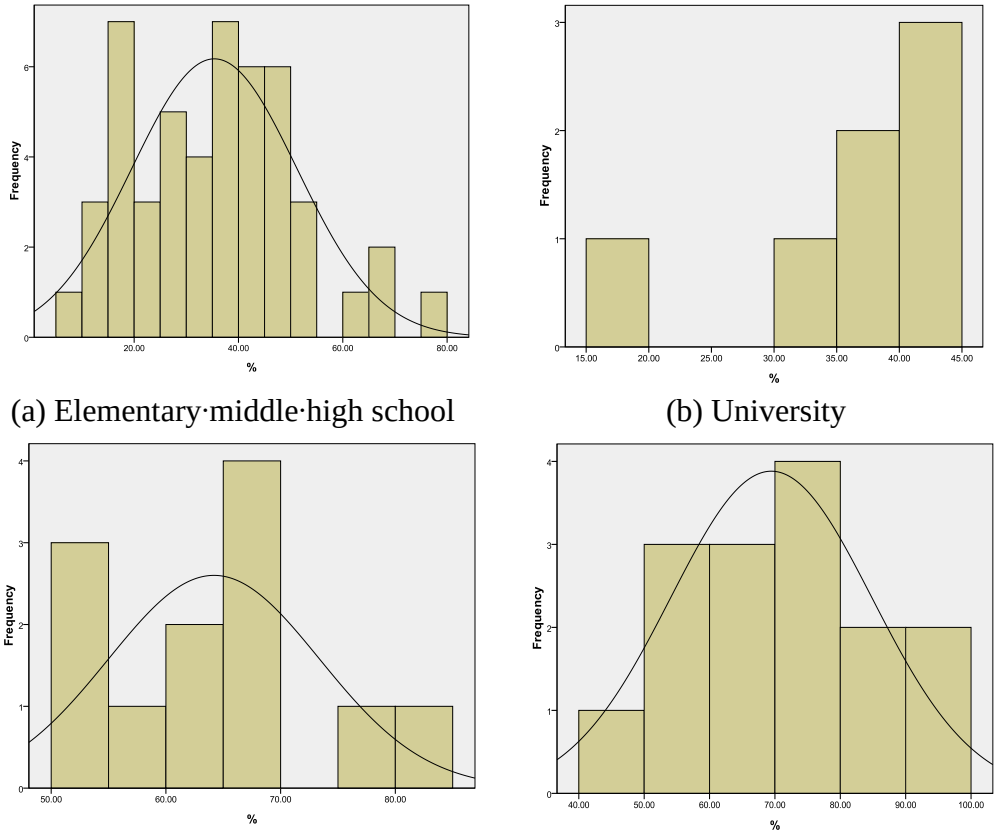
Figure 7. Relative portions of applicable unit load categories area in the golf course sites.

From the results, it is observed that NPS pollutant load can be overestimated in school, university, and apartment complex and golf course site can be underestimated. Therefore, there is a need to estimate the NPS pollutant load that each land use is subdivided by actual land use or land cover, and is applied appropriately the unit load categories to subdivide area for increasing the accuracy.

Result of statistical analysis is shown in table 6. School, apartment complex, and golf course showed a lower bound of the true significance for normal distribution; whereas, university did not show normal distribution from verification analysis due to lower significant level (less than 0.05).

Table 6. Statistical verification of normal distribution

	Kolmogorov-Smirnov Method		Shapiro-Wilk Method	
	Statistic	Sig.	Statistic	Sig.
Elementary·middle·high school	0.074	0.200*	0.973	0.311
University	0.345	0.012	0.738	0.010
Apartment complex	0.142	0.200*	0.951	0.655
Golf course	0.154	0.200*	0.941	0.399
*. This is a lower bound of the true significance.				



(a) Elementary·middle·high school (b) University

(c) Apartment complex (d) Golf course

Figure 8. Histogram of four land use

Table 7 shows the results of general statistical analysis and 5% rejection value of ground unit load category in school, university, and apartment complex, and golf course unit load category in golf course land use. The mathematical mean, standard deviation and 5% rejection values in school site were found 35.37%, 15.83 %, and 61.41% respectively. School land use needs to improve the application of unit load by applying 60 to 70% (65%) ground and 30 to 40% (35%) other unit load categories for improved method.

Table 7. Result of statistical analysis

	Elementary·middle·high school	University	Apartment complex	Golf course
Mean	35.37	35.80	64.19	69.44
Median	35.51	39.18	65.15	70.62
Variance	250.48	87.56	84.70	237.60
Std. Deviation	15.83	9.36	9.20	15.41
Minimum	8.40	16.11	51.20	45.14
Maximum	76.60	42.63	80.50	91.88
Skewness	0.390	-1.998	0.311	-0.079
Kurtosis	-0.202	4.009	-0.543	-1.139
5% rejection value	61.41%(ground)	-	79.33%(ground)	94.80% (golf course)

For the result in apartment complex, mean, standard deviation and 5% rejection values were found 64.19%, 9.20%, and 79.33% respectively. Apartment complex land use needs to improve the application of unit load applying 75 to 85% (80%) ground and 15 to 25% (20%) other land use categories considering regulation and study results. The mean, standard deviation, and 5% rejection values were found 69.44%, 15.41 %, and 94.80% in the golf course unit load category. Although golf course land use do not have regulation, there is need to improve the unit load approach, because ground land use like the club house, parking lot, and management facilities is required in golf course. From the results, golf course can be applied for 90 to 100% (95%) golf course and 0 to 10% (5%) others.

University site has poor and few data for statistical analysis therefore, it is difficult to suggest improvement method. But, ground portion of university showed convergence approximately level of 45% (figure 9). Therefore, university needs to improve the application of unit load applying 40 to 50% (45%) ground and 50 to 60% (55%) other land use categories.

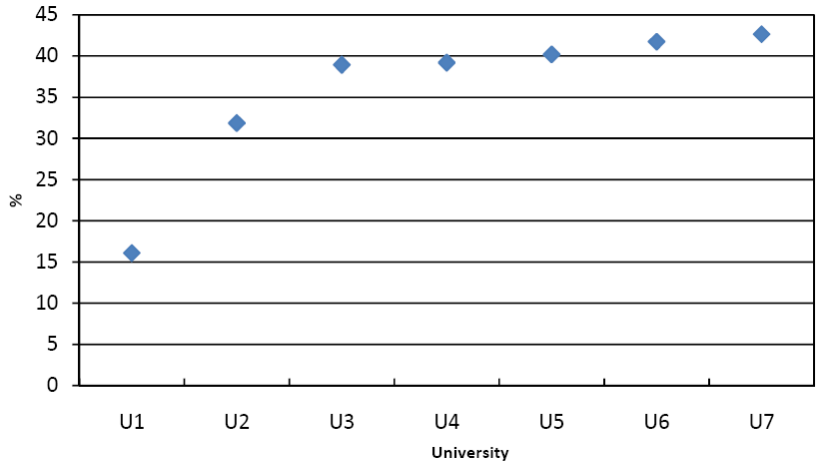


Figure 9. Ground category ratio in the university

Overall, it can be applied to improve unit load approach for 65% ground and 35% others in school, 45% ground and 55% others in university, 80% ground and 20% others in apartment complex, and 95% golf course and 5% ground in golf course land use. Table 8 shows the comparison of estimation of generation load between present method and recommended method. School, university, and apartment complex which are currently considered as 100% ground, are overestimated about 34.6%, 54.4%, and 19.8% in present method, respectively. The other hand, golf course which is considered as 100% golf course in present method, is underestimated approximately 19.1% in present method.

Results of estimation of NPS pollution generation loads, university and golf course have difference more than 50kg/day beween present and recommended method. Diffenrence of 50kg/day is similar amount of approximately 1,000 person's generation load (estimatd by TWPLMS technical guideline). It means that it is large amount that affect the development about 250 households (4 persons/household). Therefore, it should be require to improve the current unit load method considered as this study results.

Table 8. Differences of applied unit loads and estimated NPS pollution generation loads

	Present method		Recommended method	
	Applied loads	Total loads	Applied loads	Total loads
Elementary·middle·high school	Ground 100%	73.56 kg/d	Ground: 60-70%(65%) Others: 30-40%(35%)	48.11 kg/d
University	Ground 100%	108.75 kg/d	Ground: 40-50%(45%) Others: 50-60%(55%)	49.63 kg/d
Apartment complex	Ground 100%	18.92 kg/d	Ground: 75-85%(80%) Others: 15-25%(20%)	15.18 kg/d
Golf course	Golf course 100%	14.07 kg/d	Golf course: 90-100%(95%) Ground: 0-10%(5%)	73.80 kg/d

From the results, this study suggests that use of high resolution aerial images is a better way to estimate nonpoint pollutant load. If the aerial maps are not available, application of linear relationships as conversion factors of land register to image data can be a possible option to improve the unit load method for the specific categories in Kyeongan watershed. Also, if there have regular patterns of unit load categories in each land use, it can be recommended that apply the difference unit load using the portion of each unit load categories in land use. However, it should take note that recommended method is just site specific. So, it can be applied restrictively to estimate the NPS pollutant load in Kyeongan watershed.

5. Conclusion

This study is carried out to establish quantitative statistical relationship between land registration map and high resolution (0.4m) image which reflect actual land surface coverage, and to analyze the unit load categories pattern in 4 land use by the high resolution image for improvement of present unit load method.

From the results of comparison between image data and land register data, forestry, road, and river categories were found to reflect the actual land use because they showed identical relationship between image data and land register. No correlation was found for orchard, parking lot, miscellaneous land, and sports utility site categories. However, other remaining categories had significant linear correlation at 0.01 levels, particularly pond (0.987), school (0.983), and factory site (0.980) relatively had high correlation coefficient.

For the analysis of unit load categories pattern, the school, university and apartment complex currently considered as 100% ground have only 65%, 45%, and 80% of ground characteristics, respectively. Golf course, which is considered as 100% golf course, has about 5% of ground area. This indicates that the unit-load method using in TWPLMS can give overestimation of non-point pollutant loads for the school, university, and apartment complex (19.8 to 54.4%), but underestimation for the golf course (80.9%).

Overall, this study suggests that use of high resolution aerial images is a better way to estimate nonpoint pollutant load. If the image is not available, application of linear relationships as conversion factors of land register to image data can be a possible option to improve the unit load method for the specific categories in Kyeongan

watershed. Also, if there are regular patterns of unit load categories in each land use, it can be recommended that difference unit load can be applied using the portion of each unit load categories in land use. These recommendations can help to increase the TWPLMS accuracy.

References

Choi J.Y. and Shin C.M. (2002). Management of nonpoint pollution by reducing storm runoff, Korea environment institute, KEI-2002-RE-03.

Go S.H., Memon S.A., and Lee C.H. (2009). Characterization of Stormwater Pollutants and Estimation of Unit Loads for Road and Parking Lot in Gyeongan Stream Watershed, Journal of Korean society on water quality, **25**(5), 689-696.

Han River basin management committee (2007). Long term monitoring and control measures of nonpoint source pollution from major land uses.

Kang B.S. (2005). Multivariate statistics using SPSS, Hankyungsa.

Kim J.E., Ahn K.H., and Lee C.M. (2003). A study on the feasibility of residential site development in the second district unit plan area, Journal of Korea planners association, **38**(2), 203-216.

Ministry of environment (2006). Establishment guideline of Han River total water pollutant load management system.

Minitry of land transport and maritime affairs (2010). Regulation of house construction standard.

National institute of environmental research (2008). Total water pollutant load management system technical guideline.

Shin E.S., Choi J.Y., and Lee D.H. (2001). Estimation of Non - Point Source Pollutant Unit - loads in Surface Runoff Considering Land - use and Basin Characteristics, Journal of Korean society on water quality, **17**(2), 137-164.

US Environmental protection agency (2008). Handbook for developing watershed TMDLs, USA.

Water environment research foundation (2003). Navigating the TMDL process: evaluation and improvements, USA.

Wikipedia, http://en.wikipedia.org/wiki/Normal_distribution