STUDY ON THE RELATIONSHIP BETWEEN LANDSCAPE CHARACTERISTICS OF FRAGMENTED URBAN GREEN SPACES AND DISTRIBUTION OF URBAN BUTTERFLIES - APPLICATION OF OBJECT-BASED SATELLITE IMAGE ANALYSIS -

S. Yokota [a,b] *, K. Takeuchi [b]

[a] Shimizu Corporation, Institute of Technology, 3-4-17 Ettyu-jima Koto-ku Tokyo, Japan - s-yokota@shimz.co.jp
[b] Graduate School of Agricultural and Life sciences, University of Tokyo, 1-1-1 Yayoi Bunkyo-ku Tokyo, Japan - (aa67131, atake)@mail.ecc.u-tokyo.ac.jp

KEY WORDS: satellite images, Landscape indices, butterfly assemblages, Classification and Regression Trees

ABSTRACT:

The purpose of this study is to clarify the relationship between the distribution of small-scale green spaces and butterfly assemblages in urbanizing areas located north-east of Ichikawa City. Landscape indices of green cover distribution with a diversity index H of vegetation types and a number of different vegetation types, GCI (Green Cluster Index) of each vegetation type and average minimum distance from each vegetation type were calculated in different ranges from survey root units, using high-resolution green cover data made by object-based classification of Quickbird multi-spectral image. Analyses on the relationship between the richness (number of species, Shannon-Weaver's H' and combination of species types) of butterfly species in survey root units and the landscape indices around them, were executed by using Classification and Regression Trees (CART). The forest (interior) species were thought to be appropriate indicators of the width of butterfly assemblages, and the gathering of evergreen woods, deciduous woods and bamboo groves in small range guaranteed by the large-scale mosaic of those green covers were thought to be influential on their existence in the study area.

1. INTRODUCTION

In urbanized areas, which have complex landscapes that include severely fragmented green covers areas, restoration or creation of habitats are needed for the re-construction of a healthy and functional ecosystem , including private open spaces (Gaston *et al.*, 2005) or brownfield sites (Woodward *et al.*, 2003) for example. For ecological network planning based on the primal characteristics of natural environment in urban areas, an understanding of the interplay between landscape (matrix effects) and local factors (patch effects) that affect urban biodiversity (Angold *et al.*. 2006) is important. Although there has been much research on the relationship between species diversity and the structure of fragmented habitats (or landscapes) in urban or suburban area (e.g. Hobbs, 1988; Jeanneret *et al.*, 2003), the distribution of small-scale urban green covers has not been fully analyzed yet.

On the other hand, with the advance of remote sensing technologies, it is getting easier to grasp the detailed distribution and structure of small-scale green covers in urbanized areas and extract habitats at multiple scales. (e.g. Sawaya *et al.*, 2003). Therefore, it is foreseeable that high-resolution land cover monitoring will become a tool for evaluating the key elements of the distribution of fragmented urban green spaces for urban biodiversity.

Accordingly, the purpose of this study is to evaluate fine-scale spatial characteristics of the distribution of fragmented green covers in an urbanized area, and to extract the key factors for species diversity, utilizing high-resolution satellite image analysis.

As indicators of the ecological function of the distributed small-scale urban green covers, butterflies (*Lepidoptera*) were chosen for this study. Butterflies serve as good indicators of the environmental changes that occur as humans develop the landscape and as excellent indicators of urbanization (Blair & Launer, 1997), since many researchers have suggested that butterflies serve as indirect measures of environmental variation because they are sensitive to local weather, climate, and light levels (e.g. Pyle, 1980; Murphy *et al.*, 1990; Kremen, 1992).

2. STUDY AREA

The study area is within 10km^2 (3.5km in north-south, 3km in east-west) located at north of Ichikawa City, 15km northeast of central Tokyo. This area is located at the west-end of the Shimofusa-tableland, around which disordered urbanization had spread subdividing the steep forest. The main land cover in the area is residential area, mixed with agricultural land use (mainly orchards and farms), fragmented urban forests and grasslands. Small-scale greens such as small gardens and hedges are also included in the mosaic of green covers.

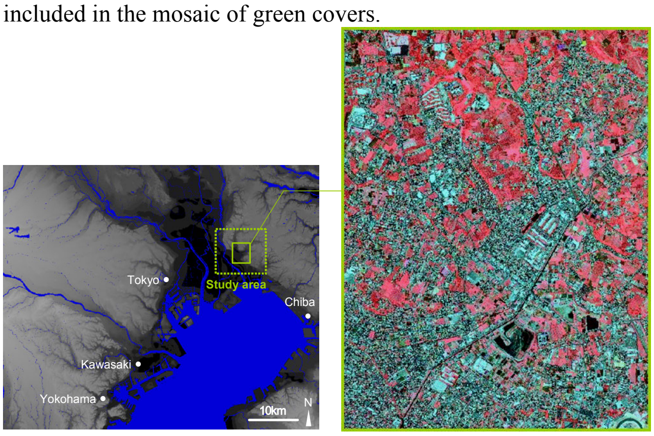

Figure 1. Study area (right: false-color image)

* Corresponding author. This is useful to know for communication with the appropriate person in cases with more than one author.

3. MATERIAL AND METHODS

3.1 Butterfly sampling

The line transect count method was used to record butterflies. The transect count is appropriate for estimating the diversity and structure of butterfly assemblage, and for monitoring changes in assemblage (Natuhara *et al.*, 1998).
The transect was composed of continuous 80 transect units, all of which were 200m long, designed to pass all types of urban green covers in the study area evenly. All the transect units were surveyed 13 times during 9 July- 7 October 2004, 25 May- 27 June 2005 and 18 April- 3 May 2006 in the daytime under fine weather conditions. Butterflies within a 10m width (5m each side of the recorder) and 5m height along the transect units were identified and their population and action were recorded. Butterflies not readily identified were netted and released immediately after identification. The appearance points of butterflies were plotted on the 1:2,000 base map, along with their environmental conditions .

3.2 Land cover classification

Satellite image data used was Quickbird multi-spectral image (spatial resolution is 2.8m) collected on 6 June, 2003. The image was high quality with no cloud cover, and ortho correlated based on a 10m Digital Elevation Model (Hokkaido-chizu Co., Ltd.). As the ground truth data, aerial photography data with 25cm resolution (Pasco Co., Ltd.) collected in May 2002 was used.

3.2.1 Extraction of green covers: In order to avoid possible confusion in classification in urban built-up areas, only green cover was extracted by the threshold of vegetation index, MRVI (Modified Ratio Vegetation Index, by Yun & Hoyano, 1998; Shirai et al, 1998).

$$MRVI = \frac{Band4(NIR)}{Band1(B) + Band2(G) + Band3(R) + Band4(NIR)} \quad (1)$$

MRVI is suited for extracting vegetation cover ratio in a pixel, in which different types of land covers are mixed, and MRVI has better accuracy than NDVI to extract green cover by its threshold, since it has a clearer linear relationship with the vegetation cover ratio in a pixel (Yun & Hoyano, 1998). By calculating MRVI with Quickbird image and comparing it with ground truth data, the green cover was extracted by the threshold of MRVI value.

3.2.2 Object-based classification
Next, object-based classification by the nearest neighbour method was applied for the extracted green cover pixels, by eCognition Elements 4.0 (Definiens A.G.). The green covers were classified into 7 types, which were evergreen woods, deciduous woods, hedges (including gardening trees and trees lining streets etc.), bamboo groves, grasses, orchards and farmland. The class hierarchy was set as in Figure 2.. All the object features of layer values, shape, texture and hierarchy were selected as standard nearest neighbor features.
As the ground truth data for sample selection, orthoimages of digital aerial photos with a spatial resolution of 0.25m, acquired in May, 2002 (Pasco Co., Ltd.) were used.

3.3 Landscape measures

In order to quantify the ecological gradient by the distribution of small-scale green covers in the ranges from root units, the four basic indices measuring the mosaic of land cover were computed within the surrounding area of root units.

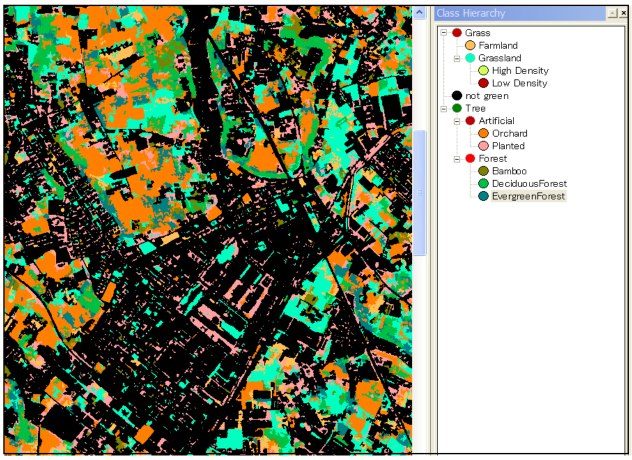

Figure 2. Class hierarchy in object-based classification

As investigating matrix effects within a landscape and selecting a set distance within which to measure land use variables is limited by the assumption that all types of land use affect species at the specific scale that has been selected (Dunford & Freemark, 2004), a multi-scale approach to extrapolating the appropriate distance into the mosaic was applied in this study.
Four different sizes of ranges from the root units were adopted to grasp the distribution of green covers around the surveyed root units, which were 25m, 50m, 100m and 200m ranges. The buffer areas with 4 types of distance from each root unit were extracted. Image analysis software mainly used to analyze landscape measures was TNT mips 7.1 (Microimages, Inc.).
The adopted indices as follows, were suitable for computing by raster data. All the indices were measured for the 4 different sizes of buffer areas from each root unit.

3.3.1 Diversity index H: The diversity index used was H, which is the application of Shannon's diversity index H' as follows:

$$\text{Diversity } H = -\sum^{s} p_k \ln p_k \quad (2)$$

where s = number of habitat types; p_k = proportion of area in habitat k (Oneil *et al*, 1988).
Diversity index H of green cover types in each range from root units was computed.

3.3.2 Edge number: Edge number is the total length of the borderlines between two different vegetation types (Gardner *et al.*, 1987).

$$\text{Edge number } E_{i,j} = \Sigma\ e_{j,j} \times l \quad (3)$$

where $e_{i,j}$ = number of horizontal and vertical interfaces between grid cells of types i and j; l = the length of the edge of a cell (Gardener *et al*, 1987).
In this study, the co-occurrence value of different green cover pixels was substituted for the edge number. The co-occurrence value is the number of heterogeneous pixels that interface horizontally, vertically and diagonally in 3 × 3 unit pixels

(Figure 3). The average co-occurrence value of the following green cover types in the different ranges from root units were computed; 1) co-occurrence of deciduous woods, evergreen woods and bamboo groves, 2) co-occurrence of grasses and woods (deciduous or evergreen) and 3) co-occurrence of grasses and bamboo groves or hedges.

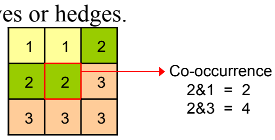

Figure 3. Co-occurrence value of different green-cover types

3.3.3 GCI(Green Cluster Index): Green Cluster Index (Hirota *et al.*, 2001), which was proposed for the pixel based raster data such as high-resolution satellite image, indicates the extent of clustering of each green-covered patch. GCI is computed by summing up GDI (Green Distance Index) values. GDI is the pixel value indicating the concentration of green-covered pixels within the surrounding quadrangular area with 15 by 15 pixels on all sides, the center of which is the GDI computing pixel (Figure 4). GDI is computed only on green-covered pixels, and computed by filtering geo-processing. GCI is the averaged value of GDI of the green-cover pixels by the area of green-cover pixels within a fixed region (e.g. grids).
As the index of green-cover clustering, GCI of each green cover type in every range from the root units was computed.

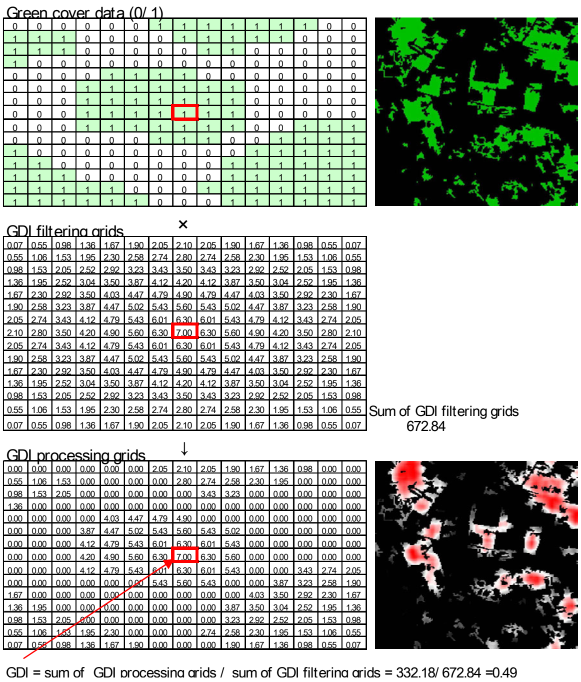

Figure 4. Process of calculating GDI

3.3.4 Average minimum distance: In order to grasp the influence of the distance from each green cover, the average minimum distance from each green-cover patch to any pixel was calculated. The distance raster from the patches of each green cover type was generated, and mean distance from each green cover type in the area of non-green pixels in each range from the root units was computed (Figure 5).

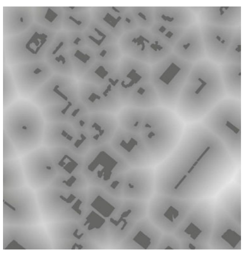

Figure 5. Distance raster from the patches of a green cover (The brighter, the further from patches)

4. DATA ANALYSIS

4.1 Relationship between the richness of species and the distribution of green covers: First, in order to certify the relationship between the number of butterfly species and the area of green covers in each range (25m, 50m, 100m and 200m), linear regression and logarithm regression was carried out as a simple correlation analysis. The dependent variable was the number of species that were present in each root unit, and the independent variables were the area of the total green covers from each root unit.
Then, a regression tree by CART(Classification and Regression Trees) (Breiman *et al*, 1984) was applied to the datasets.
As a nonparametric method, CART is robust to many of the data issues that sometimes plague parametric models, and as a recursive model, the approach is also able to capture some relationships that make sense ecologically but that are difficult to reconcile with conventional linear models in statistics (Urban, 2002).
Two dependent variables were adopted, one is the number of butterflies, and the other is Shannon-Weaver's H' as the diversity index of butterfly species. Independent variables were the four indices (Diversity index H, Edge number, GCI and Average minimum distance) of each green cover type in each different range from root units.
In CART, the regression tree was grown and pruned under the requirements that the minimum number of samples for divergence of parent node was 5, that of the child node was 2, and the maximum depth of divergence was set at 5 levels. As the software for CART analyses, SPSS Classification Trees (SPSS Inc.) was used.

4.2 Relationship between the width of butterfly species composition and the distribution of green covers:

Butterflies confirmed in the survey were classified into three types based on the bibliography; grassland species, forest-edge species and forest (interior) species. Then, the pattern of the butterfly species composition in each root unit was classified into three levels; 1) only grassland species appeared, 2) Both grassland species and forest-edge species appeared and 3) forest species appeared.
In order to examine the relationship between the indices of green-cover distribution and the width of species, classification tree of CART was executed, in which the dependent variable was the level of species composition that appeared in each root unit, and the four indices of each green cover type in every different range from root units were the independent variables. In CART, the requirements of making trees were the same as those in 3.4.1.
On the other hand, in order to extract the influential variables in each landscape indices separately, the classification tree was executed as follows; each of GCI, edge number and average minimum distance was separately adopted as independent variables, and the level of species composition in each root unit was dependent variable.

5. RESULTS

5.1 Relationship between the richness of butterfly species and the distribution of small-scale green covers

A total of 30 butterfly species of 1376 individuals were recorded in the survey (Table.1). 2 species (*Papilio Menelaides memnon thunbergii* and *Favonius saphirinus*), of which only

Table 1. Butterfly species surveyed in the study area

grassland species		forest-edge species	forest species
Family Papilionidae	*Family Lycaenidae*	*Family Pieridae*	*Family Lycaenidae*
Graphium sarpedon nipponum	*Lycaena phlaeas daimio*	*Artogeia melete melete*	*Japonica lutea*
Papilio (Papilio) machaon hippocrates	*Pseudozizeeria maha argia*	*Anthocharis scolymus*	*Narathura japonica*
Papilio (Papilio) xuthus xuthus	*Celastrina argiolus ladonides*	*Family Lycaenidae*	*Favonius saphirinus* *
Papilio (Menelaides) protenor demetrius	*Lampides boeticus*	*Curetis acuta paracuta*	*Family Nymphalidae*
Papilio Menelaides memnon thunbergii *	*Family Nymphalidae*	*Family Nymphalidae*	*Hestina japonica*
Family Pieridae	*Cynthia cardui*	*Neptis sappho intermedia*	*Family Satyridae*
Colias (Colias) erate poliographus	*Family Hesperiidae*	*Vanessa indica*	*Neope goschkevitschii*
Eurema (Terias) hecabe hecabe	*Potanthus flavum*	*Polygonia c-aureum*	*Lethe sicelis*
Artogeia rapae crucivora	*Pelopidas mathias oberthueri*	*Kaniska canace no-japonicum*	
	Parnara guttata guttata	*Family Hesperiidae*	
		Daimio tethys tethys	

* The species of which only one indivisual was observed in the survey

one individual was recorded in the survey, were excluded from the analysis in advance.

On the other hand, accuracy in the classification of satellite image was examined by comparing the output of classified image with the fine-scale land cover data in the study area made by reading aerial photography data with 25cm resolution. The correlation coefficient between the classified green-covers and the ground truth in each range is shown in Table.2.

From the result of linear and logarithm regressions, significant relationships were certified between the total butterfly species and the total area of green covers in every distance of ranges from root units (Table.3).

From the result of the regression tree by CART, in which dependent variable was the number of species, and the independent variables were 4 indices in each range with different distances from root units, the variables such as GCI of deciduous woods in 25m range (+; positive influence), average minimum distance from bamboo groves in 100m range (-;negative influence) were extracted as main factors with correlation to the number of species. Additionally, by CART, in which dependent variable was Shannon-Weaver's H', the extracted correlational indices by the regression tree were GCI of evergreen woods in 50m range (+) and GCI of deciduous woods in 100m range (+) (Figure.5).

The existence of united evergreen and deciduous woods was thought to affect the number of butterfly species. The diversity index H of green covers did not necessarily affect the number of butterfly species in the study area.

Table 2. Correlation between the classified green cover and ground truth data in each range

	Orchards	Bamboo groves	Hedges	Farmland	Evergreen woods	Deciduous woods	Grasses
200m range	0.962	0.796	0.780	0.826	0.812	0.887	0.782
100m range	0.931	0.755	0.786	0.828	0.788	0.865	0.786
50m range	0.855	0.707	0.725	0.697	0.764	0.786	0.685
25m range	0.743	0.719	0.603	0.365	0.637	0.849	0.624

Table 3. Relationship between the number of species and the area of green covers

Distance from root units	linear R^2	linear Sig.	logarithmic R^2	logarithmic Sig.
25m range	0.782	0.000	0.819	0.000
50m range	0.767	0.000	0.805	0.000
100m range	0.766	0.000	0.801	0.000
200m range	0.782	0.000	0.801	0.000

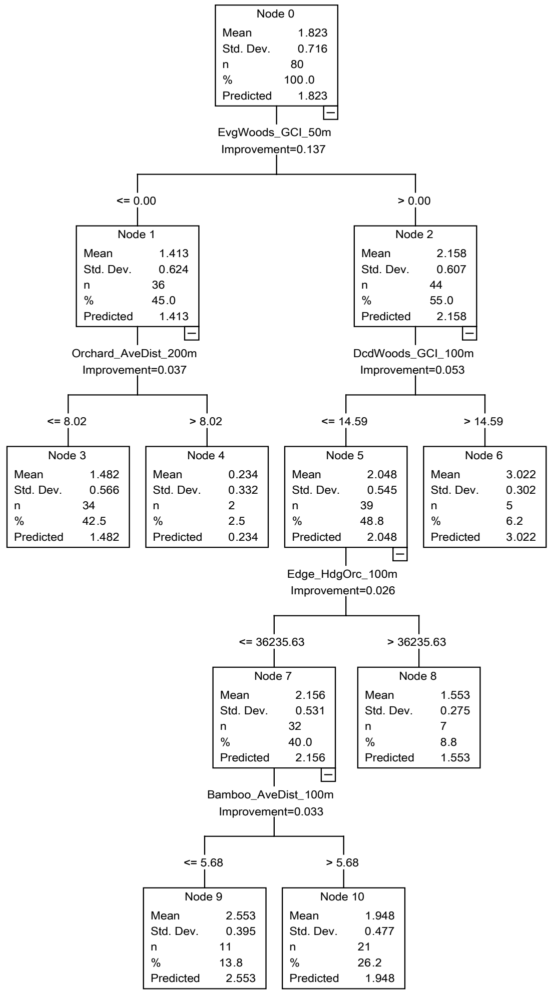

Figure 5. Output of regression tree; Shannon-Weaver's H' as dependent variable.

5.2 Relationship between the width of butterfly species composition and the distribution of green covers:

In the results from the classification tree, in which dependent variable was the combination of species types and independent variables were the landscape indices, the forest species could be the indicator to extract the requirements for the width of butterfly species in the study area.

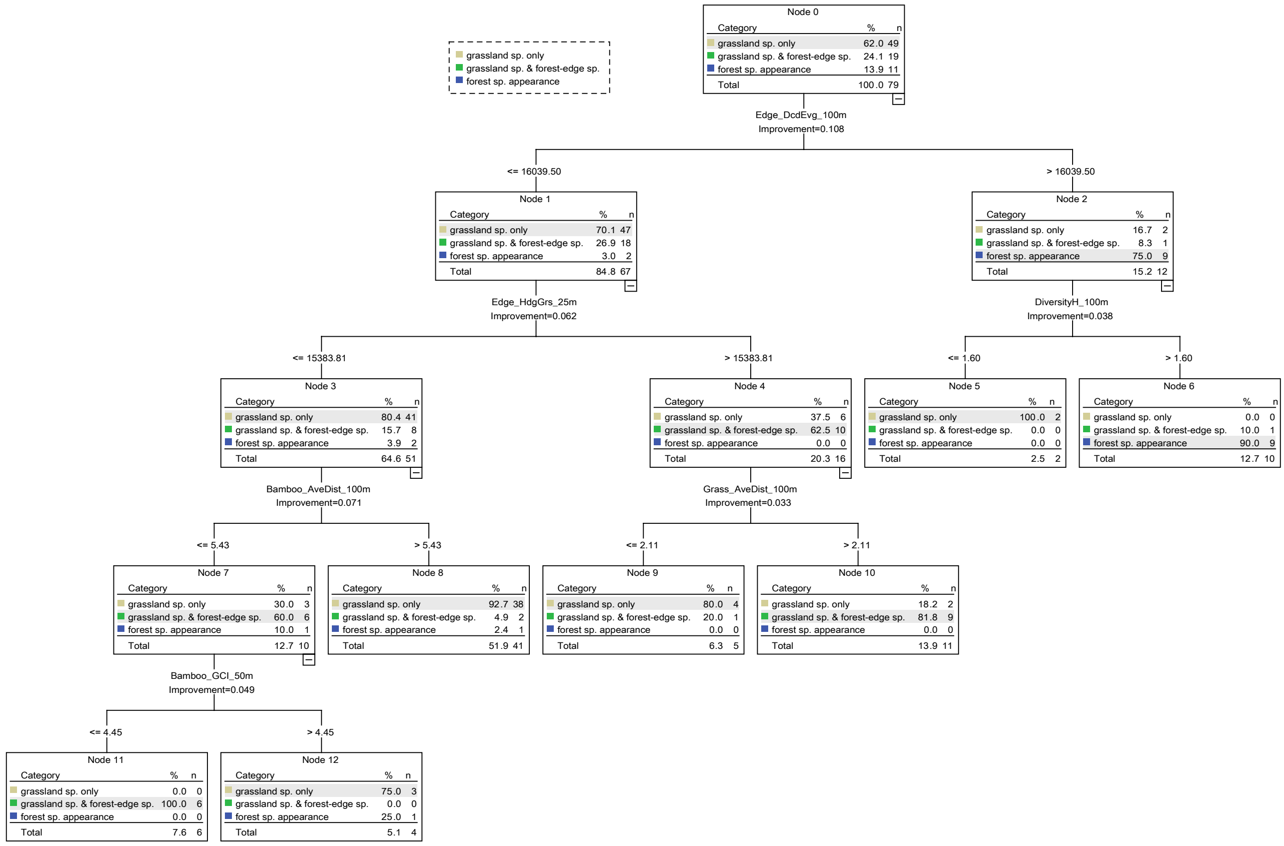

Figure 6. Output of classification tree; the types of species composition in root units as dependent variable.

The influential landscape indices for the appearance of forest species were mainly the edge number between deciduous woods and evergreen woods in 100m range (+), followed by the diversity index H of green-cover types in 100m range (+) (Figure.6).

On the other hand, from the classification tree which was executed separately by each landscape index as independent variables, the variables shown in Table.4 were extracted as influential ones for the appearance of forest butterfly species.

In GCI and average minimum distance, the deciduous woods in smaller range was important for the appearance of forest species, while in edge number, border length between deciduous woods and evergreen woods in wider range was a fundamental condition.

Table 4. The variables related to the appearance of forest

Independent Var. : Edge number		Independent Var. : GCI		Independent Var. : Average min. dist.	
Deciduous woods - Evergreen woods 100m range (+/-)		Deciduous woods 25m range (+/-)		Deciduous woods 50m range (- / +)	
Hedge - Grass 25m range (-)		Deciduous woods 25m range (+)			Bamboo groves 25m range (-)
Hedge-Bamboo grove 100m range (+)		Evergreen Woods 100m range (+)			Grass 200m range (-)
					Hedge 100m range (+)
Ratio in total 11 root units with the appearance of forest species					
18% (n=2)	82% (n=9)	27% (n=3)	64% (n=7)	73% (n=8)	18% (n=2)

6. DISCUSSION

6.1 Influence of landscape factors of each green cover type on the number of butterfly species:

By using CART, the factors of green cover distribution in different types and ranges of covers and their interrelationships were analysed.

The landscape indices related to the mosaic of deciduous woods, evergreen woods and bamboo groves were influential to the number and diversity index H' of butterfly species. Especially, the deciduous woods both in small and wide range and the mixture of evergreen woods in deciduous woods were thought to be fundamental for the number and the diversity of butterflies in the area. This would indicate that the existence of feeding plants or nectar sources for different butterfly species is assured by the mixture of deciduous and evergreen woods. On the other hand, the distance from bamboo grooves were negatively correlated to the number and the diversity index of butterfly species, which indicated that the butterflies depending on bamboo groves would locally contribute to the width of species composition.

As a factor for the number of species or the diversity index H', the diversity index H of green cover types was not necessarily influential.. In the study area, the green covers were dispersed to pieces of patches, and the influence of the variation of their sizes and the spatial distribution were supposed to be little in large scale.

6.2 Relationship between the width of butterfly species composition and the distribution of green covers

In the analyses, the appearance of forest (interior) species was thought to be the effective indicator to grasp the factors affecting the width of range of butterfly species. From the output of the classification tree, the wide-scale factors such as the edge number between deciduous woods and evergreen woods and the diversity index H of green cover types affecting the appearance of forest species. By classification trees adopting edge number, GCI and average minimum distance separately as independent variables, the scale of range which is influential to the forest species was different by the indices. In edge number, the variable in higher rank was the edge number between deciduous woods and evergreen woods in the 100m range, which is affected by the landscape characteristics in larger scale than in the variables in GCI and average minimum distance. Since the edge number between evergreen and deciduous woods was influential to the combination of types of butterfly species among the indices, the edge number was thought to be the effective variable to grasp the multi-scale function of green cover mosaics for the appearance of forest butterfly species. The local scale habitat requirements such as the patches of deciduous woods guaranteed by large scale landscape correlation (e.g. green cover mosaics with adjoining evergreen woods or grasslands, for example) were supposed to be the key factor for the width of species composition.

6.3 Effectiveness of the object-based classification

The landscape indices from the object-based classification of satellite images were applicable to extract the key factors of green-cover distribution, which was not possible using the green cover data by pixel-based classification. The classification tree model for the width of butterfly species could be applied to the green cover distribution in surrounding areas.

In this study, the relationship between the attributes of plane distribution of small-scale green covers and the richness of butterfly species was focused using landscape indices. However, as Corry *et al.* (2005) point out the limitation of landscape indices, the suitability of indices used should be cautiously investigated. Additionally, as the aspects of ecological process of urban land cover, the change of vegetation or landform in urbanizing process, should be considered in future studies.

References:

Angold, P.G., Sadler J.P., Hill, M.O., Pullin, A., Rushton, S., Austin, K., Small, E., Wood, B., Wadsworth, R., Sanderson, R. and Thompson, K., 2006. Biodiversity in urban habitat patches. *Science of the Total Environment* 360, 196-204.

Blair, R.B. and Launer, A.E., 1997. Butterfly diversity and human landuse: species assemblages along an urban gradient. *Biological Conservation* 80, 113-125.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., 1984. Classification and Regression Trees. Chapman & Hall/ CRC Press, Florida, 358 pp.

Corry, R.C. and Nassauer, J.I., 2005. Limitations of using landscape pattern indices to evaluate the ecological consequences of alternative plans and designs. *Landscape and Urban Planning* 72, 265-280.

Dunford, W. and Freemark, K., 2004. Matrix matters: effects of surrounding land uses on forest birds near Ottawa, Canada. *Landscape Ecology* 20, 497-511.

Gardner, R.H., Milne, B.T., Turner, M.G. and O'neil, R.V., 1987. Neutral models for the analysis of broad-scale landscape pattern. *Landscape Ecology* 1, 19-28.

Gaston, K.J., Smith, R.M., Thompson, K. and Warren P.H., 2005. Urban domestic gardens (II): experimental tests of methods for increasing biodiversity. *Biodiversity Conservation* 14, 395-413.

Hirota, F., Ohnishi, A., Morisugi, M. and Imura, H., 2001. A study on analysis for vegetation in urban area with high-resolution satellite data. *Journal of Environmental Systems and Engineering* 30, 91-99. (In Japanese with English summary)

Hobbs, E., 1988. Using ordination to analyze the composition and structure of urban forest islands. *Forest Ecology and Mangament* 23, 139-158.

Jeanneret, Ph., Schüpbach, B. and Luka, H., 2003. Quantifying the impact of landscape and habitat features on biodiversity in cultivated landscapes.*Agriculture, Ecosystems and Environment* 98, 311-320.

Kremen, C., 1992. Assessing the indicator properties of species assemblages for natural areas monitoring. *Ecological Applications* 2, 203-217.

Murphy, D.D., Freas, K.E. and Weiss, S.B., 1990. An environment-metapopulation approach to population viability analysis for a threatened invertebrate. *Conservation Biology* 4, 41-51.

Natuhara, Y., Imai, C. and Takahashi, M., 1998. Evaluation of community indices in seasonal assemblages of butterflies (Lepidoptera) at different frequency of transect count. *Biodiversity and Conservation* 7, 631-639.

O'neill, R.V., Krummel, J.R., Gardner, R.H., Sugihara, G.., Jackson. B., DeAngelis, D.L., Milnes, B.T., Turner, M.G., Zygmunt, B., Christensen, S.W., Dale, V.H. and Graham, R.L. ,1988. Indicies of landscape pattern. Landscape Ecology 1, 153-162.

Pyle, R.M., 1980. Butterfly eco-geography and biological conservation in Washington. *Atala* 8, 1-26.

Sawaya, K.E., Olmanson, L.G., Heinert, N.J., Brezonik, P.L. and Bauer, M.E., 2003. Extending satellite remote sensing to local scales: land and water resource monitoring using high-resolution imagery. *Remote Sensing of Environment* 88, 144-156.

Shirai, N., Setojima, M., Hoyano, A. and Yun, D., 1998. Urban vegetation cover mapping with vegetation cover. Poster session in the 19th Asian Conference on Remote Sensing.

Urban D.L., 2002. Classification and Regression Trees. In: McCune B. and Grace J.B. (eds), Analysis of Ecological Communities, MjM Software Design, Oregon, pp. 222-232.

Woodward, J.C., Eyre, M.D. and Luff, M.L., 2003. Beetles (Coleoptera) on brownfield sites in England: an important conservation resource? *Journal of Insect Conservation* 7, 223-231.

YUN, D. and Hoyano, A 1998. Index to extract vegetation cover ratio in a pixel (VCRP) in urban areas, *Jour . of the Remote. Sensing Society of Japan*, 18(3), 4-16.