Habitat Modelling for Conservation Analysis Using GIS and Remote Sensing in Lantau Island, Hong Kong

Frankie K. K. Wong^{*}, Tung Fung

Department of Geography and Resource Management, Faculty of Social Science, The Chinese University of Hong Kong, Hong Kong, China

Copyright © 2015 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract Knowledge of spatial distribution of faunal species is crucial to prioritize conservation resources. While vulnerable or endangered species are usually well-protected, common species were always overlooked. Traditional survey represented species distribution by discrete points, which is precise but often limited in applicability for conservation analysis. This study combined remote sensing; geographical information analytical tools and well-established multivariate statistical modelling to predict habitats of fifty common and representative animal species. Reliable ground data records from terrestrial biodiversity survey were collected and ecological niche factor analysis was used to identify pseudo-absence sites essential for habitat statistical models. Binary logistic regression models and generalized additive models were used to predict the faunal habitats. Species-richness map was then produced to identify biodiversity hotspots and possible conservation gaps of the current conservation system. Results from gap analysis showed that only 1% of faunal species richness hotspots coincided with existing protected sites of special scientific interest (SSSI). 45% of species-rich sites were under-protected. The results not only identified deficiency of existing protection system, where extra conservation planning efforts should be considered, but also highlighted the direction for future development to minimize the land use conflicts. Integration of geoinformatics and statistical analysis with ecological knowledge assist timely conservation policy making.

Keywords Ecological Niche Factor Analysis, Gap Analysis, Generalized Additive Model, Logistic Regression, Species Richness

1. Introduction

The spatial location of ecologically sensitive areas is fundamental for wildlife conservation. Traditionally, the distribution of species is represented by either point or ranged maps. However, these maps are criticized for not being representative enough [1] and provide inadequate

information on the distribution of species [2]. The deficiency of both information and knowledge in species distribution and abundance is likely to hinder the reliable location of ecological sensitive sites [3], which in turn can limit the effectiveness of park and reserve planning process.

Remote sensing, GIS and habitat model

Species and habitat types are non-randomly distributed. The basic needs of different species have implications on their distribution as they reflect essential environmental characteristics that related to their life support system including food, water, nesting sites, shelter, evasion of potential enemies, etc. [3]. Austin [4] and Guisan and Zimmermann [5] classified the ecological factors into resource, direct, and indirect gradients. Among various types of factors, indirect gradients are usually used as environmental predictors for species distribution modelling primarily because they can be effectively derived from remote sensing imagery and extracted geographical information database. Remote sensing identifies and retrieves land cover characteristics, particularly vegetation types at broad spatial extents [6]. In early days, vegetation map derived from aerial photos or satellite images were used as the only predictor to map species habitats. While Geographic Information System stores precise locations of land attributes related to topography, hydrology, man-made structures etc., these spatial data can be further manipulated and analyzed to generate other relevant environmental variables through various analytical tools. The observed species distribution always shows significant correlation with these environmental variables. Suitable habitat of different species can then be modeled and localized by establishing relationship between field records and environmental gradients through modelling process [3].

Over past few decades, multivariate statistical algorithms were used [7-9] to model the distribution of species by identifying suitable habitat of concerned species [10]. Regression models have been extensively applied to species distribution prediction [7,9,11,12]. Traditionally, multiple regression and multiple discriminant analysis in which data are assumed to have normal errors were used to predict species abundance. However, assumption of Gaussian

distribution error sometimes limits their applications [13,14]. New modelling paradigms, particularly, generalized linear model (GLM) and generalized additive model (GAM) were developed and intensively applied to habitats modelling of fauna [7,8,10,11,15,16]. The majority of statistical models require both presence and absence of dataset. Hirzel [17] argued even though the absence data was available, their accuracy was in doubt. The Ecological Niche Factor Analysis (ENFA) which requires presence-only data for modelling species distribution was developed. Not only would the models provide fundamental understanding of the habitat requirements of species; it can also extend to sites where no field surveys were conducted due to accessibility or manpower constraints. More crucially, if the potential habitats of a number of species overlapped, conservative measures should be considered in these sites.

Biodiversity and conservation

Parks and reserves are commonly used to conserve ecologically valuable and sensitive areas [18] and vulnerable or endangered species. While habitats of common species are usually overlooked, their habitats may coincide and perform high ecological functions. Due to the lack of reliable data to identify these habitats, these sites tend to be under-protected and may become vulnerable for development and eventually lead to the decline in both plant and animal diversity [18]. Closely linked with habitat modelling is the identification of conservation sites of high level of biodiversity. The aggregation of species distribution helps identify ecologically important sites. A biodiversity hotspot is “a site that supports a significant proportion of the species in a particular taxon or group” [19]. As one of the measurements of ecological importance, it is the composition of species richness, endemism and rarity. Species richness is commonly used as indicator of biodiversity because of its simplicity, feasibility to measure and data availability [20]. Although the discrepancy between the definition of biodiversity and its actual representation is still under many debates [21-23], biodiversity hotspots is always used as a guide to identify areas of conservation priority [24]. Gap analysis is one of the systematic conservation planning methods. It identifies geographical gaps and prioritized areas for conservation by comparing potential locations/ coverage of plant and animal habitats (species-rich areas) to those that are currently protected [1,25,26]. The method provided useful and objective information to balance conservation and development and to minimize land use conflict.

This study attempted to identify biodiversity hotspots and possible conservation gaps based on predicted distribution of common species in Lantau Island, Hong Kong Special Administrative Region (HKSAR). With the aid of remote sensing and GIS analytical techniques, the following objectives are to going to be achieved:

- to identify and map the habitats of fifty common species based on field survey data;

- to evaluate and compare the accuracy of two regression models, i.e. binary logistic regression model (BLRM) and generalized additive model (GAM) used for habitat prediction; and

- to identify and locate biodiversity hotspots and possible conservation gaps; and

- to evaluate the proposed conservation and development plan in Lantau Island.

2. Materials and Methods

Study site

Lantau Island is located at the south-western part of Hong Kong (22°16'14"N, 113°57'10"E) as shown in Figure 1. It is the biggest island in Hong Kong with an area of approximately 142km² and over half of the land areas (78.4km²) are designated as Country Parks. Lantau Island nurtures over 120 butterfly species and 63 dragonfly species which represent 50% and 60% of total butterfly and dragonfly species respectively in Hong Kong. Rare plants, birds, mammals, reptiles and amphibians are abundant with some are native and protected by laws [27]. Existing statutory country park system and SSSI of Lantau Island are captured in Figure 1. Ecological significance of Lantau is affirmed by the establishment of eight Sites of Special Scientific Interests (SSSIs), where fauna and flora with special scientific value are found. Due to surging land demand, Lantau was under the struggle between development and conservation in 2004 [28]. On one hand, the extensions of boundary of the Lantau North Country Park (LNCP) and of waters around Southwest as Marine Park were proposed [28]. On the other hand, the Island was further considered to be developed into a tourism spot supported by recreational facilities. The influx of Disneyland in 2005 was an example of mass tourism development. Although comments from public consultation were incorporated in concept plan revision in 2007, no objective and scientific evaluation was made and rendered the plan too general for further exploration and discussion until now.

Species Data

A spatial database was set up encompassing ecological surveys, remote sensing images, and GIS data from digital topographic maps.

Species data was collected from GIS database of the Biodiversity Survey of Hong Kong (version 3) published in 2002. The survey records were compiled from systematic in situ field survey of terrestrial and freshwater biodiversity, published sources and personal records [29] by the Department of Ecology and Biodiversity, the University of Hong Kong. It contains more than 95,000 point and binary locational records of more than 5000 species (from 20 taxonomic groups) and was geo-referenced to local Hong Kong 1980 Grid coordinate system. The recording resolution varies from 10 meters to 1 kilometer, but with the majority

sampled at 100 meters. Details about the survey can be found in <http://www.biosch.hku.hk/ecology/bs/>.

Fifty species from five taxonomic groups namely, amphibians, birds, butterflies, dragonflies and mammals were extracted from the survey shown in Table S1 in Supplementary Material. Except all the birds and *H. brachyura* (mammal), no species has statutory protection in Hong Kong, which means they are common to Hong Kong.

M. migrans (aves) has been listed in the protection of endangered species ordinance (Cap. 187) and *H. brachyura* is listed as ‘vulnerable’ under the IUCN Red List Status. The number of survey points acquired from the survey is shown in Table 1. Effective number of sample points after excluding data points that fell outside the *terrestrial* boundary is bracketed.

Figure 1. The boundary and statutory protection areas of Lantau Island

Table 1a. Model evaluation statistics of amphibian species

Abbrev.	No. of points from Bd. Survey	BLRM					GAM				
		No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC	No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC
Bm	130 (129)	10 (5)	55%	88.4	80.0	0.942	13 (7)	70%	92.3	91.9	0.977
Kp	100 (99)	6 (4)	79%	93.9	93.9	0.984	11 (6)	89%	97.5	97.5	0.998
Pm	271 (259)	5	49%	84.2	84.2	0.924	15 (13)	87%	91.9	91.5	0.963
Rl	48 (48)	8 (3)	84%	95.8	95.8	0.992	12 (2)	96%	100.0	100.0	1.000
Rg	209 (207)	8 (5)	51%	86.5	86.5	0.938	14 (9)	78%	93.3	93.2	0.979
Re	129 (128)	12 (6)	81%	94.5	94.5	0.988	14 (8)	97%	99.2	99.2	1.000
Mo	103 (100)	10 (5)	51%	87.5	87.5	0.930	12 (7)	87%	95.5	95.5	0.990
Rl	158 (155)	10 (7)	66%	91.0	90.3	0.965	13 (9)	92%	94.8	94.8	0.988

Table 1b. Model evaluation statistics of bird species

Abbrev.	No. of points from Bd. Survey	BLRM					GAM				
		No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC	No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC
Mm	67 (44)	11 (3)	66%	92.0	92.0	0.951	12 (6)	95%	98.9	98.9	1.000
Ps	97 (74)	10 (4)	53%	89.9	89.9	0.929	13 (6)	86%	96.6	96.6	0.996
Os	86 (69)	8 (5)	63%	89.9	89.9	0.960	12 (9)	92%	100.0	100.0	1.000
Pj	92 (72)	13 (3)	76%	95.8	95.8	0.975	13 (4)	86%	96.5	96.5	0.992
Pm	59 (52)	10 (2)	59%	88.5	88.5	0.941	12 (5)	74%	92.3	91.3	0.988
Cs	46 (32)	11 (2)	69%	93.8	90.6	0.966	11 (4)	99%	100.0	100.0	1.000
Zj	79 (64)	9 (2)	59%	89.1	89.1	0.954	12 (5)	73%	94.6	93.8	0.972
Cm	47 (36)	6 (1)	54%	87.5	87.5	0.905	13 (5)	96%	100.0	100.0	1.000
Eg	25 (19)	6 (2)	67%	92.1	92.1	0.964	8 (3)	96%	100.0	100.0	1.000
Ls	44 (31)	8 (3)	46%	83.9	83.9	0.903	13 (5)	98%	100.0	100.0	1.000
Sc	82 (64)	12 (4)	51%	88.3	88.3	0.927	13 (8)	84%	95.3	94.5	0.995
Mc	47 (33)	10 (2)	44%	81.8	81.8	0.908	12 (6)	94%	100.0	100.0	1.000
Ap	25 (19)	8 (1)	29%	76.3	76.3	0.848	10 (3)	87%	100.0	97.4	1.000

Table 1c. Model evaluation statistics of butterfly species

Abbrev.	No. of points from Bd. Survey	BLRM					GAM				
		No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC	No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC
Mll	94 (82)	11 (5)	83%	93.9	93.9	0.991	13 (6)	94%	98.8	98.8	1.000
Cnn	38 (34)	13 (4)	46%	80.9	80.9	0.903	15 (5)	93%	100.0	100.0	1.000
Ccc	82 (77)	11 (5)	27%	79.2	79.2	0.851	12 (5)	73%	93.5	93.5	0.978
Pppo	182 (168)	13 (4)	62%	90.8	90.8	0.963	14 (7)	76%	94.4	94.3	0.985
Pcc	131 (123)	10 (7)	39%	80.1	79.7	0.878	12 (6)	90%	97.6	97.6	0.998
Pma	129 (119)	12 (4)	57%	89.1	89.1	0.950	13 (6)	69%	91.6	91.6	0.974
Hgg	86 (81)	13 (3)	58%	88.9	88.9	0.943	13 (7)	84%	99.4	99.4	0.999
Zms	104 (96)	13 (6)	61%	90.6	90.6	0.952	14 (9)	92%	90.0	99.0	1.000
Pppa	143 (133)	12 (6)	76%	93.2	93.2	0.977	13 (9)	90%	98.1	97.7	0.997
Jaa	95 (87)	11 (6)	57%	86.8	86.8	0.947	14 (6)	94%	98.9	98.9	1.000
Aee	142 (128)	13 (5)	66%	91.8	91.8	0.966	14 (9)	82%	95.8	95.7	0.992
Ipp	32 (29)	8 (1)	71%	93.1	93.1	0.933	12 (6)	100%	100.0	100.0	0.879

Table 1d. Model evaluation statistics of dragonfly species

Abbrev.	No. of points from Bd. Survey	BLRM					GAM				
		No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC	No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC
Pa	46 (46)	11 (5)	78%	94.6	94.6	0.983	13 (6)	100%	100.0	100.0	1
Ed	58 (57)	13 (5)	79%	94.7	94.7	0.986	12 (6)	84%	96.5	96.5	0.994
Cci	41 (41)	10 (2)	78%	93.9	93.9	0.978	12 (3)	86%	96.4	96.3	0.995
Ccy	39 (39)	8 (3)	61%	92.3	92.3	0.948	12 (6)	91%	98.7	98.7	1
Rpp	56 (56)	12 (4)	80%	94.6	94.6	0.984	13 (6)	93%	100.0	99.1	1
Opn	71 (68)	14 (5)	44%	83.8	83.8	0.908	13 (6)	64%	89.0	89.0	0.967
Ta	60 (58)	10 (6)	53%	87.1	87.1	0.933	14 (8)	93%	99.2	99.1	1.000
Zii	44 (44)	7 (5)	68%	92.0	92.0	0.969	12 (5)	83%	97.8	97.7	0.992
Ai	26 (25)	10 (2)	58%	86.0	86.0	0.942	12 (2)	95%	100.0	100.0	1.000
Oss	70 (68)	8 (4)	38%	80.9	80.9	0.881	12 (7)	85%	97.8	97.8	0.996
Tf	67 (66)	12 (4)	72%	91.7	91.7	0.978	13 (6)	84%	98.5	98.5	0.990
Pf	107 (103)	12 (4)	43%	83.5	83.0	0.903	12 (6)	46%	85.9	85.9	0.914
Cm	53 (53)	14 (6)	58%	88.7	87.7	0.949	13 (6)	75%	94.4	94.3	0.985

Table 1e. Model evaluation statistics of mammal species

Abbrev.	No. of points from Bd. Survey	BLRM					GAM				
		No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC	No. of independent variable	% Deviance Reduction	CV accuracy (%)	LOOCV accuracy (%)	AUC
Nf	64 (63)	12 (4)	69%	91.3	91.3	0.972	13 (6)	93%	99.2	99.2	1.000
Hb	41 (41)	9 (4)	69%	92.7	92.7	0.974	13 (4)	98%	100.0	100.0	1.000
Ss	53 (53)	10 (5)	76%	95.3	95.3	0.984	11 (4)	100%	96.3	50.0	0.962
Rs	66 (63)	10 (4)	73%	94.4	94.4	0.972	11 (6)	79%	96.1	96.0	0.988

Environmental gradients

Environmental gradients were divided into four categories including: (1) types of vegetation; (2) proximity to water resource; (3) influence of topographic feature and (4) adaptability to human disturbance. Figure 2 summarizes the environmental gradients and the tools/ software used to compute them.

Figure 2. Environmental gradients extracted from SPOT-5 satellite image and digital topographic data (Analytical tools used are represented by different shades)

Figure 3. Land covers derived from supervised classification of SPOT-5 image

A SPOT 5 satellite image was acquired at GMT 03:16:46 on December 11, 2004 to extract vegetative environmental gradients. The image was ortho-rectified to same local coordinates with the overall root-mean-square-error of 0.02 pixels (0.2 meter). Supervised classification with maximum likelihood classifier was applied to classify the image into eleven land cover classes including sea water, fresh water, fish pond, mangrove, woodland, mixed shrubland, shrubby grassland, grassland, hill-fire site, bare soil and urban area. The classified map is shown in Figure 3. Classification accuracy was computed from 500 stratified random sample points (413 points fell within the land boundary). The overall classification accuracy is 72% (297/413) which is considered to be reasonable for further analysis.

Resource, landscape and human-related factors were computed based on the digital topographic map layers from the Lands Department, The Government of Hong Kong Special Administrative Region. Apart from vegetation, rivers and streams are vital resources and the data was extracted from hydrographic polygons. Landscape factors including slope, aspect, land curvature affects the potential solar radiation received of a region. Spatial layer features including spot height, contour, and coastline were used to compute these factors. Man-made structures affect integrity and cause habitat fragmentation which were represented by road and railway networks and building structures in natural

areas. Various spatial operations were conducted to compute the environmental gradients using the SPOT-5 derived vegetation maps and topographic data in IDRISI v.14.02, Biomapper v.3.1 and GRASS v.6.0. Definition, analysis and statistical description of each predictor are documented in Table S2 in Support Information.

Statistical Modelling

Habitats of individual species were predicted and mapped by forming relationship between point locations where species were recorded and environmental gradients using statistical model. Species-based modelling consists of three processes –data collection and exploration, model selection and model assessment. The dataset comprised the species occurrence points and the environmental predictors as described above. Prior to modelling, ranges of occurrence and correlation among the predictors were examined. Predictors with correlation coefficient less than or equal to 0.5 were retained.

Three multivariate statistical modelling methods including (1) Ecological Niche Factor Analysis (ENFA), (2) Binary Logistic Regression Model (BLRM) under the family of Generalized Linear Model (GLM) and (3) Generalized Additive Model (GAM) were used to predict the probability of occurrence of each of the fifty species. The modelling methods are summarized in Figure 4 and described below.

Figure 4. Methodology of multivariate statistical habitat modelling

ENFA

ENFA performed two roles. First, it provided basic description and understanding of the general habitats of the species based on global statistics. Second, it was used to generate pseudo-absences data which are essential for regression analyses. The basic idea of ENFA is to compare environmental variable or ecogeographical variable (EGV) space of those areas where species are present with the characteristic of the whole study area. Suitability functions were used to describe habitat suitability of individual species [17]. ENFA was conducted in BIOMAPPER v.3.1. The predictors were firstly normalized by the Box-Cox function. Factor analysis similar to principal component analysis (PCA) was then used to transform the predictors into uncorrelated factors explaining the same amount of variance (information) based on the variance-covariance matrix. However, unlike PCA, the transformed factors convey

ecological meanings. The first transformed factor which explains the maximum amount of variance is termed *marginality* the subsequent factors are named as *specializations* [17]. The marginality and selected specialization factors were retained with the amount of variances not less than 70%. Habitat suitability index was then computed with the distance geometric mean algorithm based on the retained factors. Three global statistics, namely, Marginality (M), Specialization (S) and Tolerance (T), describing the general habitat requirements when compared with the whole study area were computed.

Quality pseudo-absences were generated based on the habitat suitability results from ENFA. Random points were generated in areas where suitability level is less than 10. The number of pseudo-absences was the same as the recorded presences and the prevalence equals 0.5. Mann-Whitney U independent-samples test was used to check the significance

of difference between the presence and pseudo-absence dataset using each predictor. The pseudo-absences were then incorporated in BLRM and GAM modelling.

BLRM

BLRM, a category of GLM, is one of the most popular habitat modelling methods extremely suitable when the dependent variable is dichotomous or binary characterized in many field surveys [7]. The link function of logistic regression is logit transformation and the error structure is assumed to be binomial [30]. With the logit link function, the results are more closely related to ecological and biological sense [7].

The logit, $g(x)$ of BLRM is in the form of

$$g(x) = \ln \left[\frac{\pi(x)}{1-\pi(x)} \right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (1)$$

Where $\beta_i(x_i)$, $i = 1, \dots, k$ are beta values correspond to the x predictors.

The likelihood of occurrence of a species is expressed in the equation below:

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \quad (2)$$

Where $\pi(x) = P(Y = 1 | x = x_k)$ is the likelihood of occurrence of species as a function of environmental predictors x . The likelihood of occurrence ranges between 0 and 1. Model building was conducted in SPSS with backward stepwise method to select relevant predictors.

GAM

GAM is non-parametric extension of GLM and distinguished from the GLM family by allowing the predictors to be non-linear in nature [31,32]. When linking with the logistic term, the general form of the logit(p) is given as

$$g(x) = s_0 + s_1(x_1) + s_2(x_2) + \dots + s_k(x_k) \quad (3)$$

Where $s_i(x_i)$, $i = 1, \dots, k$ are smooth functions. The transformation of logit(p) into the probability of occurrence ranged between 0 and 1 was computed using (2).

Model building of GAM was implemented in Generalized Regression Analysis and Spatial Prediction (GRASP) v.3.0 in S-Plus v.6.2. As the prediction dataset was more than 250,000 observations, the prediction was required to be performed in ArcView [33].

Species models were evaluated for their reliability (goodness-of-fit) and their discriminatory ability [34] using the same dataset due to the deficiency of sample points for independent evaluation. Reliability was tested by the percentage of deviance reduction (R) while discrimination performance was described by leave-one-out cross validation (LOOCV) and Area Under ROC Curve (AUC). The logit denoted $g(x)$ of the best model was selected to compute habitat map using the image calculator in IDRISI.

Species-richness and gap Analysis

Species richness was used as a surrogate to identify biodiversity spots. It was computed by taking the average probability of occurrence of the fifty species from selected models using the equation below:

$$\text{Species Richness} = \frac{\sum_{i=1}^{50} \pi_i}{50} \quad (4)$$

where π_i is the likelihood of occurrence of a species defined in (2). The values ranged between zero (absences of species) to one (maximum species abundance). Biodiversity hotspots were defined using three cut-off thresholds, i.e. 0.5, 0.7 and 0.9. Areas with 0.9 or above are sites with extreme high species richness while 0.5 comprised sites of both extreme high and moderate richness. Hotspots defined using different thresholds were compared with protection system in 2004 including the Country Parks and eight SSSI through simple overlay analysis to identify and quantify possible conservation gaps in Lantau. The proposed extension of Country Park was evaluated.

3. Results

Geographical distribution of fifty common species

Results from statistical habitat models provided fundamental understanding of the general living conditions of various common species. Detail statistical results of individual species were presented in Table S3 – S5 in Support Information. Table S3 shows the summary statistics extracted from ENFA. Marginality (M) and Tolerance (T) computed from ENFA revealed the range of habitats. On the whole, the modelled species tend to inhabit in environment which is moderately different from the mean conditions in Hong Kong (mean M = 0.709±0.226) and adapt to wide range of environment (mean T = 0.630±0.172).

Environmental factors selected by regression models and their corresponding relationship and magnitude are summarized in Table S4 and S5 for BLRM and GAM respectively in Support Information. On average, the number of variables retained by AIC of GAM is higher than that selected by the stepwise BLRM for most species. Nonetheless, similar predictors were retained by the two regression models. The habitat characteristics were interpreted based on the summary statistics of three models in taxonomic group below.

Amphibian – ENFA showed that *P. megacephalus* is the most adaptive species reflected by the low M = 0.35. Most species tend to live in habitats which are moderately different from the mean conditions in the territory of Hong Kong (mean M = 0.771±0.451). Documentation of amphibian distribution in Hong Kong also suggested that *P. megacephalus*, *B. melanostictus*, *R. guentheri* and *R. limnocharis* are all widespread in various habitats including forest, shrublands, low-lying grasslands, freshwater marshes, wet agricultural fields, upland mountain streams, catchwaters and fish ponds at all altitudes [35]. *R. livida* (M = 1.634) and *R. exilispinosa* (M = 1.331) require an extremely special living environment. For instance, *R.*

exilispinosa is found in and near mountain streams [35]. The tolerance of most species is quite high with the maximum 0.87 (*P. megacephalus*) and mean equal to 0.713 ± 0.177 . *R. livida* is the least tolerant ($T = 0.302$). Together with the high M value, *R. livida* is likely to be restricted to a specific range of living conditions which are quite different from the background condition in Hong Kong. The regression models suggested that the amphibians inhabit in densely-vegetated areas such as woodland and shrubland and stay away from grassland. Moreover, major rivers are more favorable than small streams which reveal that most amphibian species inhabit closely to constant water sources.

Bird – Thirteen bird species tend to tolerate and adapt to a wide variety of habitats from rural to urban land uses. The summary statistics from ENFA showed that *S. chinensis* is the most adaptive species reflected by the lowest $M = 0.473$ in the group. Most species tend to inhabit in environment moderately different from the mean conditions of the territory (mean $M = 0.610 \pm 0.095$). *P. sinensis* has the highest tolerance level (0.809), which suggests that the species tends to live in a general range of living conditions. *E. garzetta* has the lowest tolerance ($T = 0.16$), which reveals that the species is restricted to a very specific range of living condition, in this case, the wetland habitats. The retained variables by the regression models suggested the bird species generally require densely-vegetated woodland and away from grassland. GAM further implies that solar radiation (AUTUMN), and water resources (RIVER) are positively related to their distribution. Eight out of thirteen species are common urban breeding birds; they adapt well to urban settings and human disturbance (BLDG).

Butterfly – ENFA revealed most butterfly species inhabit in environment moderately different from the mean environmental conditions in Hong Kong (mean $M = 0.652 \pm 0.060$). Indicated by mean global tolerance (T) of 0.677 ± 0.180 , the range of living condition is quite wide. *A. echerius echerius* has the highest tolerance level ($T = 0.811$), suggesting the species tends to live in general range of conditions. Comparatively, *I. pyrene pyrene* has the lowest tolerance ability ($T = 0.24$) and is restricted to a limited range of habitat condition. Nonetheless, local documentations suggested that *I. pyrene pyrene* is a common species which adapts to urban disturbance [36,37]. BLRM reveals that most species inhabit in densely-vegetated lands including woodland and mixed shrubland, and seldom found in open grassland and bare land. *M. leda leda*, *H. glaucippe glaucippe*, *I. pyrene pyrene* and *P. memnon agenor* are representative examples. Moreover, Species including *C. nerissa nerissa*, *P. canidia canidia*, *P. memnon agenor*, *Z. maha serica* and *J. almana almanac* adapt to proximity to human habitation such as urban parks and agricultural fields. Variables including water resources (RIVER), altitude (ELEV), distance from main roads (MAIN ROAD) and buildings (BLDG) all showed positive impact on the distribution of butterfly.

Dragonfly – Among thirteen species, *R. autumnalis*, *O. sabina sabina* and *C. marginipes* are widespread in Hong

Kong. ENFA results also show that most species tends to inhabit in environmental condition relatively different from the mean conditions in Hong Kong (mean $M = 0.709 \pm 0.103$). The living condition of *R. perforata perforata* is quite unique with $M = 0.918$. Local documentations suggested that *R. perforata perforata* are found in fast flowing streams and rivers with strong current. The mean global tolerance (T) of 0.656 ± 0.079 indicated the range of living condition is quite wide. *P. flavescens* has the highest tolerance level ($T = 0.786$), which suggests that the species tends to live in general range of living conditions while *A. immaculifrons* has the lowest tolerance ($T = 0.491$). Both BLRM and GAM revealed that proximity to water supply (RIVER) affects the distribution of dragonfly as water plays an important role in the life-cycle of dragonfly. Apart from water, they tend to inhabit in medium to highly vegetated lands (WOOD, MSHRUB, SHGRASS).

Mammal – With diversified food sources, the range of habitats of mammals is much wider that a distinctive or specific living habitat cannot be easily identified. All four species have marginality over one and tend to live in environment which is different from the mean conditions in Hong Kong (mean $M = 1.067 \pm 0.108$). The range of living condition is quite restrictive as explained by low mean global tolerance (T), 0.440 ± 0.095 . *R. sikkimensis* has the highest tolerance level ($T = 0.545$) while the tolerance of *S. scrofa* is the lowest ($T = 0.358$). Local documents also suggest that the two rats, *N. Fulvescens* and *R. sikkimensis* tend to inhabit everywhere from grassland to woodland [19]. BLRM suggests that they tend to inhabit in densely-wooded areas while GAM suggests all kinds of vegetative covers are suitable habitats. The regression models also revealed that water resources (RIVER), altitude (ELEV) and proximity to coastal locations (COAST) affects the distribution of the mammal species.

Tables 1(a – e) shows the model assessment statistics of individual species while Table 2 summarizes the accuracy statistics in taxonomic groups. The deviance reduction (R) of GAM is generally higher than that of BLRM. The mean deviance reductions are $87\% \pm 11\%$ and $61\% \pm 14\%$ for GAM and BLRM respectively. The maximum R for GAM and BLRM are 99% and 80% respectively. Discriminatory ability was evaluated by the LOOCV and AUC statistics. All species, except *S. scrofa*, show high classification accuracy in GAM in LOOCV. The mean accuracies from LOOCV are 95.9 ± 7.4 and 89.0 ± 5.0 for GAM and BLRM respectively. The AUC shows similar results. With the exception of two species – *I. pyrene pyrene* and *S. scrofa*, all species modeled with GAM show higher AUC than those modeled with BLRM. The mean AUC are 0.988 ± 0.022 and 0.946 ± 0.035 for GAM and BLRM respectively, which suggested excellent discrimination. As all three statistics suggested GAM performs better in model reliability and discriminatory ability, GAM models were used for computation of taxonomic-grouped richness maps and overall species richness map.

Table 2. Model comparison – BLRM, GAM and ENFA in terms of percentage of deviance reduction (R); Leave-one-out cross validation (LOOCV); and Area under curve (AUC) from ROC

	Amphibian		Bird		Butterfly		Dragonfly		Mammal		Overall	
	GAM	BLRM	GAM	BLRM	GAM	BLRM	GAM	BLRM	GAM	BLRM	GAM	BLRM
Deviance Reduction												
Average	87%	64%	89%	57%	86%	59%	83%	62%	93%	72%	88%	63%
SD	9%	15%	9%	12%	10%	16%	14%	15%	9%	3%	10%	12%
Maximum	97%	84%	99%	76%	100%	83%	100%	80%	100%	76%	99%	80%
Minimum	70%	49%	73%	29%	69%	27%	46%	38%	79%	69%	68%	43%
LOOCV - Classification Accuracy (%)												
Average	95.5	89.1	97.6	88.1	97.2	88.2	96.4	89.4	86.3	93.4	94.6	89.6
SD	3.2	5.5	2.9	5.0	2.8	5.4	4.3	4.9	24.3	1.8	7.5	4.5
Maximum	100.0	95.8	100.0	95.8	100.0	93.9	100.0	94.7	100.0	95.3	100.0	95.1
Minimum	91.5	80.0	91.3	76.3	91.6	79.2	85.9	80.9	50.0	91.3	82.1	81.5
Area Under Curve (AUC) from ROC												
Average	0.994	0.958	0.996	0.933	0.993	0.938	0.987	0.949	0.997	0.976	0.993	0.951
SD	0.009	0.028	0.008	0.035	0.009	0.041	0.024	0.035	0.006	0.006	0.011	0.029
Maximum	1.000	0.992	1.000	0.975	1.000	0.991	1.000	0.986	1.000	0.984	1.000	0.986
Minimum	0.975	0.924	0.973	0.848	0.976	0.851	0.914	0.881	0.988	0.972	0.965	0.895

Figure 5. Species richness derived from averaged probability of occurrence of fifty species

Biodiversity hotspots and conservation gap

Species richness map derived from the average probability of occurrence of the fifty common species is shown in Figure 5. Areas of extremely high species richness are shown in red. The boundaries of village settlements, SSSI and Country Parks were overlaid for comparison. Species-rich hotspots were found in the southern Lantau, areas around Tung Chung, Tai Ho and Nim Shue Wan in northern Lantau. These sites have the following characteristics. First, they are in close proximity to the hydrological networks. Second, they are all below 300 meters (above mean sea level). The uplands in Lantau Peak, Sunset Peak, Yi Tung Shan, Lin Fa Shan, Nei Lak Shan are in blue implying species richness is low. Third, they are mostly located at foothills and along the valleys in-between the mountainous and undulating landscapes. Finally, some of them are located in vicinity of village settlements.

Figure 6a. Gap analysis using species richness of 0.9 or above as threshold

Figure 6b. Gap analysis using species richness of 0.7 or above as threshold

Figure 6c. Gap analysis using species richness of 0.5 or above as threshold

Figure 7. Conservation gap of existing protection system and species richness using thresholds (a) 0.9, (b) 0.7, and (c) 0.5

Spatial analytical results showing the concordance between species richness and existing protection system is shown in Figure 6 and Figure 7. Species-rich sites protected by Country Parks (prior to 2004) and SSSI are represented in green and yellow respectively. Conservation gaps where high species richness sites fell outside SSSI and country park area are highlighted in red.

Species-rich areas of 0.9 and above indicated the most significant biodiversity hotspots in Lantau. As shown in Figure 6a, they located primarily near the lower bound of the Lantau South Country Park (LSCP). Corresponding, Figure 7a shows that 48% of total species-richness sites fell within the LSCP; 6% of the total was found to be coincided with the LNCP and less than one percent of the total was within the SSSI boundaries. The rest (45%) of extremely species-rich sites are unprotected by current conservation system, i.e. conservation gaps.

The total species-rich areas increased after the relaxation of threshold to 0.7 (Figure 6b) and 0.5 (Figure 6c) and they identified sites with high to moderate species richness. Spatial concordance between the species-rich areas and LNCP increases from 6% (0.9) to 8% (0.7) and 11% (0.5). However, species-rich areas spatially coincided with LSCP shrank to 47% (0.7) and 43% (0.5). The concordance with SSSI increased proportionally, but the percentage to total areas stayed around 1%. In fact, species-rich areas tended to fall outside Country Park boundaries and SSSI; and the conservation gap remained steady was around 44 to 45% as shown in Figure 7.

4. Discussion

Species-based habitat modelling

GAM is superior to BLRM in prediction accuracy and reliability, however, relationships with most environmental variables are nonlinear in GAM (Table S5). Results from ENFA and BLRM are complementarity to understand the general and specific requirement of the species.

Generally, the chosen species were more responsive to factor related to vegetation type and water resource than to other factors. Although BLRM and GAM selected different sets of environmental variables, they identified common predictors for individual species and taxonomic groups. Almost all species favors densely vegetative habitats and absent in grassland and bare soil. As suggested by local literature, bird species like *P. major*, *C. macrorhynchus*, *C. sinensis* and *S. chinensis* are commonly found in fung shui woodlands while *P. jocosus*, *C. macrorhynchus* and *L. schach* are usually absent from closed woodlands and shrublands [38,39]. Butterfly species such as *C. nerissa nerissa*, *C. clytia clytia*, *P. polytes polytes*, *P. paris paris* prefer woodland and shrubland edges. The models reflected habitats requirement of various species.

Nevertheless, a few species such as *E. garzetta*, *A. phoenicurus* and *M. caeruleus* have long been observed in

and known to well-adapt to urban environment in Hong Kong was not totally reflected by the models because field survey was mainly conducted in rural or non-urban areas. Besides, proximity to village houses was the only factor reflecting urban influence in the model.

Conservation analysis

Locations of vulnerable or endangered species are well-recognized with extensive protective measures exclude them from development consideration. For instance, SSSI is designated based on its uniqueness in terms of faunal, floral, ecological or geographical features. The emphasis of SSSI lies mainly on rare plant and animal species while the species richness map was computed based on common species. Hence, it is not unusual that species richness and SSSI did not exhibit a strong spatial relationship (shown in Figure 5) as special requirements of the rare species were not taken into account. The result was similar to other studies which also showed low correlation between species diversity and rare species [40]. In other words, the only dependency on existing protective measures cannot satisfy conservation need.

Results from gap analysis revealed conservation deficiency of current protection system and provides insights for further conservation efforts. By comparing different scenarios in Figs. 6, sites of extreme species richness (0.9) were located mainly in southern Lantau while sites of high (0.7) and moderate (0.5) were found in both LNCP and LSCP. However, the concordance of species richness with the LSCP was more significant than that with LNCP as shown in Figure 7. Although the coverage of the LSCP is extensive, extreme species-rich areas are still under-protected in southern island (shaded in red) in Figure 6a. Conservation gaps expanded to the entire southern Lantau and to an extensive part of the northern island when high to moderate level of species richness were considered in Figs. 6a and b.

The results revealed that 45% of species-rich hotspots tended to fall outside these protective borders. Such non-concordance suggested species richness provided another dimension for conservation planning and development control. In fact, no comprehensive guidelines are currently developed for the designation of SSSI or conservation areas because it is technically difficult if not infeasible in acquiring reliable data to formulate protection plans.

During the study period, an extension of the LNCP boundary (Figure 1) with the primary objective to protect unspoiled stream courses and natural woodland of high conservation and landscape value has been proposed. The results also indicated that extra conservation effort was necessary in Lantau as a high percentage of areas were unprotected. In Figure 6c, the proposed extension was overlaid on top of the results. Red areas suggest extra conservation efforts are necessary due to relatively high species richness. However, less than half of these areas

overlapped with the proposed extension and they were mainly found in the north near Sha Lo Wan and Tai Ho while the extension in north-eastern Lantau showed little coherence with species richness. This may be primarily due to the hill fire occurred at time when the study was conducted. Nevertheless, it is believed that priority of conservation should be given to southern Lantau where species richness is the highest and conservation gap is most significant.

5. Conclusions

Traditional species survey represented species distribution by discrete points. By linking indirect environmental gradients derived from timely remote sensing and GIS database and representative ground-truth survey data using multivariate statistical models, habitats of fifty common species were represented in spatially-continuous map. Three multivariate statistical models, ENFA, BLRM, and GAM were used to understand and model the living habitats of individual species. ENFA transformed the survey data into meaningful information to generally explore the habitats of species based on marginality and tolerance. Besides, results from ENFA provided reliable pseudo-absence data that were generated for BLRM and GAM regression analyses. Environmental gradients selected by stepwise BLRM and GAM further enhanced the understanding of species habitats. By comparing the two models, GAM performed better than BLRM in both reliability and discriminability. Local documentation suggested that habitat prediction showed both concordances and discrepancies.

Biodiversity hotspots were identified based on computation of species richness using the prediction results of GAM. Using various thresholds for species richness

definition, the extremely species-rich sites were found in southern Lantau. High to moderate species-rich sites extended towards northern Lantau. Comparison of species richness and current protection boundaries revealed that about 45% of land areas were under-protected. When this study was in progress, the government has proposed an extension of conservation areas and also tourism development in Lantau. The results proved to be an important justification for evaluating the validity of the proposed extension. The species richness map acts as crucial information for conservation zoning.

All in all, Hong Kong is always facing the scarcity of land resources. Continuous demand for housing and economic expansion exerted tremendous pressures on unspoiled land for development. Balancing the objectives of development and conservation is always the prime concern of government and local citizens. Species richness can be one of the ecological indicators for partitioning the areas into different levels of ecological importance, based on which the conflicts between development and conservation can be effectively minimized. The approach enables timely, efficient and effective measurement of spatial distributions of different species, which is a valuable tool for conservation management and policy making. The results acted as a scientific, quantitative foundation and provided objective guidelines for planners, government, developers and other conservation parties to identify sensitive areas and evaluate existing protection system.

Supplementary Materials

Additional Supplementary Material may be found in the online version of this article:

Table S1. Description of the fifty species, their protection status and the number of points in biodiversity survey

Class	Order	Family	Species	Abbrev.	Common English names	Statutory Protection in HK		No. of samples from Biodiversity Survey
						Cap. 170	Cap. 187	
Amphi-bia	Anura	Bufonidae	<i>Bufo melanostictus</i>	Bm	Asian Common Toad	No	No	130 (129)
	Anura	Microhylidae	<i>Kaloula pulchra</i>	Kp	Asiatic Painted Frog	No	No	100 (99)
	Anura	Rhacophoridae	<i>Polypedates megacephalus</i>	Pm	Brown Tree Frog	No	No	271 (259)
	Anura	Ranidae	<i>Rana livida</i>	Rl	Green Cascade Frog	No	No	48 (48)
	Anura	Ranidae	<i>Rana guentheri</i>	Rg	Günther's Frog	No	No	209 (207)
	Anura	Ranidae	<i>Rana exilispinosa</i>	Re	Lesser Spiny Frog	No	No	129 (128)
	Anura	Microhylidae	<i>Microhyla ornata</i>	Mo	Ornate Pigmy Frog	No	No	103 (100)
	Anura	Ranidae	<i>Rana limnocharis</i>	Rl	Paddy Frog	No	No	158 (155)
Aves	Falconiformes	Accipitridae	<i>Milvus migrans</i>	Mm	Black Kite	Yes	Yes	67 (44)
	Passeriformes	Pycnonotidae	<i>Pycnonotus sinensis</i>	Ps	Chinese Bulbul	Yes	No	97 (74)

	Passeriformes	Sylviidae	<i>Orthotomus sutorius</i>	Os	Common Tailorbird	Yes	No	86 (69)
	Passeriformes	Pycnonotidae	<i>Pycnonotus jocosus</i>	Pj	Crested Bulbul	Yes	No	92 (72)
	Passeriformes	Paridae	<i>Parus major</i>	Pm	Great Tit	Yes	No	59 (52)
	Cuculiformes	Cuculidae	<i>Centropus sinensis</i>	Cs	Greater Coucal	Yes	No	46 (32)
	Passeriformes	Zosteropidae	<i>Zosterops japonica</i>	Zj	Japanese White-eye	Yes	No	79 (64)
	Passeriformes	Corvidae	<i>Corvus macrorhynchus</i>	Cm	Jungle Crow	Yes	No	47 (36)
	Ciconiiformes	Ardeidae	<i>Egretta garzetta</i>	Eg	Little Egret	Yes	No	25 (19)
	Passeriformes	Laniidae	<i>Lanius schach</i>	Ls	Rufous-backed shrike	Yes	No	44 (31)
	Columbiformes	Columbidae	<i>Streptopelia chinensis</i>	Sc	Spotted Dove	Yes	No	82 (64)
	Passeriformes	Turdidae	<i>Myiophonus caeruleus</i>	Mc	Violet Whistling Thrush	Yes	No	47 (33)
	Gruiformes	Rallidae	<i>Amauornis phoenicurus</i>	Ap	White-breasted Waterhen	Yes	No	25 (19)
Insecta	Lepidoptera	Satyridae	<i>Melanitis leda leda</i>	Mll	Common Evening Brown	No	No	94 (82)
	Lepidoptera	Pieridae	<i>Cepora nerissa nerissa</i>	Cnn	Common Gull	No	No	38 (34)
	Lepidoptera	Papilionidae	<i>Chilasa clytia clytia</i>	Ccc	Common Mime	No	No	82 (77)
	Lepidoptera	Papilionidae	<i>Papilio polytes polytes</i>	Pppo	Common Mormon	No	No	182 (168)
	Lepidoptera	Pieridae	<i>Pieris canidia canidia</i>	Pcc	Common White	No	No	131 (123)
	Lepidoptera	Papilionidae	<i>Papilio memnon agenor</i>	Pma	Great Mormon	No	No	129 (119)
	Lepidoptera	Pieridae	<i>Hebomoia glaucippe glaucippe</i>	Hgg	Great Orange Tip	No	No	86 (81)
	Lepidoptera	Lycaenidae	<i>Zizeeria maha serica</i>	Zms	Pale Grass Blue	No	No	104 (96)
	Lepidoptera	Papilionidae	<i>Papilio paris paris</i>	Pppa	Paris Peacock	No	No	143 (133)
	Lepidoptera	Nymphalidae	<i>Junonia almana almana</i>	Jaa	Peacock Pansy	No	No	95 (87)
	Lepidoptera	Riodinidae	<i>Abisara echerius echerius</i>	Aee	Plum Judy	No	No	142 (128)
	Lepidoptera	Pieridae	<i>Ixias pyrene pyrene</i>	Ipp	Yellow Orange Tip	No	No	32 (29)
Insecta	Odonata	Protoneuridae	<i>Prodasineura autumnalis</i>	Pa	Black Threadtail	No	No	46 (46)
	Odonata	Euphaeidae	<i>Euphaea decorata</i>	Ed	Black-banded Gossamerwing	No	No	58 (57)
	Odonata	Platycnemididae	<i>Copera ciliata</i>	Cci	Black-kneed Featherlegs	No	No	41 (41)
	Odonata	Platycnemididae	<i>Coeliccia cyanomelas</i>	Ccy	Blue Forest Damsel	No	No	39 (39)
	Odonata	Chlorocyphidae	<i>Rhinocypha perforata perforata</i>	Rpp	Common Blue Jewel	No	No	56 (56)
	Odonata	Libellulidae	<i>Orthetrum pruinosum neglectum</i>	Opn	Common Red Skimmer	No	No	71 (68)
	Odonata	Libellulidae	<i>Trithemis aurora</i>	Ta	Crimson Dropwing	No	No	60 (58)
	Odonata	Libellulidae	<i>Zygonyx iris insignis</i>	Zii	Emerald Cascader	No	No	44 (44)
	Odonata	Aeshnidae	<i>Anax immaculifrons</i>	Ai	Fiery Emperor	No	No	26 (25)

	Odonata	Libellulidae	<i>Orthetrum 161 sabina sabina</i>	Oss	Green Skimmer	No	No	70 (68)
	Odonata	Libellulidae	<i>Trithemis festiva</i>	Tf	Indigo Dropwing	No	No	67 (66)
	Odonata	Libellulidae	<i>Pantala flavescens</i>	Pf	Wandering Glider	No	No	107 (103)
	Odonata	Platycnemididae	<i>Copera marginipes</i>	Cm	Yellow Featherlegs	No	No	53 (53)
	Rodentia	Muridae	<i>Niviventer fulvescens</i>	Nf	Chestnut Spiny Rat	No	No	64 (63)
Mammalia	Rodentia	Hystriidae	<i>Hystrix brachyura</i>	Hb	East Asian Porcupine	Yes	No	41 (41)
	Artiodactyla	Suidae	<i>Sus scrofa</i>	Ss	Eurasian Wild Pig	No	No	53 (53)
	Rodentia	Muridae	<i>Rattus sikkimensis</i>	Rs	Sikkim Rat	No	No	66 (63)

Table S2a. Definition and statistical description of the environmental variables

VARIABLES – DESCRIPTION	ABBREV*	SPATIAL OPERATION/ ANALYSIS	UNITS	STATISTICAL DESCRIPTION			
				MEAN	SD	MIN	MAX
Woodland —Lands covered with natural forests including both montane forest and lowland forest.	WOOD	Circular Analysis (Frequency of Occurrence) on the Boolean <i>woodland</i> map with radius of 500m	Percentage	26.53	27.89	0	100
Shrubby Grassland — Lands covered predominantly (50% or more) by grasses and contained visible woody plants covering up to 50% of the area.	SHGRASS	Circular Analysis (Frequency of Occurrence) on the Boolean <i>shrubby grassland</i> map with radius of 500m	Percentage	27.93	21.08	0	98
Grassland including Hillfire — Lands covered predominantly (50% or more) by grasses with no visible woody plants.	GRASS	Circular Analysis (Frequency of Occurrence) on the Boolean <i>grassland</i> (combined with hillfire region) map with radius of 500m	Percentage	17.09	15.19	0	93
Mixed shrubland — Lands covered with less than 50% grasses with shrubs the major woody life form.	MSHRUB	Circular Analysis (Frequency of Occurrence) on the Boolean <i>mixed shrubland</i> map with radius of 500m	Percentage	9.90	13.21	0	93
Bareland — Naturally open rock faces or disturbed lands or “badlands” denuded of vegetation.	BARE	Circular Analysis (Frequency of Occurrence) on the Boolean <i>bareland</i> map with radius of 500m	Percentage	2.31	5.70	0	87

*Abbreviations are provided for reference in the result

Table S2b. Definition and description of the environmental variables extracted from GIS database

VARIABLES – DESCRIPTION		ABBREVI*	SPATIAL OPERATION/ ANALYSIS	UNITS	STATISTICAL DESCRIPTION			
					MEAN	SD	MIN	MAX
Elevation		ELEV	The altitude recorded on the 10-meter digital elevation model	Meter	128.92	137.98	0	957
Aspect	Southness index	SOUTH	Derived from equation: $\cos[(\text{Aspect}) + 180]$	Unitless	-0.16	0.73	-1	1
	Eastness index	EAST	Derived from equation: $\sin(\text{aspect})$	Unitless	-0.10	0.71	-1	1
Slope	Slope	SLOPE	Derived from 10-meter digital elevation model	Degree	17.63	12.42	0	85.79
Curvature	Planimetric curvature	PLAN	Surface curvature in the direction of slope derived from 10-meter digital elevation model	Unitless	-0.51	2.14	-37.75	37.75
	Profile curvature	PROFILE	Surface curvature along contour derived from 10-meter digital elevation model	Unitless	0.50	1.97	-37.75	37.75
Potential Solar Radiation	Vernal Equinox	VERN	Potential solar radiation of vernal equinox (21 Mar) derived from the combination of elevation, aspects and slope from GRASS v.6.0	W/ m ²	10.06	1.12	0	12.01
	Summer Solstice	SUMM	Potential solar radiation of summer solstice (22 Jun) derived from the combination of elevation, aspects and slope from GRASS v.6.0	W/ m ²	11.23	1.18	0	13.36
	Autumnal Equinox	AUTU	Potential solar radiation of autumnal equinox (21 Sept) derived from the combination of elevation, aspects and slope from GRASS v.6.0	W/ m ²	10.07	1.12	0	12.02
	Winter Solstice	WINT	Potential solar radiation of winter solstice (22 Dec) derived from the combination of elevation, aspects and slope from GRASS v.6.0	W/ m ²	8.14	1.66	0	10.64
Topographic Distance operations	Coast	COAST	Distance from coastline derived from the topographic distance operation from BIOMAPPER v.3.1	meter	1847.01	1748.24	0	8294.28
	Railway	RAIL	Distance from railways derived from the topographic distance operation from BIOMAPPER v.3.1	Meter	8510.23	6021.78	0	26742.79
	Main Road	MAINROAD	Distance from mainroads derived from the topographic distance operation from BIOMAPPER v.3.1	Meter	3583.60	4113.92	0	20750.79
	Building	BLDG	Distance from buildings derived from the topographic distance operation from BIOMAPPER v.3.1	Meter	2098.88	3053.41	0	19243.63
Frequency of occurrence operation	River	RIVER	Frequency of occurrence operation (circular analysis) from BIOMAPPER v.3.1 with 300m radius	Percentage	1.05	1.84	0	19
	Stream	STREAM100	Frequency of occurrence operation (circular analysis) from BIOMAPPER v.3.1 with 100m radius	Percentage	4.63	5.28	0	35
		STREAM500	Frequency of occurrence operation (circular analysis) from BIOMAPPER v.3.1 with 500m radius	Percentage	2.54	3.58	0	19

*Abbreviations are provided for reference in the result

Table S3. Summary statistics of Ecological Niche Factor Analysis of the fifty species

	Abbrev.	No. of points from Bd. Survey	^Δ Number of input EGVs	^Φ Number of Ecological-niche Factors	[*] Amount of information explained (%)	[#] Amount of variance explained (%)	^Ω M	^δ T	[^] S
Amphibia	Bm	130	27	4	70.00	38.90	0.66	0.76	1.32
	Kp	100	12	7	91.00	83.00	0.52	0.77	1.30
	Pm	271	13	3	71.00	42.00	0.35	0.87	1.15
	Rl	48	27	4	91.00	82.00	1.63	0.30	3.31
	Rg	209	13	3	73.00	46.00	0.49	0.82	1.22
	Re	129	27	10	87.70	75.50	1.33	0.66	1.53
	Mo	103	13	4	81.00	63.00	0.65	0.75	1.33
	Rl	158	13	3	74.00	48.00	0.54	0.78	1.29
Aves	Mm	67	16	2	71.30	42.60	0.55	0.60	1.66
	Ps	97	16	3	71.60	43.20	0.51	0.81	1.24
	Os	86	16	10	93.70	87.50	0.70	0.68	1.48
	Pj	92	16	5	81.10	62.20	0.62	0.75	1.34
	Pm	59	16	11	95.00	90.00	0.68	0.67	1.49
	Cs	46	16	3	86.90	73.80	0.60	0.43	2.31
	Zj	79	16	4	74.70	49.40	0.53	0.76	1.33
	Cm	47	16	3	82.70	65.50	0.66	0.54	1.85
	Eg	25	16	4	98.50	97.00	0.64	0.16	6.24
	Ls	44	16	6	91.60	83.10	0.64	0.48	2.10
	Sc	82	16	10	92.20	84.50	0.47	0.77	1.30
	Mc	47	16	2	83.20	66.50	0.82	0.49	2.06
	Ap	25	16	3	93.20	86.30	0.51	0.25	3.98
Insecta (Lepidoptera)	Mll	94	26	6	78.00	56.00	0.71	0.70	1.42
	Cnn	38	26	5	89.30	78.60	0.64	0.37	2.71
	Ccc	82	26	5	73.60	47.20	0.66	0.71	1.41
	Pppo	182	26	5 (2)	71.80	43.60	0.60	0.80	1.26
	Pcc	131	26	5	72.70	45.30	0.65	0.76	1.31
	Pma	129	26	5 (2)	70.90	41.90	0.58	0.77	1.31
	Hgg	86	26	8	83.00	65.90	0.58	0.71	1.41
	Zms	104	26	5 (4)	72.50	45.10	0.61	0.76	1.32
	Pppa	143	26	5 (4)	71.50	43.00	0.65	0.80	1.25
	Jaa	95	26	5 (2)	75.30	50.60	0.76	0.70	1.44
	Aee	142	26	6 (3)	72.80	45.60	0.63	0.81	1.23
	Ipp	32	26	4	92.80	85.70	0.75	0.24	4.17
Insecta (Odonata)	Pa	46	16	5	83.50	66.90	0.68	0.64	1.56
	Ed	58	16	10	94.00	88.00	0.74	0.68	1.48
	Cci	41	16	4	82.20	64.30	0.68	0.61	1.65
	Ccy	39	16	6	88.20	76.40	0.71	0.61	1.64
	Rpp	56	26	7	82.70	65.50	0.92	0.64	1.58
	Opn	71	16	10	93.20	86.30	0.56	0.78	1.29
	Ta	60	26	8	84.40	68.80	0.73	0.64	1.56
	Zii	44	16	5	83.00	66.10	0.63	0.69	1.46
	Ai	26	16	10	96.90	93.90	0.85	0.49	2.04
	Oss	70	26	6	77.40	54.70	0.57	0.73	1.38
	Tf	67	26	6	79.00	58.00	0.75	0.67	1.49
	Pf	107	26	5 (2)	71.20	42.50	0.63	0.79	1.27
	Cm	53	26	8	86.60	73.10	0.77	0.59	1.71
Mammalia	Nf	64	16	3	85.70	71.30	1.01	0.50	2.02
	Hb	41	26	6	89.90	84.30	1.00	0.36	2.77
	Ss	53	26	6	94.30	83.90	1.23	0.36	2.80
	Rs	66	26	7	85.10	72.90	1.06	0.55	1.84

^Δ Number of Box-Cox transformed EGVs includes.

^Φ The number of factors chosen according to the broken-stick advice.

^{*} The amount of variance explained by both the marginality and specialization factors.

[#] The amount of variance explained by the specialization factors only.

^Ω Global Marginality – A summary statistic describing the difference between the species optimal living condition and the environmental conditions within the study area. Value range from 0 to 1. Low value (approaching 0) suggests that a species tends to live in average conditions throughout the study area while high value (approaching 1) shows that a species is likely to live in extreme environment.

^δ Global Tolerance – A summary statistic describing the range of living condition of a species within the study area. Value range from 0 to 1. Low value (approaching 0) depicts a species can only survive in specific range of condition while high value (approaching 1) indicates a general living condition is acceptable for a species.

[^] Global Specialization – A summary statistic which is the reciprocal of tolerance. Value range from 0 to infinity.

Table S4a. Selected predictors and corresponding β_i of BLRM for the amphibian species

Abbrev	No. of variables	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	RIVER	SLOPE	STREAM 100
Bm	5			+1.37				-0.038			+0.109	+0.281	-0.049	
Kp	4	-0.006									-0.036	+3.96	-0.167	
Pm	5	-0.006							+0.044	+0.017	+0.035	+0.591		
Rl	3	+0.04									+0.154	+1.151		
Rg	5				-0.011				+0.107	+0.034	+0.071	+0.586		
Re	6				+0.014				+0.074	+0.073	+0.129	+0.638		+0.153
Mo	5				-0.006		+0.33			+0.3			-0.093	-0.181
Rl	7			+0.969	-0.017				+0.148	+0.06	+0.104	+0.943	-0.187	

Table S4b. Selected predictors and corresponding β_i of BLRM for the bird species

Abbrev	No. of variables	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	RIVER	SLOPE	STREAM 100
Mm	3			+1.733							+0.126	-0.731		
Ps	4					+0.001			+0.075	+0.049	+0.086			
Os	5				+0.009					+0.079	+0.113	-0.52		-0.163
Pj	3							-0.141	+0.572				+0.16	
Pm	2				+0.013						+0.098			
Cs	2			+2.944				-0.268						
Zj	2									+0.077	+0.158			
Cm	1										+0.136			
Eg	2				-0.016			-0.191						
Ls	3		-0.001					-0.085	+0.137					
Sc	4								+0.104	+0.046	+0.054		+0.067	
Mc	2							-0.087					+0.129	
Ap	1							-0.123						

Table S4c. Selected predictors and corresponding β_i of BLRM for the butterfly species

Abbrev	No. of variables	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	RIVER	SLOPE	STREAM 100
Mll	5	-0.028			+0.036				+0.264	+0.073	+0.437			
Cnn	4	-0.005			+0.018			-0.001		+0.002				
Ccc	5				-0.007	-0.001			+0.11	+0.028	+0.02			
Pppo	4	-0.008							+0.191	+0.04	+0.091			
Pcc	7			+0.464	-0.003				+0.094	+0.021	+0.023	+0.867	-0.083	
Pma	4	-0.008							+0.088	+0.047	+0.096			
Hgg	3	-0.008						-0.067		-0.04				
Zms	6	-0.007		+0.894					+0.159	+0.023	+0.066	+0.552		
Pppa	6	-0.01			+0.016				+0.166	+0.029	+0.268	+0.342		
Jaa	6				-0.01	-0.002			+0.286	+0.073	+0.076		-0.065	
Aee	5	-0.009			+0.013				+0.123	+0.036	+0.144			
Ipp	1										+0.344			

Table S4d. Selected predictors and corresponding β_i of BLRM for the dragonfly species

Abbrev	No. of variables	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	RIVER	SLOPE	STREAM 100
Pa	5				-0.02				+0.308	+0.088	+0.165	+0.665		
Ed	5								+0.144	+0.092	+0.188	+0.7	-0.167	
Cci	2										+0.248	+1.22		
Ccy	3								+0.063		+0.066	+1.12		
Rpp	4								+0.252	+0.078	+0.168	+0.622		
Opn	5	-0.004							+0.062	+0.049	+0.58	+0.239		
Ta	6			+1.397					+0.146	+0.062	+0.062	+0.243	-0.085	
Zii	5	-0.006							+0.157	+0.109	+0.162		-0.135	
Ai	2						-2.385							+0.356
Oss	4									+0.044	+0.075	+0.528	-0.059	
Tf	4								+0.143	+0.054	+0.175	+0.44		
Pf	4								+0.057	+0.064	+0.083			-0.111
Cm	6	-0.005				-0.001			+0.091	+0.076	+0.104	+0.317		

Table S4e. Selected predictors and corresponding β_i of BLRM for the mammal species

Abbrev	No. of variables	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	RIVER	SLOPE	STREAM 100
Nf	4						-1.538		+0.169		+0.078	+0.64		
Hb	4		-0.001								+0.06		+0.109	+0.535
Ss	5		+0.001	-2.18	-0.041						+0.286			+0.591
Rs	4				+0.011	+0.001		-0.085			+0.058			

Table S5a. Selected predictors by AIC for generalized additive models of amphibian species

Abbrev.	No. of variables	AUTU	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	PLAN	PRO-FILE	RAIL	RIVER	SLOPE	STREAM 100
Bm	7	Non			+		Non		---	Non						Non		-
Kp	6	++	Non		Non		-		-		Non							
Pm	13	+	---	-	Non	+	Non	Non	+++		Non		Non	+	Non	Non		
Rl	2	Non							---									
Rg	9	No	-	Non	Non			-	-	Non	Non					Non		
Re	8	---		++		++	---	-	---	Non						Non		
Mo	7	Non	Non	Non				Non		Non	Non					Non		
Rl	9	Non	Non					-	---	Non	Non		Non		++	Non		

Table S5b. Selected predictors by AIC for generalized additive models of bird species

Abbrev.	No. of variables	AUTU	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	PLAN	PRO-FILE	RAIL	RIVER	SLOPE	STREAM 100
Mm	6	-	Non		Non				Non				Non			-		
Ps	6		---				+		Non	+						Non		Non
Os	9		Non		Non		++	-	---	Non	Non					Non		Non
Pj	4	-							-	+								Non
Pm	5	Non						-	Non	+	Non							
Cs	4							-	---		Non					Non		
Zj	5	-	Non		Non				Non	No								
Cm	5	Non	-	Non					---									Non
Eg	3				Non				---									Non
Ls	5		---	-						+					+	Non		
Sc	8	---						-	-	+	Non		Non	Non		Non		
Mc	6	Non		Non			Non		-	Non				Non				
Ap	3			Non			+	-										

Table S5c. Selected predictors by AIC for generalized additive models of butterfly species

Abbrev.	No. of variables	AUTU	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	PLAN	PRO-FILE	RAIL	RIVER	SLOPE	STREAM 100
Mll	6	Non	-		Non	Non				Non								Non
Cnn	5	Non						-	---				Non					Non
Ccc	5		Non					-	-	Non						Non		
Pppo	7		-			Non	Non			No	Non	Non				No		
Pcc	6		---	Non	Non					Non						Non		Non
Pma	6		-			Non			Non	Non	Non							+
Hgg	7		-		Non	Non	-		Non		Non					Non		
Zms	9		-		Non		Non	-	Non	Non	Non	Non				Non		
Pppa	9	---	-	Non		+		-	---	+	Non					Non		
Jaa	6		-				Non			Non	++	Non				Non		
Aee	9	Non	---			+	Non	-	---	+	Non					Non		
Ipp	6				++		Non			+++	+++					---		Non

Table S5d. Selected predictors by AIC for generalized additive models of dragonfly species

Abbrev.	No. of variables	AUTU	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	PLAN	PRO-FILE	RAIL	RIVER	SLOPE	STREAM 100
Pa	6		---					Non		Non					Non	Non		Non
Ed	6		-			Non	Non		---		Non					Non		
Cci	3		Non							+						No		
Ccy	6		Non	Non		Non	-								Non	No		
Rpp	6					Non	Non	Non		+	Non					Non		
Opn	6		-				Non		-	Non	Non							Non
Ta	8	Non	-			Non	Non			Non			Non		Non	Non		
Zii	5		-				Non		Non		Non					No		
Ai	2								Non	+								
Oss	7		-	Non			+	-		Non	Non					+		
Tf	6	Non		Non		+			Non	Non						Non		
Pf	6			Non				-	-	Non						+		Non
Cm	6	Non	-		Non			Non		Non						Non		

Table S5e. Selected predictors by AIC for generalized additive models of mammal species

Abbrev.	No. of variables	AUTU	BLDG	COAST	EAST	ELEV	MAIN ROAD	BARE	GRASS	MSHRUB	SH GRASS	WOOD	PLAN	PRO-FILE	RAIL	RIVER	SLOPE	STREAM 100
Nf	6		Non			Non		No	Non	+						Non		
Hb	4	Non		Non				-										Non
Ss	4						++	-		+++	---							
Rs	6	Non		Non	Non		+		Non									Non

Keys:

(No) No distinctive relationship

(Non) Non-linear relationship

(-) Slight linear negative relationship

(-) Moderate linear negative relationship

(-) Strong linear negative relationship

(+) Slight linear positive relationship

(++) Moderate linear positive relationship

(++) Strong linear positive relationship

REFERENCES

- [1] Scott, J.M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S., D'Erchia, F., Edwards, T.C. Jr., Ulliman, J. and Wright, R.G. Gap analysis: a geographic approach to protection of biological diversity. *Wildlife Monograph*, 57, pp. 1-41, 1993.
- [2] Hollander, A.D., Davis, F.W., and Stoms, D.M., Hierarchical representations of species distributions using maps, images and sighting data. In Miller, R.I. (Ed.) *Mapping the diversity of nature*, Chapman and Hall, Oxford. pp. 71-88, 1994.
- [3] Leeuw, J.D., Ottichilo, W.K., Toxopenus, A.G. and Prins, H.H.T., Application of remote sensing and geographic information systems in wildlife mapping and modelling. In A. Skidmore & H. Prins (Ed.) *Environmental modelling with GIS and Remote sensing*, Taylor and Francis. New York. pp. 121-144, 2002.
- [4] Austin, M.P. (1980). Searching for a model for use in vegetation analysis. *Vegetatio*, 42, pp. 11-21, 1980.
- [5] Guisan, A. and Zimmermann, N.E. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, pp. 147-186, 2000.
- [6] Jeganathan, P., Green, R.E., Norris, K., Vogiatzakis, I.N., Bartsch, A., Wotton, S.R., Bowden, C.G.R., Griffiths, G.H., Pain, D. and Rahmani, A.R. Modelling habitat selection and distribution of the critically endangered Jerdon's courser *Rhinoptilus bitorquatus* in scrub jungle: an application of a new tracking method. *Journal of Applied Ecology*, 41, pp. 224-237, 2004.
- [7] Osborne, P.E. and Tigar, B.J. Interpreting bird atlas data using logistic models: an example from Lesotho, Southern Africa. *Journal of Applied Ecology*, 29, pp. 55-62, 1992.
- [8] Buckland, S.T. and Elston, D.A. Empirical models for the spatial distribution of wildlife. *Journal of Applied Ecology*, 30, pp. 478-495, 1993.
- [9] Augustin, N.H., Muggleston, M.A. and Buckland, S.T. An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology*, 33, pp. 339-347, 1996.
- [10] Cowley, M.J.R., Wilson, R.J., León-Cortés, J.L., Guitierrez, D., Bulman, C.R. and Thomas, C.D. Habitat-based statistical models for predicting the spatial distribution of butterflies and day-flying moths in a fragmented landscape, *Journal of Applied Ecology*, 37, pp. 60-71, 2000.
- [11] Walker, P.A. Modelling wildlife distributions using a geographic information system: Kangaroos in relation to climate. *Journal of Biogeography*, 17, pp. 279-289, 1990.
- [12] Mladenoff, D.J., Sickley, T.A., Haight, R.G. and Wydeven, A.P. A Regional Landscape Analysis and Prediction of Favorable Gray Wolf Habitat in the Northern Great Lakes Region. *Conservation Biology*, 9, pp. 279-294, 1995.
- [13] Austin, M.P. and Meyers, J.A. Current approaches to modelling the environmental niche of eucalyptus: implications for management of forest biodiversity. *Forest Ecology and Management*, 85, pp. 95-106, 1996.
- [14] Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J. and Aulagnier, S (1996). Application of neural networks to modelling nonlinear relationships in ecology. *Ecological Modelling*, 1634, pp. 1-13, 1996.
- [15] Luoto, M., Kuussaari, M. and Toivonen, T. Modelling butterfly distribution based on remote sensing data. *Journal of Biogeography*, 29, pp. 1027-1037, 2002.
- [16] Gibson, L.A., Wilson, B.A., Cahill, D.M. and Hill, J. Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *Journal of Applied Ecology*, 41, pp. 213-22, 2004.
- [17] Hirzel, A.H., Hausser, J., Chessel, D. and Perrin, N. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, 83, pp. 2027-2036, 2002.
- [18] Western, W. and Gichohi, H. Segregation effects and the impoverishment of savanna parks: the case for ecosystem viability analysis. *African Journal of Ecology*, 31, pp. 269-281, 1993.
- [19] Dudgeon, D. and Corlett, R. *The Ecology and Biodiversity of Hong Kong*. Friends of the Country Parks and Joint Publishing (HK) Company Limited, Hong Kong, 2004.
- [20] Gaston, K.J. Species richness: measure and measurement. In K.J. Gaston (Ed.) *Biodiversity: A Biology of Numbers and Difference*. Blackwell Science, Cambridge, pp. 77-113, 1996.
- [21] Prendergast, J.R., Quinn, R.M., Lawton, J.H., Eversham, B.C. and Gibbons, D.W. Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature*, 365, pp. 335-337, 1993.
- [22] Williams, P.H. Measuring more of biodiversity for choosing conservation areas, using taxonomic relatedness. In K. Taehakkyo, P. Hanguk and K. Yönguso (Ed.) *International Symposium on Biodiversity and Conservation*. Korean Entomological Institute, Korea University, Korea, pp. 194-227, 1993.
- [23] Curnutt, J., Lockwood, J., Lue, H.-K., Nott, P. and Russel, G. Hotspots and species diversity. *Nature*, 367, pp. 326-327, 1994.
- [24] Maddock, A. and Plessis, M.A.D. Can species data only be appropriately used to conserve biodiversity? *Biodiversity Conservation*, 8, pp. 603-615, 1999.
- [25] Scott, J.M., Csuti, B., Jacobi, J.D. and Estes, J.E. Species richness, a geographic approach to protection future biological diversity. *BioScience*, 37, pp. 782-788, 1987.
- [26] Scott, J.M., Csuti, B., Smith, K., Estes, J.E. and Caicco, S. Beyond endangered species: an integrated conservation strategy for the preservation of biological diversity. *Endangered Species Update*, 5, pp. 43-48, 1998.
- [27] AFCD. Lantau Island: Its Beautiful Countryside. 2nd ed. Friends of the Country Parks, Hong Kong, 2005.
- [28] LDTF. Concept Plan for Lantau: Consultation Digest. The Government of the Hong Kong Special Administrative Region, Hong Kong, 2004.
- [29] Yip, J.Y., Corlett, R.T. and Dudgeon, D. A fine-scale gap analysis of the existing protected area system in Hong Kong,

China. Biodiversity Conservation, 13, pp. 943–957, 2004.

- [30] Hosmer, D.W. and Lemeshow, S. Confidence interval estimates of an index of quality performance based on logistic regression models. *Statistics in Medicine*, 14, pp. 2161-2172, 1995.
- [31] Hastie, T.J. and Tibshirani, R.J. *Generalized Additive Models*. Chapman and Hall, London, 1990.
- [32] Yee, T.W. and Mitchell, N.D. Generalized Additive Models in plant ecology. *Journal of Vegetation Science*, 2, pp. 587–602, 1991.
- [33] Lehmann, A., Leathwick, J.R. and Overton, J.Mc.C. GRASP v.3.0 User's Manual. Swiss Centre for Faunal Cartography, Switzerland, 2004.
- [34] Pearce, J. and Ferrier, S. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133, pp. 255-245, 2000.
- [35] Chan, K.F.S., Cheung, K., Ho, C.Y., Lam, F.N., Tang, W.S., Lau, W.N.M. and Bogadek, A. *A Field Guide to the Amphibians of Hong Kong*. Friends of the Country Parks. Cosmos Books Limited, Hong Kong, 2005.
- [36] Lo, P.Y.F. and Hui, W.L. *Hong Kong Butterflies*. 1st ed. Hong Kong: Tin Dei Book Co. Ltd, Hong Kong, 2004.
- [37] AFCD, *Butterflies of Hong Kong*. Hong Kong Biodiversity. Online available from <http://www.afcd.gov.hk/english/conservation/hkbiodiversity/hkbiodiversity.html>, 2005.
- [38] Strange, M. *Tropical Birds of Southeast Asia*. Periplus Editions (HK) Ltd, Hong Kong, 1998.
- [39] Carey, G.J., Chalmers, M.L., Diskin, D.A., Kennerley, P.R., Leader, P.J., Leven, M.R., Lewthwaite, R.W., Melville, D.S., Turnbull, M. and Young, L. *The Avifauna of Hong Kong*. Hong Kong Bird Watching Society, Hong Kong, 2001.
- [40] Williams, P., Gibbons, D., Margules, C., Rebelo, A., Humphries, C. and Pressey, R. A comparison of richness hotspots, rarity hotspots and complementary areas for conserving diversity of British birds. *Conservation Biology*, 10, pp.155-174, 1996.