

A study on the distribution of *Papilionidae* in Thailand based on statistical analysis

Turenjai Doolgindachbaporn¹, Chutinan Choosai¹,
Toshiaki Aoki², Shuhei Yamaguchi² and Shin'ichi Okamoto^{2,3*}

¹*Khon Kaen University, Khon Kaen 40002 Thailand*

²*Research Institute of Evolutionary Biology 2-4-28 Kamiyoga, Setagaya-ku, Tokyo, 158-0098 Japan*

³*Professor Emeritus of Tokyo University of Information sciences*

*Corresponding author: okamoto@rsch.tuis.ac.jp

Abstract

Many earlier studies on Thai flora and fauna showed that the richness of the ecosystem in this region arises from biogeographic situation. The distribution of Lepidoptera species is highly depended on the distribution of flora, especially food plants. Therefore, by analyzing their distribution data, we can get some insight into the flora and fauna of this area. Here, we used the *Papilionidae* family data. The raw data were composed by 0 or 1 in each cell of the spread sheet according to the presence or absence in a province. In this analysis, we could analyze the spatial differences and similarities of provinces from the view point of the butterfly's distribution. By conducting a factor analysis with the principal component analysis mode, five factors were obtained. The meanings of former four factors could be explained. This analytical result seems to support the classification of zoo-geographical sub-regions proposed by many researchers.

Keywords: Biogeography, Lepidoptera, *Papilionidae*, principal component analysis

1. Introduction

Many earlier studies on Thai flora and fauna showed that the richness of the ecosystem in this region arises from biogeographic situation; the southern region is located on the northern edge of the Sunda shelf, and the northern region is heavily influenced by China - Himalayan biosphere (Hughes et al., 2003; Gray et al., 1991).

In case of analyses on the biodiversity, insects were often used as a marker, because of their relative immobility compared to other animals such as birds and mammals. The distribution of Lepidoptera species is highly

depended on the distribution of flora, especially food plants. Therefore, more understanding about the flora and fauna of the ecosystem could be obtained by analyzing their distribution data. In addition, amateur researchers, collectors, and photographers accumulated a large number of specimens and data.

Elliot (1980) suggested an interesting hypothesis for the life zone in Thailand in the preface of volume four of the "Butterflies in Thailand" by Pinratana (1981). He mentioned that Thailand could be divided into four zoogeographic zones in which the more sedentary butterflies, characteristic of a primary forest, tend to become a distinct subspecies.

These four zones were called, (i) the Indo-Burmese zone, covering northwest and west upper 10N latitude, (ii) the Kedawian zone, covering Peninsula below 10N latitude, (iii) the Malayan zone, covering the extreme south-east confined to Yala district, and (iv) the Indo-Chinese zone, covering center and east area and extending to Vietnam. However, this concept has not been supported by quantitative analysis based on a recent database. The aim of this study is to add new insight concerning these hypotheses based on a statistical analysis using a recent database which is available from books and internet web sites.

2. Materials and Methods

Regarding to butterflies, there are many species of *Hesperiidae* and *Lycaenidae* which are difficult to distinguish, and some of the past data contains misidentified ones. Here, we believe that the easiest method to identify is the most appropriate for adopting the *Papilionidae* family data. Data used were almost the same format as our previous analysis (Okamoto, 2011 and 2013). In our previous study some factor loading values tended to be quite different in two adjacent provinces on the map. The data for these provinces might come only from a few hot spots for the collectors, and other sites within the provinces might not be investigated. This problem has the potential of becoming a serious bias. In order to mitigate these problems, we added the latest available data. Recently, there are many websites posting photos of butterflies, and their contents have been improved considerably. For this reason, the data were also extracted from the following three sites.

- Antonio Giudici, Thailand Butterflies species gallery, <https://www.thaibutterflies.com/>
- Les Day, Samuibutterflies, <http://www.samuibutterflies.com/home/>
- Butterflies of Thailand© 2015 - 2018 wingscales.com, <https://wingscales.com/Papilionidae>

Regarding to the provinces adopted, it is almost the same as our previous study. In addition, Lapkratok (2010) recently introduced

the situation of Sa Kaeo province with Pang Sida Falls. As reported by Lapkratok, Pang Sida National Park was recently registered as a World Heritage Site, and it had the highest butterfly diversity. Therefore, we added this province, and extracted the data from the WEB sites above. Moreover, we also added our original photo data as shown in Fig. 1. Almost all taxonomic methods were taken from Kimura et al (2011). However, subdivision for a few subspecies is not consistent among the databases. In this analysis the identification conducted by Inayoshi (2010) and Kimura et al. (2011) was mainly applied.

The raw data were composed by 0 or 1 in each cell of the EXCEL spread sheet. The column of the spread sheet corresponded to the province. The row was related with the species or subspecies. If a species was recorded in a province, the data "1" was inserted in the corresponding cell. If not, data "0" was inserted in the cell. A part of the data sheet is shown in Table 1. Although there are 77 provinces in Thailand, about two thirds were excluded from the analysis. Finally, the data set of 91 species (and subspecies) and 30 provinces were prepared for the statistical analysis.

In this analysis, a province was considered to be a variable; we could analyze the spatial differences and similarities of provinces in term of the butterfly's distribution. Using this data set, a correlation coefficient between two provinces was calculated. Based on a set of correlation coefficients for all combinations of the two provinces, a correlation matrix was constructed. A part of the correlation matrix is shown in Table 2.

3. Results and Discussions

Factor analysis with the principal component analysis mode was conducted by IBM-SPSS package PASW Statistics version 19. At first, communality was assumed to be 1.0 and all of the eigenvectors were calculated. Next, the eigenvector of its corresponding eigenvalues less than 1.0 was deleted. After this screening, five factors were remained and Varimax rotation was conducted. Varimax rotated factor loadings are shown in Table 3. From this table we could easily

Table 1. A part of raw data, species and subspecies used in the analysis for *Papilionidae*

	Mea Hong Son	Chiang Rai	Chiang Mai	Nan	Lampang	Phrae	Tak	Petchabun	Mukdahan	Ubon Ratchabhandi	Khon Kaen	Chaiyaphum	Nakhon Ratchasima	Kanchanaburi	Ratchaburi	Petchaburi	Prachnap Kiri Khan	Saraburi	Nakhon Nayok	Sa Kaeo	Chonburi	Chanthaburi	Trat	Ranong	Phuket	Tarang	Satun	Surat Thani	Nakhon Sri Thammarat	Yala
<i>Bhutanitis linderdallii ocellatomaculata</i>	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
<i>Troides helena cerberus</i>	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	0
<i>Troides aeacus aeacus</i>	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0	0	0	0	0	0	0
<i>Troides aeacus malaitianus</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0
<i>Troides amphrysus ruficollis</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	1
<i>Troides cuneifera pueninsulae</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
<i>Atrophaneura sycorax egerioni</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0
<i>Atrophaneura zaleucus</i>	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	0	0	0	1	0	1	0	0	0	0	0	0	0	0
<i>Atrophaneura varuna varuna</i>	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1
<i>Atrophaneura varuna astorion</i>	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 2. A part of correlation matrix

	Mea Hong Son	ChiangRai	ChiangMai	Nan	Lampang	Phrae	Tak	Phetchabun	Mukdahan
Mea Hong Son	1	0.537	0.590	0.620	0.550	0.432	0.537	0.550	0.201
Chiang Rai	0.537	1	0.592	0.636	0.434	0.373	0.471	0.434	0.036
Chiang Mai	0.590	0.592	1	0.682	0.451	0.389	0.543	0.506	0.211
Nan	0.620	0.636	0.682	1	0.564	0.569	0.537	0.508	0.260
Lampang	0.550	0.434	0.451	0.564	1	0.365	0.612	0.533	0.286
Phrae	0.432	0.373	0.389	0.569	0.365	1	0.482	0.488	0.327
Tak	0.537	0.471	0.543	0.537	0.612	0.482	1	0.493	0.218
Phetchabun	0.550	0.434	0.506	0.508	0.533	0.488	0.493	1	0.286
Mukdahan	0.201	0.036	0.211	0.260	0.286	0.327	0.218	0.286	1

Table 3 Factor loadings after Varimax rotation, variables (provinces) are rearranged according to the values of factor loadings, factor loadings higher than 0.50 are shown in bold character

province/ factor	1	2	3	4	5
Mukdahan	0.792	0.036	-0.026	-0.091	0.321
Chanthaburi	0.747	0.32	0.039	0.29	-0.01
Chonburi	0.746	0.281	0.097	0.034	0.109
Sa Kaeo	0.741	0.171	0.127	0.334	-0.131
Nakhon Nayok	0.738	0.37	0.032	0.424	0.067
Saraburi	0.726	0.118	-0.048	0.296	0.076
Trat	0.709	0.311	0.098	-0.055	-0.317
Nakhon Ratchasima	0.684	0.241	0.008	0.461	0.236
Ubon Ratchathani	0.678	-0.013	0.05	0.131	0.519
KhonKaen	0.633	0.396	0.123	0.339	0.226
Chaiyaphum	0.61	0.52	0.013	0.286	-0.127
Nan	0.211	0.791	-0.241	0.1	0.054
Chiang Rai	0.057	0.771	-0.127	0.019	-0.113
Tak	0.147	0.754	0.108	0.14	0.024
Chiang Mai	0.178	0.744	-0.327	0.042	0.067
Mea Hong Son	0.251	0.735	0.036	0.231	0.107
Lampang	0.153	0.692	0.131	0.154	0.374
Phetchabun	0.397	0.651	0.121	-0.05	0.114
Phrae	0.426	0.569	0.003	0.036	-0.312
Trang	0.065	0.001	0.854	-0.142	0.099
Ranong	-0.031	-0.016	0.795	0.028	-0.086
Yala	-0.057	-0.118	0.786	0.254	-0.031
Surat Thani	0.07	-0.085	0.777	0.069	0.208
Satun	-0.003	-0.168	0.745	-0.025	-0.145
Nakhon Sri Thammarat	0.114	0.071	0.694	-0.003	0.134
Phuket	0.157	0.193	0.653	-0.332	-0.139
Prachuap Kiri Khan	0.221	0.054	-0.044	0.845	-0.128
Petchaburi	0.3	0.156	-0.043	0.744	0.118
Kanchanaburi	0.366	0.339	0.059	0.711	0.256
Ratchaburi	0.388	0.421	0.055	0.13	0.596

Chilasa paradoxa (7 Apr. 2014)



Papilio demoleus (28 Feb. 2013)



Papilio Mahadeva (8 Sep. 2012)



Papilio helemus (8 Sep. 2012)



Papilio nepheleus (28 Feb. 2013)



Papilio polytes (7 Apr. 2014)



Papilio memnon (14 Sep. 2013)



Papilio pitmani (28 Feb. 2013)



Papilio paris (7 Apr. 2014)



Graphium aristus (28 Feb. 2013)



Graphium antiphate (28 Feb. 2013)



Graphium xenocles (19 Mar. 2012)



Graphium megaris (7 Apr. 2014)



Graphium carpedon (28 Feb. 2013)



Graphium arycles (8 Sep. 2012)



Graphium docon (8 Sep. 2012)



Graphium eurypylus (8 Sep. 2012)



Graphium agamemnon (28 Feb. 2013)



Lamproptera meges (7 Apr. 2014)



Figure 1. Some swallowtail butterflies (*Papilionidae*) of the Pang Sida national park

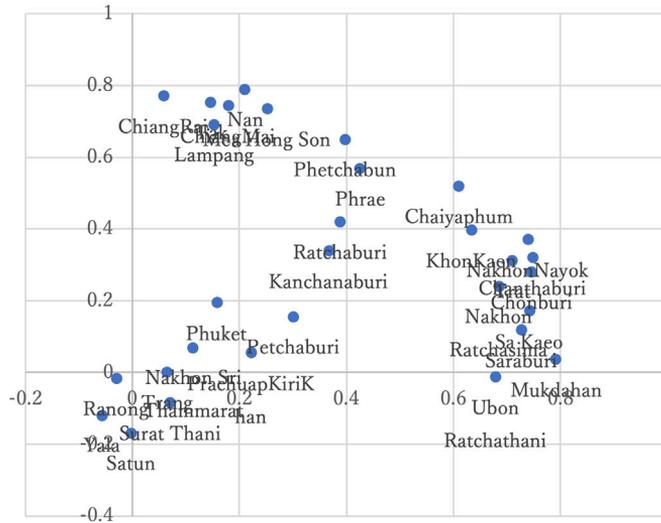


Figure 2. Scatter diagram of Varimax rotated factor loadings (factors 1 and 2)

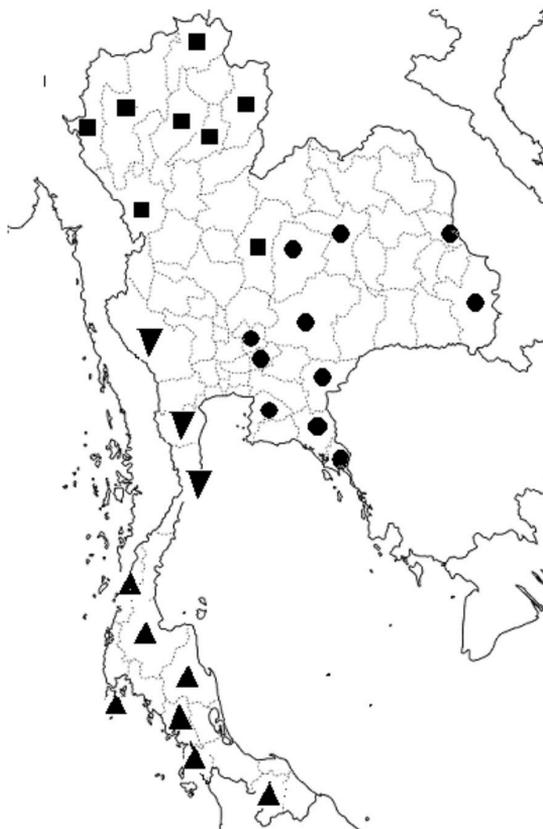


Figure 3. Spatial distributions of provinces that have higher factor loadings, circle: factor 1, rectangular: factor 2, triangle: factor 3, and triangle upside down: factor 4. (The map data was referred from the following site, http://www.freemap.jp/asia/asia_thai_all.html)

interpret that factor 1 represents the eastern region; factor 2 is the northern region; factor 3 is the peninsular region; factor 4 is the western region.

A scatter diagram for factor loadings between 1 and 2 is shown in Fig.2. The provinces which have higher factor loadings are plotted in Fig.3. As shown in Fig.3, sampling sites (provinces) which have similar positions within this scatter diagram show an adjacent position in a real geographical situation. Therefore, we can easily understand the meaning of each factor. In the scatter diagram between factors 1 and 2, almost provinces except peninsular area were allocated continuously in the graph. This result may endorse a Kimura's hypothesis that central area such as Phu Khieo, Phu Kradueng, and Khao Yai area is explained as a mixture of Indo-Myanmar and Indo-Chinese sub-regions. In the scatter diagram of factor loadings (see Figure 2), positions of the southern provinces were strongly isolated from that of northern (continental) portion. This result suggested that zoo-geographical sub-regions are consistent with the distribution for a corresponding pair of the two subspecies, one is continental, and another is peninsular.

Since factor loadings of factor 2 for the Northern provinces such as Chiang Mai, Chiang Rai, Nan, Mea Hong son and Lampang are high, we can easily understand that this factor is related with the northern species. Factor loadings of the factor 3 are large for Trang, Satan, Yala, Surat Thani, and Ranong, and this factor is related to the southern area.

Elliot (1980) and Kimura et al. (2011) divided the southernmost peninsular area into two different sub-regions, east and west side of the peninsula. However, it was found that the raw data for southern provinces was too sparse for this analysis, and therefore could not make substantial conclusions concerning this point.

4. Conclusion

In order to discuss the geographical distribution of *Papilionidae* in Thailand, a statistical analysis was conducted. By conducting a principal component analysis five factors were

obtained, and the meanings of former four factors could be explained as follows: factor 1 represents the central eastern area; factor 2 is the northern; factor 3 is the southern or peninsular; factor 4 is western. This analytical result seems to support the classification of zoo-geographical sub-regions proposed by many researchers.

References

- Corbet, A.S. & Pendlebury, H.M., (1978), The butterflies of the Malay peninsula, 3rd edition revised by Eliot, J.N., chapter IV, geographical distribution, pp.22-35, Malayan nature society, Kuala Lumpur
- Ek-Amnuay, P. (2006), Butterflies of Thailand. Fascinating insects Vol.2. (1st edition), Amarin Printing and Publishing Public Co., Bangkok.
- Elliot, J.N. (1980) Preface. Butterflies in Thailand. 4. Lycaenidae (ed. By A. Pinratana), pp. vi – vii. The Viratham Press, Bangkok.
- Gray, D., Piprell, C. and Graham, M. (1991) National parks of Thailand, Industrial Finance Corporation of Thailand, Communications Resources (Thailand) Ltd., Bangkok..
- Hughes, J.B., Round, P.D. & Woodruff, D.S., (2003), The Indochinese-Sundaic faunal transition at the Isthmus of Kra: an analysis of resident forest bird species distributions, Journal of Biogeography, 30, 569-580.
- Inayoshi, Y., (2010) A Check List of Butterflies in Indo-China, Chiefly from THAILAND, LAOS & VIETNAM, <http://yutaka.it-n.jp/>
- Kimura, Y., Aoki, T., Yamaguchi, S., Uemura, Y. and Saito, S.,(2011) The butterflies of Thailand based on Yunosuke Kimura collection, vol.1, Mokuyosha, Tokyo
- Lapkratok, S. and Suwanwaree, P., (2010), Butterflies species diversity comparison among some waterfalls in Dong Phrayayen-Khao Yai Forest Complex world heritage, Thailand, Proc. 2010 Int. Conf. Env Sci.& Develop.

- Okamoto, S., (2011) A study on the distribution of Lepidoptera in Thailand based on statistical analysis, Evolutionary Sciences 16:93-116.
- Okamoto, S., Tomita, M., Choosai, C. & Hanboonsong, Y. (2013) A study on the geographical distribution of *Papilionidae* in Thailand, Evolutionary Sciences 18:29-42.