Integration of GIS and Multivariate Statistical Analysis in Master Plan Study on Integrated Agricultural Development in Lao PDR

ラオス全国総合農業開発計画調査における GIS と多変量解析手法の統合に関する事例紹介

石塚 真*・下條哲成**・大塚恵哉*・角 幸康***

Makoto ISHIZUKA, Tetsunari GEJO, Shigeya OOTSUKA and Yukiyasu SUMI

ラオス全国総合農業開発計画調査において、筆者らは GIS（地理情報システム）と多変量解析（主成分分析およびクラスター分析）を組み合わせた意志決定のための画期的な解析手法を確立した。本手法のグルーピングにより全国141の地域を科学的、客観的に10のグループに分割し、各グループの農業の現況、開発阻害要因、開発促進重点分野を特定した。これにより、開発促進重点分野に人材および資金を集中配分し、対象地域を限定した効率的な開発戦略の策定が可能となる。本手法は、セクター・アプローチに対応したノウハウであり、農業開発計画だけでなく、さまざまな分野の**戦略的開発計画策定**への応用性が高い。また、GIS 技術や IT 技術が多用される中で、GIS の本来の活用法は空間解析もさることながら、そのデータ・ベースとしての機能および活用法にあることに着目し、基礎技術としての GIS を意志決定の流れに組み入れた。このことは、ラオス農林省のみでなく他国援助機関からも注目されることとなった。ODA は「効率化」と「透明性」を重視する動向にあり、客観的・科学的根拠に基づいた計画策定により、顧客そして被援助国に対して「**説明責任**」を提示できることは、当社の強みになる。

Key Words : *strategic planning, accountability, quantitative criteria, geographic information system (GIS), multivariate analysis, principal components analysis, cluster analysis*

1. INTRODUCTION AND OBJECTIVES

The innovative analysis technique using GIS and multivariate analysis was applied to the identification of the agricultural conditions in each district of Lao PDR and the formulation of Master Plan of Integrated Agricultural Development in Lao PDR financed by Japan International Cooperation Agency (JICA).

It is intended to use the results of analysis as **strategic development planning** to make realistic and area-specific projects/programs based on the scientific methodology applied here, and then to promote the **accountability**.

The effectiveness of GIS was optimized by the combination of GIS and statistical analysis, which integrates agricultural, social, topographical, and natural environmental data. It should be pointed out that GIS is not an isolated technology but a part of an integrated methodology of analysis.

＊　国際事業本部　農業開発部
＊＊　国際事業本部　開発計画部
＊＊＊　国際事業本部　情報・品質管理室

2. METHODOLOGY

(1) INTEGRATION OF GIS AND STATISTICAL ANALYSIS

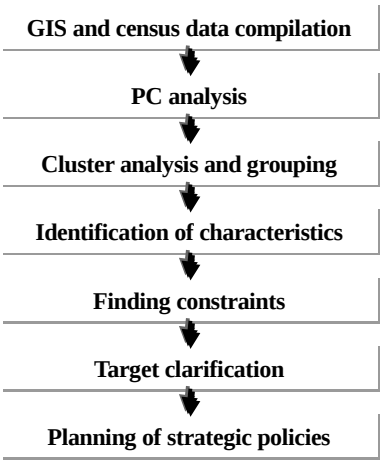
GIS derives digital data sets from an existing database and adds meaningful information to the integrated analysis. On the other hand, the statistical analysis makes readable and decipherable an immense amount of data sets (in this project about 20,000 data were used), which are difficult to understand at a glance.

The advantage of this scientific methodology derives from the fact that a strategic planning process, which makes decisions made by planners and bureaucrats accountable to the public to explain the reasoning of their policies.

(2) FLOW OF ANALYSIS

The flow of analysis is shown in Figure 1. Firstly, existing agriculture census 1998/99 and population census 1995 data and GIS data (polygon, polyline and grid data) were collected in Lao PDR. The compilation of topographical data was conducted using ESRI ArcView

Figure 1. Flow Chart of Analysis



Spatial Analyst.

After the data compilation, principal components (PC) analysis was carried out. PC analysis derives a number of PC (five of which were extracted in this project) which mainly describe the present conditions of agriculture in Lao PDR.

Then, all the 141 districts were classified into 10 groups based upon the result of PC analysis by applying a technique of cluster analysis.

Finally, through the statistical analyses the present characteristics of each group are identified in terms of various aspects of agricultural conditions and constraints. With the findings of the constraints, then, the targets of development in each group were clarified.

(3) PRINCIPAL COMPONENTS (PC) ANALYSIS

PC analysis is a statistical technique applied to plural sets of variables to discover similarities and positioning of the variables. This technique involves a mathematical procedure that transforms a number of (possibly) correlated variables into a (smaller) number of uncorrelated (independent) variables called principal components. In other words, meaningful and independent information is abstracted from an immense amount of data sets, and decision making process becomes more accountable.

Since the possible number of PC should be the total number of variables (in this analysis 141) minus 1 according to the principle of PC analysis, about 140 variables were selected as meaningful data to be applied to PC analysis.

(4) CLUSTER ANALYSIS AND GROUPING

After the identification of the main PC, the cluster analysis was applied to the grouping of the 141 districts. The cluster analysis is a multivariate analysis technique that seeks to organize information about variables so that relatively homogeneous groups, or "clusters" are formed. The number of groups is determined through a trial-and-error process which finally comes up with the meaningful division of districts in terms of agricultural conditions (in this analysis).

Several experienced Japanese and Lao experts with full understanding on the actual agricultural conditions in Lao PDR were assigned to this project and carried out GIS and statistical analyses. For successful applications of this methodology, experienced experts should verify the results of GIS and statistical analyses.

3. DATA USED

Data used in this project are shown in the following table:

Table 1. Variables Compiled and Used

Data Type	Description of Data	Data Format
Agricultural Data (950 data sets)		
District level digital data from 1998/99 agriculture census	↓#Average area of holding and number of parcels, land use and land tenure conditions ↓#Cropping pattern and major crops cultivated ↓#Purpose of production ↓#Use of production inputs ↓#Average number of livestock raised by livestock type ↓#Number of holdings with aquaculture ↓#Others	Database file
Social Data (250 data sets)		
District level digital data from 1995 population census	↓#Population density ↓#Percent distribution of population by sex ↓#Urban and rural population ↓#Percent distribution by place of birth ↓#Household size by urban and rural ↓#Population by education level and literacy rate ↓#Economically active population by occupational classification and unemployment rate ↓#Children born and deceased persons ↓#Electricity and domestic water supply conditions and availability ↓#Others	Database file
Natural Data (GIS)		
Topographical data	↓#Roads in 4 classifications	Polyline data
Elevation and slope	↓#Digital Elevation Model (DEM) of 50 and 250 m grid data ↓#Slope of 50 and 250 m grid data	Grid data (Figure 2)
Forest cover	↓#Forest cover in 1997 of 50 and 250 m grid data in all Lao PDR	Grid data (Figure 3)
Administrative boundary	↓#District boundary containing 141 districts ↓#Provincial boundary containing 18 provinces	Polygon data

Existing GIS data was obtained from National Agriculture and Forest Research Institute (NAFRI) which conducted the GIS project in cooperation with the Mekong Committee to derive Digital Elevation Model (DEM) and other essential data for the analysis (Figures 2 and 3).

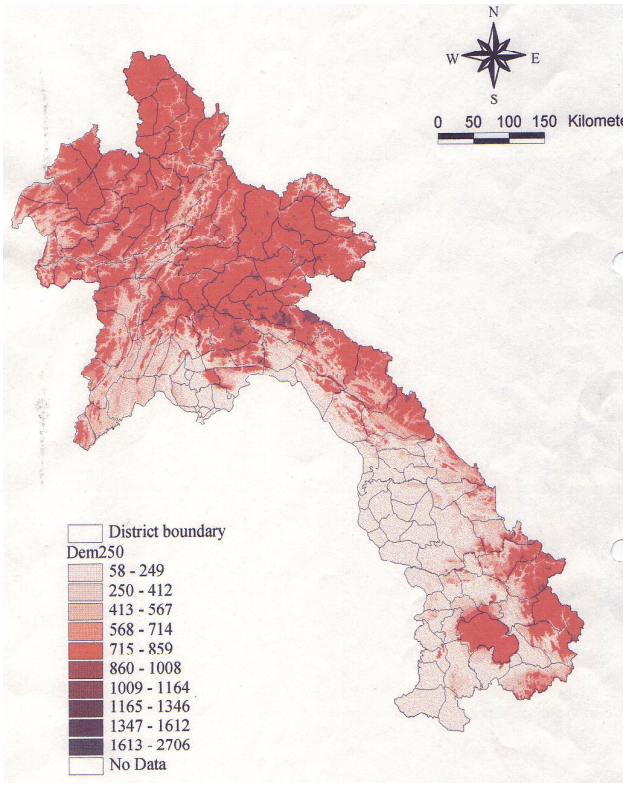


Figure 2. Digital Elevation Model (DEM), NAFRI

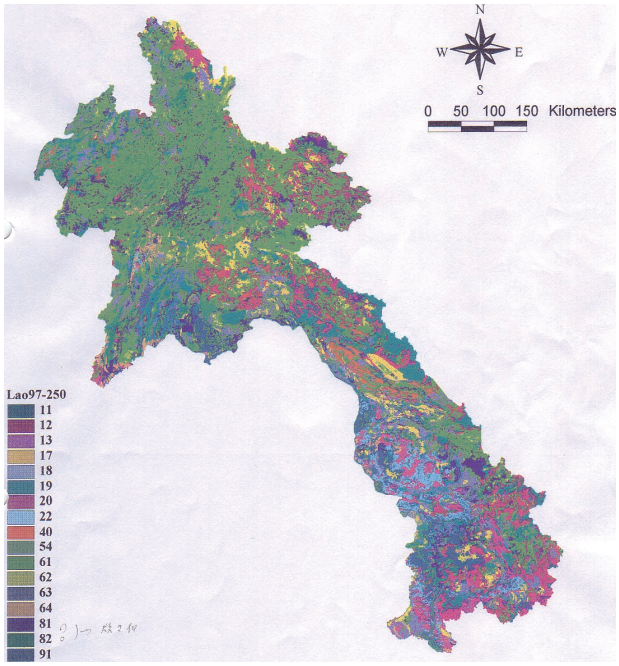


Figure 3. Forest Cover (1997), NAFRI
(ID numbers represent the classifications of forest cover.)

4. RESULTS

(1) DERIVED PRINCIPAL COMPONENTS

As a result of GIS and statistical analyses, 5 sets of principal components were identified as follows:

① **Transitional farming**: indicates a degree of farming system in transition from traditional to modernized agriculture

② **Market orientation**: designates a degree of how much farm products are targeted for sale in the market

③ **Water resource utilization**: indicates a degree to which water resources are utilized for agricultural production

④ **Farm intensity**: designates a degree to which farm resources are utilized for output of agricultural products

⑤ **Degree of diversification**: indicates the extent of diversification in cropping and livestock farming

(2) GROUPING OF DISTRICTS

All the 141 districts were classified into 10 groups by the cluster analysis in accordance with the 5 principal components of each district. Table 2 shows the combination of 5 principal component scores that **quantitatively** clarify the present agro- and socio-economic conditions of respective groups, and the result of grouping is presented in Figure 4.

Table 2. PC Scores of Each Group

Group ID	Number of Districts	Transitional Farming	Market Orientation	Water Resource Utilization	Farm Intensity	Degree of Diversification
1	41	-0.87	0.20	-0.45	0.39	0.54
2	38	-0.41	0.07	0.92	-0.09	-0.70
3	19	0.72	-0.80	-0.33	-1.05	-0.01
4	14	1.26	**-1.33**	-0.69	0.32	0.03
5	10	1.51	0.33	0.56	1.25	0.03
6	9	-0.45	-0.06	0.03	-1.64	0.97
7	5	**2.32**	3.44	0.32	-0.10	**1.19**
8	2	**-1.00**	-0.33	-1.55	**2.78**	0.55
9	2	-0.14	0.18	**0.96**	1.16	-3.17
10	1	-0.48	**3.58**	**-6.56**	**-2.53**	**-5.63**
TOTAL	**141**	-	-	-	-	-

1) The maximum and minimum scores are bolded for comparison.

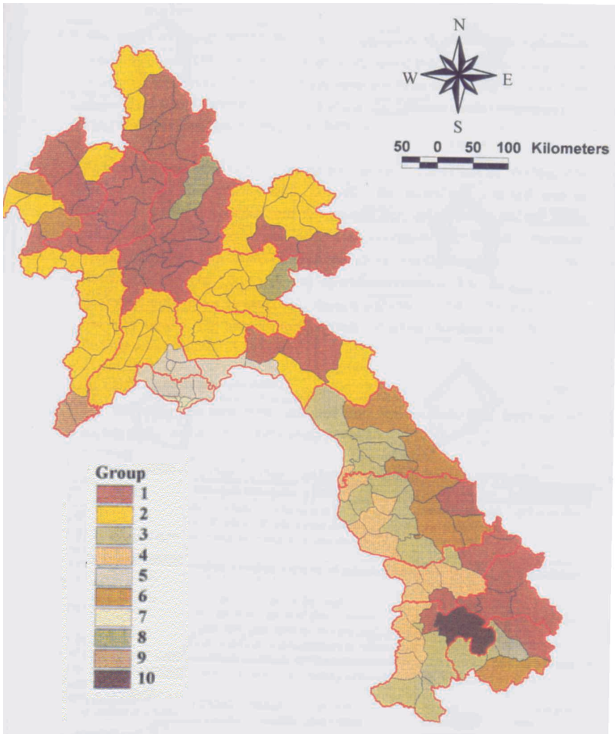


Figure 4. Grouping of Districts

(3) ASSESSMENT OF AGRICULTURAL POTENTIAL

In addition to the principal component analysis, the agriculture potential was assessed to clarify the availability of potentially new agricultural land. The data used for this analysis were based on forest cover and land use, soil, climatic condition, topographic condition, and unexploded ordinance (UXO). Figure 5 shows the result of the potential analysis for rice.

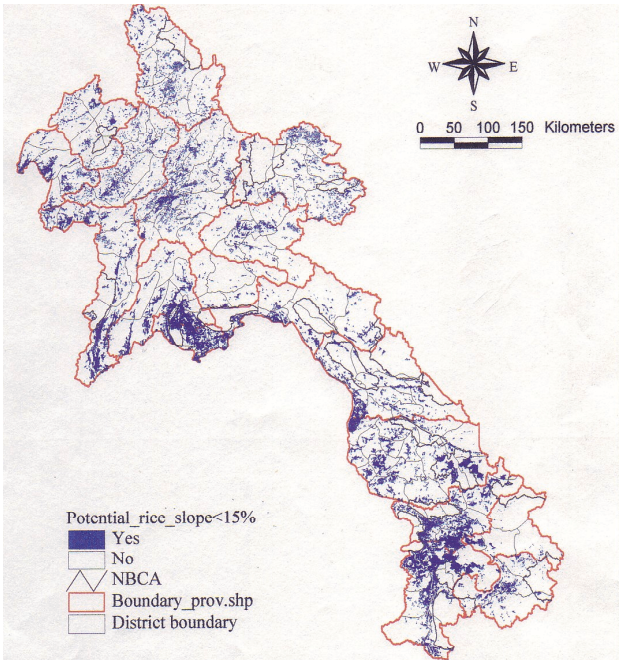


Figure 5. Potential for Rice

In the figure, high potential areas are spread along the Mekong River, while the highland areas have less potential.

5. AREA-BASED AGRICULTURAL AND RURAL DEVELOPMENT

Taking all the results of assessments as presented above, the Lao PDR was largely divided into four agricultural development areas: i) Northern Upland and Highland Areas; ii) Vientiane Plain; iii) Central and Southern Areas; and iv) Boloven Plateau, as tabulated below. The northern upland and highland areas and central and southern areas have some sub-areas. The above classification is shown in Table 3.

Table 3. Area Classification and Related Groups

Area Classification	Sub-area Classification	Related Groups
i) Northern Upland and Highland Areas	**Lowland Paddy Areas**	2, 8 and 9
	Sparsely Populated Remote Areas	1, 2 and 6
	Borderland Areas	1, 2 and 9
	Areas other than the above three areas	1 and 8
ii) Vientiane Plain		5 and 7
iii) Central and Southern Areas	**Central and Southern Mekong Corridor**	3 and 4
	Central and Southern Upland and Highland Areas	1 and 6
iv) Boloven Plateau		10

For example, the characteristics of Northern Upland and Highland Areas which include Group 1, and Vientiane Plain that embraces Group 5, are summarized in Table 4. The difference between the two groups is huge in terms of their agricultural settings.

Principal components consistently tell the conditions of agriculture of each group. As shown in the radar charts (Figure 6), the districts of Group 1 indicate low degree in terms of transition of farming. In fact, shifting cultivation is practiced on sloping land in these districts. A critical target for Group 1, then, should be to prevent expansion of shifting cultivation.

On the other hand, the district of Group 5 is generally located in the vicinity of the capital city Vientiane, where the paddy field cultivation is actively practiced. These districts have roads to access the market as well as relatively developed irrigation facilities. Farm intensity, however, shows low result because the floods occurred in rainy season in this area prevent intensification of farm land.

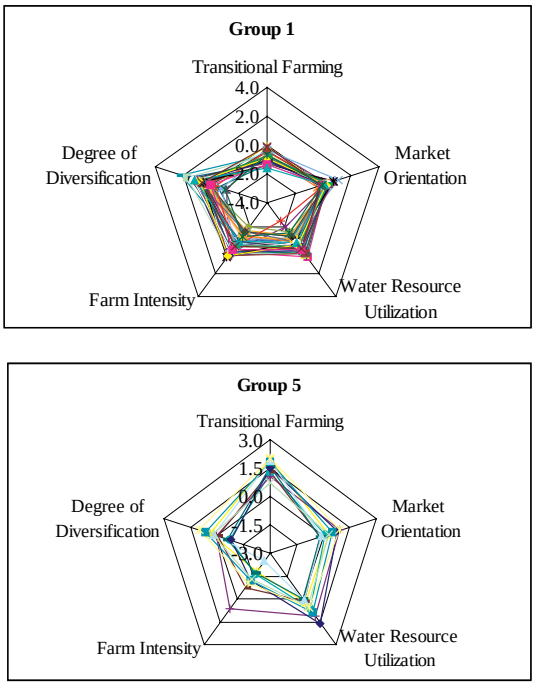


Figure 6. Examples of Radar Charts

Table 4. Description of Agricultural Settings in Groups 1 and 5

	Principal Components		Describe Characteristics
	Components	**Evaluation**	
Group 1	**Transitional Farming**	**Very Low**	• **Shifting cultivation** is widely practiced on sloping land for production of upland paddy. • In order to supplement a lower productivity, **non-paddy products** (including livestock and home manufacturing products) are produced and marketed to a certain extent. • Expansion of irrigation area mainly for lowland paddy production is at low level. • Resource management is poor and depletion is high. • Farming intensity is at mid. to low level, and diversification is at mid. to high level.
	Market Orientation	Mid. to High	
	Water Resource Utilization	Low	
	Farm Intensity	Mid. to Low	
	Degree of Diversification	Mid. to High	
Group 5	**Transitional Farming**	**High**	• Located near the capital city, Vientiane, and **production of market oriented crops** is considerably high. • Irrigation system is relatively developed and supports **crops diversification**. However, farm intensity is relatively low. • Floods occur frequently in the wet season along the Namgum River due to its topographic condition.
	Market Orientation	**High**	
	Water Resource Utilization	High	
	Farm Intensity	Low	
	Degree of Diversification	Mid.	

Table 5. Constraints and Targets Identified for Groups 1 and 5

	Find Constraints	Clarify Targets
Group 1	(1) Domination of **unsustainable shifting cultivation** which is a cause of forest cover reduction, soil erosion, etc. (2) Food crops are insufficiently produced. (3) Productivity of non-paddy crops is low, although they are important for cash income source. (4) **Production and marketing infrastructure is poorly developed**. (5) Degree of market orientation is still at mid. level.	(1) To prevent expansion of **shifting cultivation**. (2) To develop adequate production systems for sustainable use of upland. (3) To promote production of cash crops to increase farmers' income both in upland and lowland areas. (4) To provide production and **marketing infrastructure**. (5) To improve productivity of lowland paddy.
Group 5	(1) Expansion of market oriented crops is getting difficult due to small domestic market. (2) **Quality of crops is still at a low level for export**. (3) **Paddy productivity is still at a low level**. (4) **Flood damages** are considerable in the eastern part of Vientiane municipality, although the largest market is close by.	(1) To develop and introduce proper cropping pattern and production technologies so as to produce **value-added crops** throughout the year. (2) To improve **quality of products** so as to increase their competitiveness. (3) To assist farmers in **marketing development**. (4) To improve productivity of paddy by increasing inputs taking benefit/cost into account. (5) **Flood mitigation measures** are thus needed.

(1) NORTHERN UPLAND AND HIGHLAND AREAS

1) Overall Direction

Shifting cultivation, having a considerable negative impact on the land and forest, is widely practiced in this area. Accordingly, taking measures to stabilize shifting cultivation is seen as critical for agricultural development. Among others, fundamental measures to be taken for shifting cultivation stabilization include the development of alternative production systems through strengthening research efforts and the dissemination of such systems to farmers by reinforcing extension services. Another measure is to provide an adequate rural road network in shifting cultivation areas. This would improve accessibility and increase the sale of agricultural products. These measures require considerable investment over a long time period. Therefore, other measures are necessary in the short term, e.g., strengthening livestock sub-sector with intensive animal health care services, promoting feed crop production and providing micro-finance to support the procurement of farmers' production inputs.

2) Lowland Paddy Areas in Northern Upland and Highland Areas

Lowland paddy areas are developed mainly in valley bottoms. Paddy production, together with upland paddy, plays an important role in food security. Basically, the proposed directions for agricultural development for this area would be same as those for the Central and Southern Lowland Areas of Vientiane Plain, e.g. the distribution of improved seed, improvement of cultivation techniques, rehabilitation or improvement of existing irrigation systems and its effective management. In addition, it is proposed to take measures to further promote fishculture that would improve farmers' nutrition and diversify their income sources.

3) Sparsely Populated Remote Areas

There are many districts/villages in remote rural areas in which the development of rural infrastructure is usually difficult due to the low economic potential. It is, therefore, proposed to develop such areas through the pocket development approach. This approach will create growth points in remote areas by concentrating the investment for developing rural infrastructure and other services. In the proposed approach, greater attention is paid to the poverty alleviation.

4) Borderland Areas

The result of grouping reveals that two districts have a high degree of market orientation. This is mainly because of commercial cotton production grown for the Thai market. Because of its geographical advantage, it is judged that borderland areas have a high potential for commercial production for export to neighboring countries. It is, therefore, proposed to conduct a market study in order to determine promising export products. Moreover, private sector investment in agro-processing has to be promoted in parallel. In this context, market information has to be supplied to both producers and traders. This is an important step to attract commercial production and private sector investment. An efficient system of making such market information available is thus needed.

(2) VIENTIANE PLAIN

Flooding is one of the largest constraints in this area. Wet season agricultural production is difficult, although the largest market is close by. Some flood mitigation measures in the mid- to long-term are thus needed. Consequently, it is proposed to study flood mitigation measures with considering technical and economic aspects. On the other hand, measures for strengthening vegetable production in non-flood areas should be taken, They would include the introduction of new vegetable crops and varieties, development of low cost cultivation technique and quality control technologies and appropriate technologies for all year round crop production. Research should be strengthened to develop technologies covering the above fields. At the same time, a market study and a market information supply system are needed in this zone. Lowland paddy production is also practiced in this area, particularly in the suburbs of Vientiane municipality. The proposed direction for its development is to strengthen paddy production through the distribution of improved seed, improvement of plant management technology and effective use of existing irrigation systems.

6. CONCLUSION

Through this methodology, strategic development policies can be derived, which is to efficiently and effectively concentrate limited human and financial resources on a specific target and in a specific area (through grouping) to tackle poverty and to promote agricultural production.

Since the sector approach has been getting more and more important in the arena of international development, the methodology integrating GIS and statistical analyses is potential for planning and policy making processes.

As a result of the workshop for the study, governmental organizations, donors, NGOs in Lao PDR expressed their keen interest, and some of them consequently visited the study team office to request additional information. In addition, the Laotian government staff requested a series of lectures on this methodology and analysis process.

These showed the methodology applied here was epoch-making, and the result obtained was persuasive to formulate agricultural development plan in Lao PDR.

The effectiveness of integrated application of GIS and statistical analysis was realized through the course of the study. However the followings should be emphasized as remarks.

- The quality of original data is the utmost important. In other words, the quality of this methodology heavily relies on the reliability of original data sets. The availability of data

should be carefully confirmed before the adoption of this method. Therefore, the importance of both data collection and data selection cannot be overemphasized.

- Principal component analysis provides us with only principal component scores. The interpretation and naming of principal components derived through the analysis should be carried out by each expert. In this regard, the expert, if possible, plural number of experts who are familiar with the target country or area, should participate in the study.
- The influential data and the result of the analysis should be, to some extent, predicted prior to the analysis, that will facilitate the collection of data and analysis. In this study, the data on flooding was not utilized at first, which caused some difficulty in interpretation of statistical analysis.

REFERENCES

1) 菅民郎:多変量解析の実践、現代数学社 pp.128-160、1993 年
2) Kendall, M: Multivariate Analysis, Charles Griffin & Company LTD, London, Second Edition. 1980
3) 国際協力事業団:ラオス国総合農業開発計画調査主報告書、2001年