A MIXTURE MODEL APPLICATION IN THE MAPPING OF THE TOTAL ORGANIC MATTER CONTENTS OF SEDIMENTS IN THE GULF OF THAILAND: USING THE SEAFDEC DATA OF 1996

APICHART THUNYAHAN

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR
THE DEGREE OF MASTER OF SCIENCE (BIOSTATISTICS)
FACULTY OF GRADUATE STUDIES
MAHIDOL UNIVERSITY
2003

ISBN 974-04-3646-3
COPYRIGHT OF MAHIDOL UNIVERSITY

Copyright by Mahidol University

Thesis
Entitled

A MIXTURE MODEL APPLICATION IN THE MAPPING OF THE TOTAL ORGANIC MATTER CONTENTS OF SEDIMENTS IN THE GULF OF THAILAND: USING THE SEAFDEC DATA OF 1996

Apichart Thunyahan
Mr. Apichart Thunyahan
Candidate

Piangchan Rojanavipart
Assoc. Prof. Piangchan Rojanavipart,
M.H.S.(Biostatistics)
Major-Advisor

Chukiat Viwatwongkasem
Assoc. Prof.Chukiat Viwatwongkasem,
M.Sc.(Biostatistics)
Co-Advisor

Pisit Vatanasomboon
Assoc. Prof. Pisit Vatanasomboon,
M.Sc.(Environmental Health)
Co-Advisor

W Thitinanthapan
Assoc. Prof. Waraporn Thitinanthapan
D.D.S, Grad. Dip. Clin. Sc. (Endodontics)
M.Sc. (Clinical Science), Board in Endodontics
Acting Dean
Faculty of Graduate Studies

S. Singhaniyom
Assist. Prof. Sumalee Singhaniyom,
M.Sc.(Biostatistics)
Acting Chair
Master of Science Programme in Biostatistics
Faculty of Public Health

Copyright by Mahidol University

Thesis
Entitled

A MIXTURE MODEL APPLICATION IN THE MAPPING OF THE TOTAL ORGANIC MATTER CONTENTS OF SEDIMENTS IN THE GULF OF THAILAND: USING THE SEAFDEC DATA OF 1996

was submitted to the Faculty of Graduate Studies, Mahidol University
for the degree of Master of Science (Biostatistics)
on
June 9, 2003

Apichart Thunyahan
Mr. Apichart Thunyahan
Candidate

Piangchan Rojanavipart
Assoc. Prof. Piangchan Rojanavipart,
M.H.S.(Biostatistics)
Chair

Chukiat Viwatwongkasem
Assoc. Prof. Chukiat Viwatwongkasem,
M.Sc.(Biostatistics)
Member

S. Petpiroon
Lect. Saran Petpiroon,
Ph.D.
Member

Pisit Vatanasomboon
Assoc. Prof. Pisit Vatanasomboon,
M.Sc.(Environmental Health)
Member

W. Thitinanthapan
Assoc. Prof. Waraporn Thitinanthapan
D.D.S, Grad. Dip. Clin. Sc. (Endodontics)
M.Sc. (Clinical Science), Board in Endodontics
Acting Dean
Faculty of Graduate Studies
Mahidol University

C Chaikittiporn
Assoc. Prof. Chalermchai Chaikittiporn,
Dr.P.H. (Epidemiology)
Dean
Faculty of Public Health
Mahidol University

Copyright by Mahidol University

ACKNOWLEDGEMENT

I would like to express my sincere gratitude and sincere appreciation to Associate Professor Piangchan Rojanavipart, my major advisor, for her guidance, helpful comments, invaluable advice, supervision and encouragement throughout this study. I am also grateful to Associate Professor Chukiat Viwatwongkasem and Associate Professor Pisit Vatanasomboon, my co-advisor for their invaluable advice and guidance.

I am grateful to all the teachers and staff of the Department of Biostatistics, Faculty of Public Health for their guidance, knowledge and generous assistance.

I wish to thank all persons involved in this thesis for their advice with regard to English language usage.

Finally, my sincere thanks are expressed to my family, my lovely friend and my friends for kindness, support and encouragement throughout my life.

Apichart Thunyahan

Copyright by Mahidol University

A MIXTURE MODEL APPLICATION IN THE MAPPING OF THE TOTAL ORGANIC MATTER CONTENTS OF SEDIMENTS IN THE GULF OF THAILAND: USING THE SEAFDEC DATA OF 1996

APICHART THUNYAHAN 4236031 PHBS/M

M.Sc. (BIOSTATISTICS)

THESIS ADVISORS : PIANGCHAN ROJANAVIPART,M.H.S.(BIOSTATISTICS),
CHUKIAT VIWATWONGKASEM, M.Sc.(BIOSTATISTICS),
PISIT VATANASOMBOON, M.Sc.(ENVIRONMENTAL HEALTH)

ABSTRACT

The purpose of this study was to apply the Mixture Model and the conventional method for the mapping of the total organic matter contents of 0-2, 2-4, and 4-6 centimeter bottom sediments in the Gulf of Thailand. The total organic matter data were surveyed by the SOUTHEAST ASIA FISHERIES DEVELOPMENT CENTER (SEAFDEC) in 1996. Here, it is found that the Mixture approach is an appropriate model and can effectively reduce variations in mapping. The parameters were estimated using the Nonparametric Maximum Likelihood Estimator calculated with the C.A.MAN program developed by Böhning. The number of components was solved by Likelihood Ratio Statistics. The classification of the regions to construct the mapping was determined by the application of Bayes' Theorem. The conventional method was performed using percentiles.

The total organic matter contents of the 0-2, 2-4, and 4-6 centimeter bottom sediments in the Gulf of Thailand were classified into five groups, whereas those of the 0-6 centimeter bottom sediments were made up of six. The largest quantity of total organic matter contents appeared at near-shore Samui Island because of water runoff from the drainage basin of direct discharge into the water. In other areas, the quantities of total organic matter contents were acceptable. The trend with regard to total organic matter seems to show a constant but does, nevertheless, change in the central areas of the Gulf on account of the water current and geographical topography.

The Mixture Model is a practical method that identifies the distinct characteristics of geographical topography, and as such, it is most useful for ecological research.

KEY WORDS : ENVIRONMENTAL MAPPING / MIXTURE MODEL / TOTAL ORGANIC MATTER

92 P. ISBN 974-04-3646-3

Copyright by Mahidol University

การประยุกต์ Mixture Model ในการทำแผนที่ปริมาณของสารอินทรีย์ในดินตะกอนในบริเวณอ่าวไทย : ข้อมูลการสำรวจ SEAFDEC พ.ศ. 2539 (A MIXTURE MODEL APPLICATION IN THE MAPPING OF THE TOTAL ORGANIC MATTER CONTENTS OF SEDIMENTS IN THE GULF OF THAILAND: USING THE SEAFDEC DATA OF 1996)

อภิชาติ ธัญญาหาร 4236031 PHBS/M

วท.ม.(ชีวสถิติ)

คณะกรรมการควบคุมวิทยานิพนธ์ : เพียงจันทร์ โรจนวิภาต, M.H.S.(Biostatistics), ชูเกียรติ วิวัฒน์วงศ์เกษม, M.Sc.(Biostatistics), พิศิษฐ์ วัฒนสมบูรณ์, M.Sc.(Environmental Health)

บทคัดย่อ

การศึกษาครั้งนี้มีวัตถุประสงค์ คือ การประยุกต์ใช้ Mixture Model และ conventional method ในการสร้างแผนที่แสดงปริมาณของสารอินทรีย์ในดินตะกอนที่ระดับ 0-2, 2-4 และ 4-6 เซนติเมตร ในบริเวณอ่าวไทย ซึ่งสำรวจในปี พ.ศ. 2539 โดย SOUTHEAST ASIA FISHERIES DEVELOPMENT CENTER (SEAFDEC) วิธี Mixture Model เป็นวิธีที่สามารถลดความคลาดเคลื่อนของการแสดงแผนที่ได้ดี ประมาณค่าพารามิเตอร์ด้วยวิธี Nonparametric Maximum Likelihood โดยใช้โปรแกรม C.A.MAN ของ Böhning สถิติ Likelihood Ratio Statistic ใช้ทดสอบเพื่อหาจำนวน component ที่เหมาะสม การจำแนกพื้นที่เพื่อสร้างแผนที่ได้ใช้สถิติ Bayes ส่วนการศึกษา conventional method จะใช้ percentile ในการแสดงแผนที่

ผลการศึกษาพบว่าปริมาณสารอินทรีย์ที่ระดับ 0-2, 2-4 และ 4-6 เซนติเมตรของดินตะกอนในอ่าวไทยสามารถแบ่งได้ 5 กลุ่ม แต่โดยรวมที่ระดับ 0-6 เซนติเมตร จะสามารถแบ่งได้ 6 กลุ่ม ปริมาณสารอินทรีย์จะพบมากบริเวณชายฝั่งใกล้เกาะสมุย เนื่องจากเป็นบริเวณที่มีน้ำจากแผ่นดินไหลลงมามากว่าบริเวณอื่นและในบริเวณที่เหลือพบว่าปริมาณสารอินทรีย์ยังอยู่ในระดับที่น่าพอใจ และแนวโน้มของสารอินทรีย์พบว่าค่อนข้างคงที่แต่จะเปลี่ยนแปลงเล็กน้อยบริเวณกลางอ่าวเนื่องจากอิทธิพลของกระแสน้ำและลักษณะของภูมิประเทศ

Mixture Model เป็นวิธีที่เหมาะสมในทางปฏิบัติ ทำให้ทราบลักษณะความแตกต่างที่แท้จริงในบริเวณต่างๆของพื้นที่ภูมิศาสตร์ สามารถนำไปใช้ประโยชน์ทางด้านนิเวศน์วิทยา เพื่อการวางแผนพัฒนาในอนาคต

92 หน้า. ISBN 974-04-3646-3

Copyright by Mahidol University

CONTENTS

		Page
ACKNOWLEDGEMENT		iii
ABSTRACT		iv
LIST OF TABLES		viii
LIST OF FIGURES		x
CHAPTER		
I	INTRODUCTION	
	1.1 Background and Rationale	1
	1.2 Objectives	4
	1.3 Scope of the Study	5
	1.4 Definition of Terms	6
II	LITERATURE REVIEW	
	2.1 Topography of the Gulf of Thailand	8
	2.2 Total Organic Matter	9
	2.3 Mixture Model Approach	10
	2.4 Relevant Researches	18
III	MATERIALS AND METHODS	
	3.1 Study Design	22
	3.2 Data Collection and Processing	22
	3.3 Procedure for Environment Mapping	24
IV	RESULTS	
	4.1 Descriptive Study of the Total Organic Matter Contents of Bottom Sediments in the Gulf of Thailand	29
	4.2 The Data Setting for analysis	32
	4.3 The Mixture Model of the Total Organic Matter Contents of Bottom Sediments in the Gulf of Thailand	34

Copyright by Mahidol University

CONTENTS (Continues)

Page

CHAPTER

4.4 The Conventional Method of the Total Organic Matter Contents of Bottom Sediments in the Gulf of Thailand 60

V DISCUSSION

5.1 The Mixture Model Method and Conventional Method 68

5.2 The Total Organic Matter Contents of Bottom Sediments Mapping 69

VI CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion 73

6.2 Recommendations 77

REFERENCES 78

APPENDIX

Appendix A The VEM Algorithm and Stopping Rule 86

Appendix B The EM Algorithm 87

Appendix C The Location of Studied Stations in the Gulf of Thailand 88

Appendix D The Data of Total Organic Matter Measurement 90

BIOGRAPHY 92

Copyright by Mahidol University

LIST OF TABLES

Page

Table 1. Select Percentiles of $2\log\xi_n$ for Normal Distribution with Unknown (or Estimated) Variance 17

Table 2. The Total Organic Matter Contents (%Dry Weight) of Bottom Sediments in the Gulf of Thailand 1996 30

Table 3. The Mean and Variance Value of the Total Organic Matter Contents (%Dry Weight) in Each Cluster Group 32

Table 4. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments 35

Table 5. Likelihood Ratio Statistic of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments Data 36

Table 6. Classification of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments 37

Table 7. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments 41

Table 8. Likelihood Ratio Statistic of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments Data 42

Table 9. Classification of the Total Organic Matter Contents of 2-4 Centimeter Bottom sediments 43

Table 10. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments 48

Table 11. Likelihood Ratio Statistic of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments Data 49

Table 12. Classification of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments 50

Table 13. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of the Three Levels of Bottom Sediments 54

Copyright by Mahidol University

LIST OF TABLES (Continues)

Page

Table 14. Likelihood Ratio Statistic of the Total Organic Matter Contents of the Three Levels of Bottom Sediments Data 55

Table 15. Classification of the Total Organic Matter Contents of the Three Levels of Bottom Sediments 56

Table 16. The Cut of Point Value and the Number of Station of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments in Each Component 60

Table 17. The Cut of Point Value and the Number of Station of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments in Each Component 62

Table 18. The Cut of Point Value and the Number of Station of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments in Each Component 64

Table 19. Classification of the Total Organic Matter Contents of the Three Levels of Bottom Sediments by Conventional Method 66

Table 20. The Variance Value of the Total Organic Matter Contents in Each Component 69

Copyright by Mahidol University

LIST OF FIGURES

Page

Figure 1. Location of Studied Stations in the Gulf of Thailand 5

Figure 2. The Geographical Distribution of Actual Total Organic Matter Contents of Bottom Sediment in Each Assigned Cluster Groups 33

Figure 3. Map of the total organic matter contents of 0-2 centimeter Bottom Sediments Classified into Component According to the Mixture Model 39

Figure 4. Map of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments Classified into Component According to the Mixture Model 46

Figure 5. Map of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments Classified into Component According to the Mixture Model 52

Figure 6. Map of the Total Organic Matter Contents in Each Levels Classified into Component According to the Mixture Model 59

Figure 7. Map of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments Classified into Component According to the Conventional Method 61

Figure 8. Map of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments Classified into Component According to the Conventional Method 63

Figure 9. Map of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments Classified into Component According to the Conventional Method 65

Copyright by Mahidol University

CHAPTER I
INTRODUCTION

1.1 Background and Rationale

The study of environmental sciences is interdisciplinary since it is a broad area of study and it is important to develop in many aspects simultaneously. Most environmental issues are best understood when historical background is provided and political, economic, and social implications are integrated with purely scientific information. Environmental justice and equity are emerging concepts in the development of environmental policy. These concepts are related to questions on the spatial distribution of environmental contaminants in the population leading to the potential occurrence of environmental problems in different parts of the population. Environmental mapping can be defined as a method for displaying the spatial distribution of environmental problems. The environmental data can be displayed into different kinds such as description, table and histogram. Map is much used due to the fact that it is easy to understand (image and shade) and could represent the environment parameters (i.e. temperature, concentration, and density).

Generally, the map is created conventionally by using percentile, which could be calculated from observational data. It is uncomplicated and easy to understand. Nevertheless, the disadvantage of the percentile method is the data interval separation that must be determined by researchers. Therefore, in case of the data interval determination is broad, the descriptive of data will cause an error.

Recently, the remote sensing of Geographic Information System (GIS) has been used to make an environmental map. GIS is mapping software that can combine many layers of information. Implementing this analytical process with GIS can identify and locate the interested parameters. It can develop a plan and initiate control measures to management. It can also represent data by using optimal wavelength bands of reflected light from the earth. However, this method is a high technology, so users must have

Copyright by Mahidol University

special skill in this field and there is a high cost of using software, hardware and people-ware.

The Mixture Model is a statistical method and can be applied in an environmental mapping or a disease mapping. It is a model-based approach for clustering. The practical applications of Mixture Models are presented against the background of the existing literature to inference undertaken in a Mixture Model framework. It is based on the assumption that the population consists of several subpopulations with different levels of weight and parameters. Each of subpopulation assumes homogeneous population. The principal strategy for mapping with this approach corresponds an estimation of the level of observed values for each subpopulation probability. Attention is concentrated on the fitting of Mixture Models by a likelihood-based approach, using maximum likelihood where appropriate. The EM algorithm provides a convenient way for the iterative computation of solutions of the likelihood equation. Bayes' theorem could be applied to classify the individual region into one of the subpopulations of mixing distribution. The nonparametric Mixture Model approach is often preferred because it is less restrictive in its assumptions concerning the distribution of heterogeneity and it is always as good as the parametric approach can be with respect to goodness of fit (1).

Biernacki et al. (2) propose assessing a Mixture Model in a cluster analysis setting with the integrated completed likelihood. With this purpose, the observed data are assigned to unknown clusters using a maximum a posteriori operator. Then, the Integrated Completed Likelihood (ICL) is approximated using an Bayesian information criterion (BIC). Numerical experiments on simulated and real data of the resulting ICL criterion show that it performs well both for choosing a Mixture Model and a relevant number of clusters.

Therefore, Mixture Model will be chosen for making environmental map in this study. Because the nature of population consists several distributions mixed in varying distributions and we do not know number of subpopulation in population.

Natural resources and environment are important to mankind. They are indispensable to main survival and prosperity. Man needs natural resources, which can be obtained from the natural environment, to serve their well being, especially those four basic needs. The development of the industrial and agriculture as well as the rapid

Copyright by Mahidol University

increase of population in Thailand have strong impact on the existing natural resources. This results in an increase of water using and wastewater from household, industry, import, export, agriculture, aquaculture, and tourism. The wastewater does not pass the treatment process before it would be drained to channel, river and sea. The water sources and aquatic animals will be contaminated and then affect healthy environment and cause public health problem.

The Gulf of Thailand is the largest and the final water receiver. Seventy percent of the shore of the Gulf of Thailand is important an economic area because there are many natural resources. People in various activities such as consumption, trade, export and energy, etc use them. In the past, the fisheries resources are abundance and had biodiversity. Recently, the amount of them has immensely reduced and has been inadequate for consuming. The change of the ecosystem of the Gulf of Thailand cause many problems including public health. The causes of the reducing of natural resources in the Gulf of Thailand are various i.e. the unsuitability of fisheries, the improper of technological development and the draining of waste water from human activities and factories. These make the ecosystem in the Gulf of Thailand unbalance. Also the unbalances of the demand and supply of fisheries resources can even make an environment degenerate. Therefore, the study in the field of environment in the Gulf of Thailand is necessary for solving these problems and improvement.

The water quality with respect to chemistry was usually considered in term of dissolved inorganic elements or suspended particulate matter due to ease of detection and quantification. Organic constituents were seldom given equal consideration. Most of the attention, however has been directed forwards the analysis of organic compounds that are introduced into water systems by man's activities. The quantity of the organic matter in the Gulf of Thailand is an important parameter for study because it is a source of primary production in the ecosystem. Normally, the organic matter in the Gulf of Thailand is excrement, detritus that dissolve in the river and flow to the sea where flood plain. Therefore, the amount of organic matter in this area is high because there is sedimentation. That shows the sediment soil and nutrient quantity have direct relation (3). The Chao Phraya River has the biggest volume transport, followed by the Mae Klong River that flows to the Gulf of Thailand. The average runoff per year of the Chao Phraya River is 13.22×10^3 cubic meter per year and the Mae Klong River

Copyright by Mahidol University

3.26×10^3 cubic meter per year (4). The amount of nutrients is also discharged form these rivers and promotes primary productivity in the Gulf.

The organic matter contents in sediment soil are important for aquatic animal life because that is dietary of primary production (phytoplankton). Then organisms will consume phytoplankton as food for growth and replace expended energy (5). Yupa (6) found that the total organic matter is nutrient and energy of organisms in waster. The benthic fauna are especially use the organic matter contents of sediments approximate 10 - 20 percent of the total organic matter contents of the sediment soil (7). Alabaster (8) found that organic matter in the water source is chelating agent that similar EDTA (Ethylene Diamine Tetra Acetate) that can be combine with metal such as zinc, copper. The product of this reaction is complex compound matter. This activity makes the toxicity of metal reduce. It is addition of nutrient in ecosystem. Therefore, these activities are important for aquatic animals because they can produce the nutrient and the organic matter for primary productivity. However, having overmuch organic matter of sediments that will make the use of oxygen by aeromonas bacteria insufficient. The compound of sulfide will occur. If its circulation is not enough, the amount of this compound in water will over saturate. The aquatic animal may be harmful (9).

The total organic matter contents in the Gulf of Thailand was used as abundance index, evaluation of aquatic animal quantity and display or forecast problems, as well as this parameter can be used for planning and developing the natural resources in the Gulf of Thailand.

1.2 Objectives

This study focuses on the following objectives :

Investigating on environmental mapping of the total organic matter contents of bottom sediments in the Gulf of Thailand in 1996.

- Apply the Mixture Model by using Nonparametric Maximum Likelihood Estimator, Likelihood Ratio Test and Bayes' approach.
- Apply the conventional method by using percentile.

Copyright by Mahidol University

1.3 Scope of the Study

1. The Gulf of Thailand is located between latitude from 6° to 13°N and longitude from 99° to 105°E, calculated area 320,000 kilometers square and classified into 51 stations. Each station has distance of 30×30 nautically mile located along the continued and in mid water 51 stations were show in Figure 1.

2. The total organic matter contents data were surveyed in 1996 conducted by SOUTHEAST ASIA FISHERIES DEVELOPMENT CENTER (SEAFDEC).

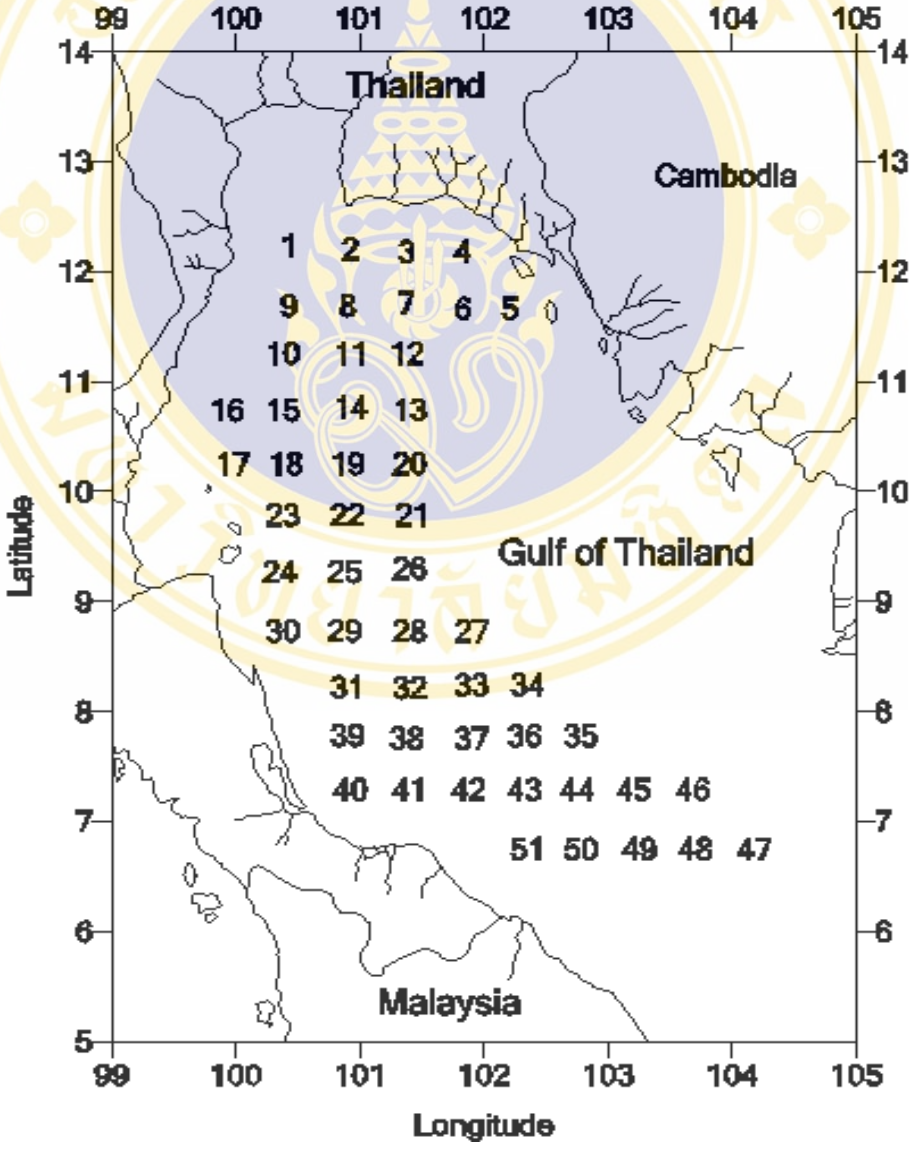


Figure 1. Location of Studied Stations in the Gulf of Thailand

Copyright by Mahidol University

1.4 Definition of Terms

The following terms have been defined to clarify their meaning in the context of this research.

1. Total organic matter represents materials of plant, animal and microbial origin that are in various stages of decomposition and are associated with the mineral fraction with different degrees of intimacy (10).
2. Environmental mapping is map for displaying the spatial distribution of parameters of interesting environment.
3. Population heterogeneity is the situation where the population consists of two or more subpopulations with different parameters interested of the overall population (11,12).
4. Mixture Model is natural models for unobserved population heterogeneity. This model is too strict to capture the variation of the parameter over a diversity of subpopulation (11).
5. Latent Variable is the variable of unobserved heterogeneity.
6. Log-likelihood Ratio Test Statistic is statistic test for an asymptotic χ^2-distribution with d degrees of freedom, where the degree of freedom, d, equal the difference between the number of parameters under the alternative and null hypothesis.
7. Nonparametric Maximum Likelihood Estimator (NPMLE) is the estimation that is done conventionally by maximum likelihood without any parametric assumption that is we have to find $\hat{P}$ which maximizes the log-likelihood $l(P) = \sum_{i=1}^{n} \log f(x_i P)$ (13).
8. Vertex Exchange Method (VEM) algorithm is suitable algorithm for computing the Nonparametric Maximum Likelihood Estimates of a mixture distribution for flexible support size case; the number of subpopulations k is not known. This algorithm is used as the standard method in the package C.A.MAN (11,14,15).
9. EM algorithm is an algorithm that was to be started with certain values for the parameters it is optimizing. This algorithm has been proposed numerous times by several authors for computing the maximum likelihood estimate in Mixture

Copyright by Mahidol University

Models for the fixed support size case; the number of subpopulations is known (11,13,14,15).

10. C.A.MAN stands for computer assisted mixture analysis. This software package provides the maximum likelihood estimates of the mixing distribution of many parameters of densities using various algorithmic approaches. Semiparametric analysis of mixtures of densities may be undertaken with flexibility of fixed supported size (number of densities in the mixture known or unknown) and individually structured data files (ungrouped, grouped, rates, SMRs etc.). Distribution include normal, Poisson and others. Results are written to a file. The program restricts a maximum of 1500 data points and 50 subpopulations (16).

Copyright by Mahidol University

CHAPTER II
LITERATURE REVIEW

2.1 Topography of the Gulf of Thailand

The Gulf of Thailand is a semi-enclosed shallow water body of South China Sea; it is a part of the Sunda continental shelf. The shoreline of the Gulf of Thailand has approximately 1,875 kilometers (17). It located between latitude from 6° to 14°N and longitude from 99° to 105°E. Northern and Western append coastal of Thailand. Eastern append coastal of Vietnam and southern append South China Sea. The overall area dimension of the Gulf is approximately 400 by 800 kilometers, or about 320,000 kilometers square (18). The average water depth is 55 meters, with a maximum of 84 meters. Northern of the Gulf of Thailand appended coastal of province that had Samut Prakan, Samut Sakhan, Samut Songkram and Bangkok. The east of the Gulf of Thailand appended coastal of province that had Trat, Chantaburi, Chon Buri, Rayong and Chachoengsao. The west of the Gulf of Thailand appended coastal of provinces that have Phetchaburi, Prachuap Khiri Khan, Surat Thani, Nakhon Si Thammarat, Songkhan, Pattani and Narathiwas.

The circulation pattern of current in the Gulf of Thailand is mainly characterized by weak clockwise circulation during the Southwest monsoon season (SW) which it prevails from May to September and strong counter-clockwise circulation during the Northeast monsoon season (NE) which it prevails from November to March (19). There are slight variations in temperature, salinity and density profiles before and after the NE monsoon season. Furthermore, salinity and density (temperature) values are slightly higher (lower) before than those after the NE monsoon season (20).

Generally surface sediments in the Gulf of Thailand are clay at the center and coarser near the coasts. Sediments along the western coast are usually composed of clay that any fewer than the sandy clay found along the eastern coast. Pockets of coarse sand can be found in certain areas, such as the sand belt across the mouth of the

Copyright by Mahidol University

upper Gulf and the sand patch off Songkha-Pattani coast. They are thought to be relate to the local physical oceanography of the areas. (21)

The accumulation of material in sediments was to near the coast of the upper Gulf of Thailand. The majority of matter of sediments is those discharged from land (22). Sedimentation rates in the Gulf of Thailand, based on the excess ^{210}Pb technique, are ranged from more than 1 gram per square centimeters per year (approximately 15 millimeters per year at sediment surface before compaction) near the mouth of Chao Phraya River to less than 0.05 gram per square centimeters per year (approximately 0.3 millimeters per year at sediment surface before compaction) near the center of the Gulf (23,24). Srisuksawad et al. (25) studied the sedimentation in upper the Gulf of Thailand are approximately 270 - 490 milligram per square centimeters per year and central of the Gulf are approximately 64 – 190 milligram per square centimeters per year. Sedimentation rates affect to the total organic matter accumulation in the Gulf of Thailand.

2.2 Total Organic Matter

Organic matter is synthesized from organic and inorganic which is composed of the element as follows: carbon 58%, hydrogen 10%, oxygen 20%, nitrogen and sulfur 1% and other elements (26,27). This material is of great importance in the marine environment such as marine ecosystem. The biogeochemistry of organic matter has less impact on the marine distribution of carbon and sulfur due to their greater abundances. Most of the hydrogen and oxygen in the ocean is present as water. Thus biological processes have little impact on the bulk distribution of these elements in the sea. Organic matter synthesized from inorganic compounds by means of autotrophs. In the marine environment, phytoplankton produced most of the particulate organic matter through the process of photosynthesis (28). The organic matter in soil and water are composed of a heterogeneous mixture of organic substances containing carbohydrate (mono-polysaccharide, cellulose, hemi-cellulose, lignin), lipid and other matter (e.g. resins, amino acid, aromatic compound and complex group of compounds etc) (29). The heterogeneity of organic material found in terms of its source (30). Patterns of organic matter in sea are organism, dissolve in the sea and colloidal organic substances and settling under the influence of gravity in water (31).

Copyright by Mahidol University

Pierce and Felbeck (32) present that organic matter in water and soil consists of 2 groups, include:

1. Non-humic substance including carbohydrates, protein and lipid which microorganism decompose rapidly the organic.
2. Humic substance compound that has unstable shapes and large molecules, known as humus. Humus is considered to be the end product of decomposition of plant material in soils by humification of organic.

Grissinger and McDowell (33); Pocklington (30); Mayers and Takeuchi (34) classified the organic matter original in sediment constitution into 2 forms including:

1. Soil erosion : soils contain organic matter that come from the accumulation of litters (the falling leaves and tree branches), and the remnant of harvest plants and the carcass (35). When organic matter in soils become saturated, usually transformed by wind and water, it has been settled onto the seabed and forms the sediments.
2. Accumulation of remains life in water (aquatic animal, water plants and phytoplanktons).

Important roles of the total organic matter in water are nutrient requirements for phytoplankton, reducing the potential toxicity of metals, and tracing elements (8, 9).

2.3 Mixture Model Approach

Mixture distribution consists of several distributions that mixed in varying distributions. It have been used extensively as model in a wide variety of important practical situations where data can be viewed as arising from two or more population mixed in varying proportions. The population consists of various subpopulations that each subpopulation is assumed to be homogeneity, we call this situation as the heterogeneous case. The population contains k components and subpopulations distributed to different parameters. We assume that $x_1, x_2, ...x_i, ..., x_n$ (for i = 1, 2, ..., n; in this study n = 51 stations) are random variables of n stations of the measure of the total organic matter. Each x_i has a discrete mixture of k conditional distributions with respective to normal probability density functions

Copyright by Mahidol University

$$f_1(x_i|\mu_1,\Sigma_1), f_2(x_i|\mu_2,\Sigma_2),...,f_j(x_i|\mu_j,\Sigma_j),...,f_k(x_i|\mu_k\ \Sigma_k),$$

with means $\mu_1,\mu_2,...,\mu_k$, with variance $\Sigma_1,\Sigma_2,...,\Sigma_k$ for j = 1, 2, ..., k, and with positive mixing probabilities $p_1, p_2,..., p_k$ with $\sum_{j=1}^{k} p_j = 1$. Let ψ have all unknown parameters in k component densities, that is,

$$\psi = (p_1, p_2,..., p_k, \mu_1, \mu_2,..., \mu_k, \Sigma_1, \Sigma_2,..., \Sigma_k).$$

Note that there are 3k-1 parameters, since $p_1 + p_2 + ... + p_k = 1$.

Therefore, the unconditional density $f(x_i)$ is the marginal density of mixture distribution that can be defined as follows:

$$f(x_i;\psi) = \sum_{j=1}^{k} f(x_i;\mu_j,\Sigma_j)p_j$$

and $f(x_i;\mu_j,\Sigma_j)$ is the probability density function of each station i and each component j,

$$f(x_i;\mu_j,\Sigma_j) = \left(\frac{1}{\sqrt{2\pi\Sigma_j}}\right)\exp\left(-\frac{(x-\mu_j)'(x-\mu_j)}{2\Sigma_j}\right),$$

where x_i is the observed in station i , i = 1, …, n

ψ denotes the vector of parameters of population in mixture model

μ_j is mean of belonging to the j-th subpopulation, j = 1, 2, ..., k

Σ_j is the variance of parameters of belonging to the j-th subpopulation, j = 1, 2, ..., k

p_j is the probability of belonging to the j-th subpopulation, j = 1, 2, ..., k

k is the number of subpopulations (components).

Copyright by Mahidol University

Consider the covariance matrices in two cases, so that

1. The homoscedastic case where the component covariance matrices Σ_j are restricted to being equal,

$$\Sigma_j = \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k = \Sigma \qquad , (j = 1,2,\ldots,k)$$

2. The heteroscedastic (unrestricted) case where the component covariance matrices Σ_j are unrestricted variances,

$$\Sigma_j = \Sigma_1 = \Sigma_2 = \ldots = \Sigma_k \qquad , (j = 1,2,\ldots,k)$$

The estimation is done by maximum likelihood method, that the unconditional likelihood of ψ for station i is given by $\sum_{j=1}^{k} f_j(x_i; \mu_j, \Sigma_j) p_j$ and for overall stations, the unconditional likelihood for ψ is

$$L(\psi) = \prod_{i=1}^{n} \left(\sum_{j=1}^{k} p_j f_j(x_i; \mu_j, \Sigma_j) \right).$$

The unconditional log-likelihood for ψ over all stations is

$$l(\psi) = \log L(\psi) = \sum_{i=1}^{n} \log \left(\sum_{j=1}^{k} p_j f_j(x_i; \mu_j, \Sigma_j) \right).$$

2.3.1 Estimation the Parameters of Mixture Density

The estimator of mixing distribution $\hat{\Psi}$ with this approach is called non-parametric maximum likelihood estimator (NPMLE). Hamza (36) estimated mixing proportions by using theoretical second moments. To maximize the log-likelihood function, it is necessary to develop an algorithm for maximizing the log-likelihood function.

Copyright by Mahidol University

However, it is unknown that the specific stations belong to which component or cluster. Some stations will belong to the same component. We speak of unobserved or latent heterogeneity that concerned "what is the component that the station belongs to". The other one enters the setting through a latent, unobserved variable Z_j The latent variable Z_{ij} are introduced $Z_{ij} = 1$, station i belong to component j and $Z_{ij} = 0$, otherwise.

Note that $\sum_{j=1}^{k} z_{ij} = 1$. If $\sum_{i=1}^{n} z_{ij} = n$, then all stations belong to component j, and we have homogeneity in this case. In other words, if $\sum_{j=1}^{k} z_{ij} < k$, then these are at least 1 station that does not belong to the same component, and there is heterogeneity. Let z_{ij} denote the value of Z_j for observation x_i. This is called complete likelihood that given for mixture model as

$$l_z(\psi) = \sum_{i=1}^{n}\sum_{j=1}^{k} z_{ij} \log\{p_j f_j(x_i; \mu_j, \Sigma_j)\}$$

It we use component indicators, Z_{ij}, the complete likelihood of these indicator variables for station i is given by

$$\prod_{j=1}^{k} \left(p_j f_j(x_i; \mu_j, \Sigma_j)\right)^{z_{ij}}.$$

Also, the complete likelihood of indicator variables over all stations is

$$L_z(\psi) = \prod_{i=1}^{n} \prod_{j=1}^{k} \left(p_j f_j(x_i; \mu_j, \Sigma_j)\right)^{z_{ij}}$$

and the complete log-likelihood of indicator variables over all stations is

$$l_z(\psi) = \log L_z(\psi) = \sum_{i=1}^{n}\left(\sum_{j=1}^{k} z_{ij} \log(\, p_j f_j(x_i; \mu_j, \Sigma_j)\,)\right).$$

Copyright by Mahidol University

Suitable algorithm consists of two methods.

1. The vertex exchange method (VEM) which is used for estimating the parameters of mixture distribution when the number of subpopulations k is not known. The basic of VEM is to exchange good vertex directions bad ones which are already in support of the current mixing distribution. The algorithm and stopping rule of VEM are shown in Appendix A.

2. EM algorithm which this method is used for estimation the parameters of mixing density when the number of subpopulations k is assumed to be known. The EM algorithm requires two log-likelihoods. The one have observed is called the incomplete log-likelihood. The other one enters the setting through a latent, unobserved variable that this is called the complete likelihood.

The EM algorithm proceeds iteratively in two step, E (for expectation) and M (for maximization).

1. E-step, which takes the conditional expectation of the complete log-likelihood $l_z(\psi)$, given the observed data x_i, using the current fit for ψ. Let $\psi^{(0)}$ be the value specified initially for ψ. It follows that on the (t+1)th iteration, the E-step requires the calculation of $Q(\psi;\psi^{(t)})$, which can be written as

$$Q(\psi;\psi^{(t)}) = E_{\Psi^{(t)}}\{l_z(\psi)|x_i\}.$$

As the complete log-likelihood $l_z(\psi)$ requires the calculation of the current conditional expectation of z_{ij} given the observed data x_i, where z_{ij} is the random variable corresponding to z_{ij}. Now

$$E_{\psi^{(t)}}\{z_{ij}|x_i\} = pr_{\psi^{(t)}}\{z_{ij}=1|x_i\}$$
$$= p_j(x_i;\psi^{(t)}),(i=1,\ldots,n;j=1,\ldots,k)$$

2. M-step is step for estimate on the (t+1)th iteration requires the maximization of $Q(\psi;\psi^{(t)})$ that

Copyright by Mahidol University

$$Q(\psi;\psi^{(t)}) = \sum_{i=1}^{n}\sum_{j=1}^{k} p_j(x_i;\psi^{(t)}) \log\{p_j f(x_i;\mu_j,\Sigma_j)\}.$$

The M-step on the (t+1)the iteration, the updated estimates $p_j^{(t+1)}$ of the mixing proportions p_j are calculated independently of the updated estimate of the mean and covariance matrices, which can be written as

$$p_j^{(t+1)} = \sum_{i=1}^{n} p_j(x_i;\psi^{(t)})/n \quad ,$$

$$\mu_j^{(t+1)} = \frac{\sum_{i=1}^{n} p_j(x_i;\psi^{(t)})x_i}{\sum_{i=1}^{n} p_j(x_i;\psi^{(t)})} \quad \text{and}$$

$$\Sigma_j^{(t+1)} = \frac{\sum_{i=1}^{n} p_j(x_i;\psi^{(t)})(x_i-\mu_j^{(t+1)})(x_i-\mu_j^{(t+1)})'}{\sum_{i=1}^{n} p_j(x_i;\psi^{(t)})}$$

The algorithm of EM are shown in Appendix B.

In this case, Normal Mixture Model that the number of components open and the variance parameters unknown (meaning to be estimated) will lead to and oversmoothed mixture model. The same research of Feng and McCulloch (37) problem in practice is to determine the number of components in a normal mixture model with unequal variances. When the likelihood ratio test statistic -2log lambda is used, it is unbounded above and fails to satisfy standard regularity conditions.

G. J. McLachlana, D. Peelb and W. J. Whitenc (38) presented the approach to clustering whereby a normal mixture model is fitted to the data by maximum likelihood. The general case of normal component densities with unrestricted covariance matrices is considered and so it extends the work of Abbas and Fahmy (39) who imposed the restriction of diagonal component covariance matrices. Attention is also focussed on the problem of testing for the number of clusters within this mixture framework, using the likelihood ratio test.

Copyright by Mahidol University

Therefore, either we fix the number of components and then estimate a common variance parameter in addition to the means and weights, or we fix the variance parameter and estimate the number of components (1). Therefore, we used both methods for estimating the parameters by estimating NPMLE with C.A.MAN. It starts to calculate solution for the flexible support size with the VEM algorithm and then computes a refined solution of the NPMLE using the EM algorithm which used the result of VEM as initial values. Furthermore, the EM algorithm is used to estimate NPMLE if the number of components is fixed. The log-likelihood values use for testing hypothesis of the number of components of mixing distribution.

2.3.2 Alternating for the Number of Components of Mixing Distribution

This is to determine the number of subpopulations. It was solved by computing the Likelihood Ratio Statistic (LR) for testing the hypothesis :

$$H_o : \text{number of components} = k$$

$$H_1 : \text{number of components} = k+1$$

and use the likelihood ratio test can be defined as follows :

$$LRS = 2 \log \xi_n$$

$$= 2\times [\, l(P_{k+1}) - l(P_k) \,]$$

where P_{k+1} is the non-parametric maximum likelihood estimator under H_1

P_k is the non-parametric maximum likelihood estimator under H_o

The log-likelihood ratio test statistic $2 \log \xi_n$ generally has an asymptotic χ^2-distribution with d degree of freedom, where the degree of freedom, d, are equal to the difference between the number of parameters under the alternative and null hypothesis. This study used critical values based on the work of Thode et al (40).

The 95-th percentile of $2\log\xi_n$ is used as the critical value of likelihood ratio test at a level of significance equal to 0.05 and the 99-th percentile of $2\log\xi_n$ is used as the critical value of likelihood ratio test at a level of significance equal to 0.01.

Copyright by Mahidol University

Table 1. Select Percentiles of $2\log\xi_n$ for Normal Distribution with Unknown (or Estimated) Variance

Sample size	95 – th percentile	99 – th percentile
100	6.68	10.33
1000	6.16	9.04
$\chi^2_{(2)}$	5.99	9.21

The decision rule is the null hypothesis H_o (the number of components = k) will be rejected if the computed value of LRS is equal to or greater than the critical value of the likelihood ratio test.

2.3.3 Classification for Map Construction

Classification is to classify the observation x into one of the subpopulations of the mixing distribution. These can be by applied Bayes' theorem and using the estimated mixing distribution as a prior distribution. The classification is done by computing the probability that each area belongs to a certain subpopulation. Let z_{ij} describe the unobserved variable that station i is in subpopulation j ($z_{ij}=1$). Therefore, the posterior probability is defined as follow:

$$\Pr(z_{ij}=1/x;\mu_{j,\Sigma_j}) = \frac{f_j(x_i;\mu_j,\Sigma_j)\hat{p}_j}{\sum_{j=1}^{k} f_j(x_i;\mu_j,\Sigma_j)\hat{p}_j}.$$

The i th station is then assigned to that subpopulation j for which it has the highest posterior probability of belonging (11).

Copyright by Mahidol University

2.4 Relevant Researches

2.4.1 Relevant Researches of the Total Organic Matter Contents in the Gulf of Thailand

Chamchoy (41) studied in Upper Gulf of Thailand near the Chao Phraya estuary that it had total organic matter contents were 6.3 - 15.7 %(dry weight)

Chumpon (42) studied organic matter contents of sediment in eastern of Gulf of Thailand and reported that the total organic matter contents in sediment were 2.0-14.6 %(dry weight). The total organic matter contents of sediment and the silt and clay contents are correspond significant and correlation coefficient was 0.643.

Vimolruk and Chitaruk (43) studied sediments in coast of Upper and Lower of the Gulf of Thailand that the total organic matter contents of sediment around coast of Upper Gulf of Thailand were 3.01 - 13.12%(dry weight). The total organic matter contents of sediments around coast of Lower Gulf of Thailand were 4.21 - 10.33%(dry weight).

Takahashi, et al. (44) studied the organic matter contents in around lower Gulf of Thailand that the total organic matter contents were 6.73±1.71 %(dry weight).

In 1981 Kasemthab (45) reported that the organic matter contents as volatile or loss on ignition in the upper Gulf of Thailand were about 3.01 – 13.12%. And there were not different between seasons. But the organic matter contents near the river mouth were higher than it was further away. The mean organic matter contents near the river mouth were 9.35 – 9.87% and far from the river mouth areas were 4.74 – 4.94%.

Rungsupa (46) measured the organic matter contents of sediment around the east coast of the Upper Gulf of Thailand. Two forms of measurement of organic matter were used, volatile organic matter or loss on ignition at high temperature and oxidisable organic matter or dichromate disgestion. The mean values of volatile organic matter in the Upper Gulf of Thailand in May and November 1991 were 10.986% and 11.595%. Mean value volatile organic matter and oxidisable organic matter in March 1992 were 11.267% and 1.8797%; in July there were 16.273 and 1.911% and in November 16.675 and 1.674%. In July 1993, the oxidisable organic matter 1993 were 1.119% to 2.8949%; in November there were 0.417% to 3.976% and

Copyright by Mahidol University

in March 1994, there were 0.986% to 4.805%. The profile distribution of oxidisable organic matter at the Bang Pakong river mouth, Bang Saen, Sichung Island and Pattaya tended to be higher near the river mouth and decreased for then away. The surface layer had a higher level than the deeper, except at the Bang Pakong river mouth. The maximum oxidisable organic matter were at the Bang Pakong river mouth: about 2.964% at 0 centimeter; 3.415% at 5 centimeter; 3.300% at 10 centimeter; and 3.128% at 15 centimeter. The minimum were at Pataya about 1.185% at 0 centimeter; 1.160% at 5 centimeter; 0.971% at 10 centimeter; and 0.774% at 15 centimeter.

Chareonpanich and Seurungreong (47) studied the total organic matter contents in Gulf of Thailand in 1996 were 3.05 - 15.4%(dry weight). The area of the Gulf can be remarkably observed to accumulate in two major zones; one in the central part of the upper gulf and another one in the western part of the Gulf.

2.4.2 Relevant Researches of Environmental Mapping

Environmental mapping is important tool to properly demonstrate geographical of parameters of environment. It can be used as management resource in the future. The data collection can be obtained by survey, experiment or remote sensing. The Training Department of SEAFDEC (48) survey supplied the data of the biological and oceanographic parameters in South China Sea (e.g. temperature, salinity, primary productivity, etc.) and presented maps by using percentile method.

Many research papers used geographic information system (GIS) or remote sensing method. Oguchia, Jarvieb and Neal (49) studied the regional water quality of the Humber catchment was mapped for key inorganic chemical determinants using a GIS system and an extensive Environment Agency and LOIS monitoring database. The resultant maps revealed the major factors affecting the general characteristics of regional water quality. Sewage inputs from industrial and domestic sources account for the high concentration of many determinants in urban areas.

Recently, the environmental mapping using the mixture model approach in studies, Basford et al (50) reported model of the distribution of stamp thickness for the printing of a given stamp issue form different types for paper that studied by Izenman and Sommer (51). They used the application of the nonparametric approach to identify components by the resulting placement of modes in the density estimate which fitting

Copyright by Mahidol University

a normal mixture model with seven components. They observed an apparent conflict when fitting a normal mixture model with three components with unequal variances. This conflict is examined further by investigating the most appropriate number of components when fitting a normal mixture of components with equal variance. For comparison with a parametric approach, Izenman and Sommer fitted a mixture of normal components with unequal variance to the data by maximum likelihood. They used the likelihood ratio statistic to test for the smallest number (g) of components compatible with the data. It yielded g=3 components, which is in apparent conflict with their nonparametric solution of seven modes. It will be seen that by restricting the normal components in the mixture model to have a common variance, the likelihood ratio test for the smallest number to components in the mixture ultimately leads to results consistent with the nonparametric approach.

Whitea et al, (52) proposed linear mixture model for map iron-oxide contents over the northern part of the Namib Sand Sea, Namibia.was that applied to Landsat Thematic Mapper data. Field samples were collected to calibrate the remotely sensed iron oxide proportions to dithionite extractable Fe in mg g-1, the in situ proportions correlating strongly with the remotely sensed estimates (r=0.91).

Metternicht and Fermont (53) presented the spectral mixture modeling that was performed for identification and mapping of land degradation features related to soil erosion processes in the Sacaba Valley, Bolivia. The model allowed use of up to five surface components to characterize the selected area, since six bands (1, 2, 3, 4, 5, and 7) of the Landsat TM sensor were used as inputs. The research has demonstrated that regional patterns of soil surface erosion features can be reliably mapped using linear spectral mixture analysis. Extrapolation of this approach to other regions where soil degradation features are correlated with spectrally distinguishable surface characteristics is feasible, provided that an optimization of the unmixing model as a function of local or regional surface component types is completed.

Carpenter et al (54) studied mapping of forest areas. While most forest maps identify only the dominant vegetation class in delineated stands, individual stands are often better characterized by a mix of vegetation types. This article examines various algorithms that use data from the Landsat Thematic Mapper (TM) satellite to estimate mixtures of vegetation types within forest stands. Included in the study are maximum

Copyright by Mahidol University

likelihood classification and linear mixture models as well as a new methodology based on the ARTMAP neural network.

Sutherlanda et al (55) reported a series of calibration experiments that were run in order to determine the effect of the colour of suspended sediment on the response of optical backscatter sensors (OBS). Munsell notation describes the properties of soil and sediment colour using the three parameters, hue, chroma, and value, which represent, respectively, the spectral position, lightness level, and amount of pure chromatic colour present. The OBS output was predicted using a linear mixture model incorporating end-member calibrations of black and white sediment.

Brown, Lewis, and Gunn (56) proposed the Mixture modeling that is becoming an increasingly important tool in the remote sensing community as researchers attempt to resolve subpixel, area information. It is shown that the constrained least squares Linear Spectral Mixture Models (LSMM) is equivalent to the linear Support Vector machines(SVM), which relies on proving that the LSMM algorithm possesses the `maximum margin' property.

Stephan (57) have sought techniques for estimating plant canopy characteristics from remote sensing data that are not site-specific. In this study, a procedure for estimating cotton (Gossypium spp.) canopy ground cover based on a linear mixture modeling approach was tested using Landsat multispectral imagery. Statistical analysis of the results suggest that the procedure produced accurate, unbiased estimates of cotton canopy ground cover.

Robertsa et al (58) studied natural deposition. Many natural deposits contain grains that have different burial histories, but reliable procedures to extract the component doses from mixed-dose samples have not been developed in optical dating. Here we present results for synthetic two- and three-component mixtures of data derived from laboratory-dosed sedimentary quartz using a single-aliquot regenerative-dose protocol and statistical models for finite mixtures. The result found that the correct number of dose components, corresponding doses, and relative proportions of each component may be estimated if the within-component dispersion is small or is known.

Copyright by Mahidol University

CHAPTER III
MATERIALS AND METHODS

3.1 Study Design

This research was a geographic environmental mapping by a way of descriptive study of the total organic matter contentss in the Gulf of Thailand in 51 stations distributed within the Gulf. The vessel, named M.V.SEAFDEC was assigned to convey the scientific data collection in the survey during 24 April to 17 May 1996.

3.2 Data Collection and Processing

3.2.1 Data Collection

The total organic matter contents of bottom sediments samples was collected from 51 stations in Gulf of Thailand. The location in latitude and longitude of each station was shown in Appendix C. These sediment samples were analyzed in the Laboratory of Aquatic Environments and Biochemistry, Department of Fishery Biology, Faculty of Fishery, Kasetsart University. The methods of analysis were described in Appendix D. The total organic matter analysis was carried out as one subproject under the Collaborative Research Project between SEAFDEC'-s Training Department (TD) in Thailand and the Marine Fisheries Resource Development and Management Department (MFRDMD) in Malaysia. The data of this study were secondary data which were reported and used under permission.

3.2.2 Data Processing

3.2.2.1 Initial Estimated of Variance for Each Cluster Group

In the initial study, we assume that the 51 stations were grouped into cluster unit of 4, 5 or 6 stations according to their reasonable range of measure data. In each of the cluster, the estimated variance of the cluster was the calculated for further analysis.

Copyright by Mahidol University

3.2.2.2 Management of Database by Recording into the Notepad Program

- The data file name t1.txt was assigned for the input data of the total organic matter contents of the 0-2 centimeter bottom sediments of 51 stations. The observe values and the estimated variance values in each group were filled into the first and second column of data file, respectively.
- The data file name t2.txt was assigned for the input data of the total organic matter contents of the 2-4 centimeter bottom sediments of 51 stations. The observe values and the estimated variance values in each group were filled into the first and second column of data file, respectively.
- The data file name t3.txt was assigned for the input data of the total organic matter contents of the 4-6 centimeter bottom sediments of 50 stations. The observe values and the estimated variance values in each group were filled into the first and second column of data file, respectively. One station (station number 45) was omitted due to unqualified specimen and unable to analyse preliminary into numerical measurement.
- The data file name tt.txt was assigned for the input data of the total organic matter contents in total levels (0-6 centimeter) of bottom sediments, which had 152 numeric values. The observe values and the variance values in each group were filled into the first and second column of data file, respectively.

3.2.2.3 Preparation of Data and Statistical Analysis

- Descriptive statistics of the total organic matter contents of three levels in sediments exploring i.e.: maximum, minimum, mean and standard deviation.
- Normality test of data in the data sets of the total organic matter contents of the 0-2 centimeter bottom sediments, 2-4 centimeter bottom sediments and 4-6 centimeter bottom sediments in the Gulf of Thailand.

Copyright by Mahidol University

3.3 Procedure for Environment Mapping

3.3.1 Mixture Model Method

As the mixture model method is assumed under the unobserved heterogeneous population assumption of each level of the total organic matter contents of bottom sediments in the Gulf of Thailand. The steps of this method are summarized for calculations as follows:

3.3.1.1 Estimation the Parameters of Mixture Distribution

The parameters of mixing distribution that are estimated for the total organic matter contents. It was chosen to be done by Nonparametric Maximum Likelihood Estimation, which is accomplished by the C.A.MAN program.

The C.A.MAN Program which can be downloaded from the homepage : http://www.medizine.fu-berlin.de/sozmed.bo1.html (11). The step of C.A.MAN program are summarized as follows:

1. Specify data-file (FIRST THING TO DO) and select specify the structure of your file:

1.1 The input data

1.2 Select specify the structure of file that it have :

1: Do you have ungrouped data ?

2: Do you have grouped data of the form: ?

VARIABLE REPLICATION FACTOR

Do you want to analyze the mixture of rates or

SMR"s with supposingly P O I S S O N (DEFAULT!!)

or BINOMIAL data grouped in the following way:

3: NUMBER OF EVENTS NUMBER OF EXPOSED (rates)

OBSERVED CASES EXPECTED CASES (smr"s)

4: NUMBER OF EVENTS NUMBER OF EXPOSED

REPLICATION FACTOR

OBSERVED CASES EXPECTED CASES

REPLICATION FACTOR

Copyright by Mahidol University

5: Mixture of supposingly normal-distributed data with
corresponding variance of the form:
VARIABLE VARIANCE

6: Do you want to read a predefined density matrix?

7: Mixture of geometric distributions:
VARIABLE REPLICATION FACTOR CENSORING VARIABLE

According to the nature of the study samples, the input data are Normal distribution, along with the corresponding variance.

2. Determination of method for computing the parameter grid from one of these choice :

1: Default parameter grid ! If your sample is less or equal 25 every data point is used as a parameter. Otherwise the number of grid points is estimated from your data and the grid is computed automatically

2: If you want to compute your grid of parameters individually

3: If you use f i x e d support size and if you want to enter starting values

In this study, we computed the grid of parameters individually due to the sample size of data being more than 25.

3. Specification of the kernel of the mixture from one of these choices.

3.1 Normal distribution with single variance.

3.2 Poisson distribution.

3.3 Laplace distribution.

3.4 Exponential distribution.

3.5 Poisson distribution for rates.

3.6 Normal distribution with different variance.

3.7 Predefined density matrix.

3.8 Binomial distribution.

3.9 Geometric-distribution.

In this study, the kernel of mixture is assumed to Normal distribution with different variance.

Copyright by Mahidol University

4. Selection of the level of accuracy to stop computing NPMLE from one of these choices.

1) 0.01
2) 0.01
3) 0.001
4) 0.0001

In this study, the level of accuracy 0.0001 is chosen because this value gives the most stable results of the iterations.

5. Computation of NPMLE when the number of components is not known.

The VEM algorithm is used for estimating NPMLE in the first step and then the results of the VEM algorithm are used as initial values for the EM algorithm to produce the fully iterated NPMLE. The results of estimation of the NPMLE consist of 4 issues, as follows :

1) The number of component : k
2) The levels of observed values for each component : $x_1, x_2, \ldots, x_j$ where j = 1, 2, …, k.
3) The probability of belonging to the j-th component : $p_1, p_2, \ldots, p_j$ where j = 1, 2, …, k.
4) The log-likelihood of NPMLE at component k.

6. Computation of NPMLE when the number of components is fixed.

This step is done to fix the number of components equal to k-1 and compute NPMLE with the EM algorithm : fix the number of components equal to k-2 and compute NPMLE with the EM algorithm : fix the number of components equal to k-3 and compute NPMLE; and so forth, until the number of components is equal to 1.

3.3.1.2 Testing for Alternatives of the Number of Components of Mixing Distribution

The determination of the appropriate number of components compatible with data would be solved by computing log-likelihood ratio statistics (LRS) for testing the number of components equal to k null hypothesis versus the number of components

Copyright by Mahidol University

equal to k+1 of alternative hypothesis. The critical values were obtained form the work of Thode et al. (39). This study used 95-th percentiles of the log-likelihood ratio test statistic for Normal distribution with unknown variance. The critical values were 6.68 and the sample size equal to 100.

3.3.1.3 Classification for Map Construction of Mixture Model

After we got the number components, this step was done to classify the value in each station into its subpopulations. This is done by computing the probability for each station of belonging to a certain component, which is also accomplished by the C.A.MAN program. Then coloring each station corresponding to the color of specified component in order to produce environmental map.

3.3.2 Conventional Method

This is a conventional way to express a random variable in term of percentile. The number of components of the total organic matter obtains by the results of the number of components of Mixture Model. The data managed into each component as the corresponding the percentile value. The steps of this method are summarized as follows:

3.3.2.1 Computation of Percentile

There are three levels (0-2, 2-4 and 4-6 centimeter) of the total organic matter contents concerned at 51, 51 and 50 stations.

1. Find a point of data in each percentile;

$$\text{percentile A} = \frac{(\max.\text{value} - \min.\text{value}) \times A}{100}$$

2. Find an interval of each component;

$$\text{component i} = \text{percentile A} + \text{min. value}$$

3. Fix the interval of each component

Copyright by Mahidol University

3.3.2.2 Classification for Mapping of Percentile

The number of components was obtained according to the method of Mixture model. The data were classified into each component based on the location of percentile. Then coloring each station corresponding to the color of it's component in order to produce environmental map.

Copyright by Mahidol University

CHAPTER IV
RESULTS

The issues in this chapter are the results of mapping with the Mixture Model approach and the conventional approach for investigating the distribution of the total organic matter contents of bottom sediments in the Gulf of Thailand in 1996. The estimation of Mixture Model under the unobserved heterogeneous population was done by Nonparametric Maximum Likelihood with the computer package C.A.MAN. The conventional method was done by percentile.

4.1 Descriptive Study of the Total Organic Matter Contents of Bottom Sediments in the Gulf of Thailand

4.1.1 The Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments

The values of the total organic matter contents of 0-2 centimeter bottom sediments in the Gulf of Thailand were illustrated in Table 2. The minimum value were 3.05% (dry weight) at station number 1 and the maximum value were 15.35%(dry weight) at station number 17. The average value of the total organic matter contents were 7.52% (dry weight) and standard deviation were 2.71.

4.1.2 The Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments

The values of the total organic matter contents of 2-4 centimeter bottom sediments in the Gulf of Thailand were illustrated in Table 2. The minimum value were 3.09% (dry weight) at station number 2 and the maximum value were 15.14%(dry weight) at station number 7. The average value of the total organic matter were 7.12%(dry weight) and standard deviation were 2.75.

Copyright by Mahidol University

4.1.3 The Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments

The values of the total organic matter contents of 4-6 centimeter bottom sediments in the Gulf of Thailand were illustrated in Table 2. These data had missing one value in station number 45. The minimum value were 2.75%(dry weight) at station number 50 and the maximum value were 14.31%(dry weight) at station number 17. The average value of the total organic matter were 6.96%(dry weight) and standard deviation were 2.76.

Table 2. The Total Organic Matter Contents (% Dry Weight) of Bottom Sediments in the Gulf of Thailand 1996

Stations	Total organic matter contents of 0-2 cm bottom sediments	Total organic matter contents of 2-4 cm bottom sediments	Total organic matter contents of 4-6 cm bottom sediments
1	3.05	3.36	3.24
2	4.04	3.09	5.02
3	5.58	5.37	4.55
4	4.39	4.39	4.84
5	9.33	8.55	8.37
6	5.93	4.86	4.75
7	12.22	15.14	13.81
8	5.17	4.90	4.66
9	4.27	4.11	4.08
10	5.87	5.63	4.76
11	5.99	7.06	7.63
12	12.73	12.36	11.54
13	8.11	7.46	5.86
14	8.36	7.40	6.89
15	5.86	5.71	6.03
16	11.47	11.13	10.27
17	15.35	13.56	14.31
18	9.71	9.54	9.45
19	7.36	5.96	6.22
20	7.75	8.03	8.13
21	7.06	5.79	4.88
22	7.88	7.17	8.00
23	14.46	14.18	13.93

Copyright by Mahidol University

Table 2. The Total Organic Matter Contents (% Dry Weight) of Bottom Sediments in the Gulf of Thailand 1996 (Continue)

Stations	Total organic matter contents of 0-2 cm bottom sediments	Total organic matter contents of 2-4 cm bottom sediments	Total organic matter contents of 4-6 cm bottom sediments
24	10.24	10.16	10.62
25	9.4	8.94	8.96
26	6.88	7.70	8.92
27	7.17	6.30	6.02
28	7.25	7.23	7.20
29	6.93	5.95	5.52
30	11.23	10.81	10.36
31	5.37	4.95	4.94
32	6.46	5.96	5.87
33	6.63	5.62	4.84
34	7.94	6.21	5.92
35	10.55	10.36	10.19
36	10.68	9.39	9.22
37	8.87	7.68	7.59
38	4.99	5.18	5.15
39	4.77	4.92	4.28
40	5.86	5.56	5.89
41	8.01	7.18	7.42
42	6.49	4.91	4.99
43	4.71	4.50	4.15
44	6.03	4.89	4.74
45	8.52	7.68	-
46	9.03	8.61	8.53
47	5.56	5.23	4.79
48	6.37	6.33	6.26
49	7.39	7.21	6.98
50	4.04	3.93	2.75
51	4.28	5.00	4.72
Minimum	3.05	3.09	2.75
Maximum	15.35	15.14	14.13
Mean	7.52	7.12	6.96
Standard deviation	2.71	2.75	2.76

Note : Technical analysis of the total organic matter contents of all sediments were done in Laboratory of Aquatic Environments and Biochemistry, Department of Fishery Biology, Faculty of Fisheries, Kasetsart University.

Copyright by Mahidol University

4.2 The Data Setting for Analysis

4.2.1 Normality Test in the Total Organic Matter Contents of Sediments

The data from the three levels sets (0-2, 2-4 and 4-6 centimeter) of the total organic matter contents were tested with the Kolmogorov-Sminov Normality test. There was not enough evidence to reject the null hypothesis of normal distribution of the data in each level sets of sediments (P-value = 0.597, 0.252 and 0.153, respectively).

4.2.2 The Classification of the Total Organic Matter Contents of Sediments

In the initial assumption, the 51 stations were roughly classified into 7 cluster groups according to their rather similarity in data distribution. It was shown in Table 3 with this initial assumption, the mean and variance in each group can be used as the initial value for estimating the parameters of the next step of the mixture procedure.

Table 3. The Mean and Variance Value of the Total Organic Matter Contents (% Dry Weight) in Each Cluster Group

Group	Station number	Mean value	Variance value
1	1, 50	3.40	0.253
2	2, 4, 9, 39, 43, 51	4.42	0.217
3	3, 6, 8, 10, 15, 21, 29, 31, 32, 33, 38, 40, 42, 44, 47, 48	5.53	0.411
4	11, 13, 14, 19, 20, 22, 26, 27, 28, 34, 37, 41, 45, 49	7.33	0.631
5	5, 18, 25, 36, 46	9.18	0.341
6	12, 16, 24, 30, 35	10.93	0.645
7	7, 17, 23	14.11	0.845

Copyright by Mahidol University

A map of geographical distribution of the total organic matter contents of bottom sediments was illustrated in Figure 2.

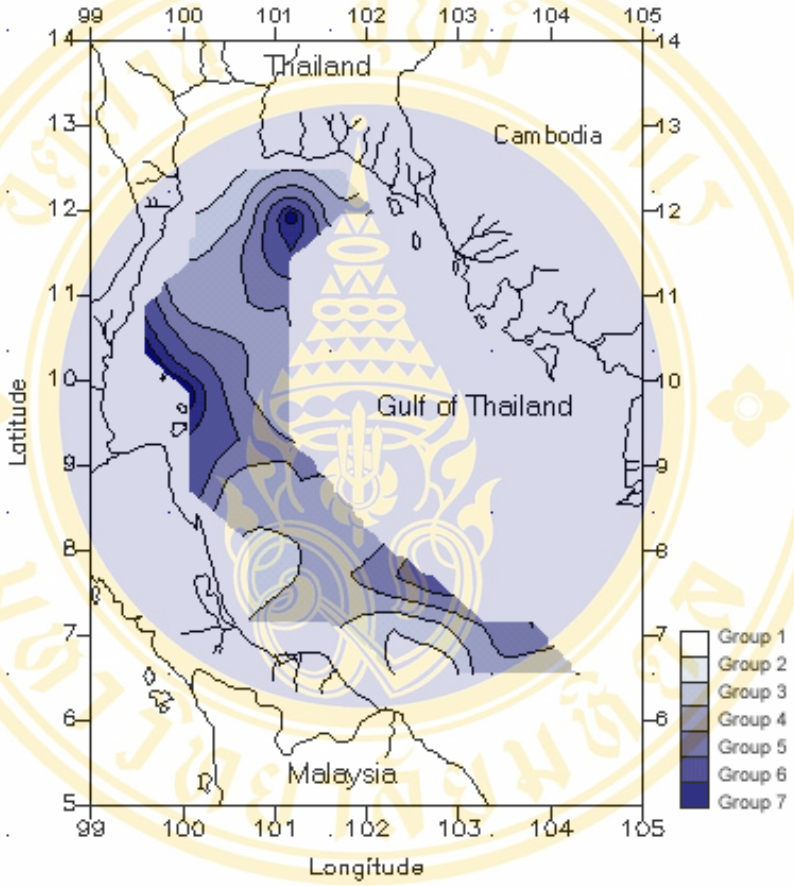


Figure 2. The Geographical Distribution of Actual Total Organic Matter Contents of Bottom Sediments in Each Assigned Cluster Groups

Copyright by Mahidol University

4.3 The Mixture Model of the Total Organic Matter Contents of Bottom Sediments in the Gulf of Thailand in 1996

4.3.1 The Mixture Model of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments

The total organic matter contents of 0-2 centimeter bottom sediments among 51 stations gave the minimum value of 3.05%(dry weight) at station number 1. The maximum value were 15.35%(dry weight) at station number 17. The average value were 7.52%(dry weight) where as the standard deviation were 2.71.

4.3.1.1 Estimation of the Parameters of Mixing Distribution

The maximum likelihood used for the parameters values estimation. The parameter values of mixing distribution are chosen which maximize the log-likelihood of mixture density. The results of estimation with the VEM algorithm indicated that the number of components k was equal to 7. The EM algorithm used to estimate the parameters, which the number of component is assumed to be 6, 5, 4, 3, 2 and 1. The parameters were calculated and corresponding weights for each subpopulation.

The NPMLE and corresponding log-likelihood of the total organic matter contentss of 0-2 centimeter bottom sediments at each fix number of components (6, 5, 4, 3, 2 and 1) were provided by C.A.MAN and shown in Table 4.

Copyright by Mahidol University

Table 4. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments

Component	Parameter	Weight	Log-likelihood
K=7	3.1348	0.0185	-116.1209
	4.5301	0.1981	
	6.4888	0.4124	
	8.9253	0.2061	
	10.9339	0.0929	
	12.2962	0.0330	
	14.8602	0.0388	
K=6	4.3660	0.2038	-116.8871
	6.4449	0.4208	
	8.9044	0.2101	
	10.9279	0.0933	
	12.2956	0.0332	
	14.8602	0.0388	
K=5	4.3668	0.2041	-117.0948
	6.4499	0.4223	
	8.9334	0.2129	
	11.2694	0.1204	
	14.8148	0.0404	
K=4	4.4033	0.2164	-121.3293
	6.6326	0.4703	
	9.7271	0.2367	
	13.6114	0.0767	
K=3	4.4489	0.2313	-142.9720
	6.7956	0.4985	
	10.6245	0.2701	
K=2	5.4199	0.6157	-162.8786
	9.8199	0.3843	
K=1	6.7872	1.0000	-206.2181

Copyright by Mahidol University

4.3.1.2 Alternating for the Number of Components of Mixing Distribution

The log-likelihood ratio statistic was used for testing the null hypothesis number of components equal to k, versus the alternative hypothesis number of components equal to k+1. The critical value of the log-likelihood ratio, with 95% simulated critical value of 6.68 (39), estimate sample size equal to 51.

The seven successive possible hypothesis testing on the number of components of the total organic matter contents of 0-2 centimeter bottom sediments by log-likelihood ratio statistic were shown in Table 5.

Table 5. Likelihood Ratio Statistic of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments Data

Hypothesis	Likelihood ratio statistic		
H_0 : k = 6 H_1 : k = 7	2*[-116.1209+116.8871]	=	1.5324
H_0 : k = 5 H_1 : k = 6	2*[-116.8871+117.0948]	=	0.4154
H_0 : k = 4 H_1 : k = 5	2*[-117.0948+121.3293]	=	8.4690*
H_0 : k = 3 H_1 : k = 4	2*[-121.3293+142.9720]	=	43.2854*
H_0 : k = 2 H_1 : k = 3	2*[-142.9720+162.8786]	=	39.8132*
H_0 : k = 1 H_1 : k = 2	2*[-162.8786+206.2181]	=	86.6790*

* statistical significance at $\alpha = 0.05$

These results indicated that the smallest number of components compatible with the total organic matter contents of 0-2 centimeter bottom sediments was five components. The parameter values and weights were shown:

Copyright by Mahidol University

Parameter	Value				
Mean	μ_1	μ_2	μ_3	μ_4	μ_5
	4.3668	6.4499	8.9334	11.2694	14.8148
Weight	p_1	p_2	p_3	p_4	p_5
	0.2041	0.4223	0.2129	0.1204	0.0404

4.3.1.3 Classification for Map Construction

The classification can be accomplished by means of the posterior distribution, which the each station is classified into one of the components for which the posterior probability is largest. The classification results of the total organic matter of 0-2 centimeter bottom sediments were shown in Table 6.

Table 6. Classification of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments

Stations	% dry weight	Classification of groups
1	3.05	1
2	4.04	1
3	5.58	2
4	4.39	1
5	9.33	3
6	5.93	2
7	12.22	4
8	5.17	1
9	4.27	1
10	5.87	2
11	5.99	2
12	12.73	4
13	8.11	3
14	8.36	3
15	5.86	2
16	11.47	4
17	15.35	5
18	9.71	3
19	7.36	2
20	7.75	2

Copyright by Mahidol University

Table 6. Classification of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments (Continue)

Stations	% dry weight	Classification of groups
21	7.06	2
22	7.88	3
23	14.46	5
24	10.24	3
25	9.4	3
26	6.88	2
27	7.17	2
28	7.25	2
29	6.93	2
30	11.23	4
31	5.37	2
32	6.46	2
33	6.63	2
34	7.94	3
35	10.55	4
36	10.68	4
37	8.87	3
38	4.99	1
39	4.77	1
40	5.86	2
41	8.01	3
42	6.49	2
43	4.71	1
44	6.03	2
45	8.52	3
46	9.03	3
47	5.56	2
48	6.37	2
49	7.39	2
50	4.04	1
51	4.28	1

Table 6 indicated the high amount of the total organic matter contents component were station number17 and 23. They were 14.46 – 15.35%(dry weight). Six and twelve stations classified into slightly high and medium amount of the total organic matter component that they were 10.55-12.73%(dry weight) and 7.88-10.24%(dry weight), respectively. The 42.23%(21 stations) and ten stations were classified into slightly low and low amount of the total organic matter component that this component discovered very much that they were 5.37-7.75%(dry weight) and 3.05-5.17%(dry weight), respectively.

Copyright by Mahidol University

A map of geographical distribution of the total organic matter contents of 0-2 centimeter bottom sediments was illustrated in Figure 3.

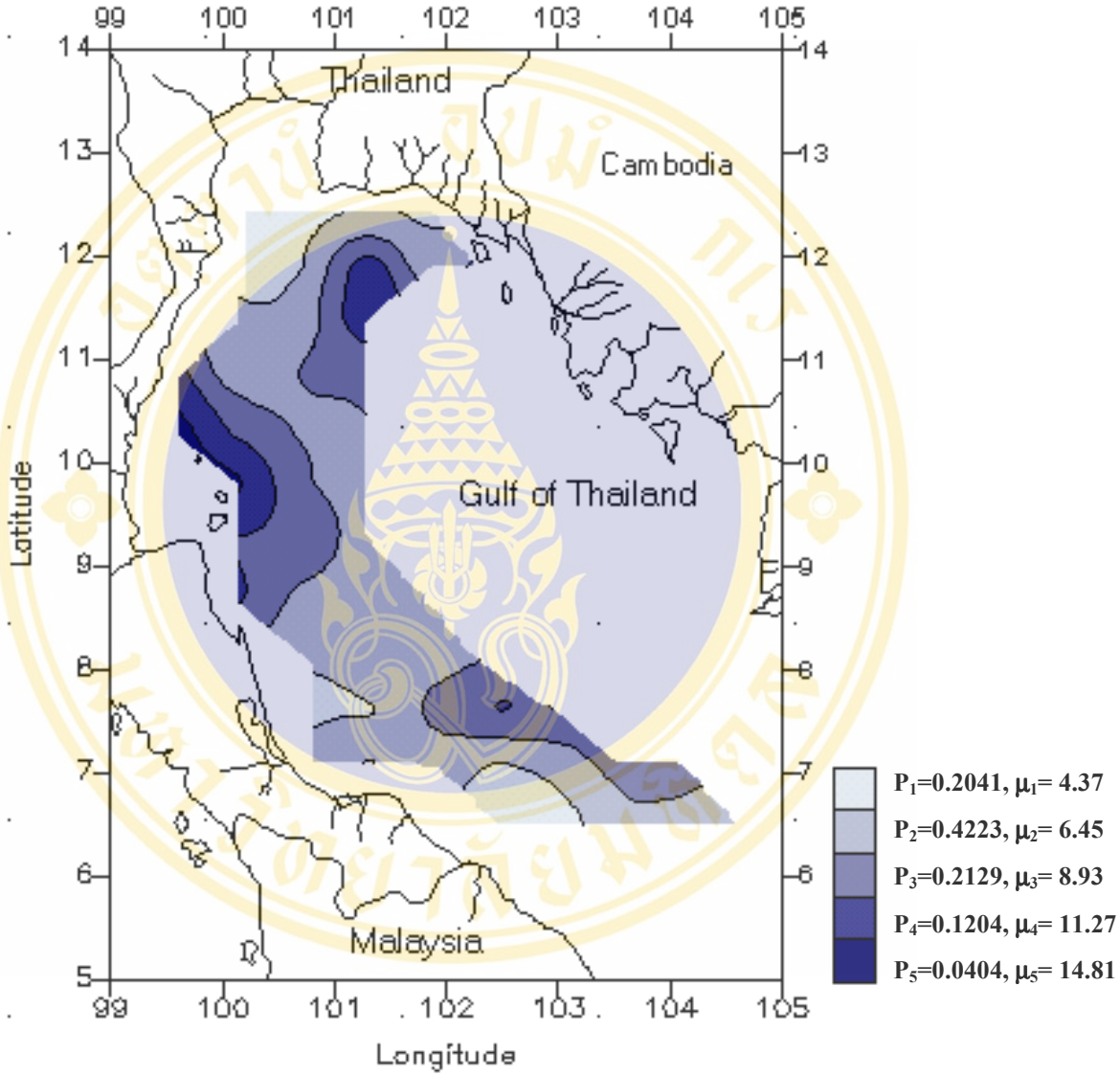


Figure 3. Map of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments Classified into Component According to the Mixture Model

Copyright by Mahidol University

4.3.2 Mixture Model of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments

The total organic matter contents of 2-4 centimeter bottom sediments among 51 stations gave the minimum valued of 3.09%(dry weight) at station number 2. The maximum value were 15.14%(dry weight) at station number 7. The average value were 7.12%(dry weight) where as standard deviation were 2.75.

4.3.2.1 Estimation of the Parameters of Mixing Distribution

The maximum likelihood used for the parameters values estimation. The parameter values of mixing distribution are chosen which maximize the log-likelihood of mixture density. The results of estimation with the VEM algorithm indicated that the number of components k was equal to 7. The EM algorithm used to estimate the parameters, which the number of component is assumed to be 6, 5, 4, 3, 2 and 1. The parameters were calculated and corresponding weights for each subpopulation.

The NPMLE and corresponding log-likelihood of the total organic matter contents of 2-4 centimeter bottom sediments at each fix number of components (6, 5, 4, 3, 2 and 1) were provided by C.A.MAN and shown in Table 7.

Copyright by Mahidol University

Table 7. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments

Component	Parameter	Weight	Log-likelihood
k=7	3.6989	0.0811	
	5.2291	0.4242	
	7.2441	0.2259	
	9.0255	0.1311	
	10.7768	0.0635	
	12.5081	0.0181	
	14.2353	0.0561	-113.6265
k=6	3.6992	0.0812	
	5.2296	0.4244	
	7.2478	0.2266	
	9.0550	0.1349	
	11.0873	0.0708	
	14.1353	0.0622	-113.6926
k=5	3.7547	0.0890	
	5.2977	0.4425	
	7.7166	0.2549	
	9.8294	0.1359	
	13.7145	0.0777	-114.8332
k=4	4.0776	0.1499	
	5.7302	0.4930	
	8.9407	0.2769	
	13.6338	0.0802	-122.0065
k=3	4.9303	0.5361	
	8.4145	0.3582	
	12.899	0.1058	-131.5993
k=2	5.2761	0.7017	
	10.0135	0.2983	-158.3135
k=1	6.4145	1.0000	-195.5051

Copyright by Mahidol University

4.3.2.2 Alternating for the Number of Components of Mixing Distribution

The log-likelihood ratio statistic was used for testing the null hypothesis number of components equal to k, versus the alternative hypothesis number of components equal to k+1. The critical value of the log-likelihood ratio, with 95% simulated critical value of 6.68 (39), estimate sample size equal to 51.

The seven successive possible hypothesis testing on the number of components of the total organic matter contents of 2-4 centimeter bottom sediments by log-likelihood ratio statistic were shown in Table 8.

Table 8. Likelihood Ratio Statistic of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments Data

Hypothesis	Likelihood ratio statistic		
H_0 : k = 6 H_1 : k = 7	2*[-113.6265+113.6926]	=	0.1322
H_0 : k = 5 H_1 : k = 6	2*[-113.6926+114.8332]	=	2.2814
H_0 : k = 4 H_1 : k = 5	2*[-114.8332+122.0065]	=	14.3466*
H_0 : k = 3 H_1 : k = 4	2*[-122.0065+131.5993]	=	19.1856*
H_0 : k = 2 H_1 : k = 3	2*[-131.5993+158.3135]	=	53.4284*
H_0 : k = 1 H_1 : k = 2	2*[-158.3135+195.5051]	=	74.3832*

* Statistical significance at $\alpha = 0.05$

Copyright by Mahidol University

These results indicate that the smallest number of components compatible with total organic matter contents of 2-4 centimeter sediments was five components. The parameter values and weights were shown :

Parameter	Value				
Mean	μ_1	μ_2	μ_3	μ_4	μ_5
	3.7547	5.2977	7.7166	9.8294	13.7145
Weight	p_1	p_2	p_3	p_4	p_5
	0.0890	0.4425	0.2549	0.1359	0.0777

4.3.2.3 Classification for Map Construction

The classification can be accomplished by means of the posterior distribution, which each station is classified into one of the components for which the posterior probability is largest. The classification results of the total organic matter contents of 2-4 centimeter bottom sediments were shown in Table 9.

Table 9. Classification of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments

Stations	% dry weight	Classification of groups
1	3.36	1
2	3.09	1
3	5.37	2
4	4.39	2
5	8.55	3
6	4.86	2
7	15.14	5
8	4.90	2
9	4.11	1
10	5.63	2
11	7.06	3
12	12.36	5
13	7.46	3
14	7.40	3

Copyright by Mahidol University

Table 9. Classification of Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments (Continue)

Stations	% dry weight	Classification of groups
15	5.71	2
16	11.13	4
17	13.56	5
18	9.54	4
19	5.96	2
20	8.03	3
21	5.79	2
22	7.17	3
23	14.18	5
24	10.16	4
25	8.94	4
26	7.70	3
27	6.30	2
28	7.23	3
29	5.95	2
30	10.81	4
31	4.95	2
32	5.96	2
33	5.62	2
34	6.21	2
35	10.36	4
36	9.39	4
37	7.68	3
38	5.18	2
39	4.92	2
40	5.56	2
41	7.18	3
42	4.91	2
43	4.50	2
44	4.89	2
45	7.68	3
46	8.61	3
47	5.23	2
48	6.33	2
49	7.21	3
50	3.93	1
51	5.00	2

Copyright by Mahidol University

Table 9 indicated the high amount of the total organic matter contents component were station number 7, 12, 17 and 23. They were 12.36 – 15.14%(dry weight). Seven and thirteen stations classified into slightly high and medium amount of the total organic matter component that they were 8.94-11.13%(dry weight) and 7.06 – 8.61% (dry weight), respectively. The 44.25%(23 stations) and four stations were classified into slightly low and low amount of the total organic matter component. There were 4.39-6.33%(dry weight) and 3.09-4.11%(dry weight), respectively.

A map of geographical distribution of the total organic matter contents of 2-4 centimeter bottom sediments is illustrated in Figure 4.

Copyright by Mahidol University

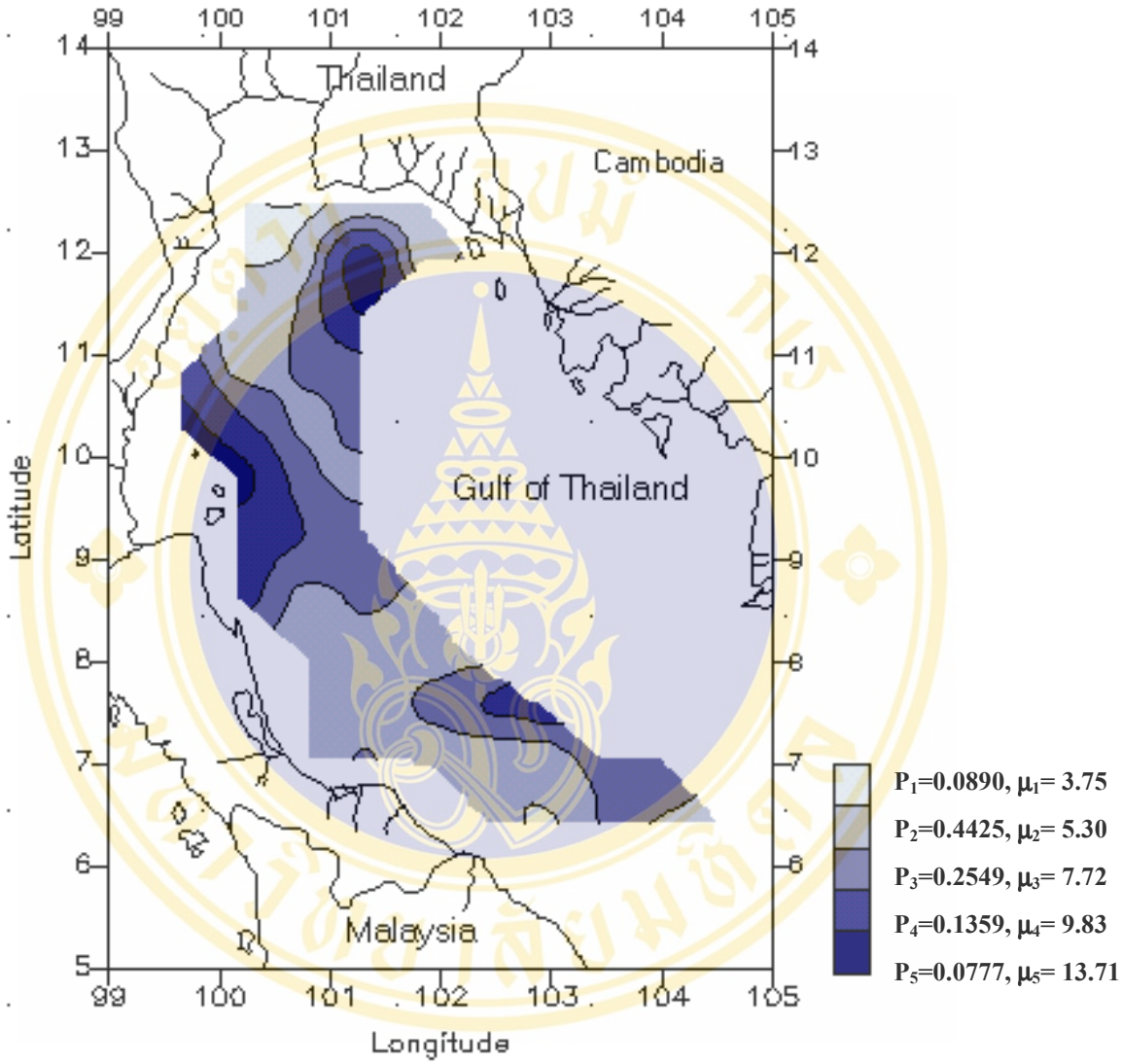


Figure 4. Map of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments Classified into Component According to the Mixture Model

Copyright by Mahidol University

4.3.3 Mixture Model of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments

The total organic matter contents of 4-6 centimeter bottom sediments among 51 stations which the station number 45 was missing data. The minimum value were 2.75%(dry weight) at station number 50 and the maximum value were 14.31%(dry weight) at station number 17. The average values were 6.96%(dry weight) where as standard deviation were 2.76.

4.3.3.1 Estimation of the Parameters of Mixing Distribution

The maximum likelihood used for the parameters values estimation. The parameter values of mixing distribution are chosen which maximize the log-likelihood of mixture density. The results of estimation with the VEM algorithm indicated that the number of components k was equal to 7. The EM algorithm used to estimate the parameters, which the number of component is assumed to be 6, 5, 4, 3, 2 and 1. The parameters were calculated and corresponding weights for each subpopulation.

The NPMLE and corresponding log-likelihood of the total organic matter contents of 4-6 centimeter bottom sediments at each fix number of components (6, 5, 4, 3, 2 and 1) were provided by C.A.MAN and shown in Table 10.

Copyright by Mahidol University

Table 10. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments

Component	Parameter	Weight	Log-likelihood
k=6	3.0327	0.0400	
	4.8351	0.4235	
	6.4687	0.1839	
	8.6774	0.2035	
	10.4894	0.0890	
	14.0013	0.0602	-108.1104
k=5	3.0359	0.0403	
	4.8634	0.4411	
	6.7601	0.2057	
	9.2733	0.2489	
	13.8200	0.0639	-110.4875
k=4	3.0789	0.0425	
	5.0646	0.5495	
	8.789	0.3301	
	13.3177	0.0778	-114.5961
k=3	4.8357	0.5787	
	8.7310	0.3429	
	13.3006	0.0784	-122.4497
k=2	4.9208	0.6191	
	9.5176	0.3809	-156.3752
k=1	6.2705	1.0000	-216.3925

4.3.3.2 Alternating for the Number of Components of Mixing Distribution

The log-likelihood ratio statistic was used for testing the null hypothesis number of components equal to k, versus the alternative hypothesis number of components equal to k+1. The critical value of the log-likelihood ratio, with 95% simulated critical value of 6.68 (39), estimate sample size equal to 51.

The seven successive possible hypothesis testing on the number of components of the total organic matter contents of 4-6 centimeter bottom sediments by log-likelihood ratio statistic were shown in Table 11.

Copyright by Mahidol University

Table 11. Likelihood Ratio Statistic of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments Data

Hypothesis	Likelihood ratio statistic		
H_0 : k = 5 H_1 : k = 6	2*[-108.1104+110.4875]	=	4.7542
H_0 : k = 4 H_1 : k = 5	2*[-110.4875+114.5961]	=	8.2172*
H_0 : k = 3 H_1 : k = 4	2*[-114.5961+122.4497]	=	15.7072*
H_0 : k = 2 H_1 : k = 3	2*[-122.4497+156.3752]	=	67.8510*
H_0 : k = 1 H_1 : k = 2	2*[-156.3752+216.3925]	=	120.0346*

* statistical significance at $\alpha = 0.05$

These results indicate that the smallest number of components compatible with the total organic matter contents of 4-6 centimeter bottom was three components. The parameter values and weights were shown :

Parameter	Value				
Mean	μ_1	μ_2	μ_3	μ_4	μ_5
	3.0359	4.8634	6.7601	9.2733	13.8200
Weight	p_1	p_2	p_3	p_4	p_5
	0.0403	0.4411	0.2057	0.2489	0.0639

Copyright by Mahidol University

4.3.3.3 Classification for Map Construction

The classification can be accomplished by means of the posterior distribution that each station is classified into one of the components for which the posterior probability is largest. The classification results of the total organic matter contents of 4-6 centimeter bottom sediments were shown in Table 12.

Table 12. Classification of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments

Stations	% dry weight	Classification of groups
1	3.24	1
2	5.02	2
3	4.55	2
4	4.84	2
5	8.37	4
6	4.75	2
7	13.81	5
8	4.66	2
9	4.08	2
10	4.76	2
11	7.63	3
12	11.54	4
13	5.86	2
14	6.89	3
15	6.03	3
16	10.27	4
17	14.31	5
18	9.45	4
19	6.22	3
20	8.13	4
21	4.88	2
22	8.00	4
23	13.93	5
24	10.62	4
25	8.96	4
26	8.92	4
27	6.02	2
28	7.20	3
29	5.52	2
30	10.36	4
31	4.94	2
32	5.87	2

Copyright by Mahidol University

Table 12. Classification of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments (Continue)

Stations	% dry weight	Classification of groups
33	4.84	2
34	5.92	2
35	10.19	4
36	9.22	4
37	7.59	3
38	5.15	2
39	4.28	2
40	5.89	2
41	7.42	3
42	4.99	2
43	4.15	2
44	4.74	2
45	-	-
46	8.53	4
47	4.79	2
48	6.26	3
49	6.98	3
50	2.75	1
51	4.72	2

Table 12 indicated the high amount of the total organic matter were station number 7, 17 and 23. There were 13.81 – 14.31%(dry weight). Thirteen and nine stations classified into slightly high and medium amount of the total organic matter component that they were 8.00 – 11.54%(dry weight) and 6.03-7.63%(dry weight), respectively. The 44.11%(23 stations) and two stations were classified into slight low and low amount of the total organic matter component. There were 4.08-6.02%(dry weight) and 2.75-3.24%(dry weight), respectively.

A map of geographical distribution of the total organic matter contents of 4-6 centimeter bottom sediments was illustrated in Figure 5.

Copyright by Mahidol University

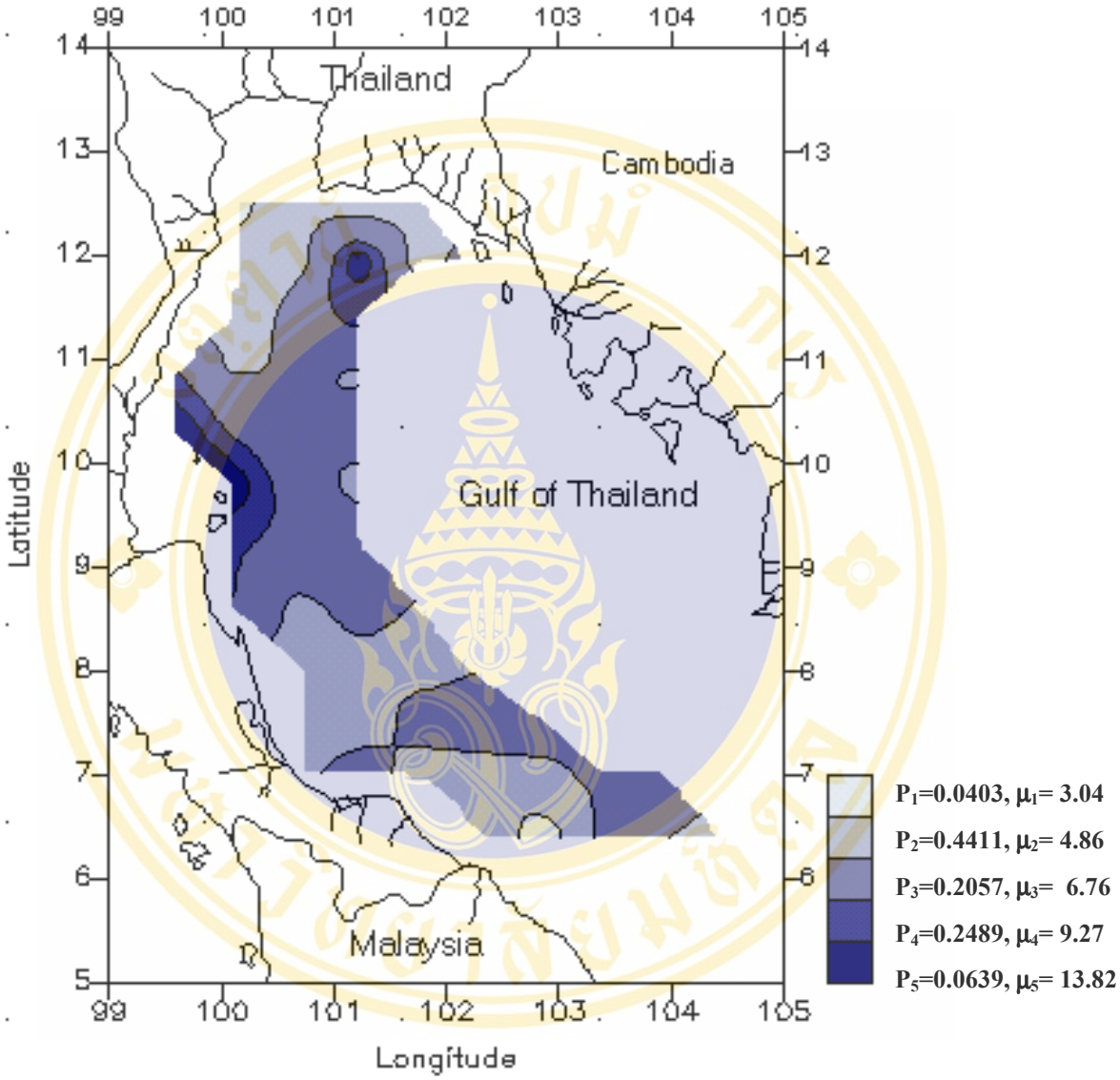


Figure 5. Map of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments Classified into Component According to the Mixture Model

Copyright by Mahidol University

4.3.4 Mixture Model for Values of the Total Organic Matter Contents of the Three Levels (0-2, 2-4 and 4-6 Centimeter) Bottom Sediments

The area of Gulf of Thailand parcels 51 stations. The data sets were three levels, which gave 150 value of the total organic matter contents. The minimum value were 2.75%(dry weight) at station number 50 of level 3 (4-6 centimeter). The maximum value were 15.35%(dry weight) at station number 17 of level 1 (0-2 centimeter). The average value of three levels were 7.20%(dry weight) where as standard deviation were 2.73.

4.3.4.1 Estimation of the Parameters of Mixing Distribution

The maximum likelihood used for the parameters values estimation. The parameter values of mixing distribution are chosen which maximize the log-likelihood of mixture density. The results of estimation with the VEM algorithm indicated that the number of components k was equal to 7. The EM algorithm used to estimate the parameters, which the number of component is assumed to be 6, 5, 4, 3, 2 and 1. The parameters were calculated and corresponding weights for each subpopulation.

The NPMLE and corresponding log-likelihood of the total organic matter contents of three levels of bottom sediments at each fix number of components (6, 5, 4, 3, 2 and 1) were provided by C.A.MAN and shown in Table 13.

Copyright by Mahidol University

Table 13. NPMLE and Corresponding Log-likelihood of the Total Organic Matter Contents of the Three Levels of Bottom Sediments

Component	Parameter	Weight	Log-likelihood
k=7	3.1869	0.0341	
	4.7394	0.2901	
	6.0561	0.2182	
	7.4202	0.1626	
	9.0538	0.1527	
	10.9486	0.0824	
	14.0715	0.0598	-342.6683
k=6	3.1989	0.0350	
	4.8049	0.3291	
	6.5764	0.3009	
	8.8589	0.1873	
	10.8991	0.0874	
	14.0613	0.0602	-343.8634
k=5	3.2214	0.0363	
	4.8701	0.3594	
	6.8356	0.3115	
	9.4934	0.2186	
	13.6719	0.0741	-351.7723
k=4	4.5233	0.3267	
	6.5541	0.3667	
	9.4200	0.2316	
	13.6442	0.0750	-361.4290
k=3	4.9279	0.5192	
	8.4200	0.3613	
	12.4433	0.1196	-403.7855
k=2	5.1658	0.6258	
	9.6119	0.3742	-478.8264
k=1	6.4917	1.0000	-617.3117

Copyright by Mahidol University

4.3.4.2 Alternating for the Number of Components of Mixing Distribution

The log-likelihood ratio statistic was used for testing the null hypothesis number of components equal to k, versus the alternative hypothesis number of components equal to k+1. The critical value of the log-likelihood ratio, with 95% simulated critical value of 6.68, estimate sample size equal to 51 (39).

The seven successive possible hypothesis testing on the number of components of the total organic matter contents of three levels of bottom sediments by log-likelihood ratio statistic were shown in Table 14.

Table 14. Likelihood Ratio Statistic of the Total Organic Matter Contents of the Three Levels of Bottom Sediments Data

Hypothesis	Likelihood ratio statistic		
H_0 : k = 6 H_1 : k = 7	2*[-342.6683+343.8634]	=	2.3902
H_0 : k = 5 H_1 : k = 6	2*[-343.8634+351.7723]	=	15.8178*
H_0 : k = 4 H_1 : k = 5	2*[-351.7723+361.4290]	=	19.3134*
H_0 : k = 3 H_1 : k = 4	2*[-361.4290+403.7855]	=	84.7130*
H_0 : k = 2 H_1 : k = 3	2*[-403.7855+478.8264]	=	150.0818*
H_0 : k = 1 H_1 : k = 2	2*[-478.8264+617.3117]	=	276.9706*

* Statistical significance at $\alpha = 0.05$

Copyright by Mahidol University

These results indicate that the smallest number of components compatible with the total organic matter contents of three levels of bottom sediments was six components. The parameter values and weights were shown :

Parameter	Value					
Mean	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
	3.1989	4.8049	6.5764	8.8589	10.8991	14.0613
Weight	p_1	p_2	p_3	p_4	p_5	p_6
	0.0350	0.3291	0.3009	0.1873	0.0874	0.0602

4.3.4.3 Classification for Map Construction

The classification can be accomplished by means of the posterior distribution which each station is classified into one of the components for which the posterior probability is largest. The classification results of the total organic matter contents of three levels of bottom sediments were shown in Table 15.

Table 15. Classification of the Total Organic Matter Contents of the Three Levels of Bottom Sediments

Stations	Classification of groups of 0-2 cm bottom sediments	Classification of groups of 2-4 cm bottom sediments	Classification of groups of 4-6 cm bottom sediments
1	1	1	1
2	2	1	2
3	2	2	2
4	2	2	2
5	4	4	4
6	3	2	2
7	5	6	6
8	2	2	2
9	2	2	2
10	3	2	2
11	3	3	3
12	6	5	5
13	4	3	3

Copyright by Mahidol University

Table 15. Classification of the Total Organic Matter Contents of the Three Levels of Bottom Sediments (Continue)

Stations	Classification of groups of 0-2 cm bottom sediments	Classification of groups of 2-4 cm bottom sediments	Classification of groups of 4-6 cm bottom sediments
14	4	3	3
15	3	2	3
16	5	5	5
17	6	6	6
18	4	4	4
19	3	3	3
20	3	4	4
21	3	3	2
22	4	3	4
23	6	6	6
24	5	5	5
25	4	4	4
26	3	3	4
27	3	3	3
28	3	3	3
29	3	3	2
30	5	5	5
31	2	2	2
32	3	3	3
33	3	2	2
34	4	3	3
35	5	5	5
36	5	4	4
37	4	3	3
38	2	2	2
39	2	2	2
40	3	2	3
41	4	3	3
42	3	2	2
43	2	2	2
44	3	2	2
45	4	3	-
46	4	4	4
47	2	2	2
48	3	3	3
49	3	3	3
50	2	2	1
51	2	2	2

Copyright by Mahidol University

Table 15 indicated six components within 51 stations (152 values) of the total organic matter contents of three levels. The twenty-eight stations (about 54.9%) were stable of allocation of total organic matter contents. The twenty-three stations were difference, the station number 7, 20 and 26 were decreased form bottom level to surface level of the total organic matter contents. The station number 6, 10, 12, 13, 14, 21, 29, 33, 34, 36, 37, 41, 42, 44, 45 and 50 were increased from bottom level to surface level of the total organic matter contents. The station number 2, 15, 22 and 40 were fluctuated into component of the total organic matter contents.

A map of geographical distribution of total organic matter contents of three levels of bottom sediments in the Gulf of Thailand was illustrated in Figure 6.

Copyright by Mahidol University

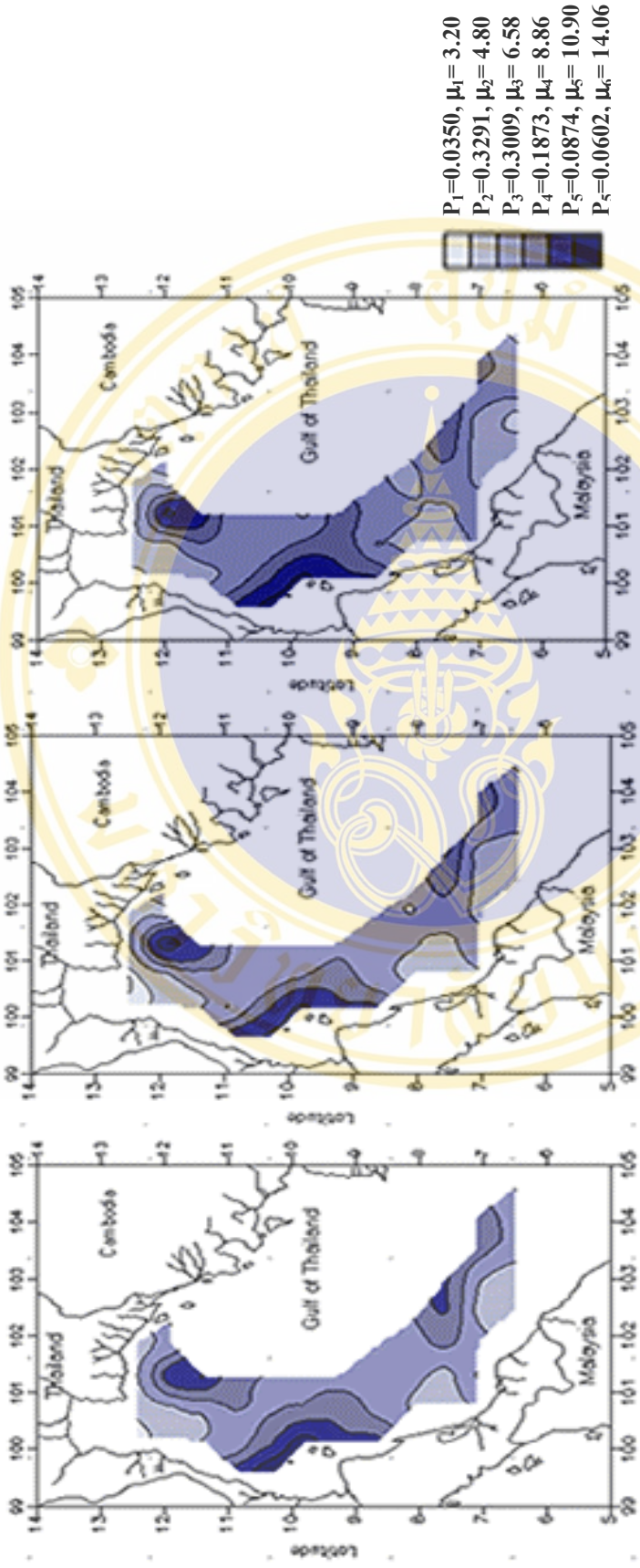


Figure 6. Map of the Total Organic Matter Contents in Each Levels Classified into Component According to the Mixture Model

Copyright by Mahidol University

4.4 The Conventional Method of the Total Organic Matter Contents of Bottom Sediments in the Gulf of Thailand

The conventional percentile method for the classification of the total organic matter contents has some limitations. First, the classification is a reliable approach (or not) depending on the experiencing of the researchers or the experts. Second, the classified stations have not been significant by the scientific method such as a statistical inference, a wavelength reflection of the high technology tools. In this study, we try to reduce some errors by using the number of components from the Mixture Model approach.

4.4.1 Conventional Method of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments

According to the Nonparametric Maximum Log-likelihood in Mixture Model, at 0-2 centimeter, there were five components. The cut of point assume with the equality of the interval of the percentile for the five components of the 0-2 centimeter level by using percentile was illustrated in Table 16.

Table 16. The Interval Value and the Number of Station of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments in Each Component

Component	The interval of the percentile	The interval of the total organic matter contents (%dry weight)	Number of stations
1	1-20	3.05-5.51	11
2	21-40	5.52-7.97	22
3	41-60	7.98-10.43	10
4	61-80	10.44-12.89	6
5	81-100	12.90-15.35	2

Table 16 indicated the high amount of total organic matter component was station number 17 and 23 station that they were 14.46- 15.35%(dry weight). Six and ten stations were classified into slightly high and medium amount of total organic matter component that they were 10.55 – 12.73%(dry weight) and 8.01-10.24%(dry weight),

Copyright by Mahidol University

respectively. Twenty-two and eleven stations were classified into slightly low and low amount of total organic matter component. They were 5.56 – 7.49%(dry weight) and 3.05-5.37%(dry weight), respectively. The results of data were illustrated in Table 19.

A map of geographical distribution of total organic matter of 0-2 centimeter bottom sediments in the Gulf of Thailand is illustrated in Figure 7.

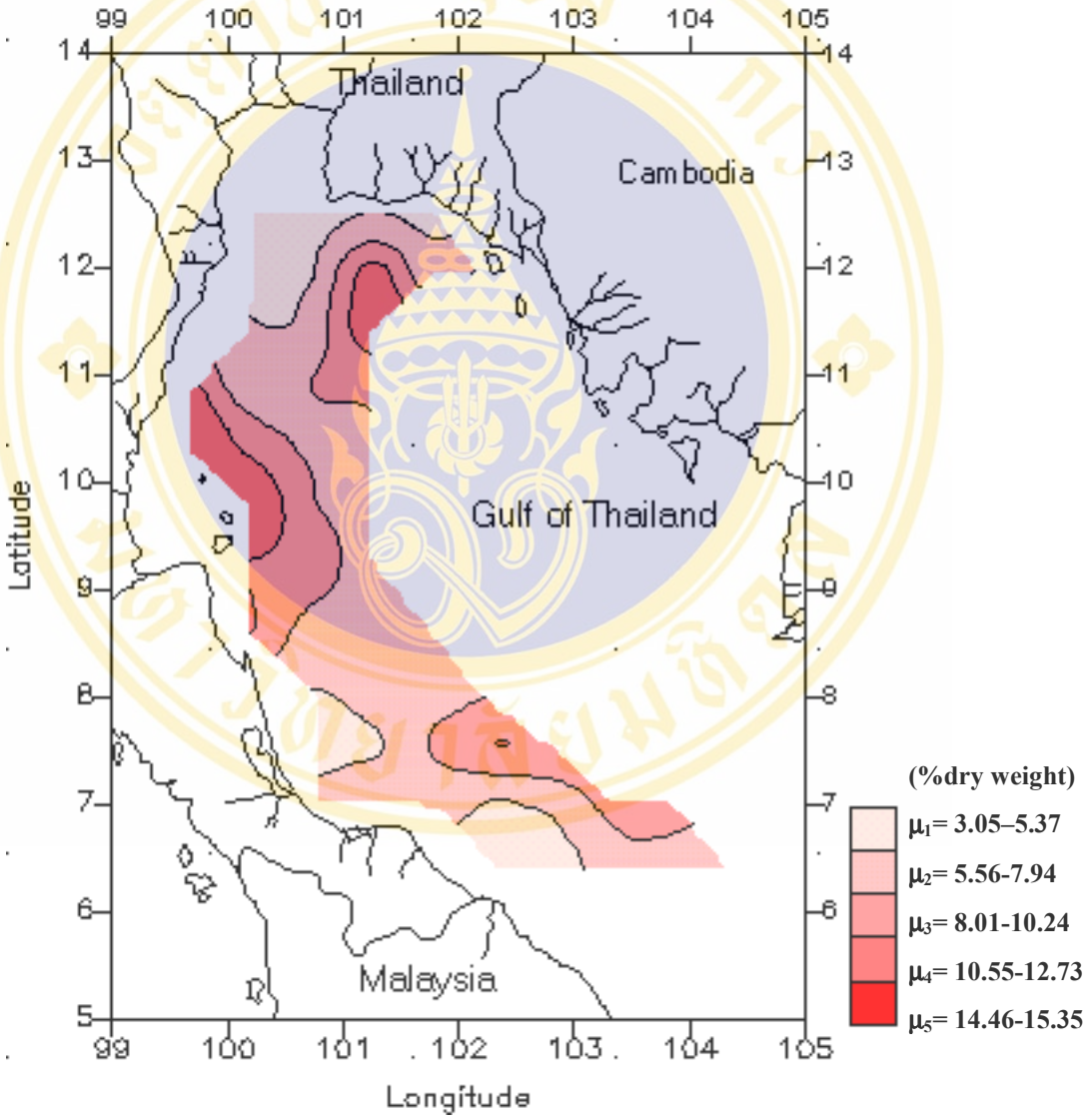


Figure 7. Map of the Total Organic Matter Contents of 0-2 Centimeter Bottom Sediments Classified into Component According to the Conventional Method

Copyright by Mahidol University

4.4.2 Conventional Method of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments

According to the Nonparametric Maximum Log-likelihood in Mixture Model, at 2-4 centimeter, there were five components The cut of point for the five components of the 2-4 centimeter level by using percentile was illustrated in Table 17.

Table 17. The Interval Value and the Number of Station of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments in Each Component

Component	The interval of the percentile	The interval of the total organic matter contents (%dry weight)	Number of stations
1	1-20	3.09-5.50	16
2	21-40	5.51-7.91	21
3	41-60	7.92-10.32	7
4	61-80	10.33-12.73	4
5	81-100	12.74-15.14	3

Table 17 indicate the high amount of total organic matter component were classified three stations that they were 13.56 – 15.14%(dry weight). Four and seven stations were classified into slightly high and medium amount of total organic matter component that they were 10.36 – 12.36%(dry weight) and 8.03 – 10.16%(dry weight), respectively. Twenty-one and sixteen stations were classified into slightly low and low amount of total organic matter component. They were 5.56 - 7.70%(dry weight) and 3.09 - 5.37%(dry weight), respectively. The results of data were illustrated in Table 19.

A map of geographical distribution of total organic matter of 2-4 centimeter bottom sediments in the Gulf of Thailand is illustrated in Figure 8.

Copyright by Mahidol University

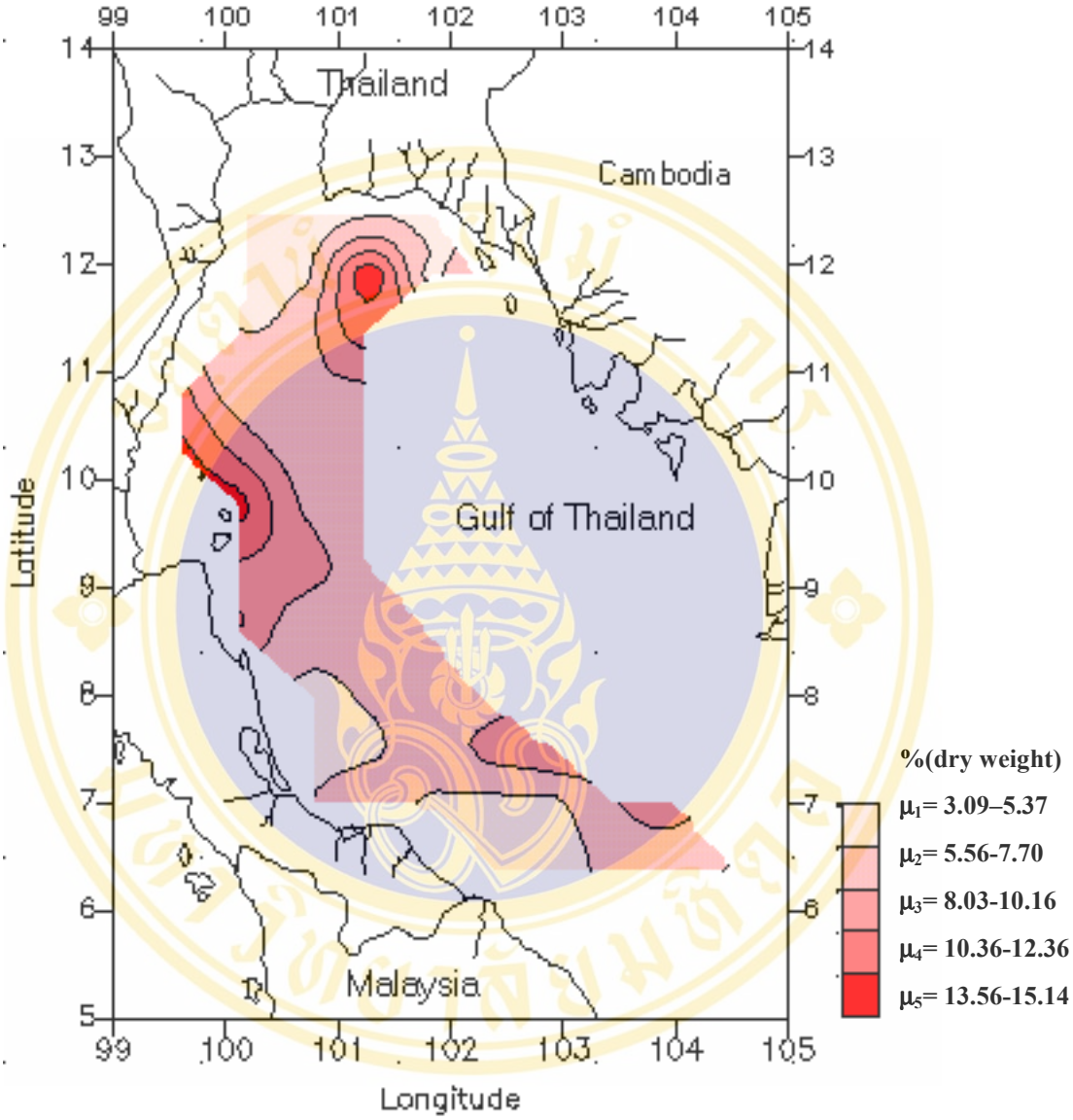


Figure 8. Map of the Total Organic Matter Contents of 2-4 Centimeter Bottom Sediments Classified into Component According to the Conventional Method

Copyright by Mahidol University

4.4.3 Conventional Method of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments

According to the Nonparametric Maximum Log-likelihood in Mixture Model, at 4-6 centimeter, there were five components The cut of point for the five components of the 0-2 centimeter level by using percentile was illustrated in Table 18.

Table 18. The Interval Value and the Number of Station of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments in Each Component

Component	The interval of the percentile	The interval of the total organic matter contents (%dry weight)	Number of stations
1	1-20	2.75-5.06	18
2	21-40	5.07-7.37	13
3	41-60	7.38-9.69	11
4	61-80	9.70-12.00	5
5	81-100	12.01-14.31	3

Table 18 indicate the conventional method of 4-6 centimeter level was classified five components. The high amount of total organic matter component were classified three stations that they were 13.81 – 14.31%(dry weight). Five and eleven stations were classified into slightly high and medium amount of total organic matter component that they were 10.19 – 11.57%(dry weight) and 7.42 – 9.45%(dry weight), respectively. Thirteen and eighteen stations were classified into slightly low and low amount of total organic matter component. They were 5.14 - 7.20%(dry weight) and 2.75 – 5.02%(dry weight), respectively. The results of data were illustrated in Table 19.

A map of geographical distribution of the total organic matter of 4-6 centimeter bottom sediments in the Gulf of Thailand is illustrated in Figure 9.

Copyright by Mahidol University

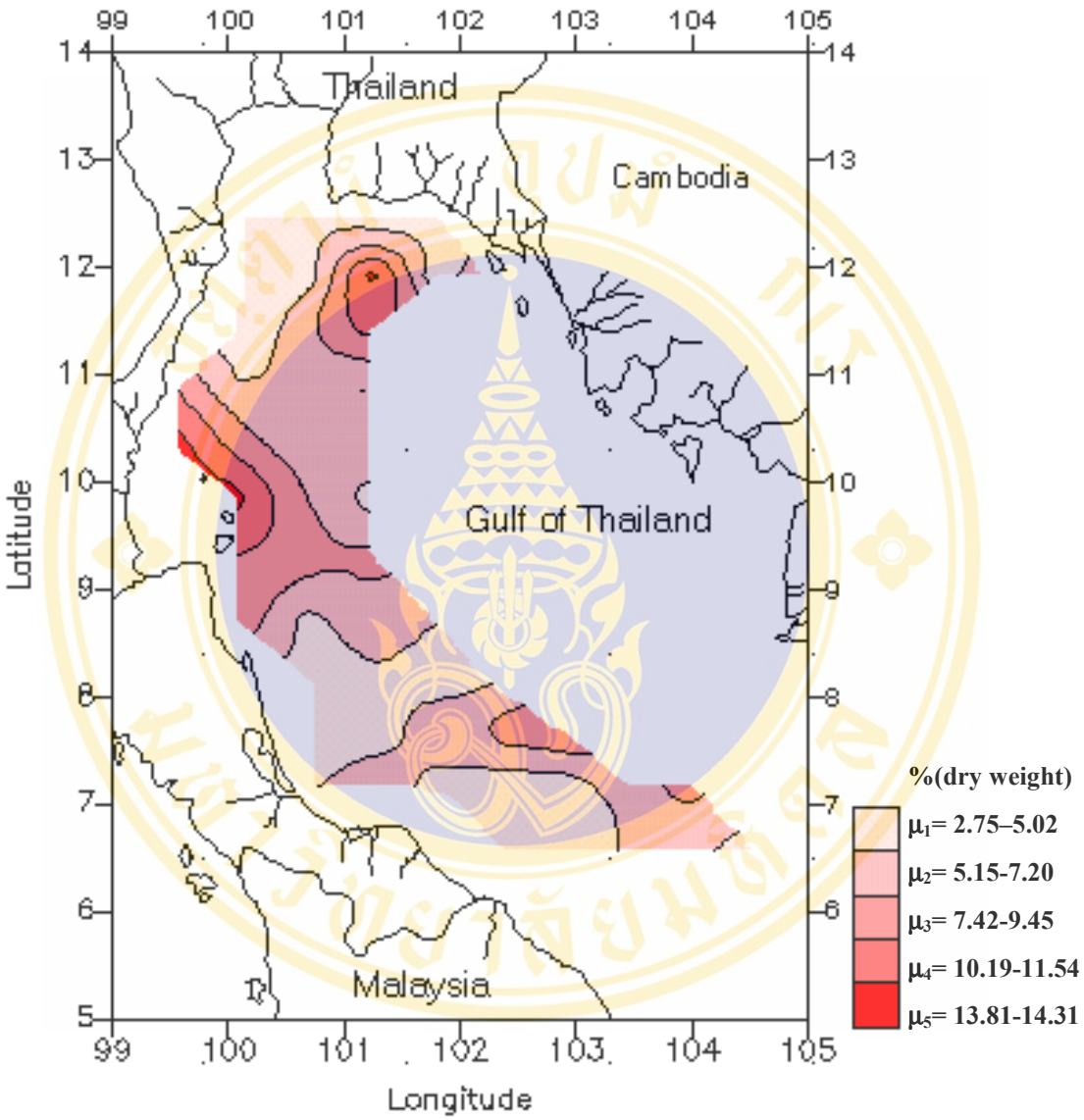


Figure 9. Map of the Total Organic Matter Contents of 4-6 Centimeter Bottom Sediments Classified into Component According to the Conventional Method

Copyright by Mahidol University

The results of the total organic matter contents for classified by conventional method were shown in Table 19.

Table 19. Classification of the Total Organic Matter Contents of the Three Levels of Bottom Sediments by Conventional Method

Stations	Classification of the total organic matter contents 0-2 cm of bottom sediments	Classification of the total organic matter contents 2-4 cm of bottom sediments	Classification of the total organic matter contents 4-6 cm of bottom sediments
1	1	1	1
2	1	1	1
3	2	1	1
4	1	1	1
5	3	3	3
6	2	1	1
7	4	5	5
8	1	1	1
9	1	1	1
10	2	2	1
11	2	2	3
12	4	4	4
13	3	2	2
14	3	2	2
15	2	2	2
16	4	4	4
17	5	5	5
18	3	3	3
19	2	2	2
20	2	3	3
21	2	2	1
22	2	2	3
23	5	5	5
24	3	3	4
25	3	3	3
26	2	2	3
27	2	2	2
28	2	2	2
29	2	2	2
30	4	4	4
31	1	1	1
32	2	2	2
33	2	2	1
34	2	2	2

Copyright by Mahidol University

Table 19. Classification of the Total Organic Matter Contents of the Three Levels of Bottom Sediments by Conventional Method (continue)

Stations	Classification of the total organic matter contents 0-2 cm of bottom sediments	Classification of the total organic matter contents 2-4 cm of bottom sediments	Classification of the total organic matter contents 4-6 cm of bottom sediments
35	4	4	4
36	4	3	3
37	3	2	3
38	1	1	2
39	1	1	1
40	2	2	2
41	3	2	3
42	2	1	1
43	1	1	1
44	2	1	1
45	3	2	-
46	3	3	3
47	2	1	1
48	2	2	2
49	2	2	2
50	1	1	1
51	1	1	1

Copyright by Mahidol University

CHAPTER V
DISCUSSION

5.1 The Mixture Model Method and Conventional Method

The Mixture Model in this study presents the spatial variation of the normal distribution with unknown components and variances (meaning to be estimated). The initial component consists of 7 groups that were classified by the maximization of likelihoods of the VEM method. Either we fix the number of components and then estimate the common variance parameters in addition to the means and weights, or we fix the variance parameters and estimate the number of components. We used both methods for estimating the parameters by estimating NPMLE with C.A.MAN. Then the non-parametric maximum likelihood estimation is used with the EM algorithm approach to estimate all parameters. The number of components was tested by the likelihood ratio statistic test and the classification of stations was compared under the posterior probability of Bay's approach. After we used C.A.MAN program, we classify 5 components of three levels of bottom sediments. This result demonstrates that the nature of the distribution of the total organic matte content of bottom sediments consists of five subpopulations.

The conventional method is a tradition approach to categorization. The disadvantages are the classification that is reliable or does not depend on the experiences of the researchers or the experts. The classified stations are not significant by the scientific methods such as a statistical inference. The summaries of two methods are as follows :

The result of Mixture Model method in Table 6 and conventional method in Table 19 of the total organic matter contents of 0-2 centimeter bottom sediments had differed 3 stations (5.88 %). The station number 22 and 34 changes from component 3 (Mixture Model method) to component 2 (conventional method). The station number 31 changes from component 2 to component 1.

Copyright by Mahidol University

The result of Mixture Model method in Table 9 and conventional method in Table 19 of the total organic matter contents of 2-4 centimeter bottom sediments had differed 25 stations (49.01 %).

The result of Mixture Model method in Table 12 and conventional method in Table 19 of the total organic matter contents of 4-6 centimeter bottom sediments had differed 30 stations (58.82 %).

The different stations of the total organic matter contents of sediments that use Mixture Model method and conventional method were closely the data values of 0-2 centimeter but in other levels were different. We estimated the variance in each component that was shown in Table 20.

Table 20. The Variance Values of the Total Organic Matter Contents in Each Component

Level (centimeter)	Method	Variance values in each component				
		1	2	3	4	5
0-2	Mixture Model	0.364	0.489	0.596	0.736	0.396
	Conventional Method	0.419	0.573	0.525	0.736	0.396
2-4	Mixture Model	0.228	0.326	0.261	0.630	1.357
	Conventional Method	0.441	0.628	0.514	0.734	0.634
4-6	Mixture Model	0.120	0.330	0.375	1.184	0.068
	Conventional Method	0.378	0.339	0.482	0.305	0.068

Table 20 indicated the variance values of Mixture Model which were less than the conventional method in each level and component. The classification of Mixture Model Method can be displayed appropriately than the conventional method, since the conventional method is not the statistical significance.

5.2 The Total Organic Matter Contents of Bottom Sediments Mapping

The mapping of the total organic matter contents of 0-2, 2-4 and 4-6 centimeter bottom sediments in the Gulf of Thailand were used by Mixture Model method and conventional method. It can be contained of 5 components. The total organic matter contents were 3.05 - 15.35 %(dry weight). The low amount of the total organic matter

Copyright by Mahidol University

contents appeared in the upper (station number 1 and 2) and lower (station number 50 and 51) of the Gulf of Thailand. The bottom sediments were particles with large grain size e.g. sand and gravel (47). The high amount of total organic matter dispersed in the western part of the Gulf, especially at the near-shore Samui Island (station number 17, 23 and 24). These areas are distinguishably altered by human activities. The total organic matter had entered the water through atmospheric deposition and runoff from the drainage basin of direct discharge into the water. The bottom sediments were mud and clay.

The total organic matter contents were changed in the central part of the Gulf, that it was related with types of sediment. The bottom sediments were particles with large grain size, such as sand and gravel. Amount of the total organic matter contents were less than the particles with fine grain size e.g. silt, clay and mud. The particles of silt and clay can be combined of higher of the total organic matter content than the particles of sandy and water (59). Chuchart (60) presented the total organic matter contents of bottom sediments in the Ping River that were depended on the grain size in bottom. The bottom deposits containing great compositions of fine particles like mud deposits show distinguished higher levels of the total organic matter content than those containing the sandy particles.

Distribution of sediment deposits was suggested to be depended upon the distribution of different velocities of water masses and currents (44) which were preliminary influenced by wind velocity and direction and topography of the areas. Chareonpanich and Seurungreong (47) reported some former studies showing that the velocity of water movement in the central Gulf seems to be weak even during the monsoon season. Some illustrate the oceanic fronts of two different water masses occurring in the water off Samui Island. These phenomena are well responded for the pattern of organic accumulation in those areas. Therefore, the factors of the distribution of the total organic matter content were grain size, deposition of the total organic matter, geological topography and current water (9).

Waleerat et al. (61) represent the daily primary production at 50 meter in $mgC/m^3/day$, at station number 7 to 18 and station number 33 to 36. It was higher than at other areas witch the estimation of daily primary production was greatly due to the chlorophyll-a concentration. These stations were the total organic matter content in

Copyright by Mahidol University

medium and high components (component 3 to component 5). The variability in daily primary production is closely related to change with the total organic matter content.

Windom et al. (23) reported the sedimentation in upper areas of the Gulf that are approximately 0.3 (mid of the Gulf) to 15 (near shore) millimeter per year (average 7.5 millimeter per year). Therefore, we are approximately 15 years time periods which the sedimentation rate is approximately 2 centimeter (equal to level samples of sediment in this study). In summary, these studied indicate that 0 to 2 centimeter, 2 to 4 centimeter and 4 to 6 centimeter bottom sediment has been time periods of the sedimentation are approximately 15 years (1981 to 1996), 30 years (1965 to 1980) and 45 years (1949 to 1964), respectively. The time periods 45 years, the trend of the total organic matter content in the Gulf of Thailand seems to show an decrease at central of the Gulf, while at other areas, it seems to show a constant.

Comparatively high densities of suspended particles, together with phytoplankton, zooplankton or small fish larvae, etc. could be reasonably observed in the water column. Accordingly, several kinds of phytoplankton feeding fish may aggregate due to the comparatively high density of living phytoplankton. In this aspect, Takahashi et at (44) have revealed that the planktonic organisms and fishes apparently congregate in the zone near Samui Island at the optimum temperature and salinity. Fishes, particularly demersal species, are attracted to the areas of weak movement of water and they avoid places where strong currents prevail. Therefore, such typical pattern of particle accumulation (related to the total organic matter content) of the bottom sediment in such stagnant area can illustrate the mode of biological fishery resources in the water column as well. In conclusion, the physical and chemical characteristics of bottom sediments of organic accumulation have been revealed to provide appropriate aquatic conditions which are able to promote the growth and successive development of several kinds of organisms from the primary producers to those in the trophic levels.

Romig (62) reported the soil health properties. It was derived from farmer interviews with the permission of the Soil and Water Conservation Society of America that organic matter is as high as possible, at soil's potential is more than 3%. The organic matter contents in the Gulf were 3.05 - 15.35 %(dry weight). Therefore, they

Copyright by Mahidol University

had good health properties. Thus we can used these results for basic data to plan and develop the natural resources in the Gulf of Thailand.

Copyright by Mahidol University

CHAPTER VI
CONCLUSION AND RECOMMENDATIONS

6.1 Conclusion

The study is presented to apply the Nonparametric Mixture Model for identifying the spatial distribution of the total organic matter contents in the Gulf of Thailand. The estimation of parameters is conducted by using the Nonparametric Maximum Likelihood of mixing distribution by C.A.MAN program. The determination of the number of components can be solved by computing the Log-likelihood Ratio Statistics for the testing hypothesis. The classification of regions for mapping construction can be accomplished by the means of the posterior probability of Bayes' approach. The another method is the conventional method by using the percentile. The results of the distribution of the total organic matter content in the Gulf of Thailand can be summarized as follows:

6.1.1 The Distribution of the Organic Matter Contents in the Gulf of Thailand by Using the Mixture Model Method

The results from the investigation of the total organic matter contents of 0-2 centimeter bottom sediments show that the level of the amount of total organic matter consists of five components. It can be summarized as follows:

- The high amount of the total organic matter component : station numbers 17 and 23 (near Samui Island) were 4.04% of the whole areas. The average value were 14.81%(dry weight).
- The slightly high amount of the total organic matter component : six stations were 12.04% of the whole areas. The average value were 11.27%(dry weight).
- The medium amount of the total organic matter component : twelve stations were 21.29% of the whole areas. The average value were 8.93%(dry weight).

Copyright by Mahidol University

- The majority was classified into slightly low of the amount of total organic matter component : twenty-one stations (center of the Gulf) that were 42.23% of the whole areas. The average value were 6.45%(dry weight).
- The low amount of the total organic matter component : ten stations (upper and lower of the gulf) were 20.41% of the whole areas. The average value were 4.37% (dry weight).

The results from the investigation of the total organic matter contents of 2-4 centimeter bottom sediments show that the level of the amount of total organic matter consists of five components. It can be summarized as follows:

- The high amount of the total organic matter component : station numbers 7, 12, 17 and 23 were 7.77% of the whole areas. The average value were 13.71%(dry weight).
- The slightly high amount of the total organic matter component : seven stations were 13.59% of the whole areas. The average value were 9.83%(dry weight).
- The medium amount of the total organic matter component : thirteen stations were 25.49% of the whole areas. The average value were 7.72%(dry weight).
- The majority was classified into slightly low of the amount of total organic matter component : twenty-three stations (center of the Gulf) that were 44.25% of the whole areas. The average value were 5.30%(dry weight).
- The low amount of the total organic matter component : stations numbers 1, 2, 9 and 50 were 8.90% of the whole areas. The average value were 3.75%(dry weight).

The results from the investigation of the total organic matter contents of 4-6 centimeter bottom sediments show that the level of the amount of total organic matter consists of five components. It can be summarized as follows:

- The high amount of the total organic matter component : station numbers 7, 17 and 23 were 6.39% of the whole areas. The average value were 13.82%(dry weight).
- The slightly high amount of the total organic matter component : five stations were 24.89% of the whole areas. The average value were 9.27%(dry weight).

Copyright by Mahidol University

- The medium amount of the total organic matter component : eleven stations were 20.57% of the whole areas. The average value were 6.76%(dry weight).
- The majority was classified into slightly low of the amount of total organic matter component : twenty-three stations (center of the gulf) that were 44.11% of the whole areas. The average value were 4.86%(dry weight).
- The low amount of the total organic matter component : station numbers 1 and 51 were 4.03% of the whole areas. The average value were 3.04%(dry weight).

The results from the investigation of the total organic matter contents of 0-6 centimeter bottom sediments show that the level of the amount of total organic matter consists of six components. It can be summarized as follows:

- The high amount of the total organic matter component was station numbers 17 and 23 (near Samui Island).
- The twenty-eight stations (about 54.9%) were stable of the allocation of the total organic matter contents.
- The station numbers 7, 20 and 26 decreased form the bottom level to the surface level, whereas sixteen stations increased from the bottom level to the surface level. There are four stations that fluctuated into each level.

6.1.2 The Distribution of the Organic Matter Contents in the Gulf of Thailand by Using the Conventional Method

The total organic matter contents of 0-2 centimeter bottom sediments result consists of five components. Then we classify it by using percentile. It can be summarized as follows:

- The high amount of the total organic matter component : station number 17 and 23 (near Samui Island) were 3.92% of the whole areas. The average value were 14.90%(dry weight).
- The slightly high amount of the total organic matter component : six stations were 11.76% of the whole areas. The average value were 11.48%(dry weight).
- The medium amount of the total organic matter component : ten stations were 25.49% of the whole areas. The average value were 8.96%(dry weight).

Copyright by Mahidol University

- The majority was classified into slightly low of the amount of total organic matter component : twenty-two stations (center of the gulf) that were 43.14% of the whole areas. The average value were 6.65%(dry weight).
- The low amount of the total organic matter component : eleven stations (upper and lower of the Gulf) were 21.56% of the whole areas. The average value were 4.46%(dry weight).

The total organic matter contents of 2-4 centimeter bottom sediments result consists of five components. Then we classify it by using percentile. It can be summarized as follows:

- The high amount of the total organic matter component : three stations were 5.88% of the whole areas. The average value were 14.29%(dry weight).
- The slightly high amount of the total organic matter component : four stations were 7.84% of the whole areas. The average value were 11.17%(dry weight).
- The medium amount of the total organic matter component : seven stations were 13.73% of the whole areas. The average value were 9.03%(dry weight).
- The majority was classified into slightly low of the amount of total organic matter component : twenty-one stations that were 41.18% of the whole areas. The average value were 6.61%(dry weight).
- The low amount of the total organic matter component : sixteen stations were 31.37% of the whole areas. The average value were 4.60%(dry weight).

The total organic matter contents of 4-6 centimeter bottom sediments result consists of five components. Then we classify it by using percentile. It can be summarized as follows:

- The high amount of the total organic matter component : station numbers 7, 17 and 23 were 5.88% of the whole areas. The average value were 14.02%(dry weight).
- The slightly high amount of the total organic matter component : five stations were 9.80% of the whole areas. The average value were 10.60%(dry weight).
- The medium amount of the total organic matter component : eleven stations were 21.57% of the whole areas. The average value were 8.38%(dry weight).

Copyright by Mahidol University

- The slightly low amount of the total organic matter component : thirteen stations were 25.49% of the whole areas. The average value were 6.13%(dry weight).
- The majority was classified into low of the amount of total organic matter component: eighteen stations that were 35.29% of the whole areas. The average value were 4.50%(dry weight).

6.2 Recommendations

6.2.1 The purpose of this research is to collect updated sediment information concerning on the marine fishery resources and the present oceanographic conditions in the sea areas for the right development of fishery resource schemes of the sea areas.

6.2.2 The further studies could be focused on the other application to construct environmental maps for displaying the distribution of other interested parameters.

6.2.3 The problem with C.A.MAN program is the number values of the data restrict of a maximum of 500 data points. Some parameters (especial normal distribution) are very large sample sizes. Therefore, the limitation is only a problem for continuous quantities. In that case we recommend grouping the observations into very small intervals and using the replication option (1).

6.2.4 It is suggested to use other environmental mapping programs to confirm the results e.g.

- The program EMMIX was developed as a general tool to fit mixtures of normal of t components by maximum likelihood via the EM algorithm to continuous multivariate data (http://www.maths.uq.edu.au/~gjm/emmix/emmix.html.) (63).
- The program MCLUST/EMCLUST is a software package for hierarchical clustering on the basis of mixtures of normal components under various parameterizations of the component - covariance matrices. The EM algorithm is used in the fitting process, and BIC is used for the determination of the number of components (http://www.stat.washington.edu/fraley/software.html) (64).
- The program NORMIX fits mixtures of normals or Bernoulli distributions from specified initial values of the parameters or from initial partitions obtained by various hierarchical clustering methods (http://www.alumnus.caltech.edu/~wolfe/) (65). etc.

Copyright by Mahidol University

REFERENCES

1. Böhning D. Computer-Assisted Analysis of Mixtures and Applications: meta-analysis, disease mapping and others. New York: Chapman & Hall/CRC; 1999.
2. Biernacki C, Celeux G, Govaert G. Assessing a mixture model for clustering with the integrated completed likelihood. IEEE Transactions on Pattern Analysis and Machine Intelligence 2000; 22(7): 719-25.
3. Chongprasith P, Srinetr V. Marine Water Quality and Pollution of the Gulf of Thailand. SEAPLO Intergrated Studies of the Gulf of Thailand Vol.1. Bongkok: Innomedia; 1998.
4. Siripong A. The causes of circulation in the Gulf of Thailand : Proceeding of the First ASEAMS Symposium on Southeast Asian Marine Science and Environmental Protection. UNEP Regional Seas Reports and Studies No.116; 1990. p.21-35.
5. Duursma EK, Dawson R. Marine Organic Chemistry. Amsterdam: Elsevier Scientific; 1981.
6. ยุพา ศรีนาค. ผลกระทบต่อคุณภาพน้ำ. ในการสัมมนาวิทยาศาสตร์สิ่งแวดล้อม เรื่องการพังทลายและการทับถมของดินตะกอนกับสิ่งแวดล้อม. โครงการบัณฑิตศึกษา สาขาวิทยาศาสตร์สิ่งแวดล้อม : บัณฑิตวิทยาลัย มหาวิทยาลัยเกษตรศาสตร์, กรุงเทพฯ ; 2527.
7. Mare MF. A study of a marine benthic community with special reference to the microorganism. J Mar Biol Ass 1942; 25: 517-54.
8. Alabaster JS. Water Quality Cliteria for Freshwater Fish. London: Butterworthy; 1980.

Copyright by Mahidol University

9. กฤษฎา หน่อเนื้อ. องค์ประกอบทางกายภาพและทางเคมีบางประการของดินตะกอนในอ่าวไทย [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์การประมง]. กรุงเทพฯ : คณะประมง มหาวิทยาลัยเกษตรศาสตร์ ; 2541.

10. Sumner ME. Handbook of Soil Science. Boca Raton: CRC Press; 2000.
11. Böhning D, Schlattmann P. Disease mapping with hidden structure using mixture models. In: Lawson AB et al, editors. Disease mapping and risk assessment for public health. New York: John Wiley & Sons; 1999. p.49-60.
12. McLachlan GL, Basford KE. Mixture model: inference and applications to clustering. New York: Marcel Dekker; 1998.
13. Laird N. Nonparametric maximum likelihood estimation of a mixing distribution. J Am Stat Assoc 1978; 73(364): 805-11.
14. Schlattmann P, Dietz E, Böhning D. Covariate adjusted mixture models and disease mapping with the program DISMAPWIN. Stat Med 1996; 15: 919-29.
15. Lesperance ML, Kalbfleish JD. An algorithm for computing the nonparametric MLE of a mixing distribution. J Am Stat Assoc 1992; 87(417): 120-6.
16. Schlattmann P, Böhning D. Computer packages C.A.MAN. (computer assisted mixture analysis) and Dismap. Stat Med 1993; 12: 1965.
17. สำนักนโยบายและแผนสิ่งแวดล้อม. รายงานสถานการณ์คุณภาพสิ่งแวดล้อม พ.ศ. 2542. ปทุมธานี : พี.พริ้นติ้ง กรุ๊ป ; 2543.
18. Suvapepun S. Long term ecological changes in The Gulf of Thailand. Marine Pullution Bulletin 1991; 23: 213-17.
19. ผุสดี ศรีพยัตต์. การเปลี่ยนแปลงแพลงค์ตอนสัตว์ในอ่าวไทย พ.ศ.2519-2537. เอกสารวิชาการ ฉบับ 4. กองสำรวจแหล่งประมง กรมประมง, กรุงเทพฯ; 2539.

Copyright by Mahidol University

20. Nasir MS, Rojana-anawat P, Snidvongs A. Physical Charcteristics of Watermass in the South China Sea, Area I: Gulf of Thailand and East Coast of Peninsular Malaysia. Proceedings of the First Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area I Gulf of Thailand and East Coast of Peninsular Malaysia, 24-26 November 1997, Bangkok, Thailand: 1999. p.1-5.
21. Snidvongs A, Vongprommek R. Tertiary and Quaternary sediments in Thai seas. In: Status of Marine Science and Oceanographic Research Chemical Oceanography and Geological Oceanography. Bangkok: National Research Council of Thailand; 1997. p.129-46.
22. มนูวดี หังสพฤกษ์. สมุทรศาสตร์เคมี. กรุงเทพฯ : ภาควิชาวิทยาศาสตร์ทางทะเล คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ; 2526.
23. Windom HL, Paulsen M, Hungspreugs M. Rate of sedimentation in the Upper Gulf of Thailand using the ^{210}Pb method. Bangkok: Research Faculty of Science, Chulalungkorn Univ 1982; 7: 58-65.
24. Snidvongs A. Sedimentary calcium carbonate dissolution of the net current in the Gulf of Thailand under different monsoon regimes. Report submitted in the SEAFDEC First Technical Seminar of the Interdepartmental Collaborative Research Program, 24-26 November 1997, Bangkok, Thailand. SEAFDEC/CRP/97/WP.10; 1993.
25. Srisuksawad S, Porntepkasemsan B, Nouchpramool S, Yamkate P, Carpenter R, Peterson ML, Hamiltons T. Radionuclide activities, geochemistry and accumulation rates of sediments in Gulf of Thailand. Continental Shelf Research 1997; 17(8): 925-65.
26. Black CA, Evans DD, White SL, Ensminger LE, Clark FF. Method of Soil Analysis. New York: American Society of Agronomy; 1965.
27. Jackson ML. Organic Matter Determinations for Soils in the Soil Chemical Analysis. New York: Prentile-Hall; 1958.
28. Susan M L. An Introduction to Marine Biogeochemistry. New York: John Wiley&Sons ; 1992. p.129-41.

Copyright by Mahidol University

29. Wenner S, Morgan JJ. Aquatic Chemistry: An Introduction Emphasizing Chemical Equilibrium in Natural Water. New York: John Wiley & Sons; 1981.

30. Pocklington R. Terrigenous organic matter in surface sediments from the Gulf of St. Lawrence. J Fish Res Board Can 1976; 33: 93-7.

31. เปี่ยมศักดิ์ เมนะเศวต. แหล่งน้ำกับปัญหามลภาวะ. กรุงเทพฯ : สำนักพิมพ์จุฬาลงกรณ์มหาวิทยาลัย; 2536.

32. Pierce RH, Felbeck GT. A comparison of three methods of extracting organic matter from soils and marine sediments, In Povoledo, D. and H.L. Golterman (eds.). Humic Substances, Their Structure and Function in the Biosphere. Wagenigen: Center for Agriculture Publishing and Documentaion; 1972. p.32.

33. Grissinger EH, McDowell LL. Sediment in relation to water quality. Water Resources Bulletin 1970; 6: 7-14.

34. Mayers PA, Takeuchi N. Environmental changes in Saginaw Bay, Lake Huron recorded by geolipid contents of sediment deposited since 1800. Environ Geol 1981; 3: 257-66.

35. พงศ์ศักดิ์ ลาภอุดมเลิศ, ประชุม สันทัดการ และเกษม จันทร์แก้ว. ความผันแปรของปริมาณสารอินทรีย์วัตถุภายหลังการแผ้วถางป่าดิบเขาดอยปุย เชียงใหม่. การวิจัยลุ่มน้ำที่ห้วยคอกม้า เล่มที่ 18. กรุงเทพฯ : ภาควิชาอนุรักษ์วิทยา คณะวนศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ ; 2517.

36. Hamza Erol. A practical method for construction the mixture model for a spectral class. International Journal of Remote Sensing 2000; 21: 823-830.

37. Feng ZD, McCulloch CE. On the likelihood ratio test statistic for the number of components in a normal mixture of unequal variances. Biometrics 1994; 50 (4): 1158-62.

38. McLachlan GJ, Peel D, Whiten WJ. Maximum likelihood clustering via normal mixture models. Signal Processing: Image Communication 1996; 8: 105-11.

Copyright by Mahidol University

39. Abbas HM, Fahmy MM. Neural networks for maximum likelihood clustering. Signal Processing 1994; 36: 111-26.

40. Thode HC, Finch SJ, Mendell NR. Simulated percentage points for the null distribution of the likelihood ratio test for a mixture of two normals. Biometrics 1988; 44: 1195-201.

41. แช่มช้อย ฐานพงษ์. สัตว์พื้นทะเลและสภาวะแวดล้อมพื้นทะเล บริเวณปากแม่น้ำเจ้าพระยา-ศรีราชา (ปี 2524-2525). เอกสารวิชาการฉบับที่ 19/2530, กองสำรวจแหล่งประมง กรมประมง; กรุงเทพฯ. 41 น.

42. จุมพล สงวนสิน. สัตว์พื้นทะเลบริเวณอ่าวไทยฝั่งตะวันออก(ช่องแสมสาร-ตราด). เอกสารวิชาการฉบับที่ 18/2532, ศูนย์พัฒนาประมงทะเล ฝั่งตะวันออก กองประมงทะเล กรมประมง, กระทรวงเกษตรและสหกรณ์ ; กรุงเทพฯ. 50 น.

43. วิมลรัตน์ เกษมทรัพย์ และณิฏฐารัตน์ ปภาวสิทธิ์. ปริมาณความจุรวมของซัลไฟด์ในดินตะกอนจากอ่าวไทยและทะเลอันดามัน. ในการสัมมนาวิทยาศาสตร์ทางทะเลแห่งชาติ ครั้งที่ 2 เรื่องนโยบายวิทยาศาสตร์ทางทะเลแห่งชาติ. กองโครงการประสานงานวิจัย สำนักงานคณะกรรมการวิจัยแห่งชาติ, กรุงเทพฯ; 2525. หน้า 278-89.

44. Takahashi K, Ruangsirakul N, Rassmee S. A comparative study on the fisheries oceanographic conditions in the central Gulf of Thailand. SEAFDEC Current Technical Paper, TD/CTP/31; 1984.

45. Kasemthab, V. Total organic matter and sulfide capacity of sediment from the Gulf of Thailand. [Master Thesis of Marine Science Department], Chulalongkorn University; 1981.

46. Sompop Rungsupa. Organic matter content in sediment and suspended solid around the eastcoast at the Upper Gulf of Thailand: 1991-1992. The 5th Seminar on Marine Fisheries Research. Eastern Marine Fisheries Development Center and Japan International Cooperation Agency: August 3-5, 1994.

Copyright by Mahidol University

47. Charumas Chareonpanich and Siriporn Seurungreong. Some Physical and Chemical Characteristics of Bottom Sediments in the South China Sea, Area I: Gulf of Thailand and East Coast of Peninsular Malaysia. Proceedings of the First Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area I Gulf of Thailand and East Coast of Peninsular Malaysia, 24-26 November 1997, Bangkok, Thailand: 1999. p.12 – 33.
48. Training Department of SAFDEC. Proceedings of the First Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area I Gulf of Thailand and East Coast of Peninsular Malaysia, 24-26 November 1997, Bangkok, Thailand: 1999.
49. Takashi Oguchia, Helen PJ, Colin Neal. River water quality in the Humber catchment: an introduction using GIS-based mapping and analysis .The Science of The Total Environment 2000; 251-252: 9-26.
50. Basford KE, McLachlan GJ, York MG. Modelling the distribution of stamp paper thickness via finite normal mixtures: The 1872 Hidalgo stamp issue of Maxico revisited; 1997.
51. Izenman AJ, Sommer CJ. Philatelic mixtures and multimodal densities. Journal of the American Statistical Associatin 1988; 83: 941-53.
52. Kevin Whitea, et al. Mapping the Iron Oxide Content of Dune Sands, Namib Sand Sea, Namibia, Using Landsat Thematic Mapper Data. Remote Sensing of Environment 1997; 62: 30-9.
53. Metternicht GI, Fermont A. Estimating Erosion Surface Features by Linear Mixture Modeling. Remote Sensing of Environment 1998; 64: 254-65.
54. Carpenter GA, Gopal S, Macomber S, Martens S, Woodcock CE. A Neural Network Method for Mixture Estimation for Vegetation Mapping. Remote Sens of Environ 1999; 70: 138-52.
55. Sutherlanda TF, Laneb PM, Amosc CL, Downingd J. The calibration of optical backscatter sensors for suspended sediment of varying darkness levels Marine Geology 2000; 162 (2-4): 587-97.
56. Brown M, Lewis HG, Gunn SR. Linear spectral mixture models and support vector machines for remote sensing. IEEE Transactions on Geoscience and Remote Sensing 2000; 38(5): 2346-60.

Copyright by Mahidol University

57. Stephan JM. Linear Mixture Modeling Approach for Estimating Cotton Canopy Ground Cover using Satellite Multispectral Imagery. Remote Sensing of Environment 2000; 72(3): 304-8.

58. Robertsa RG, Galbraithb RF, Yoshida H, Laslettc GM, Olleyd JM. Distinguishing dose populations in sediment mixtures: a test of single-grain optical dating procedures using mixtures of laboratory-dosed quartz. Radiation Measurements 2000; 32(5-6): 459-65.

59. Montira Piamthipmanus. Temporal Changes in the Abundance of Macrobenthos in the South China Sea, Area I: Gulf of Thailand and East Coast of Peninsular Malaysia. Proceedings of the First Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area I Gulf of Thailand and East Coast of Peninsular Malaysia, 24-26 November 1997, Bangkok, Thailand: 1999. p.156 - 71.

60. ชูชาติ หุตะเจริญ. การศึกษาปริมาณสารอินทรีย์ในน้ำและดินตะกอนบริเวณลุ่มน้ำแม่ปิง. [วิทยานิพนธ์ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาศาสตร์สิ่งแวดล้อม]. กรุงเทพฯ : มหาวิทยาลัยเกษตรศาสตร์ ; 2527.

61. Waleerat, et al. Primary Production Determination in the South China Sea, Area I: Gulf of Thailand and East Coast of Peninsular Malaysia. Proceedings of the First Technical Seminar on Marine Fishery Resources Survey in the South China Sea, Area I Gulf of Thailand and East Coast of Peninsular Malaysia, 24-26 November 1997, Bangkok, Thailand: 1999. p.135-146

62. Romig DE, Garlynd MJ, Harris RF, McSweeney K. How farmers assess soil health and quality. J. Soil Water Conserv 1995; 50: 229-36.

63. McLachlan GJ, Peel E, Basford KE, and Adams P. Fitting of mixtures of normal and t-components. Journal of Statistical Software 1999; 4(2).

64. Fraley C and Raftery AE. MCLUST: Software for model-based cluster analysis. Journal of Classification 1999; 16: 297-306.

65. Wolfe JH. NORMIX: Computational methods for estimating the parameters of multivariate normal mixtures of distributions. Research Memo. SRM 68-2. San Diego: U.S.Naval Personnel Research Activity; 1967.

Copyright by Mahidol University

APPENDIX

Copyright by Mahidol University

APPENDIX A : The VEM Algorithm and Stopping Rule

Let P and Q be two discrete probability distribution in Ω. Then, with $\alpha \in [0,1]$ the convex combination $(1-\alpha)P + \alpha Q$ is also in Ω.

VEM Algorithm

Let P_0 be any initial value, for example $P_0 = Q_{\bar{x}}$.

Step 1 : Find μ_{max} such that $D_{P_n}(\mu_{max}) = \sup_{\mu} D_{P_n}(\mu)$ and

μ_{min} such that $D_{P_n}(\mu_{min}) = \min\{D_{P_n}(\mu) | \mu \in$ support of P_n

Step 2 : Find α_{max} such that

$$I(P_n + \alpha_{max} P_n(\mu_{min})\{Q\mu_{max} - Q\mu_{min}\} = \sup_{\alpha} I(P_n + \alpha P_n(\mu_{min})\{Q\mu_{max} - Q\mu_{min}\}).$$

Step 3 : Set $P_{n+1} = P_n + \alpha_{max} P_n(\mu_{min})(Q\mu_{maz} - Q\mu_{min}\}$, n = n+1 and go to step 1.

This algorithm need to be stopped at some iteration n with associated Ψ_n close to $\hat{\Psi}$. The stopping rule for algorithm of VEM can be defined as :

Stopping rule of VEM

Fix $\varepsilon > 0$ to some small number such as 0.01, 0.001, 0.0001, or 0.00001. Stop algorithm if $D_{P_n}(\mu_{max}) < \varepsilon$

Copyright by Mahidol University

APPENDIX B : The EM Algorithm

The EM algorithm for mixture models is defined as following :

Stop 0 : Let p be any initial vector of weights and μ any initial vector of component parameters.

Step 1 : Compute the E-step.

Step 2 : Compute the M-step leading to $p^{(new)}$ and $\mu^{(new)}$.

Step 3 : Set $p = p^{(new)}$, and $\mu = \mu^{(new)}$ and go to step 1.

Copyright by Mahidol University

APPENDIX C : The Location of Studied Stations in the Gulf of Thailand

Station number	Latitude	Longitude
1	12_19.24N	100_16.40E
2	12_19.24N	100_46.00E
3	12_19.24N	101_15.30E
4	12_19.12N	101_44.54E
5	11_49.48N	102_14.12E
6	11_48.54N	101_45.60E
7	11_48.48N	101_12.36E
8	11_50.12N	100_45.42E
9	11_49.30N	100_14.54E
10	11_21.00N	100_15.18E
11	11_20.00N	100_46.32E
12	11_19.30N	101_15.30E
13	10_50.40N	101_16.18E
14	10_50.24N	100_44.54E
15	10_50.10N	100_16.50E
16	10_50.30N	99_46.70E
17	10_19.20N	99_45.30E
18	10_19.45N	101_15.58E
19	10_20.00N	100_46.31E
20	10_20.70N	101_16.70E
21	09_49.45N	101_16.40E
22	09_49.70N	100_44.50E
23	09_49.40N	100_15.32E
24	09_20.40N	100_15.59E
25	09_20.00N	100_45.55E
26	09_19.50N	100_15.40E
27	08_49.30N	100_45.36E
28	08_49.15N	101_15.40E
29	08_49.15N	100_44.50E
30	08_50.20N	100_15.50E
31	08_19.00N	100_45.30E
32	08_18.50N	101_16.80E
33	08_49.45N	101_46.00E
34	08_19.32N	102_14.25E
35	07_49.55N	102_44.40E
36	07_51.30N	102_14.50E
37	07_50.59N	101_45.20E
38	07_50.10N	102_15.35E
39	07_51.25N	101_44.37E
40	07_19.60N	100_44.22E

Copyright by Mahidol University

APPENDIX C : The Location of Studied Stations in the Gulf of Thailand (Continue)

Station number	Latitude	Longitude
41	07_19.30N	100_16.30E
42	07_20.45N	101_45.40E
43	07_20.50N	102_15.51E
44	07_19.30N	102_40.00E
45	07_20.10N	103_15.50E
46	07_18.50N	103_46.20E
47	06_50.20N	104_16.28E
48	06_48.43N	104_44.37E
49	06_48.20N	103_15.50E
50	06_48.30N	102_45.45E
51	06_50.36N	102_15.30E

Copyright by Mahidol University

APPENDIX D : The Data of Total Organic Matter Measurement

Total Organic Matter Measurement

1. Collection of sediment samples

The sediment samples were collected from 51 stations in the Gulf of Thailand by M.V. SEAFDEC during 24 April to 17 May 1996. Use the corer which can collect a cross section of sediments.

The corer consists of a hollow metal with ca 100 cm in length and ca 6 cm in diameter in which contains easily removable plastic liners (core tubes; 85 cm long and 5.2 cm inner diameter) which fit into a core barrel and retains the sediment sample. The piston mounted on the top of the core barrel was opened and allows water to flow through the barrel during descent, but shut upon penetration of the corer into the sediment. The tip was a core cutter for better penetration of the sampler and the stabilizing fins situated in the upper part to assure vertical descent of the corer.

During application, the gravity corer was lowered to the depth of about 2 m above the sediment surface. From that depth, it was let to sink freely to penetrate the bottom. At least two sedimental cores were sampled station. In addition, sediment grab was applied in stead of the corer in few stations in which the bottom was mainly composed of sand and gravels that was unable to retrieve by the corer.

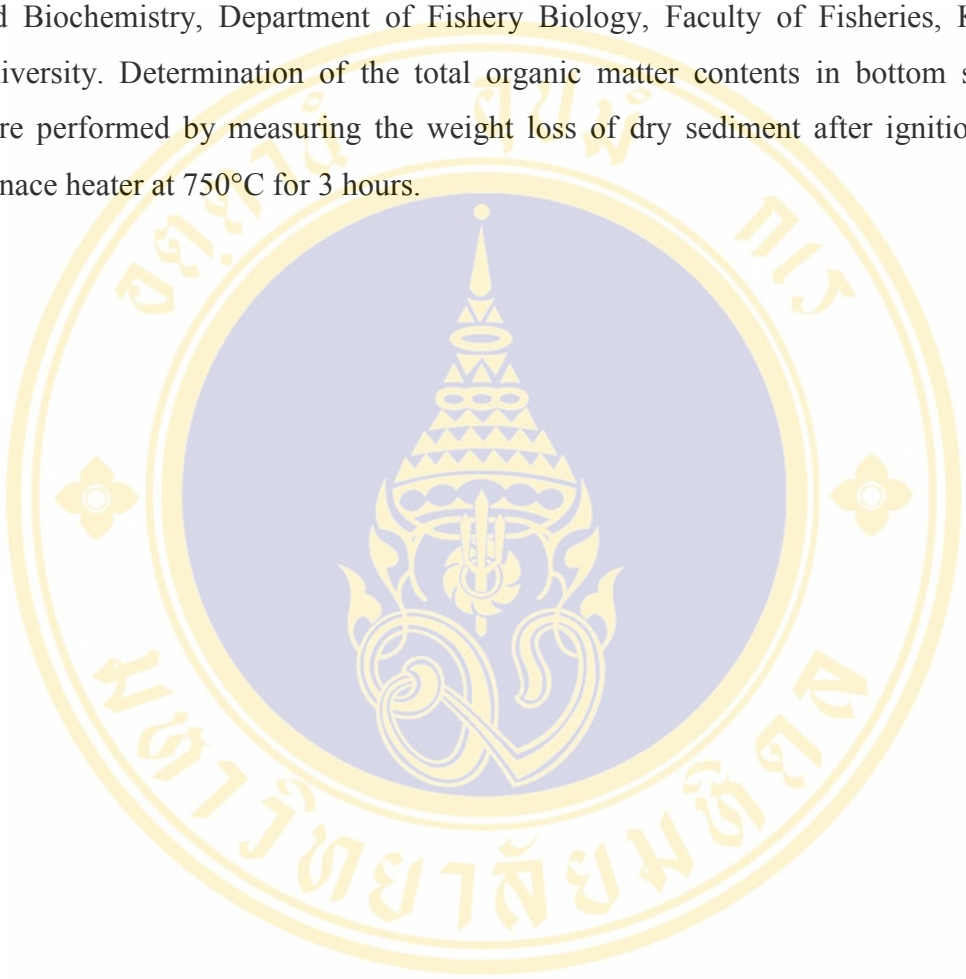
2. Subsampling of sedimental cores

The cores was subsampled as soon as possible after retrieval of the core. After the overlying water was calmly suctioned off, the sediment core was extruded out of the liner and subsampled into 2 cm sections for the upper 10 cm depth and 3 cm sections for the depth of more than 10 cm. The sectioned sediment samples were packed tightly in prepared containers and immediately frozen under - 60°C until further analysis.

Copyright by Mahidol University

3. Analysis of sediment samples

Laboratory analyses were performed in the Laboratory of Aquatic Environments and Biochemistry, Department of Fishery Biology, Faculty of Fisheries, Kasetsart University. Determination of the total organic matter contents in bottom sediment were performed by measuring the weight loss of dry sediment after ignition in the furnace heater at 750°C for 3 hours.



Copyright by Mahidol University

BIOGRAPHY

NAME	Mr. Apichart Thunyahan
DATE OF BIRTH	14 October 1977
PLACE OF BIRTH	Bangkok, Thailand
INSTITUTIONS ATTENDED	Kasetsart University, 1994-1998: Bachelor of Science (Fisheries) Mahidol University, 1999-2003: Master of Science (Biostatistics)
RESEARCH-GRANT	Support in part by the Thesis Grant, Faculty of Graduate Studies, Mahidol University
POSITION AND OFFICE	2002-Present, Inland Aquatic Animal Health Research Institute, Fisheries Department Position: Fisheries Scientist 3

Copyright by Mahidol University