Statistical Modeling of Aquatic Environmental Data Application to Macrobenthic Fauna Organism Counts

Uraiwan Sampantarak

A Thesis Submitted in Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in
Research Methodology
Prince of Songkla University
2010

Thesis Title Statistical Modeling of Aquatic Environmental Data
Application to Macrobenthic Fauna Organism Counts
Author Mrs. Uraiwan Sampantarak
Major Program Research Methodology

Major Advisor

Don McNeil
....

(Emeritus Prof. Dr. Don McNeil)

Co-advisor

Saowapa Ang
....

(Prof. Dr. Saowapa Angsupanich)

Examining Committee:

Yupadee Chaisuksant
..................................Chairperson

(Asst. Prof. Dr. Yupadee Chaisuksant)

Don McNeil
....

(Emeritus Prof. Dr. Don McNeil)

Saowapa Ang
....

(Prof. Dr. Saowapa Angsupanich)

....

(Asst. Prof. Dr. Rakpong Petkam)

The Graduate School, Prince of Songkla University, has approved this thesis as fulfillment of the requirements for the Doctor of Philosophy in Research Methodology.

....

(Assoc. Prof. Dr. Krerkchai Thongnoo)
Dean of Graduate School

ชื่อวิทยานิพนธ์ การสร้างตัวแบบทางสถิติของข้อมูลด้านสิ่งแวดล้อมทางน้ำประยุกต์กับข้อมูลแจงนับของสัตว์หน้าดิน
ผู้เขียน นางอุไรวรรณ สัมพันธารักษ์
สาขาวิชา วิธีวิทยาการวิจัย
ปีการศึกษา 2552

บทคัดย่อ

วิทยานิพนธ์ฉบับนี้เป็นการนำตัวแบบทางสถิติมาประยุกต์ใช้ในการวิเคราะห์ข้อมูลสัตว์หน้าดินที่มีแหล่งอาศัยบริเวณอ่าวไทย และทะเลสาบสงขลาตอนกลาง โดยแยกการศึกษาออกเป็น 2 การศึกษา สำหรับการศึกษาแรกเป็นการศึกษาความชุกของการเปลี่ยนเพศในหอยฝาเดียวในอ่าวไทย ปรากฏการณ์นี้ถูกใช้เป็นตัวบ่งชี้ทางชีวภาพที่สำคัญของสารมลพิษไตรบิวทิลทินที่ตกค้างในสิ่งแวดล้อมทางน้ำ ส่วนการศึกษาที่สองเป็นการศึกษาตัวแบบสำหรับใช้ประเมินความสัมพันธ์ระหว่างความชุกชุมของสัตว์หน้าดินขนาดใหญ่กับปัจจัยสิ่งแวดล้อมในทะเลสาบสงขลาตอนกลาง ซึ่งสัตว์หน้าดินเหล่านี้มีบทบาทสำคัญต่อการจัดการระบบนิเวศชายฝั่ง

การศึกษาที่ 1 การเปลี่ยนเพศของหอยฝาเดียวในอ่าวไทย จำนวน 22 ชนิด 5 วงศ์ ได้ถูกนำมาพิจารณาเพื่อเปรียบเทียบให้เห็นถึงความชุกของการเปลี่ยนเพศหอย ณ พื้นที่ต่าง ๆ หลังจากปรับลดอิทธิพลของความแปรปรวนของของตัวแปรที่เกิดจากชนิดหอยที่แตกต่างกัน ตัวแบบการถดถอยแบบลอจิสติก (Logistic regression model) เป็นเครื่องมือสำคัญที่ถูกนำมาใช้ในการวิเคราะห์ข้อมูลทางสถิติ จากการเปลี่ยนเพศในหอยบางชนิดสำรวจพบได้น้อยจนถึงไม่พบเลยในหลายพื้นที่ จึงได้พัฒนาตัวแบบเพื่อเป็นการชดเชยส่วนต่างของความเข้มข้นของสารทีบีทีของแต่ละตัวแปรให้อยู่ในระดับที่ใกล้เคียงกัน ผลการศึกษาพบว่า ระดับของสารพิษไตรบิวทิลทินในอ่าวไทยมีอยู่ในระดับสูง โดยอ้างอิงจากจำนวนการเปลี่ยนเพศในหอยเพศเมียที่สำรวจพบ โดยเฉพาะบริเวณอำเภอศรีราชา (71.9 ± 3.8 เปอร์เซ็นต์) และอำเภอพัทยา (48.4 ± 4.5 เปอร์เซ็นต์) จังหวัดชลบุรี นอกจากนี้เพื่อตอบข้อสงสัยที่ว่าชนิดหอยศึกษาที่แตกต่างกันอาจมีความไวต่อสารพิษไตรบิวทิลทินแตกต่างกันได้ถูกทำให้น่าเชื่อมากยิ่งขึ้น โดยพบว่า *Lataxiena blosvillei* (Deshayes, 1832) มีโอกาสเปลี่ยนเพศสูงถึง 69.7 ± 4.7 เปอร์เซ็นต์ รองลงมา ได้แก่ *Murex occa* Sowerby II, 1824 (46.4 ± 4.8 เปอร์เซ็นต์) *Nassarius siquijorensis* (A. Adams, 1852) (38.5 ± 6.7

เปอร์เซ็นต์) และ *Thais lacera* (Von Born, 1778) (32.2 ± 5.2 เปอร์เซ็นต์) ขณะที่ *Morula musiva* (Kiener, 1835) มีโอกาสเปลี่ยนเพศน้อยที่สุด (1.1 ± 1.6 เปอร์เซ็นต์)

สำหรับการศึกษาที่ 2 เป็นการศึกษาความชุกชุมของสัตว์หน้าดินขนาดใหญ่ในทะเลสาบสงขลาตอนกลางโดยมีวัตถุประสงค์เพื่อบ่งบอกถึงปัจจัยสิ่งแวดล้อมที่เกี่ยวข้อง ส่งผลให้มีความเข้าใจในสภาพปัจจุบันของทะเลสาบเพื่อรองรับการพัฒนาที่เหมาะสมต่อไป ข้อมูลที่ใช้ได้ถูกเก็บรวบรวมทุก 2 เดือน ตั้งแต่เดือนเมษายน 2548 ถึงกุมภาพันธ์ 2549 ณ จุดเก็บในทะเลสาบสงขลาตอนกลาง 9 จุด สืบเนื่องจากลักษณะเฉพาะของข้อมูล การวิเคราะห์ทางสถิติจึงใช้หลักการสร้างตัวแบบการถดถอยแบบพหุที่มีตัวแปรตามมากกว่าหนึ่ง (Multivariate multiple regression model) ร่วมกับการวิเคราะห์ปัจจัย (Factor analysis) เพื่อสกัดปัจจัยสิ่งแวดล้อมหลัก สำหรับตัวแปรตามของการศึกษานี้ ได้แก่ ปริมาณสัตว์หน้าดินที่มีความถี่ที่พบมากที่สุด จำนวน 24 วงศ์ จากทั้งหมด 81 วงศ์ แปลงข้อมูลโดยใช้ลอการิทึมฐานธรรมชาติ (Natural logarithms) ประโยชน์ของการนำตัวแบบนี้มาประยุกต์ใช้ ได้แก่ (1) สามารถอธิบายความสัมพันธ์ของตัวแปรทางด้านสิ่งแวดล้อมที่มีอิทธิพลต่อตัวแปรตาม โดยการกำหนดตัวแปรด้านสิ่งแวดล้อมต่าง ๆ ให้เป็นตัวแปรปัจจัยหลัก (2) ได้ค่าความคลาดเคลื่อนมาตรฐานที่ถูกต้องตามหลักสถิติ และ (3) ตัวแบบที่ได้มีความสามารถในการทำนายความสัมพันธ์ระหว่างปริมาณของสัตว์หน้าดินขนาดใหญ่และปัจจัยสิ่งแวดล้อม

แม้ว่าเครื่องมือที่ใช้ในการวิเคราะห์ข้อมูลทางสถิติในวิทยานิพนธ์ฉบับนี้ (ตัวแบบการถดถอยแบบลอจิสติก ตัวแบบการถดถอยแบบพหุที่มีตัวแปรตามมากกว่าหนึ่ง และการวิเคราะห์ปัจจัย) จะไม่ใช่วิธีการวิเคราะห์ที่ใหม่ แต่วิธีการเหล่านี้ได้ถูกนำมาประยุกต์ใช้เพื่อให้ได้ผลการวิเคราะห์ที่สามารถอธิบายเชื่อมโยงความสัมพันธ์ของสัตว์น้ำกับสิ่งแวดล้อมได้อย่างเหมาะสม และมีเหตุมีผลในเชิงนิเวศวิทยา อีกทั้งขั้นตอนในการวิเคราะห์ข้อมูลในแต่ละการศึกษายังง่ายต่อการทำความเข้าใจ อันจะเป็นประโยชน์สำหรับนักวิจัยทั้งหลายผู้ที่ต้องการสร้างสรรค์งานวิจัยอันมีคุณประโยชน์ต่อประเทศในอนาคตต่อไป

Thesis Title	Statistical Modeling of Aquatic Environmental Data Application to Macrobenthic Fauna Organism Counts
Author	Mrs. Uraiwan Sampantarak
Major Program	Research Methodology
Academic Year	2009

ABSTRACT

This thesis, statistical models were applied to analyse data of macrobenthic fauna. The thesis is based on two studies carried out in different habitats, coastal (Gulf of Thailand) and coastal lagoon (Middle Songkhla Lake). With respect to environmental linkages, imposex in gastropods is a significant biomarker of tributyltin (TBT) pollution whereas macrobenthic fauna abundance is important for coastal ecosystem management.

The first study, imposex in gastropods was considered to compare its prevalence in various species at different locations of the Gulf of Thailand after adjustant for species-dependent TBT-sensitivity. A logistic regression model was an important tool for this study because some species were in low numbers or absent at several areas. The model was developed to compensate for differences in TBT-concentrations for each determinant. The findings suggested that the TBT levels in the gulf were so high that imposex in female gastropods was recorded, there are, Si Racha (71.9 ± 3.8 percent) and Pattaya (48.4 ± 4.5 percent) in Chon Buri province. In addition, the suspicion that the local species may differ in sensitivity to TBT has been strengthened by the present study. *Lataxiena blosvillei* (Deshayes, 1832) was so high sensitivity prone to imposex (69.7 ± 4.7 percent), followed by *Murex occa* Sowerby II, 1824 (46.4 ± 4.8 percent); *Nassarius siquijorensis* (A. Adams, 1852) (38.5 ± 6.7 percent); and *Thais lacera* (Von Born, 1778) (32.2 ± 5.2 percent) whereas *Morula musiva* (Kiener, 1835) was the lowest sensitivity (1.1 ± 1.6 percent).

The second study was concerned with macrobenthic fauna abundances that aimed to specify the complex relational factors of environment for better understanding the current conditions and the suitable further developments in the Songkhla Lake. The data were obtained from nine sampling stations in Middle Songkhla Lake, Thailand at bimonthly intervals from April 1998 to February 1999. With respected to the data characteristics, the analysis based on a multivariate multiple regression model (MMR) involving factor analysis to extract the main environmental factors. The density outcome variables of 24 families of macrobenthic fauna selected as having the most coverage from 81 families which were observed at nine sampling stations during six bimonthly periods were transformed using natural logarithms. The advantages of the MMR application are that (1) it separates the effects of observed environmental variables on organism outcomes from unobserved factors; (2) it gives standard errors for these estimated effects, and thus provides a firmer statistical basis for clustering; and (3) it provides a predictive model for the outcomes.

Although, the main analysis methods (logistic regression, multivariate multiple regression, and factor analysis) are not new, they were applied suitably for analysing these datasets and gave reasonable results. Steps used for data analysis here could be useful guides for other scientists undertaking further research in aquatic environments.

Acknowledgements

I would like to express my sincere gratitude and deep appreciation to my supervisor, Emeritus Prof. Dr. Don McNeil, for his encouragement, guidance, and advice on this thesis. Thanks to my co-supervisor, Prof. Dr. Saowapa Angsupanich, for allowing me to use her data of macrobenthic fauna in the Middle Songkhla Lake and for giving me helpful suggestions throughout the completion of the thesis. I also wish to thank Dr. Cornelis Swennen for allowing me to use his data of imposex in gastropods in the Gulf of Thailand including share his experience for preparing the manuscript to submit in the international journal.

Thanks to Asst. Prof. Dr. Chamnein Choonpradub for her kindness, Dr. Phattrawan Tongkumchum for her encouragement, Dr. Nittaya McNeil for her comments and recommendations, Asst. Prof. Dr. Yupadee Chaisuksant and Asst. Prof. Dr. Rakpong Petkam for their work on the examining committee and giving me helpful comments. Thank you, Greig Rundle and David Broadfoot for editing my English writing. Thanks to all the staff of Prince of Songkla University for their cooperation, the Graduate School of Prince of Songkla University for supporting the scholarship. Thanks to my PhD friends for their encouragement and memorable learning experiences.

I am very grateful to Department of Fisheries, Ministry of Agricultural and Cooperatives, Thailand for allowing me to study for this PhD full-time. Finally, I would like to express my deepest appreciation to my mother, Jaree Chamnanwech, and my husband, Yuttakan Sampantarak, for their encouragement throughout my study.

Uraiwan Sampantarak

Contents

Page

Chapter

Contents (cont.)

Contents (cont.)

List of Tables

Table Page

List of Figures

Figure Page

List of Figures (cont.)

List of Appendices

Chapter 1

Introduction

1.1 Background and rationale

Macrobenthic fauna plays a significant role in the food chain in marine and aquatic ecosystems, acting as a link between the detritus deposited on the bottom, and the higher trophic level. These animals have limited mobility and many of them are unable to avoid adverse conditions brought about by natural stressors or human impacts. With their life span exceeding two years, they are useful indicators of environmental factors. They are sensitive indicators of environmental conditions (Bergen et al 2000, Dauer et al 2000, Summers 2001, Hyland et al 2003) and are used in integrated environmental monitoring where chemical data, toxicity testing and detailed ecology are measured, to provide a complete summary of the health of the ecosystem being investigated (Hagger and Galloway 2009).

Detection of imposex in gastropods is one such indicator. The occurrence of imposex in gastropods has been recognized as a very important biomarker of tributyltin (TBT) contamination used for monitoring on a global scale (Fent 2005). This is a serious issue in many countries world-wide because over 200 gastropods species have been shown to be affected by imposex (Shi et al 2005).

The aim of the study of imposex in gastropods in the Gulf of Thailand is to see how the prevalence of imposex varies by location in the Gulf, and also the sensitivity of each gastropod species. But in this survey, no species were found in every study area.

This made it difficult to use simple data analysis tools to compare the data. With respect to this limitation, a contamination model was used for analysis. The results indicated the current status of imposex in gastropods for each area in the Gulf, including which species were most prone to imposex.

For ecosystem management purposes, it is very important to know the environmental status in order to be able to detect any changes. Benthic animal activity has a wide influence on the physico-chemical conditions in the sediment. In addition, its diversity has a corollary for the functioning of the aquatic ecosystem. In our second study, macrobenthic fauna abundance in the Middle Songkhla Lake was studied with respect to both spatial and seasonal patterns.

In the Songkhla Lake, quantitative studies have been conducted including abundance, diversity, distributions and variations (Predalumpaburt and La-ongsiriwong 1997, Angsupanich and Siripech 2001, Ruensirikul et al 2007). However, due attention has not been given to the environmental factors affecting organism abundance. So this work has been done to study how efficient data analysis methods can be used to extract relevant new information and thus better understand the main processes affecting the current condition and possible future developments as well.

1.2 Definition of terms

Gastropods: Wallace (2009) gave the following definition, “A gastropod is a single valve (shell), soft-bodied animal belonging to the mollusk phylum. Gastropods, which are also known as univalves or gasteropods, are the largest class belonging to the mollusk family. Estimates of how many species of gastropod are alive today range from 65,000 to 90,000.”

Macrobenthic fauna: Vitaliano et al (2006) defined macrobenthic fauna as "invertebrates living in or on the sediments that are quantitatively sampled by a 0.1 m^2 Smith McIntyre grab and are retained on a 1.0 mm or 0.5 mm sieve."

1.3 Review of literature

1.3.1 Imposex in gastropods

The phenomenon of induced male sex characteristics superimposed on normal female gastropods was first named "Masculinization" or "Pseudo-hermaphroditism" and soon became commonly known as "Imposex" (Swennen et al 2009). The first observation was in the early 1970s for *Nucella lapillus* found along the coastline of the United Kingdom (Gibbs et al 1987). The occurrence of imposex correlated significantly with TBT levels (Hagger et al 2006) which are elevated close to harbour (Blaber 1970) and shipping traffic (Vos et al 2000).

Additional research showed that these morphological changes were caused by TBT (Smith 1981, Gibbs and Bryan 1986, Oehlmann et al 1993, Mensink et al 1996), and was confirmed by long-term field and laboratory experiments (Nishikawa 2006).

TBT is defined as an endocrine-disrupting or neurotoxic chemical reagent (Nakatsu et al 2006). It has been used in marine paints in order to control fouling by organisms on ships' hulls. It is leached from the surface of the paint and thus prevents attachment (Law et al 2005).

de Mora et al (1995) indicated that TBT degradation occurred with first order kinetics and that TBT in marina sediments had a half-life of about 2.5 years. In addition, TBT degradation in sediment is very slow under anaerobic conditions and remobilization occurs via mixing and dredging (Hagger et al 2006).

Bryan et al (1989) studied effects of TBT pollution on the mud snail, *llyanassa obsoleta*, from the York River and Sarah Creek, Chesapeake Bay and reported that the exposure to trace levels of 1-2 ng l^{-1} induce imposex.

Gibbs et al (1987) described six successive stages of imposex in *Nucella lapillus.* Stage 1 is formation of a deferens between the penis and the prostate, penis development stage 2 to 4, block-age of the oviduct stage 5 and 6.

Several mechanisms have been suggested to explain imposex induction. One, that TBT increases androgen levels by inhibiting the enzyme activity that metabolizes testosterone. Two, that TBT acts as a neurotoxin to abnormally release the peptide hormone penis morphogenic factor (Féral and Le Gall 1983). However, this has not yet been confirmed. There must be something else that directly interacts with TBT in the initial step of imposex induction.

Determining the occurrence of TBT in seawater is a complicated chemical procedure. However, if TBT pollution has occurred it can easily been shown by biological indicator. Even in very low concentrations, TBT creates a male/female hormonal imbalance in many organisms, which generates abnormalities in the reproductive organs, and induces change in behavior, growth, mortality, and sex ratio shifts in favour of males (Blaber 1970, Bryan et al 1987, Swennen et al 2001, Hagger et al 2006).

Imposex has been investigated in Southern Thailand, along both the western and eastern coast, since 1996 (Bech 1999, Swennen et al 1997). Several species had a high imposex incidence at several areas in the Gulf of Thailand (Swennen et al 1997).

Bech (2002) reported that imposex in *Thais distinguenda* from the east coast of Phuket Island increased significantly from 2002 to 2004 (p-value < 0.001: Chi Square test for trend). Increasing intensity of imposex was also observed for *Thais bitubercularis.* These two species are recommended as indicators of TBT in Southeast Asia because of their sensitivity and wide distribution in the region.

Kan-Atireklap et al (1997) surveyed butyltin compounds (BTs) including TBT, dibutyltin (DBT) and monobutyltin (MBT) in sediments from coastal water of Thailand in 1995. The finding suggested that the major TBT pollution in Thailand probably originated from antifouling paints used mainly on the hulls of far seas commercial vessels.

1.3.2 Macrobenthic fauna variation

Macrobenthic fauna are recognized as sensitive indicators of environmental disturbance (Pearson and Rosenberg 1978, Rygg 1985, Engle et al 1994, Weisberg et al 1997, Borja et al 2000, Ranasinghe et al 2004). The observed distribution of macrobenthic fauna has been useful in diagnostic studies and environmental monitoring (Warwick 1986).

Tropical areas of the world have some features that differ from temperate regions, where most of the techniques for assessing the biological effects of pollution on communities have been developed (Warwick and Clarke 1995).

Nizzoli et al (2002) studied the sediment characteristics and evaluated the effect of the dominant macrobenthos on sediment respiration, benthic nitrogen exchanges and nitrification rates in two sheltered sites located at opposite poles of the Sacca di Goro, a coastal lagoon of the Pole River delta. The results showed that dissolved inorganic nitrogen fluxes correlated with the biomass of the dominant species. Ammonium fluxes were significantly correlated with the biomass of macrofauna.

Armah et al (2005) studied macrofauna variation in the south-western part of the Keta Lagoon, Ghana. The physico-chemical parameter distributions were found to exhibit a high degree of similarity, but their spatio-temporal spread was mainly influenced by salinity and turbidity, which again are mainly under the influence of water depth, which by itself is probably a function of water discharge into the lagoon from streams feeding the lagoon.

Various statistical methods are useful for clustering according to patterns of variation in space and time (Hawkins et al 2000, Joy and Death 2000, Frédou et al 2006). To describe patterns of species abundance, multivariate analytical techniques such as cluster analysis and dendrograms based on similarity matrices, multidimensional scaling, and correspondence analysis, have been used extensively. Numerous benthic metrics are commonly used in assessments, and impacts to any single metric could be due to combinations of chemical and physical water characteristic parameters, such as the gradient of salinity from the sea to the river as well as the often associated sedimentary changes or variations in turbidity, including sediment physico-chemical features, making their identification complicated (Melwani and Thompson 2007).

Clarke and Warwick (1994) outlined the basic methods now commonly used by biological scientists for analysis of their data. For descriptive studies, these methods include data transformation using square roots, fourth roots, or logarithms (after adding 1 to cell counts or densities to handle zeros) to remove skewness, principal components analysis of covariance matrices, and ordination procedures to cluster taxa in space and time, as well as more complex multivariate analytical techniques such as dendrograms based on similarity matrices and multidimensional scaling. Measures of association in assemblage data such as the Bray-Curtis similarity index are preferred to Pearson correlation coefficients "for sound biological reasons" (Clarke et al 2006), but such measures do not satisfy the positive-definiteness assumptions that underpin conventional multivariate statistical analysis. Furthermore, the traditional methods used by biologists to correlate taxa abundances with environmental determinants do not generally provide standard errors for estimated parameters.

1.3.3 The Gulf of Thailand

The Gulf of Thailand is a semi-enclosed tropical sea located between UTM 520000E and 1130000E in the west-east direction and between UTM 1500000N and 680000N in the north-south direction (Figure 1.1).

The Gulf of Thailand is a gulf that borders, but is not part of the South China Sea. The average depth is 45 m and the maximum depth is 80 m. The overall area is approximately 320,000 km^2. The mouth of the Gulf is defined by a line connecting the Cape of Camau (the southernmost tip of Vietnam) and the coastal town of Tumpat in north Malaysia - near the Thailand-Malaysia border (Johnston 1998).

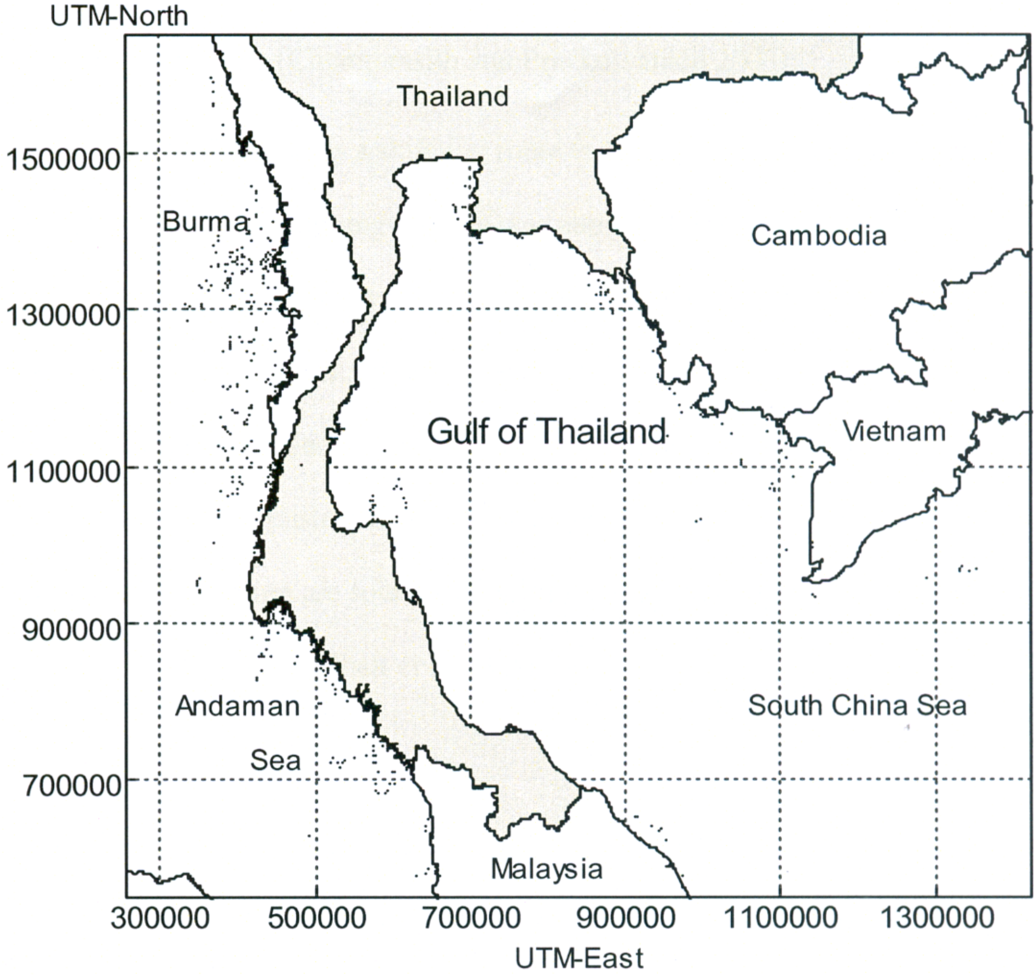

Figure 1.1: The Gulf of Thailand

The Gulf of Thailand is one of the most productive regions of the world and thus marine fisheries are an important industry for the Gulf's coastal countries. Marine fishery products have long been a significant percentage of the diet for people in this region (UNESCO 1997).

The pollutants in the Gulf come from either land-based or sea-based activities and from both point sources, such as industrial discharge, oil spill incidents, and domestic sewage, and non-point sources like agriculture runoff, all of which affect coastal water quality, marine sediment conditions, and particular organisms, mangrove, seagrass and coral reefs (Chongprasith and Praekulvanich 2003).

Using the Gulf of Thailand as a shipping route is of immense value to the coastal states. For example, the shipping volume was about 140 million tons of cargo and 2.4 million TEUs of container by over 5,000 vessels called at Thai ports in 1997. These numbers were predicted to increase by about 12% annually (UNESCO 1997). In 2004, the Gulf ports handled more than 100 million tons of cargo. The largest port (Laem Chabang on the Eastern Seaboard) handled 63 million tons. Port operations and marine transportation are sources of pollution that have impacts on marine environments and animals (Bhatt et al 2006).

1.3.4 The Songkhla Lake

The Songkhla Lake (Thai name: Thale Sap Songkhla) is the largest lake in Thailand. It is located in the southern region, on the western coast of the Gulf of Thailand. The lake covers an area of 1,040 km^2, with a width of 20 km and length 75 km approximately. It is divided into three parts: the Upper Songkhla Lake in Phatthalung, the Middle Songkhla Lake between borders of Songkhla and Phatthalung, and the Lower Songkhla Lake around Songkhla being a harbor, which is the mouth of the lake (Figure 1.2).

Some canals pour fresh water into the lake. The water around the outlet of those canals is fresh water. The salinity slowly increases where the fresh water and the sea water met. Thus, the water in the central area of the lake is brackish, and become saltier in the area near the lake's mouth.

The Songkhla Lake is known as an important source of aquatic animals in Thailand. Several hundred species of fish have been found in the lake. Benthic fauna affect the biodiversity of aquatic animals (Angsupanich 2007). Over 160 species of macrobenthos have been recorded (Angsupanich and Kuwabura 1995). The abundance and species composition are generally higher during the late southwest monsoon (October) than during the mid-northeast monsoon (December).

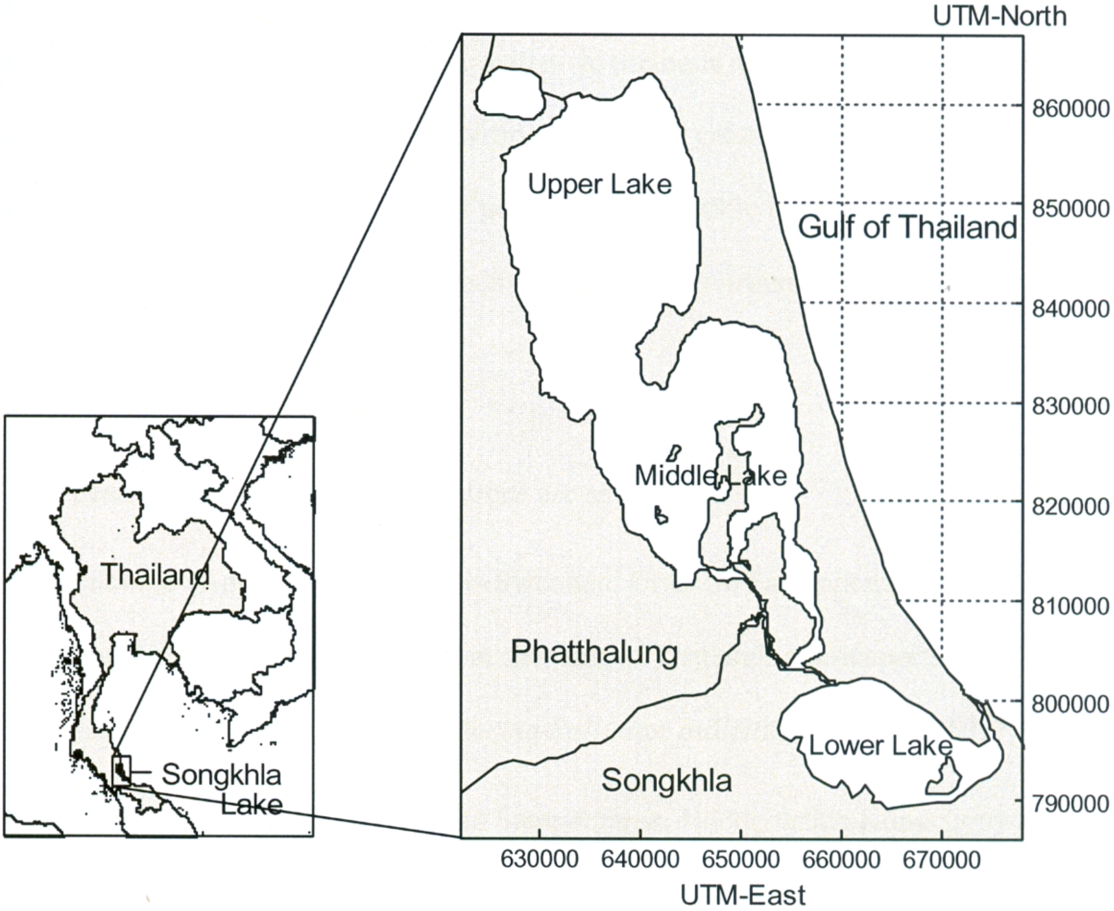

Figure 1.2: The Songkhla Lake and three zones of the lake

The lake has combined freshwater and estuarine complex of high productivity representing an extraordinary combination of environmental resources. Conflicting land use changes have degraded the watershed area and have changed the rainfall pattern and runoff, which eventually impacted the salinity patterns and water quality of the lake (Chufamanee et al 2003).

1.4 Objective

The aim of these studies was to investigate methods for modeling the health and abundance of coastal and estuarine animals in Thailand. Such statistical analysis can provide a better understanding of the effect of environmental change on such living organisms in their environment. The statistical methods apply to two outcomes: (1) the prevalence of imposex in gastropods of a size greater than 10 mm in coastal environments, and (2) the density of general macrobenthic fauna (polychates, crustacean, pelecypods and gastropods) in estuarine environments.

1.5 Publications of the thesis

Two peer-reviewed original publications are as follows:

Publication 1 (Appendix 1): Cornelis Swennen, Uraiwan Sampantarak and Nukul Ruttanadakul. 2009. TBT-pollution in the Gulf of Thailand: a re-inspection of imposex incidence after 10 years. *Marine Pollution Bulletin*, 58 (4): 526-532.

Publication 2 (Appendix 2): Uraiwan Sampantarak, Noodchanath Kongchouy and Saowapa Angsupanich. 2009. Regression-based modeling of macrobenthic fauna density in Middle Songkhla Lake, Thailand. (Manuscript)

Publication 3 (Appendix 3): Noodchanath Kongchouy and Uraiwan Sampantarak. 2010. Confidence Intervals for Adjusted Proportions Using Logistic Regression. *Modern Applied Science*, 4 (6).

1.6 Expected advantages of the thesis

The major scientific findings presented in this thesis are as follows:

(1) To provide analytical techniques in order to show how to analyse biological data collected from a large area, each with a different assemblage of species.

(2) To analyse and describe community structure of macrobenthic fauna and their relation with environmental variables.

Publication 1 identifies the prevalence of imposex in gastropods in the Gulf of Thailand. The main objectives were to (1) establish the present levels of imposex in the area; and (2) clarify possible differences in sensitivity between species. Logistic regression was used to model the effects of multiple determinants on the prevalence of imposex and also for adjusting a frequency that varies with a determinant of interest for a covariate determinant.

Publication 2 uses a multivariate multiple regression modeling to describe the variation in density over space and time of the 24 families of macrobenthic fauna. For model fitting, the response variable was taken as $\log(1 + c \times \text{density})$ with the multiplier c chosen to approximate symmetry of error distributions. The environmental variables were defined as environmental factors based on factor analysis, which were then used as predictor variables in the model.

Publication 3 presents confidence intervals for adjusted proportions using logistic regression with weighted sum contrasts. The methods are applied to data from two studies, (1) imposex percentages among female gastropods at different locations in the Gulf of Thailand adjusted for different species, and (2) complication-based neonatal morbidity risk for births at a major hospital adjusted for demographic factors.

1.7 Structure of the thesis

The chapters following the introduction are organized as follows. Reviews of the methodology in the both publications are presented in Chapter 2. The application of the logistic regression model to imposex gastropods prevalence is presented in Chapter 3. Chapter 4 presents the linear regression model with application to macrobenthic fauna density. Finally, Chapter 5 summarizes the work and includes discussion and conclusions.

Chapter 2

Methodology

The studies examined two datasets from coastal and estuarine environments:

(1) imposex prevalence in female gastropods in the Gulf of Thailand in 2006; and

(2) density of macrobenthic fauna in the Middle Songkhla Lake during 1998-1999.

Section 2.1 describes the methodology details for the prevalence of imposex (study 1). Section 2.2 describes the methodology details for the macrobenthic fauna abundances (study 2). Section 2.3 describes data management. Section 2.4 details the statistical methods used for the two studies.

All the graphical displays, map creations, statistical model fitting, and goodness-of-fit assessments were carried out using the R program (R development core team 2009, Venables and Ripley 2002, Murrell 2006).

2.1 Prevalence of imposex

The prevalence of imposex is studied by examining its distribution in various species of gastropods with the objective of determining its association with TBT pollution in the Gulf of Thailand. Gastropod samples were obtained from 56 sampling sites located in the Gulf over two periods in 2006 (June 8 to June 27; and September 19 to October 10).

2.1.1 Data source

Some specimens were hand-collected in the intertidal zone during low tide, but most were obtained from the by-catch of small commercial fishing boats and classified into species immediately. A small sample of shells belonging to species that were not immediately recognised was preserved in alcohol for later identification. A part of the shell was removed by hammering to identify sex.

Males of gastropods can easily be identified by the presence of a penis and vas deferens. These organs are on the right side of the body near the right rhinophore and cannot be withdrawn into the body.

Females have an inconspicuous vaginal pore next to the anus. Thus the presence of a prominent penis is the crucial factor. However, due to TBT pollution females may currently show a small penis and a small vas deferens at the same external body position as in the males. Therefore, a clear internal characteristic has to be used and that was found in a preliminary study to be the presence or absence of a good recognizable capsule gland in females. In practice the internal examination was only necessary in the beginning of a series of a species. The female penis reaches lengths that are 10% or less of that of males of the same shell height. Presence or absence of a penis whether or not with a vas deferens in females was conclusive for the qualification imposex or no imposex. To avoid errors, we did not use an imposex index such as used by Mensink et al (2002), since sizes of the specimens and sizes and shapes of the male penis among the several species studied varied too much for correctly using the index on all species under local field conditions. Thus females with stage 1 and 2 were considered as showing no imposex.

2.1.2 Path diagram and research questions

The effect of TBT on the prevalence of imposex in the Gulf of Thailand could be addressed simply by examining the prevalence of imposex at different locations in the

Gulf, and seeing if higher prevalence was associated with TBT contamination. However, if the effect of TBT in causing imposex differs between gastropods species, any such association could be distorted by variation of species distribution with location, so it is necessary to take species into account when developing a contamination model.

Figure 2.1 shows the relevant variables as a path diagram. A total of 13 areas comprised the determinant while 16 species groups comprised the covariate. The outcome of interest is imposex (normal female or imposex female).

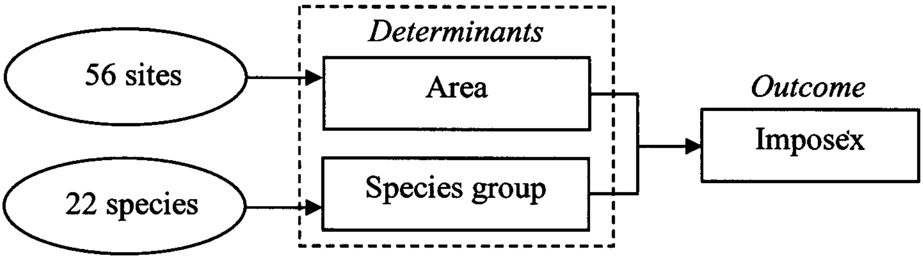

Figure 2.1: Path diagram showing roles of variables

Two research questions were considered, indicated by the dotted rectangles in the path diagram. In order of scientific importance, they are as follows.

(1) How does the prevalence of imposex in gastropods vary with area?

(2) How does the prevalence of imposex vary with species?

2.2 Macrobenthic fauna abundance

2.2.1 Data source

The datasets that were used for this work (Figure 2.2) consisted of biotic samples (macrobenthic fauna densities per square meter for each identifiable family) and other datasets containing information on environmental variables (water quality parameters and sediment characteristics). All were collected from nine sampling stations located in the Middle Songkhla Lake, labeled 1-9 (Figure 2.3) at bimonthly intervals from April 1998 to February 1999. The details of each dataset are as follows:

(1) Macrobenthic fauna were collected via a Tamura's grab (0.05 m^2). The assemblage was conducted with eleven replications for each station. The samples were sieved consecutively through three orders of screen residue (5, 1, and 0.5 mm of sieve mesh size, respectively) and fixed in 10% Rose Bengal formalin solution for later taxa identification.

(2) Environmental variables (water depth, salinity, temperature, water pH, dissolved oxygen, total suspended solid, organic matter contents, organic carbon contents and total nitrogen contents, together with pH of sediment, and soil structure as sand, silt and clay percentages) were obtained in triplicate on the same occasions as the data of macrobenthic fauna.

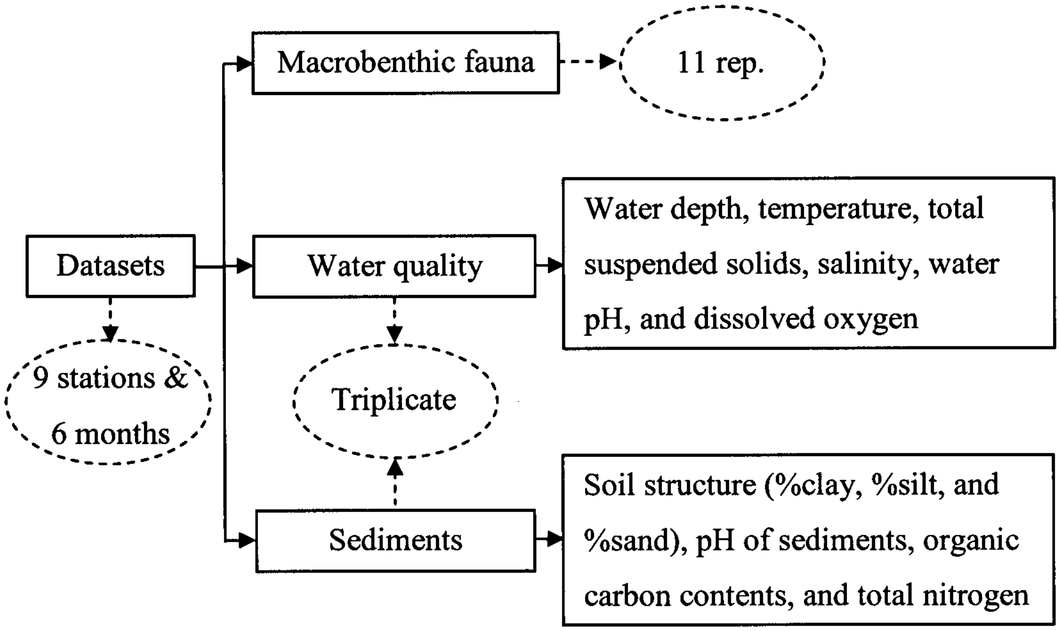

Figure 2.2: Dataset of the study of macrobenthic fauna variation

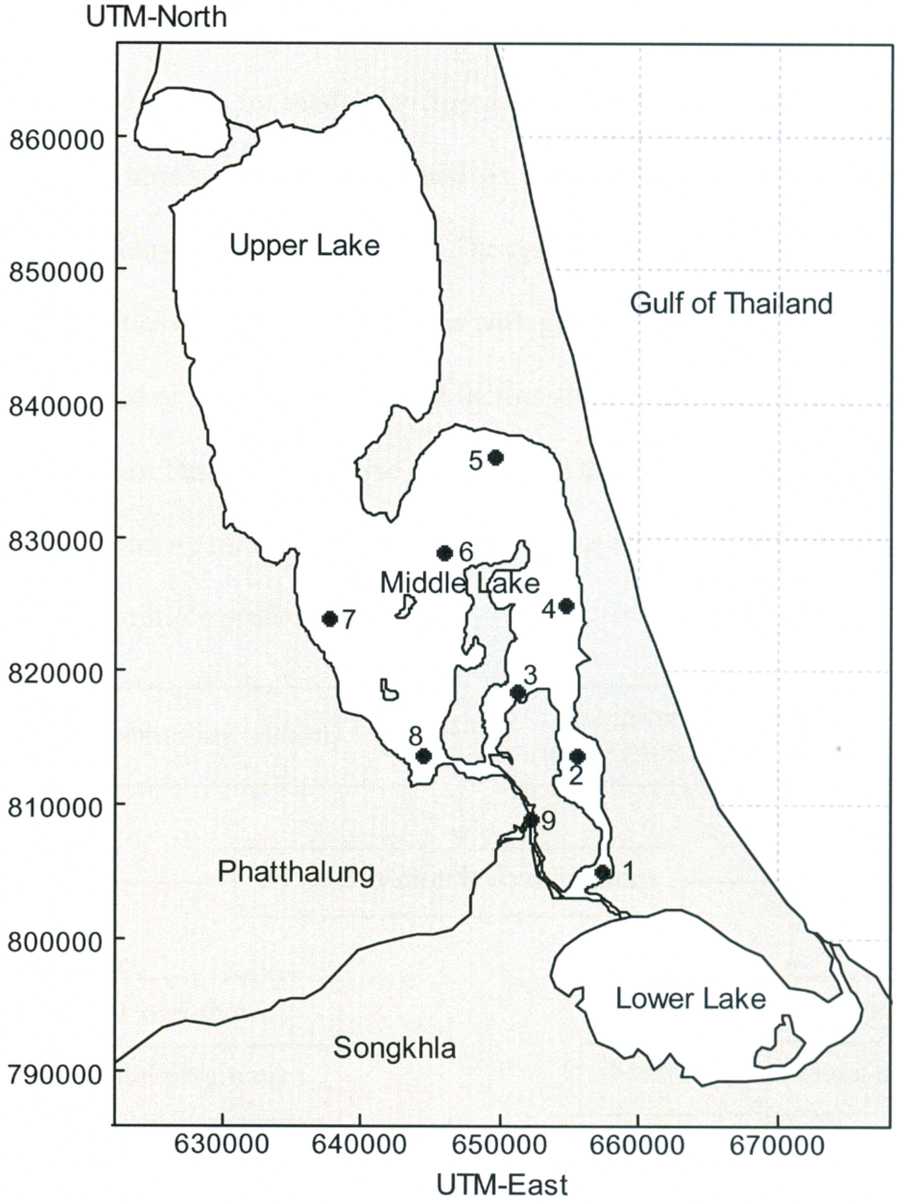

Figure 2.3: The Songkhla Lake and nine sampling stations (1: Ban Laem Chak, 2: Ban Koh Nang Khum, 3: Ban Nak Ka Rat, 4: Ban Koh Khob, 5: Ban Ta Kura, 6: Ban Laem Kruad, 7: Ban Hat Kai Tao, 8: Ban Ta Wa, and 9: Ban Bang Tan) within the Middle Songkhla Lake

2.2.2 Steps for model development

Figure 2.4 shows the steps for model development to assess macrobenthos distributions. The nine sampling stations and six bimonthly periods were combined into fifty four station-month combinations. The response variables were the log-transformed densities of 24 selected families with greater than 35% occurrence. The predictors consisted of three environmental factors and one unique variable derived from factor analysis. Finally, to analyse and describe the relationships between these variables, a multivariate multiple regression model was used to relate these multiple responses to the multiple predictors.

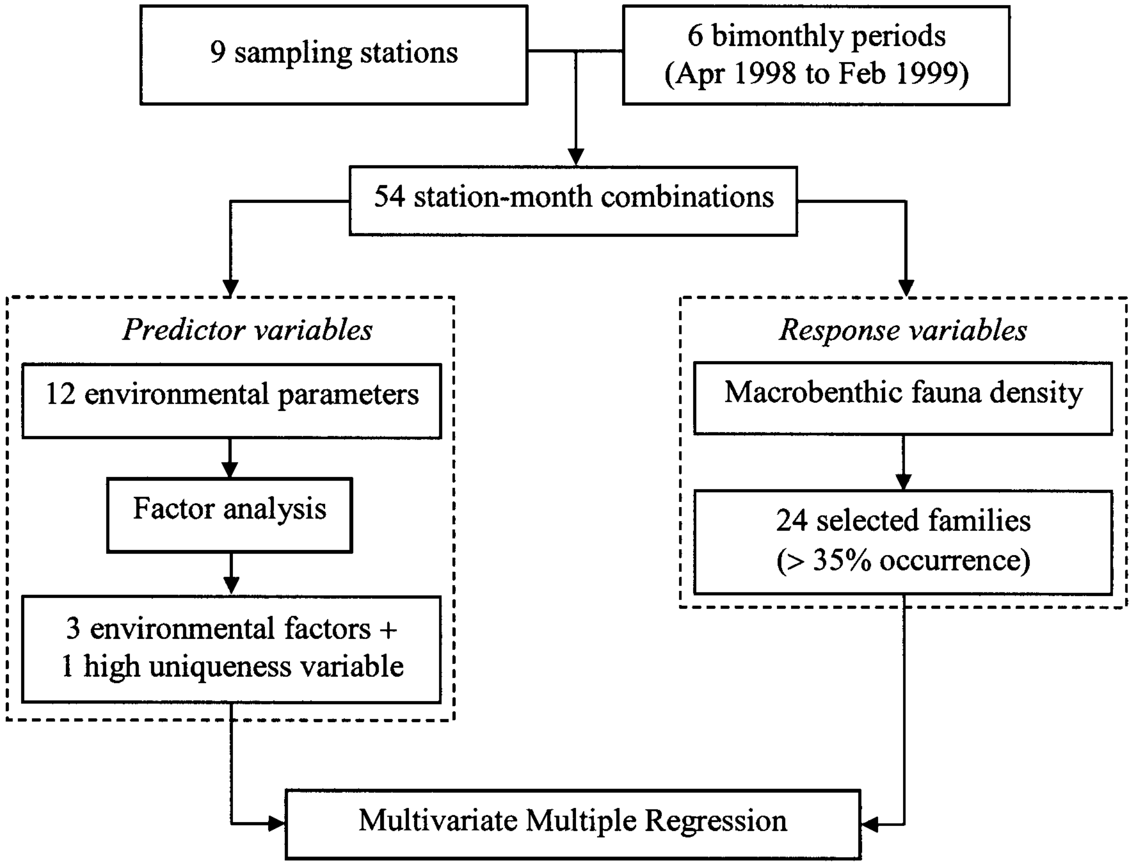

Figure 2.4: Steps for development of model to assess macrobenthic fauna variations

2.3 Data managements

Both datasets were loaded into a Microsoft Excel spreadsheet file and exported to WebStat (a set of programs for graphical and statistical analysis of data stored in an SQL database, written in HTML and VBScript) for cleaning. Then they were restructured into a simple format suitable for using the R program for further graphical displays, map creations, and statistical analysis.

2.4 Statistical methods

Several statistical methods were used in this thesis. Logistic regression modeling is well-suited to the analysis of the binary outcome in the gastropod dataset. The statistical methods for the analysis of macrobenthic fauna abundances were based on a multivariate multiple regression model involving factor analysis, which was used to define the factors in the environment used to link benthic fauna community structure and environmental predictors.

2.4.1 Logistic regression

Logistic regression (Hosmer and Lemeshow 2000, Kleinbaum and Klein 2002) is a statistical method widely used to model the association between a binary outcome probability - the probability of a specific adverse outcome - and a set of fixed determinants. When the determinants are categorical factors, these factors can be structured as a multi-way contingency table of counts and the data for analysis comprise the proportions of adverse outcomes in the cells of this table.

In our first study (Publication 1), we investigated how imposex in female gastropods can be predicted by one or more predictor variables. A sample of 8,757 gastropods belonging to 16 species groups that were sampled from 13 areas in the Gulf of Thailand in 2006 was considered. We employed logistic regression to model the effects of multiple determinants on the prevalence of imposex. If p is the prevalence of outcomes with a specific characteristic in a sample of size n, an asymptotically valid (for large n) formula for its standard error is

$$SE = \sqrt{\frac{p(1-p)}{n}}. \tag{2.1}$$

An asymptotically valid 95% confidence interval (95% CI) for the prevalence is thus given by

$$p - 1.96 \times SE,\ p + 1.96 \times SE. \tag{2.2}$$

When p_{ij} denotes the probability of outcome in taxa j on area i, the simplest such model takes the additive form

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = a_i + b_j. \tag{2.3}$$

Formula 2.3 can be inverted to give an expression for the prevalence p as

$$p_{ij} = \frac{1}{1+\exp(-a_i - b_j)}. \tag{2.4}$$

To avoid over specification of the parameters, one of the categories of the second determinant is taken as the reference and the corresponding parameter is zero. The model is fitted by maximizing the likelihood of the observed data given the parameters, providing standard errors for the fitted parameters. For each determinant, the method gives a p-value based on a chi-squared statistic for testing the null hypothesis that the outcome prevalence is the same for each of its component categories. The method also gives separate p-values for comparing each parameter in the model with the mean (for the determinant of interest) or reference (for the covariate determinant).

Logistic regression provides a straightforward method for adjusting a prevalence that varies with a determinant of interest for the effect of a covariate determinant. To calculate the adjusted prevalence for category i of the determinant of interest, the term b_j in (2.4) is replaced by a constant b^*, that is,

$$p_i^* = \frac{1}{1+\exp\left(-a_i - b^*\right)}. \tag{2.5}$$

The value of b^* in (2.5) is chosen to ensure that sum of the expected number of adverse outcomes is equal to the sum of the observed number, that is,

$$\sum p_i^* n_i = \sum p_i n_i, \tag{2.6}$$

where n_i is the sample size in category i of the determinant of interest. This method extends straightforwardly to additional covariates.

2.4.2 Multivariate multiple regression

In the second study, the multivariate multiple regression model was used to evaluate the effects of multiple predictor variables (the three environmental factors and the unique environmental parameter) on multiple response variables (the densities of the twenty four families of macrobenthic fauna), both the predictors and the response variables were observed at the fifty four occasions. The multivariate multiple regression model (Mardia et al 1979) is expressed in a matrix form, that is,

$$\mathbf{Y}_{(n \times p)} = \mathbf{X}_{(n \times q)} \mathbf{B}_{(q \times p)} + \mathbf{E}_{(n \times p)} \tag{2.7}$$

In this formulation $\mathbf{Y}_{(n \times p)}$ is an observed matrix of p response variables on each of n occasions, $\mathbf{X}_{(n \times q)}$ is the matrix of q predictors (including a vector of 1s) in columns and n occasions in rows, $\mathbf{B}_{(q \times p)}$ contains the regression coefficients (including the intercept terms), and $\mathbf{E}_{(n \times p)}$ is a matrix of unobserved random errors with mean zero and common covariance matrix $\mathbf{\Sigma}$. Ordinary (univariate) multiple regression arises as the special case when $p = 1$. If $q - 1$ environmental predictors $f_i^{(k)}$ $(k = 1, 2, \ldots, q - 1)$ are available, the predict model for outcome j occasion i model may be expressed as

$$y_{ij} = \mu_j + \sum_{k=1}^{p} \beta_j^{(k)} f_i^{(k)} + z_{ij} \, , \tag{2.8}$$

where y_{ij} is the observed abundance for family j on occasion i, μ_j is the mean abundance associated with family j, $\beta_j^{(k)}$ is the effect of environmental variable k on family j, and z_{ij} are the random errors.

The model fit may be assessed by plotting the residuals against normal quantiles (Venables and Ripley 2002), and also by using the set of r-squared values for the response variables to see how much of the variation in each is accounted for by the model.

The method also provides standard errors for each of the $p \times q$ regression coefficients thus providing p-values for testing their statistical significance after appropriate allowance for multiple hypothesis testing. The multivariate analysis of variance (MANOVA) decomposition is also used to assess the overall association between each environmental predictor and the set of outcomes by the likelihood ratio, Pillai's trace criterion (Olson 1976, Johnson and Wichern 1998).

2.4.3 Factor analysis

Factor analysis is a mathematical model that tries to explain the correlation between a large set of variables in terms of a small number of underlying factors. A major assumption of the analysis is that it is not possible to observe these factors directly: the variables depend upon the factors but are also subject to random errors (Mardia et al 1979).

In our second study, factor analysis is performed on the environmental variables with the aim of substantially reducing correlations between them that could mask their associations with the outcome variables. Each factor identifies correlated groups of variables. Ideally each group (which must contain at least two variables to contribute to the factor analysis) contains variables with small correlations with variables in other groups. To achieve this, any variable uncorrelated with all other variables is omitted from the factor analysis. Each factor comprises weighted linear combinations of the variables, and these factors are rotated to maximize the weights of variables within the factor group and minimize the weights of variables outside the group. The resulting weights are called "loadings". Variables omitted from the factor analysis due to low correlation with all other variables (high "uniqueness") are treated as separate predictors, so predictors include single variables as well as factors.

The number of factors selected was based on obtaining an acceptable statistical fit using the chi-squared test, and these factors were fitted using maximum likelihood with promax rotation in preference to varimax, which requires the rotation to be orthogonal (Browne 2001, Abdi 2003).

Chapter 3

Using a logistic regression to model the prevalence of imposex

This chapter presents a preliminary data analysis and the results from fitting the logistic regression model for prevalence of imposex in gastropods in the Gulf of Thailand in 2006. Section 3.1 is the description of variables. Section 3.2 is data characteristics. Section 3.3 is the logistic regression model to imposex outcome. The results presents in this chapter also appeared in Swennen et al (2009) and Khongchouy and Sampantarak (2010).

3.1 Description of variables

The roles of variables are classified as the determinants and outcome. The outcome of interest is imposex (normal female/or female showing imposex), which is binary data type. The determinant (sometimes called the "study factor") is thirteen areas in the Gulf of Thailand where imposex in gastropods are measured. There is also a covariate of determinant: sixteen species groups of gastropods assemblages. Both of them (determinant and covariate) are the nominal data type. The target population is all gastropod assemblages that were collected in the Gulf of Thailand in 2006. Only female gastropods were used for the statistical analysis. These variables and their roles, data types, and categories are listed in Table 3.1.

Table 3.1: Variables and their role, data type, and categories

Role	Variable	Type	Categories
Determinant	Area	Nominal	A: Rayong, B: Namrin, C: Pattaya, D: Si Racha, E: Phet Buri, F: Ban Khau, G: Songkhla, H: Bang Tawa, I: Rusamilae, J: Laem Nok, K: Pattani Bay Mouth, L: Panare, and M: Tak Bai
Covariate	Species group	Nominal	1: *Murex trapa*, 2: *Murex altispira*, 3: *Murex occa*, 4: *Lataxiena blosvillei*, 5: *Semiricinula muricoides*, 6: *Thais bitubercularis*, 7: *Thais lacera*, 8: *Morula musiva*, 9: *Babylonia areolata*, 10: *Nassarius jacksonianus*, 11: *Nassarius siquijorensis*, 12: *Nassarius stolatus*, 13: *Pugilina cochlidium*, 14: *Hemifusus ternatanus*, 15: *Turricula javana*, and 16: other
Outcome	Imposex	Binary	Normal female, Imposex female

3.2 Characteristics of the data

A total of 8,757 specimens were collected in various species and sampling sites in the Gulf of Thailand; 5,044 specimens were female, and 25.2% (n = 3,713) of total females were imposex. Specimens were classified into twenty two species belonging to five families (Muricidae, Buccinidae, Nassariidae, Melongenidae, and Turridae). The female gastropods heights varied between 11.0-131.0 mm. The minimum height of 11.0 mm was found in *Nassaria pusilla* while the maximum height of 131.0 mm was found in *Hemifusus ternatanus*. The names of the species and families are presented in Table 3.2, which also summaries the number and percentage of females, and also percentage of females showing imposex.

Table 3.2: Lists of families and species of the gastropods, with a summary of total number of sample (nTotal), number of females (nFem), percentage of females (%Fem), and percentage of females showing imposex (%Imp) in 2006

Families and species	nTotal	nFem	%Fem	%Imp
Family Buccinidae				
Babylonia areolata (Link, 1807)	314	158	50.3	1.3
Nassaria pusilla (Röding, 1798)	67	53	79.1	1.9
Phos senticosus (Linnaeus, 1758)	2	1	50.0	0.0
Family Muricidae				
Chicoreus banksii (Sowerby II, 1841)	58	39	67.2	12.8
Lataxiena blosvillei (Deshayes, 1832)	721	366	50.8	88.8
Morula musiva (Kiener, 1835)	324	161	49.7	0.6
Murex altispira Ponder & Vokes, 1988	359	286	79.7	0.7
Murex occa Sowerby II, 1824	888	417	47.0	40.8
Murex trapa Röding, 1798	1,156	698	60.4	22.9
Rapana rapiformis (Von Born, 1778)	85	32	37.6	6.3
Semiricinula muricoides (De Blainville, 1832)	753	521	69.2	50.9
Thais bitubercularis (Lamarck, 1822)	427	214	50.1	14.0
Thais clavigera (Kuester, 1860)	75	36	48.0	25.0
Thais lacera (Von Born, 1778)	587	306	52.1	21.9
Thais rufotincta Tan & Sigurdsson, 1996	22	10	45.5	10.0
Family Melongenidae				
Hemifusus ternatanus (Gmelin, 1791)	368	225	61.1	15.6
Pugilina cochlidium (Linnaeus, 1758)	923	413	44.7	9.4
Family Nassariidae				
Nassarius jacksonianus (Quoy & Gaimard, 1833)	233	166	71.2	10.8
Nassarius livescens (Philippi, 1849)	62	29	46.8	6.9
Nassarius siquijorensis (A. Adams, 1852)	260	202	77.7	28.7
Nassarius stolatus (Gmelin, 1791)	533	306	57.4	4.9
Family Turridae				
Turricula javana (Linnaeus, 1767)	540	405	75.0	15.6
Total	8,757	5,044	57.6	25.2

The preliminary results showed that the highest proportion of imposex was found in *Lataxiena blosvillei* (88.8%) followed by *Semiricinula muricoides* (50.9%), Murex occa (40.8%). Other species were as follows: *Nassarius siquijorensis* (28.7%), *Thais clavigera* (25.0%), *Murex trapa* (22.9%), *Thais lacera* (21.9%), *Hemifusus ternatanus* (15.6%), *Turricula javana* (15.6%), *Thais bitubercularis* (14.0%), *Chicoreus banksii* (12.8%), *Nassarius jacksonianus* (10.8%), *Thais rufotincta* (10.0%), *Pugilina cochlidium* (9.4%), *Nassarius livescens* (6.9%), *Rapana rapiformis* (6.3%), *Nassarius stolatus* (4.9%), *Nassaria pusilla* (1.9%), *Babylonia areolata* (1.3%), *Murex altispira* (0.7%), *Morula musiva* (0.6%), and *Phos senticosus* (which had no imposex females).

3.3 Logistic regression model to imposex outcome

Logistic regression was used to fit a model where the outcome event was the allocation to the treatment group and determinant (area) including covariate (species group) prior to treatment. A total of 56 sampling sites were grouped into 13 geographical areas. They are… A: Rayong, B: Namrin, C: Pattaya, D: Si Racha, E: Phet Buri, F: Ban Khau, G: Songkhla, H: Bang Tawa, I: Rusamilae, J: Laem Nok, K: Pattani Bay Mouth, L: Panare, and M: Tak Bai. (see more details in Appendix 4).

Gastropod species with fewer than 100 individuals (*C. banksii* (*n* = 58), *T. clavigera* (*n* = 75), *T. rufotincta* (*n* = 22), *R. rapiformis* (*n* = 85), *Nassaria pusilla* (*n* = 67), *P. senticosus* (*n* = 2), and *N. livescens* (*n* = 62)) were combined into a single group refer to as "other" because logistic regression cannot handle grouped data where any cell has zero count. *M. Trapa* (*n* = 1,156) is the species that had the largest sample size, so it was chosen as referent categories.

Table 3.3 gives the results from fitting a logistic regression model. The results show the evidence of the difference between the two treatment groups with respect to area or species group. The model gave a residual deviance of 171.45 with 55 df based on the grouped data (the 84 combinations of area and species where samples were obtained). Chi-squared statistics for testing a common imposex frequency between areas and species groups were found to be 438.90 (12 df) and 453.04 (15 df) respectively, with the correspondence being highly statistically significant (Table 3.3).

The coefficients from the model can be related to the differential severity risks for each area and each species group. Imposex varies by area: area D (Si Racha) shows the highest risk with the maximum of coefficient equals to 1.13; second is area C (Pattaya) with the coefficient equals to 0.138. The minimum risk (coef = − 3.253) is found in area E (Phet Buri). As for variation by species, many species of Muricidae are very sensitive prone to imposex e.g. *L. blosvillei* appears to have the most imposex (coef = 1.674), the second is *M. occa* (coef = 0.697). *Morula musiva* seems to have the lowest sensitivity (coef = − 3.646).

From the results as presents in Table 3.3, the adjusted imposex frequency was calculated using formula 2.5 (Chapter 2) with the corresponding 95% confidence intervals (formula 2.2, Chapter 2). Table 3.4 presents imposex frequency before and after adjusting for species.

Before adjusting for the different sensitivities of the species, areas C (Pattaya) and D (Si Racha) had a high imposex frequency (62.7% and 73.6%, respectively), areas F (Ban Khau), H (Bang Tawa), I (Rusamilae), J (Laem Nok), and K (Pattani Bay Mouth) had medium frequency (18.5%, 17.4%, 15.6%, 31.0%, and 18.8%, respectively), whereas areas A (Rayong), B (Namrin), E (Phet Buri), G (Songkhla), L (Panare), and M (Tak Bai) had low frequency (4.7%, 5.2%, 3.1%, 0.6%, 6.4%, and 1.3%, respectively).

Table 3.3: Results of the logistic model with outcome "Imposex"

Determinant	Estimate	Std. Error	*p*-value
Area			
A: Rayong	− 2.119	0.389	< 0.001
B: Namrin	− 1.429	0.369	< 0.001
C: Pattaya	0.138	0.231	0.551
D: Si Racha	1.138	0.186	< 0.001
E: Phet Buri	− 3.253	0.289	< 0.001
F: Ban Khau	− 1.139	0.164	< 0.001
G: Songkhla	− 2.815	1.110	0.011
H: Bang Tawa	− 1.322	0.214	< 0.001
I: Rusamilae	− 1.491	0.154	< 0.001
J: Laem Nok	− 0.721	0.190	< 0.001
K: Pattani Bay Mouth	− 1.140	0.230	< 0.001
L: Panare	− 1.253	0.462	0.007
M: Tak Bai	− 1.987	1.251	0.112
Species			
01: *Murex trapa* (reference)	0		
02: *Murex altispira*	− 2.961	0.834	< 0.001
03: *Murex occa*	0.697	0.186	< 0.001
04: *Lataxiena blosvillei*	1.674	0.271	< 0.001
05: *Semiricinula muricoides*	− 0.215	0.211	0.308
06: *Thais bitubercularis*	− 0.494	0.289	0.087
07: *Thais lacera*	0.097	0.203	0.633
08: *Morula musiva*	− 3.646	1.069	0.001
09: *Babylonia areolata*	− 2.398	1.243	0.054
10: *Nassarius jacksonianus*	− 0.700	0.291	0.016
11: *Nassarius siquijorensis*	0.373	0.253	0.141
12: *Nassarius stolatus*	− 1.350	0.308	< 0.001
13: *Pugilina cochlidium*	− 1.602	0.231	< 0.001
14: *Hemifusus ternatanus*	− 1.504	0.247	< 0.001
15: *Turricula javana*	− 0.323	0.208	0.120
16: other	− 0.775	0.327	0.018

Table 3.4: Prevalence of imposex (%) by area before and after adjusting for species

Area	Total females	Number of imposex	Unadjusted imposex (%)	Adjusted imposex (%)	
				Prevalence	95% CI
A: Rayong	190	9	4.7	9.0	4.9 - 13.0
B: Namrin	249	13	5.2	16.4	11.8 - 21.0
C: Pattaya	474	297	62.7	48.4	43.9 - 52.9
D: Si Racha	537	395	73.6	71.9	68.1 - 75.7
E: Phet Buri	417	13	3.1	3.1	1.4 - 4.7
F: Ban Khau	286	53	18.5	20.8	16.1 - 25.5
G: Songkhla	171	1	0.6	4.7	1.5 - 7.8
H: Bang Tawa	735	128	17.4	17.9	15.1 - 20.7
I: Rusamilae	916	143	15.6	15.6	13.2 - 17.9
J: Laem Nok	507	157	31.0	28.5	24.5 - 32.4
K: Pattani Bay Mouth	266	50	18.8	20.8	15.9 - 25.6
L: Panare	140	9	6.4	19.0	12.5 - 25.4
M: Tak Bai	156	2	1.3	10.1	5.4 - 14.8

CI: Confidence interval

After adjusting, ignoring species differences, areas C (Pattaya) and D (Si Racha) showed a lower frequency than before, but still had the highest values (48.4%, CI: 43.9%-52.9% and 71.9%, CI: 68.1%-75.7%, respectively), while areas B (Namrin) and L (Laem Nok) moved up to the medium severity level of imposex frequency (16.4%, CI: 11.8%-21.0% and 19.0%, CI: 12.5%-25.4%, respectively). Figure 3.1 shows the areas and the sampling sites with the adjusted imposex prevalence.

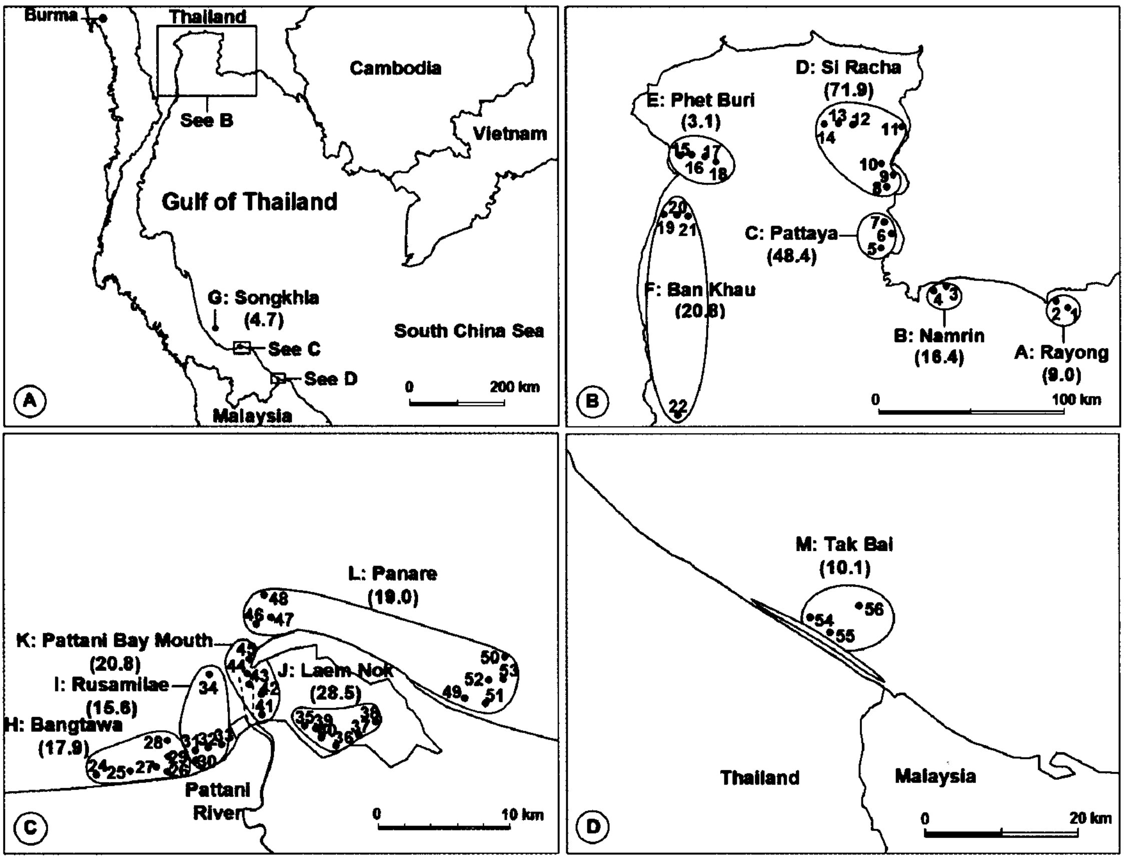

Figure 3.1: The study areas and sampling sites: Ⓐ is Overview; Ⓑ is northern part of the Gulf with study areas A-F; Ⓒ is Pattani Bay and surroundings with study areas H-L; Ⓓ is Narathiwat province with study area M. Dots with their number corresponds with the sampling sites in the detailed of Appendix 4. Value between brackets gives the mean of imposex frequency in the area

Table 3.5 presents the prevalence of imposex by species before and after adjusting for area. The sensitivity to imposex varied with species: *L. blosvillei* was found to be most prone to imposex (88.8%), followed by *S. muricoides* (50.9%), *M. occa* (40.8%) whereas *M. musiva* is least sensitive (0.6%). After adjusting for areas, many species showed a higher sensitivity than before adjusting except for *L. blosvillei* (69.7%, CI: 65.0%-74.4%), *S. muricoides* (25.8%, CI: 22.0%-29.6%), *P. cochlidium* (8.0%, CI:

5.4%-10.6%), and *H. ternatanus* (8.7%, CI: 5.1%-12.4%). Although *L. blosvillei* showed a lower value than before adjusting, it still was the highest sensitivity. In addition, *M. musiva* still was the lowest sensitivity prone to imposex.

Table 3.5: Prevalence of imposex (%) by species before and after adjusting for area

Species groups	Total females	Number of imposex	Unadjusted imposex (%)	Adjusted imposex (%)	
				Prevalence	95% CI
01: *Murex trapa*	698	160	22.9	30.1	26.7 - 33.5
02: *Murex altispira*	286	2	0.7	2.2	0.5 - 3.9
03: *Murex occa*	417	170	40.8	46.4	41.6 - 51.2
04: *Lataxiena blosvillei*	366	325	88.8	69.7	65.0 - 74.4
05: *Semiricinula muricoides*	521	265	50.9	25.8	22.0 - 29.6
06: *Thais bitubercularis*	214	30	14.0	20.8	15.4 - 26.3
07: *Thais lacera*	306	67	21.9	32.2	27.0 - 37.4
08: *Morula musiva*	161	1	0.6	1.1	0.0 - 2.7
09: *Babylonia areolata*	158	2	1.3	3.8	0.8 - 6.7
10: *Nassarius jacksonianus*	166	18	10.8	17.6	11.8 - 23.4
11: *Nassarius siquijorensis*	202	58	28.7	38.5	31.8 - 45.2
12: *Nassarius stolatus*	306	15	4.9	10.1	6.7 - 13.4
13: *Pugilina cochlidium*	413	39	9.4	8.0	5.4 - 10.6
14: *Hemifusus ternatanus*	225	35	15.6	8.7	5.1 - 12.4
15: *Turricula javana*	405	63	15.6	23.8	19.6 - 27.9
16: other	200	20	10.0	16.6	11.4 - 21.7

CI: Confidence interval

Chapter 4

A multivariate multiple regression model for macrobenthic fauna density

This chapter presents a preliminary analysis and a model fitting macrobenthic fauna distribution in the Middle Songkhla Lake. Section 4.1 is the descriptions of the variables. Section 4.2 is data characteristics. Section 4.3 is the multivariate multiple regression (MMR) model fitted to the density of twenty-four families of macrobenthic fauna (see more details, Publication 2, in Appendix 2).

4.1 Description of the variables

The roles of variables are classified as determinants and outcome. These variables, their roles and data type as shown in Table 4.1. The outcome of interest is the densities of the twenty-four families of macrobenthic fauna from nine sampling stations at six bimonthly periods. These densities are of continuous data type. The predictor variables (determinants) consisted of the environmental factors and the unique variable derived from factor analysis.

Table 4.1: Variables, their roles and data type

Variable	Role	Type
Environmental factors (3)	Determinant	Continuous
Unique environmental parameter (1)	Determinant	Continuous
Density of 24 macrobenthic fauna families	Outcomes	Continuous

4.2 Characteristics of the data

4.2.1 Occurrence and abundance of macrobenthic fauna

A total of 24 families were classified into three phyla of Annelida (class Polychaeta), Arthropoda (class Crustacea) and Mollusca (classes Gastropoda and Bivalvia), which comprised the most diverse groups (35.2-98.2% of occurrence).

The Polychaeta was represented by nine families (Capitellidae, Goniadidae, Hesionidae, Nephtyidae, Nereididae, Pilargiidae, Pholoidae, Spionidae and Terebellidae). The Crustacea was also represented by nine families (Aoridae, Isaeidae, Melitidae, Oedicerotidae, Apseudidae, Pseudotanaidae, Anthuridae, Cirolanidae and Alpheidae). Marginellidae, Retusidae, Skeneopsidae and Stenothyridae were in the Gastropoda whilst the two remaining families (Tellinidae and unidentified species were in the Bivalvia).

Nereididae was the highest occurrence family over all station and month combinations with 98.2% of occurrence, whereas the families of Terebellidae and Stenothyridae had the lowest occurrence (35.2%). Apseudidae was the most abundant family with average density of 40,083.6 ind m^{-2}, while Alpheidae was the least abundant, with average density of 98.2 ind m^{-2}.

The taxonomy, percentages of coverage and densities in ind m^{-2} of the 24 families of macrobenthic fauna in the Middle Songkhla Lake from April 1998 to February 1999 are shown as Table 4.2.

Table 4.2: The taxonomy, percentages of occurrence (%occ) and densities (ind m^{-2}) of the 24 macrobenthic fauna families in the Middle Songkhla Lake from April 1998 to February 1999. [* note the unidentified species in Bivalvia]

Phylum	Class	Order	Family	%occ	Density
Annelida	Polychaeta	Capitellida	Capitellidae	87.0	1,227.3
	Polychaeta	Phyllodocida	Goniadidae	37.0	443.6
	Polychaeta	Phyllodocida	Hesionidae	55.6	698.2
	Polychaeta	Phyllodocida	Nephtyidae	87.0	2,218.2
	Polychaeta	Phyllodocida	Nereididae	98.2	8,507.3
	Polychaeta	Phyllodocida	Pilargiidae	70.4	1,625.5
	Polychaeta	Phyllodocida	Pholoidae	59.3	658.2
	Polychaeta	Spionida	Spionidae	92.6	5,056.4
	Polychaeta	Terebellida	Terebellidae	35.2	1,136.4
Arthropoda	Crustacea	Amphipoda	Aoridae	59.3	2,421.8
	Crustacea	Amphipoda	Isaeidae	87.0	6,900.0
	Crustacea	Amphipoda	Melitidae	94.4	4,438.2
	Crustacea	Amphipoda	Oedicerotidae	66.7	667.3
	Crustacea	Tanaidacea	Apseudidae	90.7	40,083.6
	Crustacea	Tanaidacea	Pseudotanaidae	37.0	4,265.5
	Crustacea	Isopoda	Anthuridae	75.9	3,816.4
	Crustacea	Isopoda	Cirolanidae	37.0	427.3
	Crustacea	Decapoda	Alpheidae	40.7	98.2
Mollusca	Gastropoda	Neogastropoda	Marginellidae	85.2	3,963.6
	Gastropoda	Cephalaspidea	Retusidae	55.6	5,536.4
	Gastropoda	Mesogastropoda	Skeneopsidae	38.9	956.4
	Gastropoda	Mesogastropoda	Stenothyridae	35.2	581.8
	Bivalvia	Unidentified	Unidentified*	44.4	338.2
	Bivalvia	Veneroida	Tellinidae	81.5	17,134.5

4.2.2 Environmental variables

Abiotic data (wDep: water depth, wTemp: water temperature, Sal: salinity, spH: sediment pH, wpH: water pH, DO: dissolved oxygen, TSS: total suspended solids, TN: total nitrogen contents, OC: organic carbon contents, and soil structure as sand, silt, and clay) were used as environmental variables. Their values varied substantially between site and period.

Figure 4.1 plots the water characteristics in the Middle Songkhla Lake from April 1998 to February 1999. The water depth varied to a lesser extent, varying with location from an average of less than 1 m at stations four and nine to more than 2 m at station eight. It was also higher during the rainy season. The water temperature showed decreased values in the rainy season, with a range of 27-34 °C. The salinity increased from close to zero during the rainy season (December to February) to an average close to 20 in the other months. The amount of carbonic acid in rainfall caused the pH of water to decrease, and thus shows the lowest value in December.

Figure 4.2 plots sediment characteristics in the Middle Songkhla Lake from April 1998 to February 1999. The total nitrogen contents at each station was very low (0.02%) during October-February, possibly due to the fact that organic volume is lower during the rainy season since the current sweeps away the organic matter that is spread over the lake. The organic carbon content was relatively constant with respect to month, but showed the highest value at station nine in every month except August. The lake bed at station six was mostly characterized by sand (mean = 84.6%) and station 9 was mostly characterized by clay (mean = 53.2%), also with high values of organic carbon. Note that sand, silt and clay percentages sum to 100%.

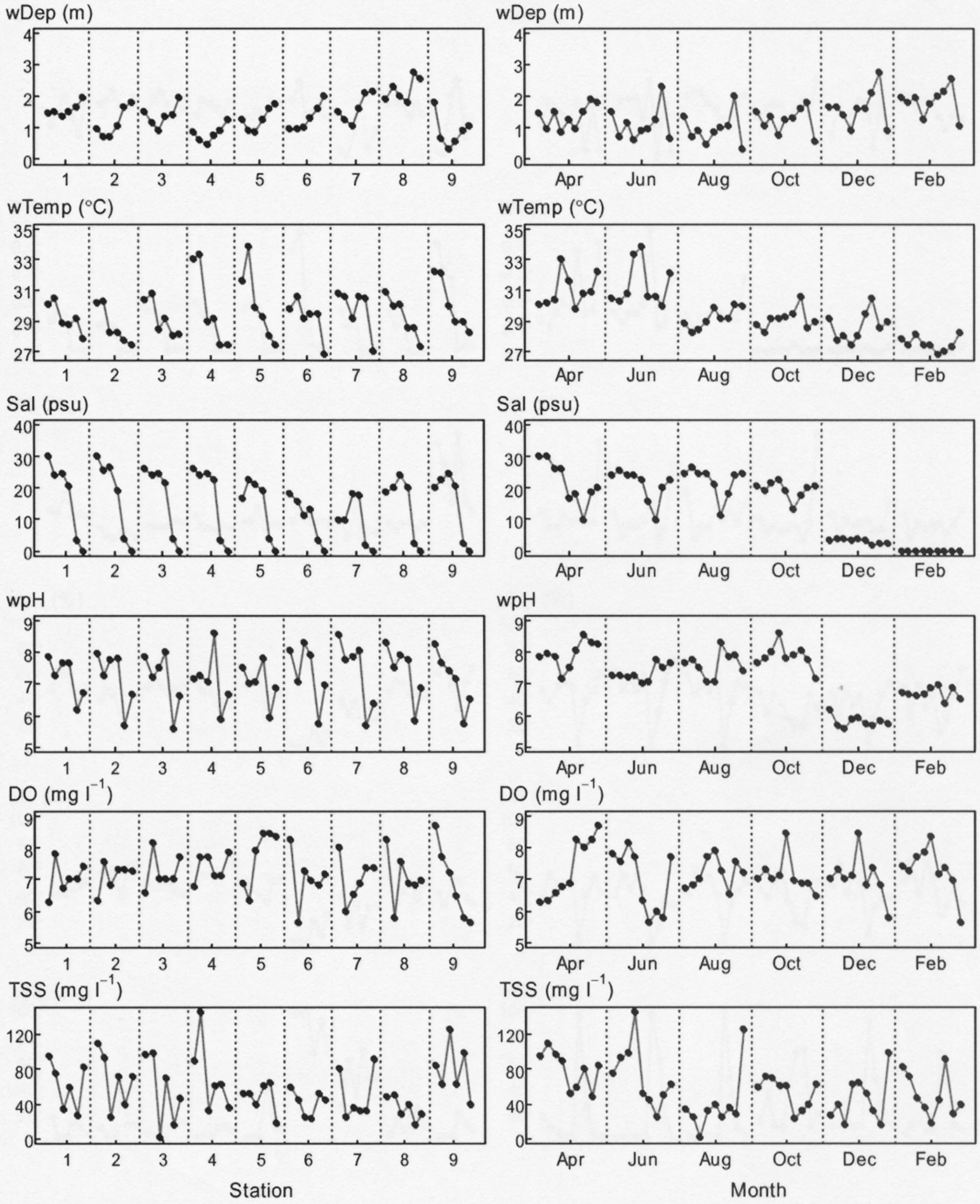

Figure 4.1: The water characteristics: water depth (wDep), water temperature (wTemp), salinity (Sal), water pH (wpH), dissolved oxygen (DO), total suspended solids (TSS) in the Middle Songkhla Lake from April 1998 to February 1999 by station (left panel) and month (right panel)

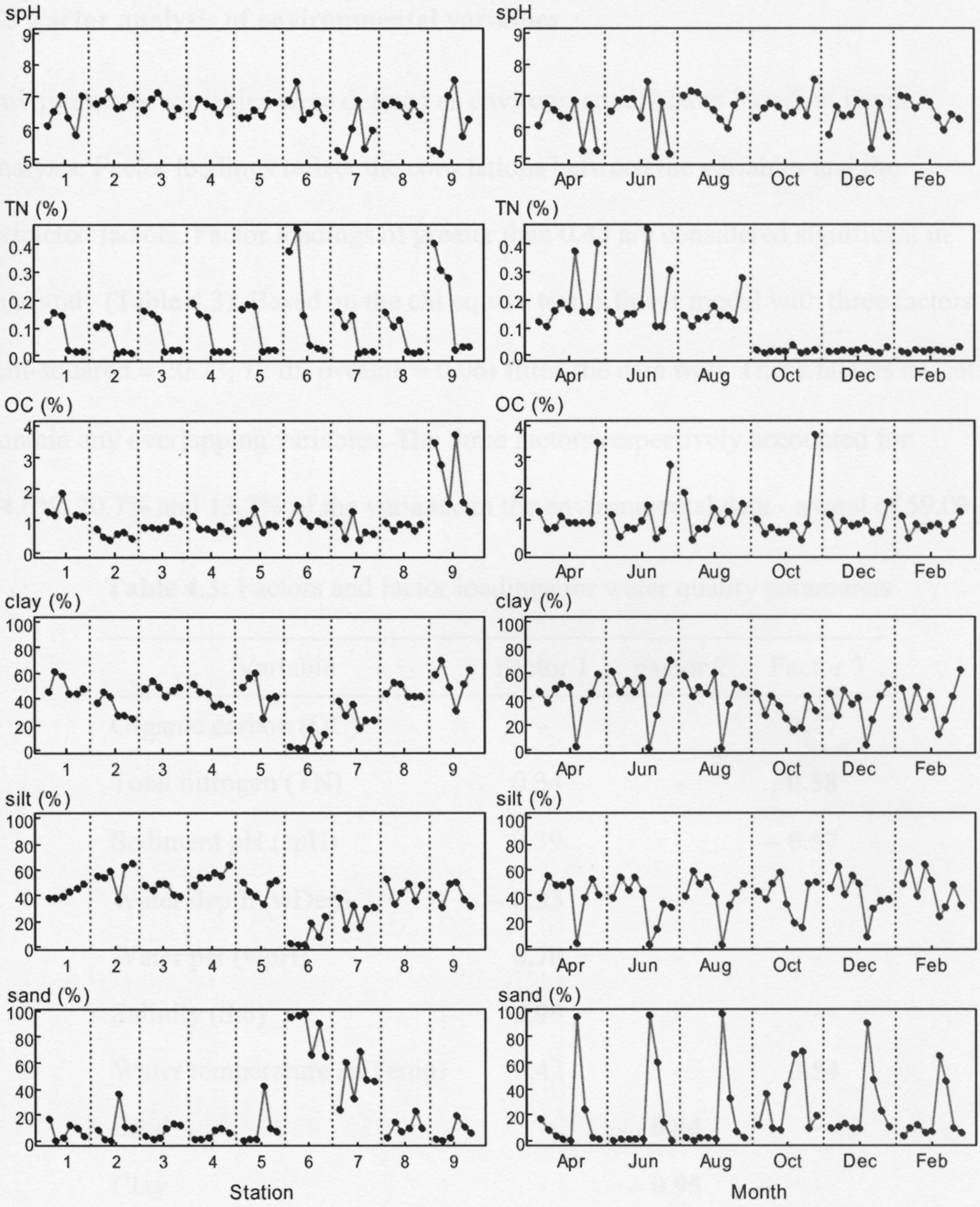

Figure 4.2: The sediment characteristics: sediment pH (spH), total nitrogen contents (TN), organic carbon contents (OC), percentage of clay, silt, and sand in the Middle Songkhla Lake from April 1998 to February 1999 by station (left panel) and month (right panel)

4.3 Factor analysis of environmental variables

Environmental variables were defined as environmental factors based on factor analysis. Factor loadings reflect the correlations between the variables and the extracted factors. Factor loadings of greater than 0.45 are considered significant in this study (Table 4.3). Based on the chi-square test, a factor model with three factors (chi-squared = 20.23, 12 df, *p*-value = 0.06) fitted the data well. These factors do not contain any overlapping variables. The three factors respectively accounted for 24.6%, 20.7% and 13.7% of the variance in the environmental data - a total of 59.0%.

Table 4.3: Factors and factor loadings for water quality parameters

Variable	Factor 1	Factor 2	Factor 3
Organic carbon (OC)	-	-	**0.47**
Total nitrogen (TN)	0.34	-	**0.58**
Sediment pH (spH)	0.39	-	**– 0.57**
Water depth (wDep)	**– 0.53**	-	-
Water pH (wpH)	**0.70**	-	-
Salinity (Sal)	**0.99**	-	-
Water temperature (wTemp)	0.42	-	**0.54**
Sand	-	**0.94**	-
Clay	-	**– 0.95**	-
% Total variance	24.6	20.7	13.7
% Cumulative variance	24.6	45.3	59.0

Interpreting the results, Factor 1 encompasses salinity, containing positive loadings for salinity (Sal) and water pH (wpH), and a negative loading for water depth (wDep) as expected, with deeper water during the rainy season. Factor 2 represents the effect of sediment characteristics in sand-clay habitat, consisting of a positive loading for sand and a similar negative loading for clay. Factor 3 characterizes physical and chemical compositions in the lake, comprising positive loadings for total nitrogen (TN), organic carbon (OC), and water temperature (wTemp); and a negative loading for sediment pH (spH). Each factor was defined as follows:

$$\text{Factor 1} = -0.53 \times \text{wDep} + 0.70 \times \text{wpH} + 0.99 \times \text{Sal};$$

$$\text{Factor 2} = 0.94 \times \text{Sand} - 0.95 \times \text{Clay};$$

$$\text{Factor 3} = 0.47 \times OC + 0.58 \times \text{T}N - 0.57 \times spH + 0.54 \times wTemp.$$

The three factors were included in the MMR model as predictors together with the two singleton variables omitted from the factor analysis (total suspended solids (TSS) and dissolved oxygen (DO)), with each of these five predictor variables scaled to have mean 0 and standard deviation 1.

4.4 Multivariate multiple regression model

The density was taken as log (1 + c × density) with the multiplier c chosen to approximate normality of the error distribution. The choice $c = 100$ gave residuals satisfying the normality assumption. A total of 24 families of macrobenthic fauna, which were the most diverse groups (> 35% of occurrence) over 54 station-month combinations, were selected. The densities of the selected families were used as the outcome of interest for model fitting.

The environmental factors together with the two singleton variables omitted from the factor analysis (TSS and DO) were first fitted to MMR model. The adequacy of the model can be tested by the MANOVA approach. The model-fitting resulted in the omission of dissolved oxygen, as shown in Table 4.4.

Figure 4.3 is a plot of residuals against corresponding normal quantiles. The plot shows the residuals in the y axis against normal quantiles in the x axis. It was used for checking the normality assumption and the adequacy of model fit to the data, and demonstrates that the points in the residual plot are randomly dispersed around the straight line, thus satisfying the normality assumption, and confirming the appropriateness of the model.

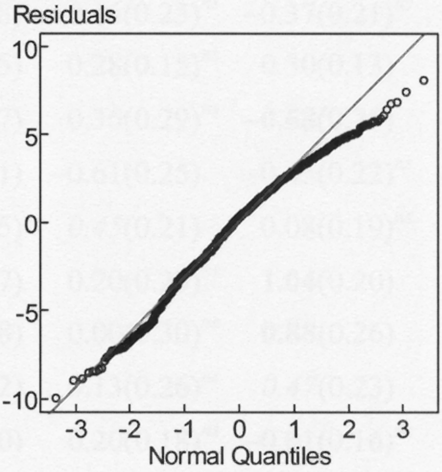

Figure 4.3: Residuals plotted against corresponding normal quantiles for the model

Table 4.4 shows the corresponding individual regression coefficients and standard errors and r-squared values for each family after fitting the MMR model with all four environmental predictors included. The coefficients listed are the ones statistically significant at 5% and 1% (in bold).

Since there are 96 regression coefficients in all, and 5% of these would be expected to have p-values < 0.05 even if all their corresponding parameters was zero, the five largest p-values < 0.05 are italicized to indicate failure to "honest" significance.

Table 4.4: Regression coefficients and standard errors (in parenthesis) from fitting MMR model with the four environmental predictors; coefficients with p-values < 0.05 are shown; those adjudged not honestly statistically significant are shown in italics and those with p-values < 0.01 are shown in bold; the coefficients that were not statistically significant are labeled "*ns*". [* note the unidentified species in Bivalvia]

Family	Intercept	Factor 1	Factor 2	Factor 3	TSS	r^2
Capitellidae	6.27(0.37)	0.16(0.23)^{ns}	0.06(0.20)^{ns}	−0.20(0.28)^{ns}	−0.47(0.41)^{ns}	0.05
Goniadidae	2.59(0.41)	**1.01**(0.25)	0.04(0.22)^{ns}	−0.76(0.31)	−1.10(0.45)	0.31
Hesionidae	3.87(0.45)	**0.76**(0.28)	−0.45(0.24)^{ns}	0.11(0.34)^{ns}	−0.85(0.49)^{ns}	0.23
Nephtyidae	6.58(0.38)	−0.16(0.23)^{ns}	−0.37(0.21)^{ns}	0.25(0.28)^{ns}	0.35(0.42)^{ns}	0.11
Nereididae	8.44(0.25)	0.28(0.15)^{ns}	0.30(0.13)	0.42(0.19)	0.01(0.27)^{ns}	0.25
Pilargiidae	5.17(0.47)	0.36(0.29)^{ns}	**−0.68**(0.25)	−0.18(0.35)^{ns}	−0.06(0.51)^{ns}	0.17
Pholoidae	3.97(0.41)	−0.61(0.25)	0.43(0.22)^{ns}	−0.54(0.30)^{ns}	−0.05(0.44)^{ns}	0.32
Spionidae	7.38(0.35)	*0.45*(0.21)	0.08(0.19)^{ns}	0.17(0.26)^{ns}	−0.51(0.38)^{ns}	0.12
Terebellidae	2.48(0.37)	0.20(0.23)^{ns}	**1.04**(0.20)	*0.60*(0.28)	−0.91(0.40)	0.44
Aoridae	4.48(0.48)	0.00(0.30)^{ns}	**0.88**(0.26)	0.54(0.36)^{ns}	*1.08*(0.53)	0.25
Isaeidae	7.16(0.42)	0.13(0.26)^{ns}	*0.47*(0.23)	0.54(0.31)^{ns}	−0.42(0.46)^{ns}	0.15
Melitidae	7.68(0.30)	0.20(0.18)^{ns}	−0.01(0.16)^{ns}	0.58(0.22)	0.24(0.32)^{ns}	0.24
Oedicerotidae	4.62(0.47)	−0.11(0.29)^{ns}	−0.36(0.25)^{ns}	−0.34(0.35)^{ns}	−0.42(0.51)^{ns}	0.08
Apseudidae	8.91(0.49)	−0.26(0.30)^{ns}	−0.31(0.26)^{ns}	−0.12(0.36)^{ns}	0.10(0.53)^{ns}	0.05
Pseudotanaidae	2.79(0.45)	−0.64(0.28)	**0.81**(0.24)	0.27(0.34)^{ns}	−0.63(0.49)^{ns}	0.34
Anthuridae	5.58(0.45)	0.11(0.28)^{ns}	**0.69**(0.24)	−0.40(0.34)^{ns}	−0.15(0.49)^{ns}	0.19
Cirolanidae	2.38(0.39)	0.42(0.24)^{ns}	0.51(0.21)	0.67(0.29)	0.46(0.42)^{ns}	0.29
Alpheidae	2.39(0.40)	−0.19(0.25)^{ns}	−0.23(0.22)^{ns}	0.21(0.30)^{ns}	0.46(0.44)^{ns}	0.07
Marginellidae	7.02(0.41)	−0.38(0.25)^{ns}	0.42(0.22)^{ns}	0.11(0.31)^{ns}	−0.33(0.45)^{ns}	0.16
Retusidae	4.32(0.52)	−0.16(0.32)^{ns}	−0.31(0.28)^{ns}	−0.42(0.39)^{ns}	**−1.59**(0.57)	0.21
Skeneopsidae	2.58(0.48)	0.20(0.29)^{ns}	−0.14(0.26)^{ns}	−0.05(0.35)^{ns}	−0.28(0.52)^{ns}	0.02
Stenothyridae	2.32(0.40)	0.35(0.25)^{ns}	−0.42(0.22)^{ns}	0.65(0.30)	−0.26(0.44)^{ns}	0.25
Unidentified*	2.98(0.44)	0.71(0.27)	0.04(0.24)^{ns}	−0.10(0.33)^{ns}	*−0.99*(0.48)	0.16
Tellinidae	7.02(0.52)	0.57(0.32)^{ns}	−0.17(0.28)^{ns}	0.10(0.39)^{ns}	−0.39(0.56)^{ns}	0.09

The densities of Hesionidae, Spionidae, unidentified species in Bivalvia, and Goniadidae were positively related with Factor 1 (mainly salinity) while Pholoidae and Pseudotanaidae were negatively related with this Factor 1.

Eight families (Pseudotanaidae, Pilargiidae, Isaeidae, Anthuridae, Nereididae, Cirolanidae, Terebellidae, and Aoridae) were associated with Factor 2 (sand-clay excess), only Pilargiidae related to the clay habitat.

Six families (Goniadidae, Nereididae, Cirolanidae, Terebellidae, Melitidae, and Stenothyridae) were associated with Factor 3 (Physico-chemical properties); only Goniadidae was negatively related with Factor 3.

Five families (unidentified species in Bivalvia, Goniadidae, Terebellidae, Aoridae, and Retusidae) were associated with total suspended solids (TSS), only Aoridae being positively related with TSS.

Finally, eight families (Capitellidae, Nephtyidae, Oedicerotidae, Apseudidae, Alpheidae, Marginellidae, Skeneopsidae, and Tellinidae) showed no evidence of any environmental predictors.

Goniadidae and Terebellidae could be predicted by three environmental factors (r-squared statistics of 31% and 44%, respectively). Nereididae, Aoridae, Pseudotanaidae, Cirolanidae, and unidentified species in Bivalvia could be predicted by two factors, having r-squared values ranging from 16% to 34%. Hesionidae, Pilargiidae, Pholoidae, Spionidae, Isaeidae, Melitidae, Anthuridae, Stenothyridae, and Retusidae could be predicted by one factor, having r-squared values ranging from 12% to 32%.

The observed densities and fitted values from model fitting are plotted by the periods (Figure 4.4) and by station (Figure 4.5). Capitellidae, Nephtyidae, Oedicerotidae, Apseudidae, Marginellidae, and Tellinidae could be widespread for each observed period and station. Whereas, Alpheidae and Skeneopsidae could not be predicted because the data was underestimate.
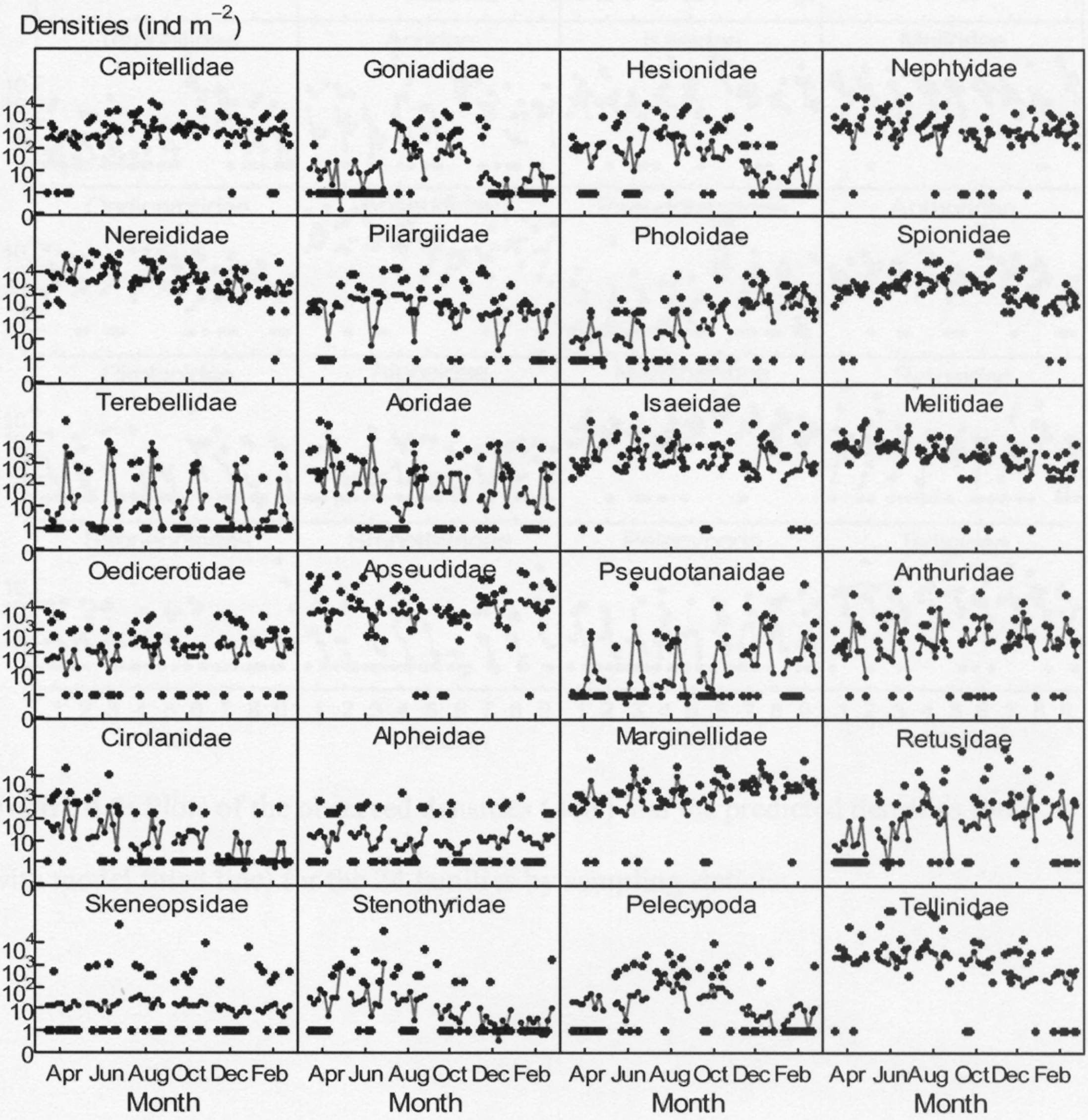
Figure 4.4: Plots of the observed densities (dots) and the predicted densities (dots with model fitted line) for the 24 families by the periods of observation
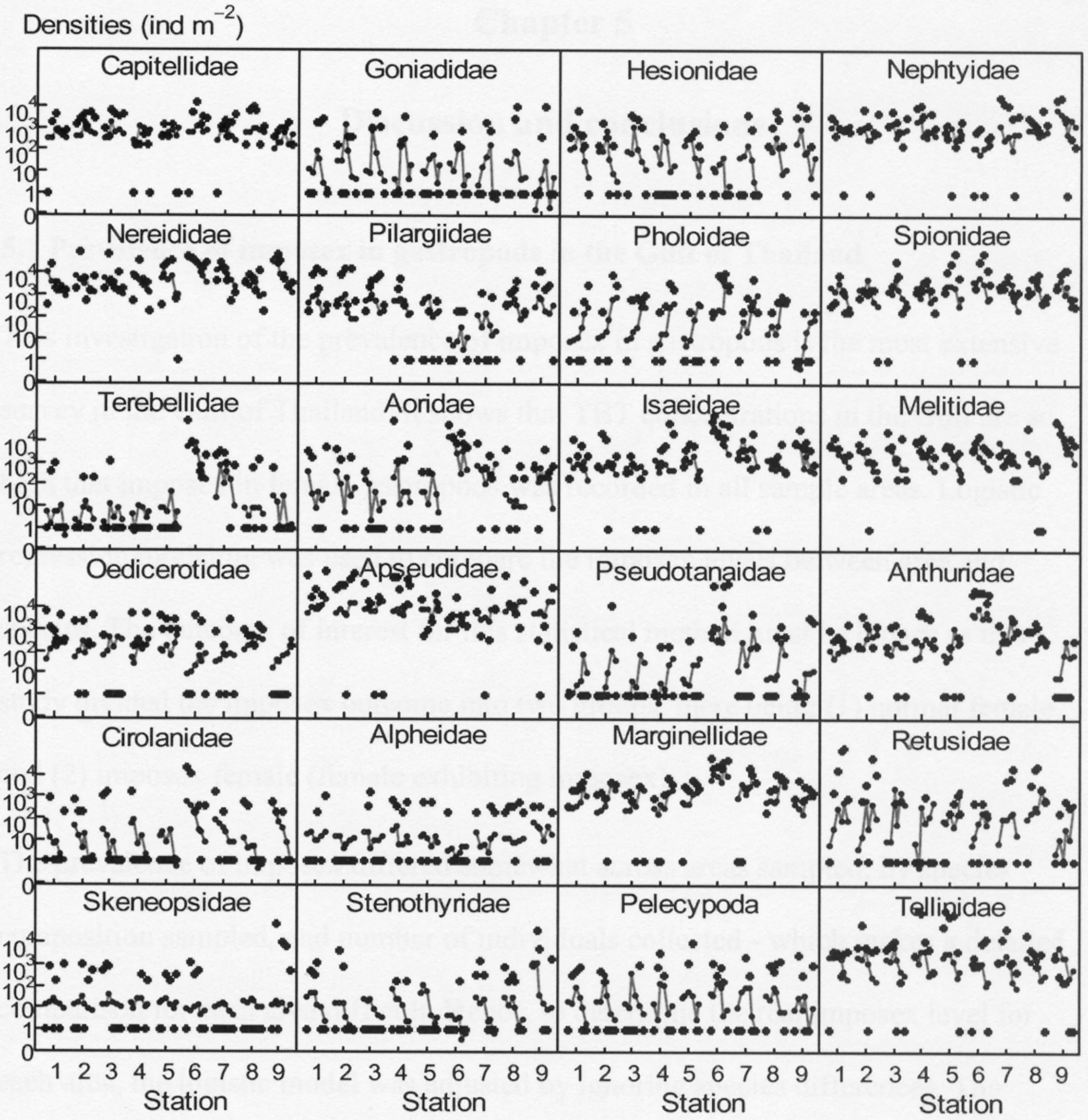
Figure 4.5: Plots of the observed densities (dots) and the predicted densities (dots with model fitted line) for the 24 families by sampling stations

Chapter 5

Discussion and conclusions

5.1 Prevalence of imposex in gastropods in the Gulf of Thailand

This investigation of the prevalence of imposex in gastropods is the most extensive survey in the Gulf of Thailand. It shows that TBT concentrations in the Gulf are so high that imposex in female gastropods was recorded in all sample areas. Logistic regression modeling was used to compare the imposex levels between area and species. The outcome of interest for this statistical method must be binary as this study divided the imposex outcome into two groups, there being (1) normal female, and (2) imposex female (female exhibiting imposex).

The prevalence of imposex differed somewhat across areas sampled, by species composition sampled, and number of individuals collected - which makes a detailed comparison for each area difficult. Hence, to determine the real imposex level for each area, the logistic model was adjusted by ignoring species differences. The findings showed that the highest imposex levels occurred at the eastern part of the Bight of Bangkok. This area had the largest port, the most intensive boating activity, and the biggest far seas commercial vessel harbor in the Gulf of Thailand (Kan-Atireklap et al 1997, Bhatt et al 2006). On the other hand, other areas of concern were found in many locations along the coastal line in Pattani province. This could be related to dumping of highly polluted harbour sediments in the shallow coastal area (see more details, Publication 1, in Appendix 1).

Since we considered gastropod females with stage 1 and 2 of imposex as showing no imposex, this might be have caused underestimation in areas B (Namrin), G (Songkhla), L (Panare), and M (Tak Bai), see more details, Publication 3, in Appendix 3.

The suspicion that the local species may differ in sensitivity to TBT (Swennen et al 1997) has been strengthened by this study, which involved larger samples at each site. In order to study imposex prevalence by species, a simplified model was used that ignored areas. The results indicated that *Morula musiva*, *P. cochlidium*, and *B. areolata* seem less sensitive for developing imposex while *T. lacera*, *M. occa*, and *L. blosvillei* seem most sensitive among the species studied (see more details, Publication 1, in Appendix 1).

5.2 Macrobenthic fauna density in the Middle Songkhla Lake

This study used data collected for a previous study to develop a statistical regression model for data analysis. The data structure is one commonly used in ecology and involves measurements of densities of aquatic organisms collected on various occasions at various locations, together with measurements of levels of environmental determinants taken at corresponding times and places. The objective of this second study is to account for variation in organism abundances using the environmental determinants. Multivariate multiple regression (MMR) was used to model these multiple outcomes and multiple determinants. The advantages of its application are that (a) it separates the effects of observed environmental predictor variables on density outcomes; (b) it gives standard errors for these estimated effects, thus providing a firmer statistical basis for clustering; and (c) it provides a predictive model for the outcomes.

To reduce the problem of outcome with too many zeros, we used the data at the family level instead of using species level. Hence, the outcome variables were the log-transformed densities of the 24 families of macrobenthic fauna selected as having the most coverage (seen on at least 35% of occasions) from 81 families observed at 9 sites during 6 bimonthly periods.

The environmental variables (12 physico-chemical characteristics of water and sediments) were defined as environmental factors. In our second study, factor analysis was used to represent a set of variables by a smaller number of variables (called "factor"). The analysis gave three main factors (Factor 1: mainly salinity, Factor 2: sand-clay excess, and Factor 3: physico-chemical properties of sediments and water) and also a high uniqueness variable (TSS) that were allowed to model fitting as the multiple predictor variables also.

Of the 96 regression coefficients in the regression model (24 families × 4 predictor variables), 8 were statistically significant at the 0.01 level and 17 were significant at the 0.05 level.

These environmental factors and total suspended solids were all found to be associated with the densities of polychaetes, crustaceans and mollusks. The finding is consistent with results reported for a coastal lagoon in Ghana (Lamptey and Armah 2008), where salinity, percent clay, pH and turbidity were significant variables structuring the macrobenthic fauna.

The salinity factor was positively associated with the abundance of Goniadidae, Hesionidae, Spionidae, and unidentified species in Bivalvia, while negatively associated with Pholoidae and Pseudotanaidae. In general, salinity is an important factor affecting the distribution of macrobenthic fauna in estuary. Although the Middle Songkhla Lake is not connected to the sea directly, this zone receives the effect of salinity from the saltwater inflow through the Lower Lake which is open to the Gulf of Thailand. Salinity is often regarded as a primary descriptor in estuarine ecosystems (Gaston 1988, Lamptey and Armah 2008).

A sedimentary habitat contains information mirroring the functional biodiversity and activity patterns of macrobenthic fauna (Rosenberg et al 2009). The main characteristics at the bottom of the Middle Songkhla Lake are clay and silt (Angsupanich et al 2005a) except for station six, which is mainly sand (84.6%). We found that sand/clay excess was positively associated with the densities of Pseudotanaidae, Anthuridae, Isaeidae, Aoridae, Terebellidae, Cirolanidae, and Nereididae at station six. In contrast, sand-clay excess was associated with low densities of Pilargiidae at station six.

A typical genus *Sigambra* within Pilargiidae (Angsupanich et al 2005a), was found to be negatively related with sand-clay excess, a finding supported by a study in the southeastern Gulf of California reporting that *Sigambra* was dominant in the areas where sand percentage was 1%, or mud was 60-70% (Méndez 2007). The dominant genus *Cyathura* within Anthuridae (Angsupanich et al 2005a) was found to be positively related with sand habitat, in agreement with an evaluation of the physical and biological conditions of sea bed habitats at the Charleston Offshore Dredge Material Disposal Site, South Carolina (Zimmerman et al 2002). In addition, and Salazar-Vallejo 2003) but also in brackish water such as occurs in the Middle Songkhla Lake. Fourteen species of Nereididae were reported in a former study (Angsupanich et al 2005a) and it seems that Nereididae is widespread in the Middle Songkhla Lake, where it had the highest species richness. No evidence of Nereididae variation with salinity was found, possibly due to species diversity within this family. Some species, such as *Ceratonereis hircinicola*, were widely spread in the high salinity areas (Angsupanich and Kuwabara 1995), whereas *Namalycastis indica* has been found to inhabit fresh to slightly brackish water in cisterns, pools and lagoons (Glasby 1999). As for Nephtyidae, which was not correlated with any factors, it can survive in a wide range of salinity and environmental conditions, and thus may be relatively insensitive to the environmental factors considered in this study.

This study has provided information regarding the influence of environmental conditions and seasonal variations on the macrobenthic fauna assemblages of the Middle Songkhla Lake. A data analysis procedure for monitoring associations has been provided, and this method can be used generally in estuaries to determine the complex interaction between biotic and abiotic factors. The results clearly showed relationships between macrobenthic fauna density and the major environmental factors of salinity, sand-clay excess, total nitrogen and total suspended solids.

This knowledge is useful for the natural resource management that needs to be conducted on this lake in the future, because the Songkhla Lake nowadays suffers from the use of coastal land and water resources for uncontrolled shrimp farming, the destruction of both mangrove areas and peat swamp forest, construction of intake and outfall structures, and the construction of a deep sea port (Chufamanee et al 2003).

References

Abdi, H. 2003. Factor rotations in factor analyses. In Lewis-Beck, M., Bryman, A., and Futing, T. (Editors), Encyclopedia of Social Sciences Research Methods. Thousand Oaks, CA, Sage, pp. 1-8.

Angsupanich, S. 2007. Songkhla Lake - Why is it important? Available at: http://www.khlong-u-taphao.com. [Accessed date: August 20, 2007].

Angsupanich, S., and Kuwabara, R. 1995. Macrobenthic fauna in Thale Sap Songkhla, a brackish lake in southern Thailand. *Lakes and Reservoirs: Research and Management*, 1(2): 115-125.

Angsupanich, S., and Siripech, A. 2001. Role and distribution of the dominant benthic fauna, *Apseudes sapensis* Chilton 1926 (Crustacea: Tanaidacea) in Songkhla Lake, southern Thailand. *Songklanakarin Journal of Science and Technology*, 23(4): 515-525.

Angsupanich, S., Siripech, A., and Charoenpornthip, M. 2005a. Macrobenthic fauna community in the Middle Songkhla Lake, Southern Thailand. *Songklanakarin Journal of Science and Technology*, 27: 365-390.

Angsupanich, S., Somsak, S., and Phrommoon, J. 2005b. Stomach contents of the catfishes *Osteogeneiosus militaris* (Linnaeus, 1758) and *Arius maculatus* (Thunberg, 1792) in the Songkhla Lake. *Songklanakarin Journal of Science and Technology*, 27: 391-402.

Armah, K.A., Ababio, D.S., and Darpaah, A.G. 2005. Spatial and temporal variations in water physicochemical parameters in the South-Western of sector of the Keta Lagoon, Gnana. Proceedings of the 14th Biennial Coastal Zone Conference (July 17-21, 2005), New Orleans, Louisiana.

Bech, M. 1999. Increasing levels of tributyltin-induced imposex in muricid gastropods at Phuket Island, Thailand. *Applied Organometallic Chemistry*, 13: 799-804.

Bech, M. 2002. A survey of imposex in muricids from 1996 to 2000 and identification of optimal indicators of tributyltin contamination along the east coast of Phuket Island, Thailand. *Marine Pollution Bulletin*, 44: 887-896.

Bergen, M., Cadien, D., Dalkey, A., Montagne, D.E., Smith, R.W., Stull, J.K., Velarde, R.G., and Weisberg, S.B. 2000. Assessment of benthic infaunal condition on the mainland shelf of Southern California. *Environmental Monitoring and Assessment*, 64: 421-434.

Bhatt, N., Chockanapitaksa, P., Heijden, A., Hirunwatsiri, W., Illangovan, P., Phetmanee, T., Pinnoi, N., Shah, J., Tansanguanwong, P., Vichittrananda, S., and Unkulvasapaul, M. 2006. *Thailand environment monitor 2006.* The International Bank for Reconstruction and Development, World Bank.

Blaber, S.J.M. 1970. The occurrence of a penis-like outgrowth behind the right tentacle in spent females of *Nucella lapillus* (L.). *Proceedings of the Malacological Society of London*, 39: 231-233.

Borja, A., Franco, J., and Pérez, V. 2000. A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments. *Marine Pollution Bulletin*, 40: 1100-1114.

Browne, M.W. 2001. An overview analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research*, 36: 111-150.

Bryan, G.W., Gibbs, P.E., Burt, G.R., and Hummerstone, L.G. 1987. The effects of tributyltin (TBT) accumulation on adult dog-whelks, *Nucella lapillus*: long-term field and laboratory experiments. *Journal of the Marine Biological Association UK*, 67: 525-544.

Bryan, G.W., Gibbs, P.E., Huggett, R.J., Curtis, L.A., Bailey, D.S., and Dauer, D.M. 1989. Effects of tributyltin pollution on the mud snail, *Ilyanassa obsoleta*, from the York River and Sarah Creek, Chesapeake Bay. *Marine Pollution Bulletin*, 20: 458-462.

Chongprasith, P., and Praekulvanich, E. 2003. Coastal pollution management in Thailand. Poster Papers, Diffuse Pollution Conference, Dublin, pp. 77-83.

Chufamanee, P., Boromthanarat, S., and Lønholdt, J. 2003. Case story 1: partnership policy in Songkhla Lake basin, a case experience on integrated environment and water management towards people's livelihoods. Report from the Songkhla workshop on linking management of catchment and coastal ecosystems. Royal Thai ministry of natural resources and environment and the United Nations Environment Program. Songkhla, Thailand, pp. 15-32.

Clarke, K.R., and Warwick, R.M. 1994. *Change in marine communities: an approach to statistical analysis and interpretation.* UK: Plymouth Marine Laboratory.

Clarke, K.R., Somerfield, P.J., and Chapman, M.G. 2006. On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages. *Journal of Experimental Marine Biology and Ecology*, 330: 55-80.

Dauer, D.M., Ranasinghe, J.A., and Weisberg, S.B. 2000. Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. *Estuaries*, 23(1): 80-96.

de Mora, S.J., Stewart, C., and Phillips, D. 1995. Sources and rate of degradation of Tri (n-butyl) tin in marine sediments near Auckland, New Zealand. *Marine Pollution Bulletin*, 30: 50-57.

Engle, V.D., Summers, J.K., and Gaston, G.R. 1994. A benthic index of environmental condition of Gulf of Mexico estuaries. *Estuaries*, 17: 372-384.

Fent, K. 2005. Worldwide occurrence of organotins from antifouling paints and effects in the aquatic environment. *The Handbook of Environmental Chemistry*, 5: 71-100.

Féral, C., and Le Gall, S. 1983. The influence of a pollutant factor (tributyltin) on the neuroendocrine mechanism responsible for the occurrence of a penis in the females of Ocenebra erinacea. In Lever, J., and Boer, H.H. (Editors), *Molluscan neuro-endocrinology*. North Holland Publ Co, Amsterdam, pp 173-175.

Frédou, T., Ferreira, B.P., and Letourneur, Y. 2006. A univariate and multivariate study of reef fisheries off northeastern Brazil. *Journal of Marine Science*, 63: 883-896.

Gaston, G.R., Lee, D.A.L., and Nasci, J.C. 1988. Estuarine macrobenthos in Calcasieu Lake, Louisiana: community and trophic structure. *Estuaries and Coasts*, 11: 192-200.

Gibbs, P.E., and Bryan, G.W. 1986. Reproductive failure in populations of the dog-whelk, *Nucella lapillus*, caused by imposex induced by tributyltin from antifouling paints. *Journal of the Marine Biological Association UK*, 66: 767-777.

Gibbs, P.E., Bryan, G.W., Pascoe, P.L., and Burt, G.W. 1987. The use of the dog-whelk, Nucella lapillus, as an indicator of tributyltin (TBT) contamination. *Journal of the Marine Biological Association UK*, 67: 507-523.

Glasby, C.H. 1999. The Namanereidinae (Polychaeta: Nereididae). Part 1, Taxonomy and phylogeny. Records of the Australian Museum, pp. 1-129.

Gonzalez-Escalante, L.E., and Salazar-Vallejo, S.I. 2003. A new estuarine species, *Nereis garwoodi* (Polychaeta: Nereididae), from Bahia Chetumal, Macican Caribbean coast. *Revista de Biología Tropical*, 51: 155-164.

Hagger J.A., Norris, R.H., Hogue, J.N., and Galloway, T.S. 2006. Is there a casual association between genotoxicity and the imposex effects? *Environmental Health Perspectives*, 114(1): 2026.

Hagger, J.A., and Galloway, T.S. 2009. Biological markers of exposure and effect for water pollution monitoring. In Gonzalez, C., Quevauviller, P., and Greenwood, R., (Editors), *Rapid chemical and biological techniques for water monitoring*. ISBN: 978-0-470-05811-4, John Wiley & Sons Ltd., pp. 221-239.

Hawkins, C.P., Norris, R.H., Hogue, J.N., and Feminella, J.W. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. *Ecological Applications*, 10: 1456-1477.

Hosmer, D.W., and Lemeshow, S. 2000. *Applied logistic regression.* The second edition. John Wiley and Sons, Inc.

Hyland, J.L., Balthis, W.L., Engle, V.D., Long, E.R., Paul, J.F., Summers, J.K., and Van Dolah, R.F. 2003. Incidence of stress in benthic communities along the U. S. Atlantic and Gulf of Mexico Coasts within different ranges of sediment contamination from chemical mixtures. *Environmental Monitoring and Assessment*, 81: 149-161.

Johnson, R.A., and Wichern, D.W. 1998. *Applied multivariate statistical analysis.* The fourth edition. Prentice-Hall, USA, 816 pp.

Johnston, D.M. 1998. *SEAPOL Integrated studies of the Gulf of Thailand (Volume 1).* Southeast Asian Programme in Ocean Law, Policy and Management (SEAPOL).

Joy, M.K., and Death, R.G. 2000. Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 34: 241-252.

Kan-Atireklap, S., Tanabe, S., and Sanguansin, J. 1997. Contamination by butyltin compounds in sediments from Thailand. *Marine Pollution Bulletin*, 34: 894-899.

Khongchouy, N., and Sampantarak, U. 2010. Confidence intervals for adjusted proportions using logistic regression. *Modern Applied Science*, 4(6).

Kleinbaum, D.G., and Klein, M. 2002. *Logistic regression: a self-learning text.* The second edition. Springer-Verlag, New York, Inc.

Lamptey, E., and Armah, A.K. 2008. Factors affecting macrobenthic fauna in a tropical hypersaline coastal lagoon in Ghana, West Africa. *Estuaries and Coasts*, 31: 1006-1019.

Law, R., Hustwayte, G., and Sims, D. 2005. Monitoring of the quality of the marine environment, 2002-2003. Centre for Environment, Fisheries and Aquaculture Science. Lowestoft. *Aquatic Environment Monitoring*, 57.

Layman, C.A., and Silliman, B.R. 2002. Preliminary survey and diet analysis of juvenile fishes of an estuarine creek on Andros Island, Bahamas. *Bulletin of Marine Science*, 70: 199-210.

Mardia, K.V., Kent, J.T., and Bibby, J.M. 1979. *Multivariate analysis*. Academic Press, London and New York.

Melwani, A., and Thompson, B. 2007. The Influence of chemical and physical factors on macrobenthos in the San Francisco Estuary. San Francisco Estuary Institute, Oakland, CA. *Contribution*, 543: 1-42.

Méndez, N. 2007. Relationships between deep-water polychaete fauna and environmental factors in the southeastern Gulf of California, Mexico. *Scientia Marina*, 71: 605-622.

Mensink, B.P., Everaarts, J.M., Kralt, H., Ten Hallers-Tjabbes, C.C., and Boon, J. 1996. Tributyltin exposure in early life stages induces the development of male characteristics in the common whelk, *Buccinum undatum. Marine Environmental Research*, 42: 151-154.

Mensink, B.P., Kralt, H., Vethaak, A.D., Ten Hallers-Tjabbes, C.C., Koeman, J.H., van Hattum, B., and Boon, J.P. 2002. Imposex induction in laboratory reared juvenile *Buccinum undatum* by tributyltin (TBT). *Environmental Toxicology and Pharmacology*, 11: 49-65.

Murrell, P. 2006. *R graphics*. Florida: Chapman & Hall/CRC.

Nakatsu, Y., Kotake, Y., and Ohta, S. 2006. Tributyltin-induced cell death is mediated by calpain in PC12 cells. *Neurotoxicology*, 27(4): 587-593.

Nishikawa, J. 2006. Imposex in marine gastropods may be caused by binding of organotins to retinoid X receptor. *Marine Biology*, 149: 117-124.

Nizzoli, D., Castaldelli, G., Bartoli, M., Welsh, D.T., Gomez, P.A., Fano, A.E., and Viaroli, P. 2002. Benthic fluxes of dissolved inorganic nitrogen in a coastal lagoon of the Northern Adriatic Sea: an interpretation of spatial variability based on sediment features and infauna activity. *Marine Ecology*, 23(1): 297-306.

Oehlmann, J., Stroben, E., Bettin, C., and Fioroni, P. 1993. Hormonal disorders and tributyltin-induced imposex in marine snails. In *Quantified phenotypic responses in morphology and physiology*. Proceedings of the 27th European Marine Biology Symposium (September 7-11, 1992), Dublin Ireland, pp. 301-305.

Olson, C.L. 1976. On choosing a test statistic in multivariate analyses of variance. *Psychological Bulletin*, 83: 579-586.

Pearson, T.H., and Rosenberg, R. 1978. Macrobenthic succession in relation to organic enrichment and pollution of the marine environment. *Oceanography and Marine Biology: An Annual Review*, 16: 229- 311.

Pillay, D., and Perissinotto, R. 2009. Community structure of epibenthic meiofauna in the St. Lucia Estuarine Lake (South Africa) during a drought phase. *Estuarine, Coastal and Shelf Science*, 81: 94-104.

Predalumpaburt, Y., and La-ongsiriwong, N. 1997. Variations and relationship between sediment qualities and macrobenthos in Songkhla Lake. Technical paper No. 3/1997. National Institute of Coastal Aquaculture (NICA), Songkhla, Thailand.

R Development Core Team. 2009. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http: //www.R-project.org.

Ranasinghe, J.A., Thompson, B., Smith, R.W., Lowe, S., and Schiff, K.C. 2004. Evaluation of benthic assessment methodology in Southern California Bays and San Francisco Bay. Southern California Coastal Water Research Project. Technical Report, 432: 1-70.

Rosenberg, R., Magnusson, M., and Nilsson, H.C. 2009. Temporal and spatial changes in marine habitats in relation to EU Framework Directive: The use of sediment profile imagery. *Marine Pollution Bulletin*, 58: 565-572.

Ruensirikul, J., Angsupanich, S., and Phongdara, A. 2007. Abundance and diversity of amphipod crustaceans in the Upper Songkhla Lagoon. *Songklanakarin Journal of Science and Technology*, 29: 1225-1249.

Rygg, B. 1985. Distribution of species along pollution-induced diversity gradients in benthic communities in Norwegian fjords. *Marine Pollution Bulletin*, 16: 469-474.

Shi, H.H., Huang, C.J., Zhu, S.X., Yu, X.J., and Xie, W.Y. 2005. Generalized system of imposex and reproductive failure in female gastropods of coastal waters of mainland China. *Marine Ecology Progress Series*, 304: 179-189.

Smith, B.S. 1981. Tributyltin compounds induce male characteristics on female mud snails, *Nassarius obsoletus = Iliyanassa obsoleta*. *Journal of Applied Toxicology*, 1: 141-144.

Summers, J.K. 2001. Ecological condition of the estuaries of the Atlantic and Gulf Coasts of the United States. *Environmental Toxicology and Chemistry*, 20(1): 99-106.

Swennen, C., Moolenbeek, R.G., Ruttanadakul, N., Hobbelink, H., Dekker, H., and Hajisamae, S. 2001. The molluscs of the southern Gulf of Thailand. *Thai Studies in Biodiversity*, 4: 1-210.

Swennen, C., Ruttanadakul, N., Ardseungnern, S., Singh, H.R., Mensink, B.P., and Ten Hallers-Tjabbes, C.C. 1997. Imposex in sublittoral and littoral gastropods from the Gulf of Thailand and Strait of Malacca in relation to shipping. *Environmental Technology*, 18: 1245-1254.

Swennen, C., Sampantarak, U., and Ruttanadakul, N. 2009. TBT-pollution in the Gulf of Thailand: a re-inspection of imposex incidence after 10 years. *Marine Pollution Bulletin*, 58: 526-532.

UNESCO. 1997. IOC/ WESTPAC Workshop on co-operative study in the Gulf of Thailand: a science plan in Bangkok, Thailand (25-28 February 1997). Intergovernmental Oceanographic Commission. IOC Workshop Report, 141: 1-36.

Venables, W.N., and Ripley, B.D. 2002. *Modern applied statistics with S.* The fourth edition. New York, USA: Springer Science + Business Media.

Vitaliano, J.J., Fromm, S.A., and Guida, V.G. 2006. Macrobenthos (polychaetes, crustaceans, mollusks, other). In Link, J.S., Griswold, C.A., Methratta, E.T., and Gunnard, J., (Editors), *Documentation for the energy modeling and analysis exercise (EMAX).* US Dep. Commer., Northeast Fisheries Science Center Reference Document 06-15, pp. 37-46.

Vos, J.G., Dybing, E., Greim, H.A., Ladefoged, O., Lambré, C., Tarazona, J.V., Brandt, I., and Vethaak, A.D. 2000. Health effects of endocrine-disrupting chemicals on wildlife, with special reference to the European situation. *Critical Reviews in Toxicology*, 30(1): 71-133.

Wallace, O. 2009. What is a gastropod? Available at: http://www.wisegeek.com/what-is-a-gastropod.htm. [Accessed date: November 5, 2009].

Warton, D.I. 2005. Many zeros does not many zero inflation: comparing the goodness-of-fit of parametric model to multivariate abundance data. *Environmentrics* 16: 275-289.

Warwick, R.M. 1986. A new method for detecting pollution effects on marine macrobenthic communties. *Marine Biology*, 92: 557-562.

Warwick, R.M., and Clarke, K.R. 1995. Multivariate measures of community stress and their application to marine pollution studies in the East Asian region. *Phuket Marine Biological Center Research Bulletin*, 60: 99-113.

Weisberg, S.B., Ranasinghe, J.A., Dauer, D.M., Schaffner, L.C., Díaz, R.J., and Frithsen, J.B. 1997. An estuarine benthic index of biotic integrity (B-BI) for Chesapeake Bay. *Estuaries*, 20: 149-158.

Zimmerman, L.E., Jutte, P.C., and Van Dolah, R.F. 2002. An environmental assessment of the Charleston Ocean dredged material disposal site and surrounding areas: physical and biological conditions after partial completion of the Charleston Harbor Deepening Project. Marine Resources Research Institute, South Carolina Department of Natural Resources: pp. 1-72.

Appendix 1

Publication 1: TBT-pollution in the Gulf of Thailand: a re-inspection of imposex incidence after 10 years

Marine Pollution Bulletin 58 (2009) 526–532

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: www.elsevier.com/locate/marpolbul

TBT-pollution in the Gulf of Thailand: A re-inspection of imposex incidence after 10 years

Cornelis Swennen *, Uraiwan Sampantarak, Nukul Ruttanadakul

Faculty of Science and Technology, Prince of Songkla University, Pattani 94000, Thailand

ARTICLE INFO

Keywords:
Tributyltin
Imposex
Neogastropods
Gulf of Thailand
Southeast Asia

ABSTRACT

Imposex in neogastropods was used to determine the relative TBT distribution in the Gulf of Thailand in 2006. To identify the imposex prevalence, 8757 specimens, belonging to 22 species from five families of neogastropods, were collected at 56 sites in 13 sample areas. These areas were located between the Bight of Bangkok in the north and the Malaysian border in the south. A contamination model was developed to compensate for differences in TBT sensitivity among species and to make comparisons among areas. At every area imposex was found in neogastropods. The highest incidence was in the east side of the Bight of Bangkok off Si Racha and Pattaya and in the southern part around Pattani. The same areas showed the highest frequency of imposex in 1996. While the frequency of imposex appeared to have slightly decreased in these areas, increases were found elsewhere. In Pattani Province this could be related to dumping of highly polluted harbour sediments in the shallow coastal area. The overall frequency of imposex in the Gulf of Thailand significantly increased from 1996 to 2006.

1. Introduction

In the late 1960s, tributyltin (TBT) came in use as biocide in anti-fouling paints on ships and installations at sea. TBT provides a long lasting protection against fouling organisms such as barnacles, mussels, hydrozoans and algae. Adverse bio-impacts to organisms in the environment were not expected. However, gradually severe negative effects to organisms in non-target areas in the marine environment were reported over the next 40 years.

Some years after the introduction of TBT use, Blader (1970) discovered that females of the snail *Nucella lapillus* (Linnaeus, 1758) developed a small penis that was followed by a decline in the affected populations (Bryan et al., 1986, 1987). Subsequently, more snail species were discovered where females showed a small penis and pallial vas deferens. Additional research showed that these morphological changes were caused by TBT (Smith, 1981; Gibbs and Bryan, 1986; Oehlmann et al., 1993; Mensink et al., 1996). The phenomenon was first named masculinization or pseudo-hermaphroditism. However, the imposition of a mini-penis and mini pallial vas deferens in gastropods soon became commonly known as imposex. In the late 1970s, abnormalities in the shells of cultured oysters in Arcachon Bay cost the French oyster culture industry billions of US$, shown to be caused by TBT (Alzieu et al., 1981, 1986). Similar cases were found at other sites and countries (Phelps and Pace, 1997).

All these cases were close to harbours and large marinas. In particular yachts, moored for long periods in harbours, were suspected to increase TBT level in the water sufficiently to result in negative effects. This forced governments to ban the use of TBT containing anti-fouling paint on all vessels smaller than 25 m. France instituted this ban in 1982, followed by several other countries (Ten Hallers-Tjabbes, 1997). However, few localities have shown any positive effect of this ban (Dowson et al., 1993; Evans et al., 1996).

Identification of TBT (and other organotin compounds) in water is an expensive process when the distribution of TBT has to be established over large areas in the marine environment. The occurrence of imposex has commonly been accepted as a good bio-indicator of TBT pollution (Gibbs et al., 1987; Oehlmann et al., 1996). Monitoring of gastropods indicated that the abnormality is found on all continents close to harbours (Ellis and Pattisima, 1990; Horiguchi et al., 1991; Stewart et al., 1992). Most studies focused on gastropods that were collected by hand in or just below the intertidal zone. However, TBT also impacts gastropods in the open sea. The occurrence of imposex in *Buccinum undatum* Linnaeus, 1758 and the decline of this species in the North Sea could be related to shipping routes (Ten Hallers-Tjabbes et al., 1994). Similarly, imposex in 22 gastropod species appears to be related to shipping in south-east Asian waters (Strait of Malacca, Gulf of Thailand) up into the deepest possible sampling sites (Swennen et al., 1997). TBT has become a global pollution problem in marine habitats. This moved the Marine Environmental Protection Committee (MEPC) in 1996 to start discussions about a reduction of the use of TBT on large ships within five years with a total ban thereafter (Ten

---

* Corresponding author.
E-mail address: swennen@wxs.nl (C. Swennen).

0025-326X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.marpolbul.2008.11.028

Hallers-Tjabbes, 1997). The ban has been postponed, but in the meantime several vessels already no longer use anti-fouling paints containing TBT.

The present study is undertaken to determine if positive effects already occur in gastropods, specifically a decline in the frequency of imposex in the Gulf of Thailand. The main objectives were to: (1) establish the present levels of imposex in the area, (2) clarify possible differences in sensitivity among species, and (3) compare the results with data collected 10 years ago.

2. Materials and methods

The study was limited to key areas in the Gulf of Thailand (Fig. 1). Travel time and costs were reduced by dropping the Strait of Malacca from the sampling design, but this allowed collection of larger samples per site in the Gulf of Thailand. Neogastropods were collected from 56 sites in 13 areas in the Gulf of Thailand between June and November 2006. Specimens belonging to the neogastropod families Columbellidae and Volutidae were excluded, because absence of the imposex response has been reported in the former family (Gibbs et al., 1997), and females of the local genera *Cymbiola* and *Melo* of the Volutidae always show a small penis and vas deferens, (pseudo-imposex; Swennen and Horpet, 2008).

Samples were collected in both sublittoral and littoral environments. Specimens were hand-collected in the intertidal zone during low tide, but most were obtained from the by-catch of small, commercial fishing boats. The geographical positions of the sampling sites were determined by a hand-held GPS apparatus, while for those fishing trips in which we did not participate, the position was determined with help of the sea map on instruction of the skipper. The taxonomic nomenclature followed Swennen et al. (2001). A few species from the northern part not treated in Swennen et al. (2001) were identified by the curator of the Department of Malacology, University of Amsterdam (The Netherlands). The lengths (mm) of the gastropod shells were measured, and a part of the shell was removed to identify the sex and to check for imposex in females. Females with a short, curved penis and at least a part of a vas deferens were noted as showing imposex. An imposex index such as used by Mensink et al. (2002) was not used for avoiding errors, because the sizes of the specimens and the sizes and shapes of the male penis among the several species studied varied too much for correctly using the index on all species under local field conditions. Thus females with stage 1 and 2 were considered as showing no imposex. However, these stages were rarely found in specimens checked in the laboratory.

3. Statistical methods

Surveying a large area has the disadvantage that several habitats are included, each with a different assemblage of neogastropod species. No species was found at all sites. In the 1996 study, this imbalance was solved by calculating the imposex frequency over all females of species known to be sensitive for showing imposex. Since indications were found that species may differ in sensitivity

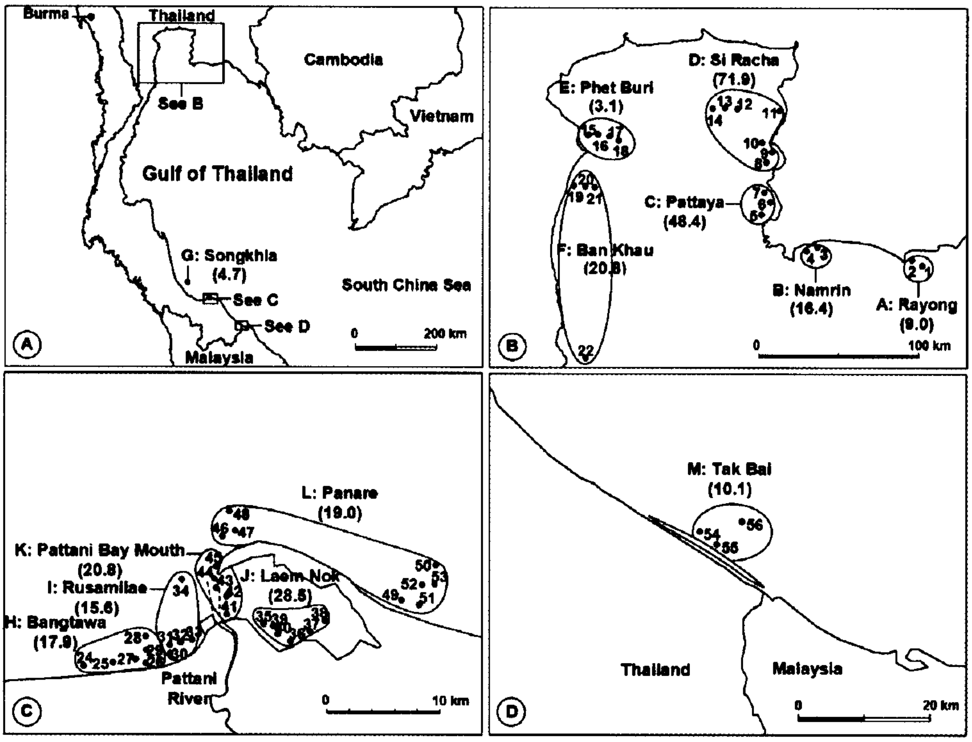

Fig. 1. The Gulf of Thailand with the study areas and sampling sites: A. overview; B. northern part of the Gulf of Thailand with study areas A–F; C. Pattani Bay and surroundings with study areas H–L; D. Narathiwat province with study area M. Dots show the sampling sites. Their number corresponds with the number in the detailed list appended to the online version. Value between brackets gives the adjusted imposex incidence in the area.

(Swennen et al., 1997), a contamination model is used that takes species into account for comparing imposex occurrence between areas.

The R statistical system (R Development Core Team, 2007; Venables and Ripley, 2002) was used for statistical model fitting. If p is the frequency of outcomes with a specific characteristic ("adverse" outcome) in a sample of size n, for large n, formula (1) is an asymptotically valid formula for the standard error. Formula (2) gives the 95% confidence interval (95% CI) for the frequency.

$$SE = \sqrt{\frac{p(1-p)}{n}} \tag{1}$$

$$p - 1.96 \times SE, p + 1.96 \times SE \tag{2}$$

For comparing two frequencies, we computed a p-value using Pearson's chi-squared test. We employed logistic regression (Hosmer and Lemeshow, 2000; Kleinbaum and Klein, 2002) to model the effects of multiple determinants on the occurrence of imposex. If there are two categorical determinants, and p_{ij} denotes the probability of the adverse outcome in categories i and j of these determinants, respectively, the simplest such model takes the additive form of formula (3). The method also gives a measure of the residual deviation between the model and the data and its corresponding number of degrees of freedom (df), defined as the number of cells into which the data are grouped minus the number of fitted parameters (Venables and Ripley, 2002 Chapter 7).

$$\ln\left(\frac{p_{ij}}{1 - p_{ij}}\right) = a_i + b_j \tag{3}$$

Formula (4) is used to express the frequency of the imposex outcome as a function of the two determinants. To eliminate redundant parameters, we took one of the categories of the second determinant as the referent and the corresponding parameter was set to 0. We fitted the model by using the maximum likelihood method with the observed data to estimate the parameters, providing standard errors for the fitted parameters. For each determinant, the method gives a p-value based on a chi-squared statistic for testing the null hypothesis that the outcome frequency is the same for each of its component categories. The method also gives separate p-values for comparing each parameter in the model with the mean (for the determinant of interest) or referent (for the covariate determinant).

$$p_{ij} = \frac{1}{1 + \exp(-a_i - b_j)} \tag{4}$$

Logistic regression provides a straightforward method for adjusting a frequency that varies with a determinant of interest for a covariate determinant. To calculate the adjusted frequency for category i of the determinant of interest, the term b_j in formula (3) is replaced by a constant b, giving formula (5). The value of the constant b is chosen to ensure that the sum of the adjusted number of adverse outcomes is equal to the observed number, as expressed in formula (6), where n_i is the sample size in category i of the determinant of interest. This method extends straightforwardly to additional covariates.

$$p_i^* = \frac{1}{1 + \exp(-a_i - b)} \tag{5}$$

$$\sum p_i^* n_i = \sum p_i n_i \tag{6}$$

4. Results

A total of 8757 snail specimens were collected in the Gulf of Thailand in 2006. The specimens comprised 22 species belonging to five families of neogastropods. The full names of the species

Table 1
Names of families and species of the neogastropods studied, with a summary of the size range of males and females, number of females, percentage of females, and percentage of females showing imposex in 2006.

Species and families	Sample size	Shell height (mm)		No. of females	Female (%)	Imposex (%)
		Males	Females			
Family Muricidae						
1. *Murex altispira* Ponder and Vokes, 1988	359	40–97	41–125	286	79.7	0.7
2. *Murex occa* Sowerby II, 1824	888	24–90	26–98	417	47.0	40.8
3. *Murex trapa* Röding, 1798	1156	34–118	31–125	698	60.4	22.9
4. *Chicoreus banksii* (Sowerby II, 1841)	58	44–81	37–84	39	67.2	12.8
5. *Lataxiena blosvillei* (Deshayes, 1832)	721	21–44	22–45	366	50.8	88.8
6. *Semiricinula muricoides* (De Blainville, 1832)	753	13–22.5	13–28	521	69.2	50.9
7. *Thais bituberculaiis* (Lamarck, 1822)	427	18–35	19–37	214	50.1	14.0
8. *Thais clavigera* (Kuester, 1860)	75	20–34	21–34	36	48.0	25.0
9. *Thais lacera* (Von Born, 1778)	587	15–57	15–70	306	52.1	21.9
10. *Thais rufotincta* Tan and Sigurdsson, 1996	22	15–26	18–22	10	45.5	10.0
11. *Morula musiva* (Kiener, 1835)	324	13–29	14–30	161	49.7	0.6
12. *Rapana rapiformis* (Von Born, 1778)	85	35–88	33–90	32	37.6	6.3
Family Buccinidae						
13. *Babylonia areolata* (Link, 1807)	314	38–64	20–70	158	50.3	1.3
14. *Nassaria pusilla* (Röding, 1798)	67	17.6–23	11–24	53	79.1	1.9
15. *Phos senticosus* (Linnaeus, 1758)	2	26	33	1	50.0	0.0
Family Nassariidae						
16. *Nassarius jacksonianus* (Quoy and Gaimard, 1833)	233	15–21	13–22	166	71.2	10.8
17. *Nassarius livescens* (Philippi, 1849)	62	16–29	16–25	29	46.8	6.9
18. *Nassarius siquijorensis* (A. Adams, 1852)	260	16–28	13–23	202	77.7	28.7
19. *Nassarius stolatus* (Gmelin, 1791)	533	11–22	13–23	306	57.4	4.9
Family Melongenidae						
20. *Pugilina cochlidium* (Linnaeus, 1758)	923	35–93	23–98	413	44.7	9.4
21. *Hemifusus ternatanus* (Gmelin, 1791)	368	38–105	27–131	225	61.1	15.6
Family Turridae						
22. *Turricula javana* (Linnaeus, 1767)	540	18–42	19–64	405	75.0	15.6
Total	8757			5044	57.6	25.2

Table 2
Results of the logistic model for imposex outcome among females in 2006.

Determinants		Coefficient	Standard error	*p*-Value
Area	A: Rayong	−2.119	0.389	<0.001
	B: Namrin	−1.429	0.369	<0.001
	C: Pattaya	0.138	0.231	0.551
	D: Si Racha	1.138	0.186	<0.001
	E: Phet Buri	−3.253	0.289	<0.001
	F: Ban Khau	−1.139	0.164	<0.001
	G: Songkla	−2.815	1.111	0.011
	H: Bang Tawa	−1.322	0.214	<0.001
	I: Rusamilae	−1.491	0.154	<0.001
	J: Laem Nok	−0.721	0.190	<0.001
	K: Pattani Bay Mouth	−1.140	0.230	<0.001
	L: Panare	−1.253	0.462	0.007
	M: Tak Bai	−1.987	1.251	0.112
Species	01: *Murex trapa* (referent)	0		
	02: *Murex altispira*	−2.961	0.834	<0.001
	03: *Murex occa*	0.697	0.186	<0.001
	04: *Lataxiena blosvillei*	1.674	0.271	<0.001
	05: *Semiricinula muricoides*	−0.215	0.211	0.308
	06: *Thais bituberculari*s	−0.494	0.289	0.087
	07: *Thais lacera*	0.097	0.203	0.633
	08: *Morula musiva*	−3.646	1.069	0.001
	09: *Babylonia areolata*	−2.398	1.243	0.054
	10: *Nassarius jacksonianus*	−0.700	0.291	0.016
	11: *Nassarius siquijorensis*	0.373	0.253	0.141
	12: *Nassarius stolatus*	−1.350	0.308	<0.001
	13: *Pugilina cochlidium*	−1.602	0.231	<0.001
	14: *Hemifusus ternatanus*	−1.504	0.247	<0.001
	15: *Turricula javana*	−0.323	0.208	0.120
	16: Other	−0.775	0.327	0.018

Table 3
Prevalence (%) of imposex by area before and after adjusting for species in the sampling of 2006.

Area	Total number of females	Number with imposex	Imposex (%)	Adjusted imposex (%)	
			(unadjusted)	Prevalence	95% CI
A: Rayong	190	9	4.7	9.0	4.9–13.0
B: Namrin	249	13	5.2	16.4	11.8–21.0
C: Pattaya	474	297	62.7	48.4	43.9–52.9
D: Si Racha	537	395	73.6	71.9	68.1–75.7
E: Phet Buri	417	13	3.1	3.1	1.4–4.7
F: Ban Khau	286	53	18.5	20.8	16.1–25.5
G: Songkla	171	1	0.6	4.7	1.5–7.8
H: Bang Tawa	735	128	17.4	17.9	15.1–20.7
I: Rusamilae	916	143	15.6	15.6	13.2–17.9
J: Laem Nok	507	157	31.0	28.5	24.5–32.4
K: Pattani Bay Mouth	266	50	18.8	20.8	15.9–25.6
L: Panare	140	9	6.4	19.0	12.5–25.4
M: Tak Bai	156	2	1.3	10.1	5.4–14.8

Table 4
Prevalence (%) of imposex per species before and after adjusting for areas in the sampling of 2006.

Species groups	Total females	No. of imposex	Imposex (%) unadjusted	Adjusted imposex (%)	
				Prevalence	95% CI
1. *Murex trapa*	698	160	22.9	30.1	26.7–33.5
2. *Murex altispira*	286	2	0.7	2.2	0.5–3.9
3. *Murex occa*	417	170	40.8	46.4	41.6–51.2
4. *Lataxiena blosvillei*	366	325	88.8	69.7	65.0–74.4
5. *Semiricinula muricoides*	521	265	50.9	25.8	22.0–29.6
6. *Thais bituberculari*s	214	30	14.0	20.8	15.4–26.3
7. *Thais lacera*	306	67	21.9	32.2	27.0–37.4
8. *Morula musiva*	161	1	0.6	1.1	0.0–2.7
9. *Babylonia areolata*	158	2	1.3	3.8	0.8–6.7
10. *Nassarius jacksonianus*	166	18	10.8	17.6	11.8–23.4
11. *Nassarius siquijorensis*	202	58	28.7	38.5	31.8–45.2
12. *Nassarius stolatus*	306	15	4.9	10.1	6.7–13.4
13. *Pugilina cochlidium*	413	39	9.4	8.0	5.4–10.6
14. *Hemifusus ternatanus*	225	35	15.6	8.7	5.1–12.4
15. *Turricula javana*	405	63	15.6	23.8	19.6–27.9
16 Other*	200	20	10.0	16.6	11.4–21.7

* *Chicoreus banksii*, *Thais clavigera*, *Thais rufotincta*, *Rapana rapiformis*, *Nassaria pusilla*, *Phos senticosus* and *Nassarius livescens* combined.

and families are presented in Table 1, which also summarises the size range of males and females, number of females, sex ratio, and percentage of females showing imposex. A list of the 56 sampling sites with location, depth, species, number and details of specimens per site is added to the online version.

The logistic regression model (formula (3)) for the imposex frequency used the 13 areas as the determinant of interest and the 16 species groups as the covariate. Species with fewer than 100 individuals were combined into a single group referred to as "other". The logistic regression model gave a residual deviance of 171.4 with 55 df based on the grouped data (the 84 combinations of area and species where samples were obtained). The chi-squared statistics for testing a common imposex frequency between areas and species groups were found to be 438.9 (12 df) and 453.0 (15 df) respectively, with corresponding highly statistically significant *p*-values (Table 2). From the results (Table 2), the adjusted imposex frequency was calculated using formula (5) with the corresponding 95% confidence intervals (formula (2)). Imposex frequency before and after adjusting for species is given in Table 3. When all species were combined, areas C and D had a high (>50%) imposex frequency, areas F, H, I, J, and K had medium (15–49%) frequency, whereas A, B, E, G, L, and M had low (<15%) frequency. After adjusting for the different sensitivities of the species, C and D still

Table 5
Number of female gastropods showing imposex and total number studied (imposex/ females) per area in the Gulf of Thailand in 1996 (Cited from Swennen et al., 1997).

Species	Area*								
	C	D	E	F	I	J	K	L	M
Murex trapa	–	57/59	–	–	–	–	–	–	–
Thais lacera	–	–	–	–	–	–	1/2	–	–
Babylonia areolata	13/15	–	0/27	0/28	1/42	1/1	–	–	0/71
Nassarius stolatus	–	–	–	–	–	–	1/1	–	–
Hemifusus ternatanus	6/7	3/5	–	–	–	1/3	1/5	–	1/34
Thais hippocastaneum	–	–	–	–	–	–	–	0/139	–
Murex sp.	28/136	19/19	–	0/28	–	–	–	–	0/11
Other	1/1	5/10	–	–	–	7/13	5/22	–	0/29
Total	48/159	84/93	0/27	0/56	1/42	9/17	8/30	0/139	1/145
Crude prevalence (%)	30.2	90.3	0.0	0.0	2.4	52.9	24.1	0.0	0.7
Adjusted prevalence (%)	33.5	84.8	–	–	0.3	57.8	26.7	–	0.1

* C: Pattaya; D: Si Racha; E: Phet Buri; F: Ban Khau; I: Rusamilae; J: Laem Nok; K: Pattani Bay Mouth; L: Panare; M: Tak Bai.

showed the highest values, but B and L moved up to the medium severity level of imposex frequency (Table 3). The location of the areas and the sampling sites are shown in Fig. 1 where the mean frequency is also given per area.

Sensitivity to imposex varied among species, *Lataxiena blosvillei* was found to be most prone to imposex (mean 69.7 ± 4.7%; 95% CI), whereas *Morula musiva* is least sensitive (mean 1.1 ± 1.6%; 95% CI) (Table 4). The 1996 data collected by Swennen et al. (1997) were grouped to the same areas as the present study to compare the results of this study with those of 1996 using the logistic regression model (Table 5). The results show an increase in five areas C (Pattaya), E (Pet Buri), F (Ban Khau), I (Rusamilae), and L (Panare) and a decrease in three areas D (Si Racha), J (Laem Nok), and K (Pattani Bay Mouth). The overall frequency significantly increased from 21.3% to 25.2% (Pearson chi-squared = 4.95 with 1 df, $p = 0.026$).

5. Discussion

Our investigation is currently the most extensive survey of imposex in the Gulf of Thailand. It shows that the TBT concentrations in the Gulf of Thailand are so high that imposex in female neogastropods was recorded in all 13 areas sampled. The overall imposex frequency has increased compared to 1996 (Swennen et al., 1997). Both studies identified the same locations with the highest occurrence of imposex, namely the eastern part of the Bight of Bangkok (C, D) followed by Pattani Province (H–K). Increases were also reported around Phuket in the Andaman Sea (Bech, 1999). Likewise, in the North Sea, the imposex frequency increased between surveys in 1991 and 1999 (Ten Hallers-Tjabbes et al., 2003).

The 1996 and 2006 studies differ somewhat in areas sampled, species composition sampled, and number of individuals collected, which makes a detailed comparison for each area difficult. Some changes can still be noted. The western part of the Bight of Bangkok (areas E, F), where no imposex was found in 1996, showed 3.1% imposex off Phet Buri (area E) and 20.8% imposex off Ban Khau (area F) in 2006. During both surveys not a single large ship was seen in these areas. The observed increase may be caused by a gradual dispersion of TBT from the eastern side. In Pattani Province in the southern part of Thailand, the species-adjusted imposex occurrence decreased slightly in Pattani Bay Mouth (area K), which is along the shipping route, and Laem Nok (area J) in Pattani Bay, but increased in Rusamilae (area I) and Bang Tawa (area H). All the areas in Pattani Province (H–K) are only accessible to small ships, mainly wooden fishing boats. Large commercial ships pass a long distance away from the coast. Contrary to the situation in 1996, no TBT containing paints were found on and around the local dry docks. A switch to vinyl copper paints may explain the decrease of imposex in J and K. The increase in H and I may have been caused by extensive suction dredging in the mouth of the Pattani River and Pattani Bay to construct a larger harbour and deepening the shipping route for large fishing boats to the Gulf of Thailand. The sediment has been deposited on the intertidal area west of the excavated channel against the mainland. That area was surrounded by a low stone dam that kept most of the sand fraction but allowed the heavily polluted fine material in suspension to disperse. Tidal movements spread the mud cloud that settled over the intertidal and subtidal area to the west. We speculate that the TBT contaminated mud from the shipping lane increased the imposex incidence in Rusamilae (area I) from only 0.3% during the 1996 study to 15.6% in 2006. A similar situation was found in the adjacent area Bang Tawa (area H) where no imposex was found in 1996 but 17.9% in 2006, and no imposex was found further westward in 1995 (C. Swennen pers. obs.). Considering the high incidence in the western sampling sites, TBT contamination could be present in areas that were off-limits in 2006 due to safety risk by local insurgents. These findings show that regulation for the safe disposal of contaminated sediments and of anti-fouling paints after cleaning ships in dockyards is urgently needed in Thailand. For the most southern part of the Gulf of Thailand (Tak Bai, area M), the difference between both surveys is negligible (one individual showed imposex in 145 females in 1996 and two individuals in 156 females in 2006). The observed frequency for imposex from Tak Bai was 1.3% (2 of 156 females sampled). According to the contamination model a much higher occurrence can be expected (Fig. 1D). The Tak Bai samples contained 156 females, but all but three were from species with very low sensitivity to imposex. If the samples from Tak Bai had the typical representation of the species sampled from the Gulf in this study, the risk of imposex would be higher than 1.3%, because the sample would then contain species with a higher sensitivity to imposex. Therefore, an estimated 10.1% risk at Tak Bai is quite plausible (see also Tables 3 and 4, and the basic data added to the online version).

The suspicion that the local species may differ in sensitivity to TBT (Swennen et al., 1997) has been strengthened by the present study involving larger samples per site. Although the local TBT contamination is the dominant factor, *M. musiva*, *Pugilina cochlidium*, and *Babylonia areolata* seems less sensitive for developing imposex while *Thais lacera*, *Murex occa*, and *L. blosvillei* seem most sensitive among the species used in this study (Table 4). The lower sensitivity in *Pugilina* and *Babylonia* was also found in the first survey. Bech (2002) found also that *M. musiva* is a less sensitive species. It is not clear what factor induces the difference in sensitivity to TBT. Imposex appears wide-spread in Southeast Asia and has been reported in intertidal gastropods from various Southeast Asian areas, such as the Andaman coast of Thailand (Bech, 1999; Bech et al., 2002), Singapore (Tan, 1999), Hong Kong (Blackmore, 2000; Leung et al., 2006), China (Shi et al., 2005), and the Asia-Pacific Mussel Watch showed sites with high TBT values in all countries between China and India (Sudaryanto et al., 2002).

TBT not only kills larvae of fouling organisms that will settle on ships and other installations at sea, but also affects a broad range of non-target organisms in the wider surroundings. TBT residues have been found in algae (Maguire et al., 1984), crabs (Lee, 1985), fish (Lee, 1985; Kannan et al., 1996; Kannan and Falandysz, 1997), sea birds (Guruge et al., 1997; Kannan and Falandysz, 1997), seals (Kim et al., 1996; Berge et al., 2004), dolphins (Iwata et al., 1995; Kannan and Falandysz, 1997; Berge et al., 2004), and in humans (Zaucke and Krug, 1997; Whalen et al., 2002; Grün and Blumberg, 2006) and is far from harmless in humans (Aluoch et al., 2006). In sublethal concentrations, it induces the growth of male characters such as a penis and a pallial vas deferens (imposex) in most neogastropods. As result of this the females of some affected species die and populations become extinct in shipping routes and around harbours. In about 200 gastropod species imposex has already been found (Shi et al., 2005). In bivalves it can induce shell abnormalities and economic damage is reported from oyster cultures (Alzieu et al., 1981, 1986; Phelps and Pace, 1997). Other effects usually remain hidden, but laboratory studies showed effects of sublethal concentrations on the growth and development of bivalve larvae (Stenalt et al., 1998; Coelho et al., 2001), meiobenthos (Gustafsson et al., 2000). In fish, a low TBT level induces masculinization (Shimasaki et al., 2003), disturbs visual and olfactory functions (Wang and Huang, 1999) and causes chromosomal aberrations (Cipriano et al., 2004). It means that the anti-fouling paints seriously disturb the food chains in the marine environment.

Regrettably, the International Maritime Organization (IMO) did not recommend a total ban of TBT on ships and structures in sea until September 2008, while discussions about this began in 1996. Preferably, the Thai government should have taken precautionary measures to reduce pollution of the Gulf of Thailand by banning the use of TBT containing anti-fouling paint on ships and structures at sea such as landing piers, gas and oil installations, in aqua-culture in their part of the Gulf. Furthermore, the Thai government could have requested appropriate compensation for using mooring facilities along the Thai coast from foreign vessels with these paints. The extensive TBT pollution in the Gulf of Thailand has been known since the 1990s (Swennen et al., 1996, 1997); even Thai newspapers paid attention to it. Thousand of local fishermen depend on marine products from these waters, and local people and tourists eat these sea foods. TBT contaminated food can impose a risk to both the local population and impact the tourist industry.

Acknowledgements

We like to thank the several anonymous fishermen that allowed us to collect the gastropods from their catches, Ni-hasem Waesalae, Somsak Buatip, Masakree Sani and Uma Suetae for their help in collecting and processing the gastropods. We thank Professor Don McNeil for suggesting the use of logistic regression to adjust imposex frequency for different species sensitivity, Dr. A.K.A. Holthuijzen and the anonymous referees for detailed comment that helped us to improve the final text.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.marpolbul.2008.11.028.

References

Aluoch, A.O., Odman-Ghazi, S.O., Whalen, M.M., 2006. Alteration of an essential NK cell signalling pathway by low doses of tributyltin in human natural killer cells. Toxicology 224, 229–237.

Alzieu, C., Héral, M., Thibaud, Y., Dardignac, M.J., Feuillet, M., 1981. Influence des peintures antisalissures a base d'organostanniques sur la calcification de l'huitre- *Crassostrea gigas*. Revue des Travaux de l'Institut des Pêches Maritimes 45, 100–116.

Alzieu, C., Sanjuan, J., Deltriel, J.P., Borel, M., 1986. Tin contamination in Arcachon Bay: effects on oyster shell anomalies. Marine Pollution Bulletin 17, 494–498.

Bech, M., 1999. Increasing levels of tributyltin-induced imposex in muricid gastropods at Phuket Island, Thailand. Applied Organometallic Chemistry 13, 799–804.

Bech, M., 2002. A survey of imposex in muricids from 1996 to 2000 and identification of optimal indicators of tributyltin contamination along the east coast of Phuket Island, Thailand. Marine Pollution Bulletin 44, 887–896.

Bech, M., Strand, J., Jacobsen, J.A., 2002. Development of imposex and accumulation of butyltin in the tropical muricid *Thais distinguenda* transplanted to a TBT contaminated site. Environmental Pollution 119, 253–260.

Berge, J.A., Brevik, E.M., Bjorge, A., Folsvik, N., Gabrielsen, G.W., Wolkers, H., 2004. Organotins in marine mammals and seabirds from Norwegian territory. Journal of Environmental Monitoring 6, 108–112.

Blackmore, G., 2000. Imposex in *Thais clavigera* (Neogastropoda) as an indicator of TBT (tributyltin) bio-availability in coastal waters of Hong Kong. Journal of Molluscan Studies 66, 1–8.

Blader, S.J.M., 1970. The occurrence of a penis-like outgrowth behind the right tentacle in spent females of *Nucella lapillus* (L.). Proceedings of the Malacological Society of London 39, 231–233.

Bryan, G.W., Gibbs, P.E., Hummerstone, L.G., Burt, G.R., 1986. The decline of the gastropod *Nucella lapillus* around south-west England: evidence for the effect of tributyltin from anti-fouling paints. Journal of the Marine Biological Association UK 66, 767–777.

Bryan, G.W., Gibbs, P.E., Burt, G.R., Hummerstone, L.G., 1987. The effects of tributyltin (TBT) accumulation on adult dog-whelks, *Nucella lapillus*: long-term field and laboratory experiments. Journal of the Marine Biological Association UK 67, 525–544.

Cipriano, R.R., Ribeiro, C.A.O., Cestari, M.M., Fenocchio, A.S., 2004. Evaluation of the effects of tributiltin (TBT) on chromosomes of the neotropical fish *Astyanax* sp. (Pisces, Tetragonopterinae). Cytologia 69, 187–190.

Coelho, M.R., Fuentes, S., Bebianno, M.J., 2001. TBT effects on the larvae of *Ruditapes decussatus*. Journal of the Marine Biological Association of the UK 81, 259–265.

Dowson, P.H., Bub, J.M., Lester, J.N., 1993. Temporal distribution of organotins in the aquatic environment: five years after the 1987 UK retail ban on TBT based antifouling paints. Marine Pollution Bulletin 26, 487–494.

Ellis, D.V., Pattisima, L.A., 1990. Widespread neogastropod imposex: a biological indicator of global TBT contamination? Marine Pollution Bulletin 21, 248–253.

Evans, S.M., Evans, P.M., Leksono, T., 1996. Widespread recovery of dogwhelks, *Nucella lapillus* (L.), from tributyltin contamination in the North Sea and Clyde Sea. Marine Pollution Bulletin 32, 263–269.

Gibbs, P.E., Bryan, G.W., 1986. Reproductive failure in populations of the dog-whelk, *Nucella lapillus*, caused by imposex induced by tributyltin from antifouling paints. Journal of the Marine Biological Association UK 66, 767–777.

Gibbs, P.E., Bryan, G.W., Pascoe, P.L., Burt, G.W., 1987. The use of the dog-whelk, *Nucella lapillus*, as an indicator of tributyltin (TBT) contamination. Journal of the Marine Biological Association UK 67, 507–523.

Gibbs, P.E., Bebianno, M.J., Coelho, M.R., 1997. Evidence of the differential sensitivity of neogastropods to tributyltin (TBT) pollution, with notes on a species (*Columbella rustica*) lacking the imposex response. Environmental Technology 18, 1219–1224.

Grün, F., Blumberg, B., 2006. Environmental obesogens: organotins and endocrine disruption via nuclear receptor signaling. Endocrinology 147 (Suppl.), 50–55.

Guruge, K.S., Iwata, H., Tanaka, H., Tanabe, S., 1997. Butyltin accumulation in the liver and kidney of seabirds. Marine Environmental Research 44, 191–199.

Gustafsson, M., Dahllöfb, I., Blanck, H., Hallb, P., Molanderd, S., Nordberga, K., 2000. Benthic foraminiferal tolerance to tri-*n*-butyltin (TBT) pollution in an experimental mesocosm. Marine Pollution Bulletin 40, 1072–1075.

Horiguchi, T., Shiraishi, H., Shimizu, A., Morita, S., 1991. Imposex and organotin compounds in *Thais clavigera* and *T. bronni* in Japan. Journal of the Marine Biological Association UK 74, 651–669.

Hosmer, D.W., Lemeshow, S., 2000. Applied Logistic Regression, second ed. John Wiley and Sons, Inc. 375 p.

Iwata, H., Tanabe, S., Mizuno, T., Tatsukawa, R., 1995. High accumulation of toxic butyltins in marine mammals from Japanese coastal waters. Environmental Science and Technology 29, 2959–2962.

Kannan, K., Corsolini, S., Focardi, S., Tanabe, S., Tatsukawa, R., 1996. Accumulation pattern of butyltin compounds in dolphin, tuna and shark collected from Italian coastal waters. Archives of Environmental Contamination and Toxicology 31, 19–23.

Kannan, K., Falandysz, J., 1997. Butyltin residues in sediment, fish, fish-eating birds, harbour porpoise and human tissues from the Polish coast of the Baltic Sea. Marine Pollution Bulletin 34, 203–207.

Leung, K.M.Y., Kwong, P.Y., Ng, W.C., Horiguchi, T., Qiu, J.W., Ruiqiang, Y., Maoyong, S., Guibin, J., Zheng, G.J., Lam, P.K.S., 2006. Ecological risk assessments of endocrine disrupting organotin compounds using marine neogastropods in Hong Kong. Chemosphere 65, 922–938.

Kim, G.B., Lee, J.S., Tanabe, S., Iwata, H., Tatsukawa, R., Shimazaki, K., 1996. Specific accumulation and distribution of butyltin compounds in various organs and tissues of the Stellar Sea Lion (*Eumetoplas jubatus*): comparison with organochlorine accumulation pattern. Marine Pollution Bulletin 32, 558–563.

Kleinbaum, D.G., Klein, M., 2002. Logistic Regression: a Self-Learning Text, second ed. Springer-Verlag, New York, Inc.. 513 p.

Lee, R.F., 1985. Metabolism of tributyltin oxide by crabs, oysters and fish. Marine Environmental Research 17, 145–148.

Maguire, R.J., Wong, P.T.S., Rhamey, J.S., 1984. Accumulation and metabolism of tri-n-butyltin cation by a green alga, *Ankistrodesmus falcatus*. Canadian Journal of Fisheries and Aquatic Sciences 41, 537–540.

Mensink, B.P., Everaarts, J.M., Kralt, H., Ten Hallers-Tjabbes, C.C., Boon, J., 1996. Tributyltin exposure in early life stages induces the development of male characteristics in the common whelk, *Buccinum undatum*. Marine Environmental Research 42, 151–154.

Mensink, B.P., Kralt, H., Vethaak, A.D., Ten Hallers-Tjabbes, C.C., Koeman, J.H., van Hattum, B., Boon, J.P., 2002. Imposex induction in laboratory reared juvenile *Buccinum undatum* by tributyltin (TBT). Environmental Toxicology and Pharmacology 11, 49–65.

Oehlmann, J., Stroben, E., Bettin, C., Fioroni, P., 1993. Hormonal disorders and tributyltin-induced imposex in marine snails. In: Quantified Phenotypic Responses in Morphology and Physiology. Proceedings of the 27th European Marine Biology Symposium, Dublin Ireland, 7–11 September 1992, pp. 301–305.

Oehlmann, J., Stroben, E., Schulte-Oehlmann, U., Bauer, B., Fiorini, P., Mardert, B., 1996. Tributyltin biomonitoring using prosobranchs as sentinel organisms. Fresenius Journal of Analytical Chemistry 354, 540–545.

Phelps, H.L., Pace, S.S., 1997. Tributyltin biomonitoring in Portuguese estuaries with the Portuguese Oyster (*Crassostrea angulata*). Environmental Technology 18, 1269–1276.

R Development Core Team, 2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria http://www.R-project.org.

Shi, H.H., Huang, C.J., Zhu, S.X., Yu, X.J., Xie, W.Y., 2005. Generalized system of imposex and reproductive failure in female gastropods of coastal waters of mainland China. Marine Ecology Progress Series 304, 179–189.

Shimasaki, Y., Kitano, T., Oshima, Y., Inoue, S., Imada, N., Honjo, T., 2003. Tributyltin causes masculinization in fish. Environmental Toxicology and Chemistry 22, 141–144.

Smith, B.S., 1981. Tributyltin compounds induce male characteristics on female mud snails, *Nassarius obsoletus* = *Ilyanassa obsoleta*. Journal of Applied Toxicology 1, 141–144.

Stenalt, E., Johansen, B., Lillienskjold, S.V., Hansen, B.W., 1998. Mesocosm study of *Mytilus edulis* larvae and postlarvae, including the settlement phase, exposed to a gradient of tributyltin. Ecotoxicology and Environmental Safety 40, 212–225.

Stewart, C., De Mora, S.J., Jones, M.R.L., Miller, M.C., 1992. Imposex in New Zealand neogastropods. Marine Pollution Bulletin 24, 204–209.

Sudaryanto, A., Takahashi, S., Monirith, I., Ismail, A., Muchtar, M., Zheng, J., Richardson, B.J., Subramanian, A., Prudente, M., Hue, N.D., Tanabe, S., 2002. Asia-Pacific mussel watch: monitoring of butyltin contamination in coastal waters of Asian developing countries. Environmental Toxicology and Chemistry 21, 2119–2130.

Swennen, C., Ruttanadakul, N., Ardseungnern, S., Singh, H.R., Tan, K.S., Mensink, B.P., Ten Hallers-Tjabbes, C.C., 1996. Sexual aberrations in gastropods in Gulf of Thailand and Strait of Malacca in relation to shipping. Wallaceana 78, 1–13.

Swennen, C., Ruttanadakul, N., Ardseungnern, S., Singh, H.R., Mensink, B.P., Ten Hallers-Tjabbes, C.C., 1997. Imposex in sublittoral and littoral gastropods from the Gulf of Thailand and Strait of Malacca in relation to shipping. Environmental Technology 18, 1245–1254.

Swennen, C., Moolenbeek, R.G., Ruttanadakul, N., Hobbelink, H., Dekker, H., Hajisamae, S., 2001. The molluscs of the southern Gulf of Thailand. Thai Studies in Biodiversity 4, 1–210.

Swennen, C., Horpet, P., 2008. Pseudo-imposex; male features in female volutes not TBT-induced (Gastropoda: Volutidae). Contributions to Zoology 77, 13–20.

Tan, K.S., 1999. Imposex in *Thais gradata* and *Chicoreus capucinus* (Mollusca, Neogastropoda, Muricidae) from the Straits of Johor: a case study using penis length, area and weight as measures of imposex severity. Marine Pollution Bulletin 39, 295–303.

Ten Hallers-Tjabbes, C.C., 1997. Tributyltin and polices for anti-fouling. Environmental Technology 18, 1265–1268.

Ten Hallers-Tjabbes, C.C., Kemp, J.F., Boon, J.P., 1994. Imposex in whelks (*Buccinum undatum*) from the open North Sea: relation to shipping traffic intensities. Marine Pollution Bulletin 28, 311–313.

Ten Hallers-Tjabbes, C.C., Wegener, J.W., Van Hattum, A.G.M., Kemp, J.F., Ten Hallers, E., Reitsema, T.J., Boon, J.P., 2003. Imposex and organotin concentrations in *Buccinum undatum* and *Neptunea antiqua* from the North Sea: relationship to shipping density and hydrographical conditions. Marine Environmental Research 55, 203–233.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, fourth ed. Springer Science + Business Media, New York, USA. 495 p.

Wang, D.-Y., Huang, B.-Q., 1999. TBT (tributyltin) toxicity to the visual and olfactory functions of Tigerperch (*Terapon jarbua* Forsskål). Zoological Studies 38, 189–195.

Whalen, M.M., Ghasi, S., Loganathan, B.G., Hatcher, F., 2002. Expression of CD16, CD18 and CD56 in tributyltin-exposed human natural killer cells. Chemico-Biological Interactions 139, 159–176.

Zaucke, F., Krug, H.F., 1997. Induction of apoptosis in human Hl-60 granulocytes by the organic tin compound tri-*n*-butyltin (TBT). European Journal of Cell Biology 72, 145.

Appendix 2

Publication 2 (Manuscript): Regression-based modeling of macrobenthic fauna density in the Middle Songkhla Lake, Thailand

Regression-Based Modeling of Macrobenthic Fauna Density in Middle Songkhla Lake, Thailand

Uraiwan Sampantarak*[a], Noodchanath Kongchouy [b], and Saowapa Angsupanich [c]

[a] Pattani Inland Fisheries Research and Development Center, Department of Fisheries,
Pattani 94160, Thailand

[b] Department of Mathematics, Faculty of Science, Prince of Songkla University,
Songkhla 90112, Thailand

[c] Department of Aquatic Science, Faculty of Natural Resources, Prince of Songkla University,
Songkhla 90112, Thailand

* Author for correspondence; e-mail: uraiwan111@hotmail.com, Tel: +66 8 4653 9391

ABSTRACT

This study examined distributional patterns of macrobenthic fauna assemblages in relation to environmental characteristics in the middle of Songkhla Lake, Thailand. Macrobenthic fauna and water quality parameters including sediment characteristics were obtained from nine sampling sites at bimonthly intervals from April 1998 to February 1999. Factor analysis was used to define five predictors including three composite variables based on salinity, physical sediment characteristics, and physico-chemical properties of water and sediment, together with total suspended solids and dissolved oxygen as single variables. A multivariate multiple regression model (MMR) was used to examine relationships between these predictors and the densities of twenty four selected macrobenthic families with greater than 35% occurrence. To remove skewness, the densities were log-transformed before fitting the model. Results were compared with those obtained using canonical correspondence analysis. MMR can be used as additional or alternative method to analyse relationship between environmental variables and abundance of benthic organisms in coastal ecosystem.

Keywords: Macrobenthic fauna, Multivariate multiple regression model, Factor analysis, Canonical correspondence analysis, Tropical lagoon

1. INTRODUCTION

Macrobenthic fauna are recognized as sensitive indicators of environmental disturbance (Weisberg et al. [1], Borja et al. [2], Ranasinghe et al. [3]). They have limited mobility. Many of them are unable to avoid adverse conditions brought about by natural stresses or human impacts. Moreover, their relative longevity, with many species having life spans in excess of two years, allows them to integrate responses to environmental processes over extended time periods (Gray et al. [4]). In addition, observed distribution of macrobenthic fauna are useful in diagnostic studies and environmental monitoring (Warwick [5]).

Clarke and Warwick [6] outlined the basic methods now commonly used by biological scientists for analysis of their data. For descriptive studies these methods include data transformation using square roots, fourth roots or logarithms (after adding 1 to cell counts or densities to handle zeros) to remove skewness, principal components analysis of covariance matrices, and ordination procedures to cluster taxa in space and time, as well as more complex multivariate analytical techniques such as dendrograms based on similarity matrices and multidimensional scaling. Measures of association in assemblage data such as the Bray-Curtis similarity index are preferred to Pearson correlation coefficients "for sound biological reasons" (Clarke et al. [7]), but such measures do not satisfy the positive-definiteness assumptions that underpin conventional multivariate statistical analysis.

For comparative studies to assess associations between species abundance outcomes and environmental predictor variables, canonical correspondence analysis (Ter Braak [8]) is now used extensively in the biological literature (von Wehrden et al. [9]). Some important studies using this method include those reported by Rakocinski et al. [10], Hawkins et al. [11], Joy and Death [12], Guerra-García et al. [13], Hajisamae and Chou [14], Morrisey et al. [15], Ysebaert et al. [16], Quintino et al. [17], Anderson [18], Glockzin and Zettler [19].

Although exceptions exist such as studies by Liang et al. [20] using structural equation modeling and by Warton and Hudson [21] using multivariate analysis of variance, multivariate multiple regression analysis is not commonly used in the biological literature for analyzing species abundance patterns. However, this method would appear to be an ideal statistical method for assessing relationships between species abundance outcomes and their environmental predictors, for the simple reason that it is the natural extension of ordinary regression analysis involving a single outcome to any number of mutually correlated outcomes such as species abundances. It is thus of interest to compare this method with its biologically preferred counterpart using common sets of biological data relating taxonomic abundances to environmental determinants, and this is the object of our study.

For this comparison we used data from a study involving macrobenthic fauna abundances and various water and sediment characteristics collected at specified locations in an estuarine lake over a period of one year reported by Angsupanich et al. [22]. The methods compared are canonical correspondence analysis (CCA) using CANOCO Version 4.5 (Ter Braak and Šmilauer [23]) and multivariate multiple regression (MMR) using R Version 2.10.0 (R Development Core Team [24]).

2. METERIALS

Songkhla Lake is a shallow coastal lagoon, located in a tropical coastal ecosystem in Southern Thailand. It covers an area of 1,040 km^2 with 20 km width and 75 km length, approximately. It is divided into three parts as the Upper Lake in Phatthalung Province, the Middle Lake between borders of Songkhla Province and Phatthalung Province, and the Lower Lake in Songkhla Province connected to the Gulf of Thailand. Some canals pour fresh water into the lake. The salinity slowly increases where the freshwater and seawater meet. Thus, the water in the Middle Lake is brackish, and becomes saltier in the area around the lake mouth (Lower Lake). The zone of interest for this study covers an area of 390 km^2 located between UTM 635000E and 660000E in the west-east direction and between UTM 840000N and 805000N in the north-south direction (Figure 1).

Angsupanich et al. [22] collected macrobenthic fauna using a Tamura's grab (0.05 m^2) from the nine sampling stations. The assemblages were conducted with 11 replications for each station at bimonthly intervals from April 1998 to February 1999. The samples were sieved consecutively through the screens and fixed in 10% Rose Bengal-formalin for later identification.

The densities of macrobenthic fauna were recorded as the number of individuals per square meter (ind m^{-2}) for each species. A total of 161 taxa of macrobenthic fauna were found and classified into 81 families. In many cases the species could not be identified exactly, so in our model the outcomes were classified by family instead of species. With nine locations and six bimonthly data study periods, we defined the occurrence for a specified family as the proportion of these 54 occasions on which at least one organism was found. We then selected the 24 families with greater than 35% occurrence (93.2% total assemblages) for data analysis.

Environmental variables comprised water depth (wDep), water temperature (wTemp), salinity (Sal), water pH (wpH), dissolved oxygen (DO), total suspended solids (TSS), with sediment pH (spH), total nitrogen content (TN), organic carbon content (OC), and soil structure (percentages of sand, silt, and clay). These were measured with three replications on the same occasions as the biotic data.

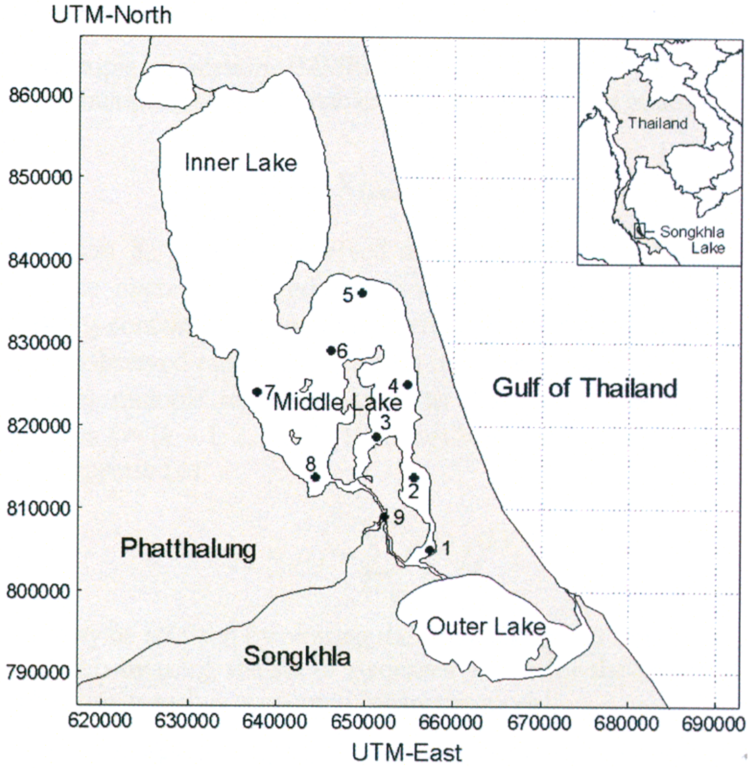

Figure 1. Songkhla Lake and sampling sites (labeled 1-9)

3. METHODS

The nine sampling stations and six bimonthly periods were combined as 54 station-month measurement occasions. The response variable was taken as log (1 + *c* × density) with the multiplier *c* chosen to approximate normality of error distributions. The predictors comprised environmental components derived from a factor analysis together with unique variables not accommodated by the factor analysis.

FACTOR ANALYSIS

Factor analysis is performed on the environmental variables with the aim of substantially reducing correlations between them that could mask their associations with the outcome variables. Each factor identifies correlated groups of variables. Ideally each group (which must contain at least two variables to contribute to the factor analysis) contains variables having small correlations with variables in other groups. To achieve this, any variable uncorrelated with all other variables is omitted from the factor analysis. Each factor comprises weighted linear combinations of the variables and these factors are rotated to maximize the weights of variables within the factor group and minimize the weights of variables outside the group. The resulting weights are called "loadings". Variables omitted from the factor analysis due to low correlation with all other variables (high "uniqueness") are treated as separate predictors, so predictors include single variables as well as factors.

The number of factors selected was based on obtaining an acceptable statistical fit using the chi-squared test, and these factors were fitted using maximum likelihood with promax rotation in preference to varimax, which requires the rotation to be orthogonal (Browne [25], Abdi [26]).

MULTIVARIATE MULTIPLE REGRESSION

Multivariate multiple regression (MMR) is used to evaluate the effects of multiple predictor variables on multiple response variables. The model (Mardia et al. [27]) may be defined in matrix form as

$$\mathrm{Y}_{(n\times p)} = \mathrm{X}_{(n\times q)}\mathrm{B}_{(q\times p)} + \mathrm{E}_{(n\times p)}. \tag{1}$$

In this formulation $\mathbf{Y}_{(n\times p)}$ is an observed matrix of p response variables on each of n occasions, $\mathbf{X}_{(n\times q)}$ is the matrix of q predictors (including a vector of 1s) in columns and n occasions in rows, $\mathbf{B}_{(q\times p)}$ contains the regression coefficients (including the intercept terms), and $\mathrm{E}_{(n\times p)}$ is a matrix of unobserved random errors with mean zero and common covariance matrix Σ. Ordinary (univariate) multiple regression arises as the special case when $p = 1$. If $q-1$ environmental predictors $f_i^{(k)}$ ($k = 1, 2,\ldots, q-1$) are available, the prediction model for outcome j on occasion i may be expressed as

$$y_{ij} = \mu_j + \sum_{k=1}^{q-1}\beta_j^{(k)} f_i^{(k)}. \tag{2}$$

The model fit may be assessed by plotting the residuals against normal quantiles (Venables and Ripley [28]), and also by using the set of r-squared values for the response variables to see how much of the variation in each is accounted for by the model.

The method also provides standard errors for each of the $p \times q$ regression coefficients thus providing p-values for testing their statistical significance after appropriate allowance for multiple hypothesis testing. The multivariate analysis of variance (MANOVA) decomposition is also used to assess the overall association between each environmental predictor and the set of outcomes by the likelihood ratio, Pillai's trace criterion (Olson [29], Johnson and Wichern [30]).

CANONICAL CORRESPONDENCE ANALYSIS

Assuming that the data structure comprises the **Y** and **X** matrices with rows corresponding to measurements of outcomes and predictors taken on the same occasions, canonical correspondence analysis (Ter Braak [8]) produces a two-dimensional *biplot* comprising arrows of variable lengths and directions (*gradients*) emanating from a common origin representing the predictor variables, together with superimposed points denoting the outcome variables. The relative lengths of the arrows and the angles between them are based on the correlation matrix of the predictor variables, and the coordinates of the points are planar projections of the density outcomes, computed in such a way that their positions relative to the arrows portray their associations with the environmental predictors. The method also produces coordinate scores and p-values for the overall associations based on Monte Carlo permutation tests.

4. RESULTS

ENVIRONMENTAL PARAMETERS

Figure 2 plots the water characteristics in Middle Songkhla Lake from April 1998 to February 1999. The water depth varied to a lesser extent, but was also higher during the rainy season, varying with location from an average of less than 1 m at stations four and nine to more than 2 m at station eight. The water temperature showed decreased values in the rainy season, with range 27-34°C. The salinity increased from close to zero during the rainy season (December to February) to an average close to 20 in other months. The pH of water was also lowest in December.

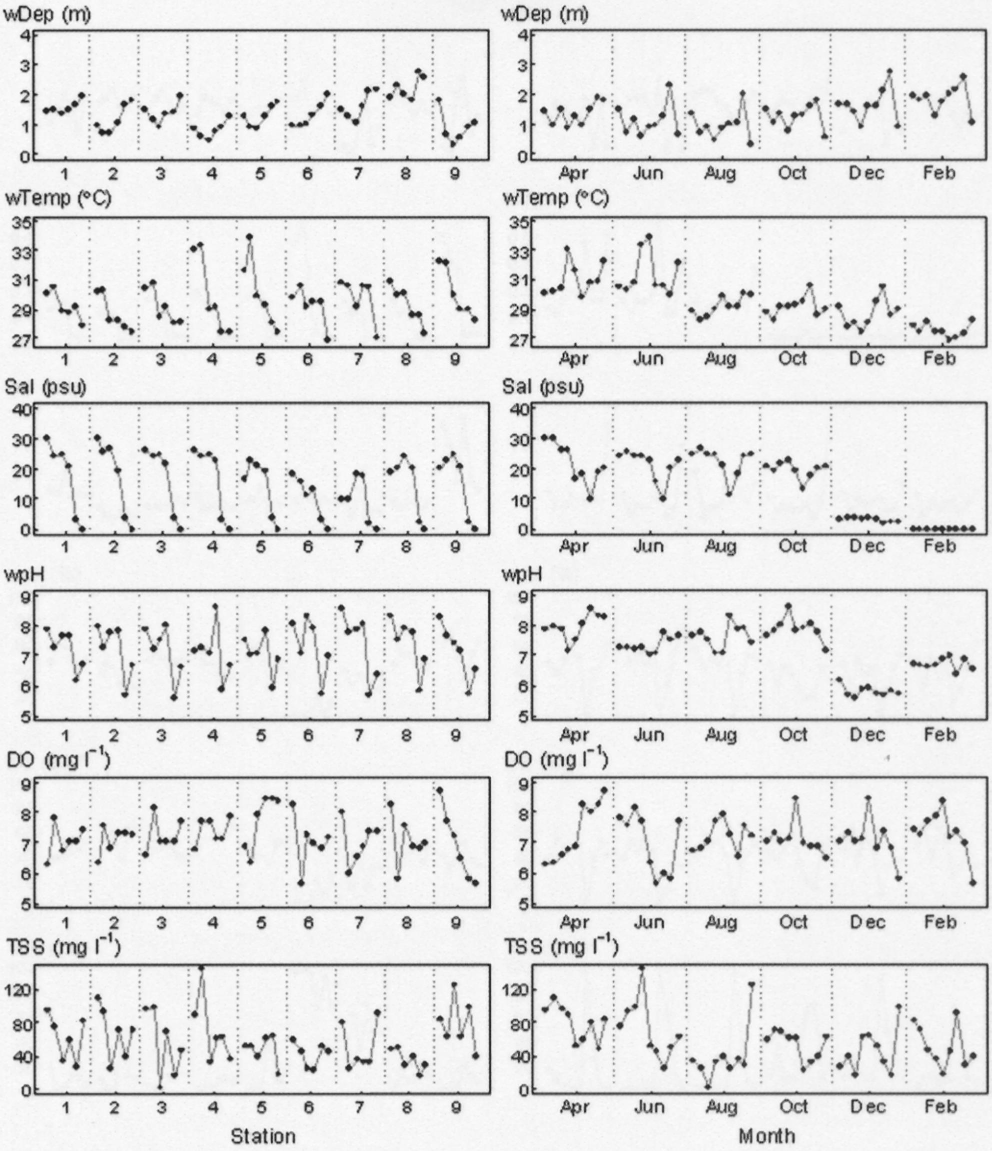

Figure 2. Water characteristics in Middle Songkhla Lake from April 1998 to February 1999 by station (left panel) and month (right panel)

Figure 3 plots sediment characteristics measured on the same occasions as the water characteristics. The total nitrogen content at each station was very low (0.02%) from October to February. The organic carbon content was relatively constant with respect to month, but showed the highest value at station nine in every month except August. The lake bed at station six was mostly characterized by sand (mean = 84.6%) and station 9 was mostly characterized by clay (mean = 53.2%), also with high values of organic carbon. Note that the sand, silt, and clay percentages sum to 100%.

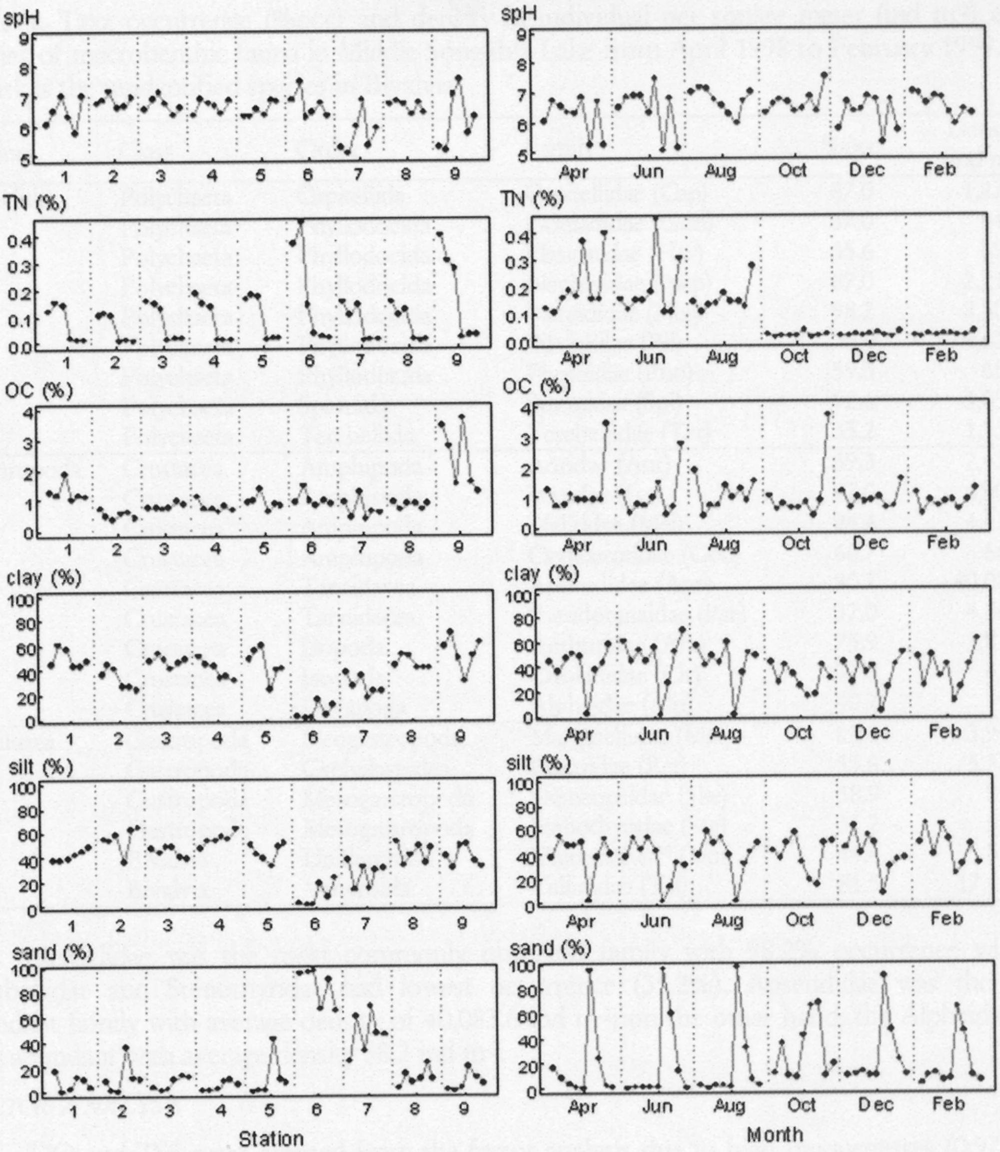

Figure 3. Sediment characteristics in Middle Songkhla Lake from April 1998 to February 1999 by station (left panel) and month (right panel)

OCCURRENCE AND ABUNDANCE OF MACROBENTHIC FAUNA

Table 1 shows the taxa percentages of occurrence and density in individuals per square meter of the 24 families of macrobenthic fauna measured with the water characteristics. A total of 24 families were classified in three phyla of Annelida (Polychaeta), Arthropoda (Crustacea) and Mollusca (Gastropoda and Bivalvia), which comprised the most diverse group (35.2-98.2% of occurrence). Polychaeta was represented by nine families (Capitellidae, Goniadidae, Hesionidae, Nephtyidae, Nereididae, Pilargiidae, Pholoidae, Spionidae and Terebellidae). Crustacea was also represented by nine families (Aoridae, Isaeidae, Melitidae, Oedicerotidae, Apseudidae, Pseudotanaidae, Anthuridae, Cirolanidae and Alpheidae). Marginellidae, Retusidae, Skeneopsidae and Stenothyridae were in the Gastropoda whilst the two remaining families (Tellinidae and unidentified species were in the Bivalvia).

Table 1. Taxa occurrence (%occ) and density in individual per square meter (ind m^{-2}) of 24 families of macrobenthic fauna in Middle Songkhla Lake from April 1998 to February 1999. The asterisk is the unidentified species in Bivalvia

Phylum	Class	Order	Family	%occ	Density (ind m^{-2})
Annelida	Polychaeta	Capitellida	Capitellidae (Cap)	87.0	1,227.3
	Polychaeta	Phyllodocida	Goniadidae (Gon)	37.0	443.6
	Polychaeta	Phyllodocida	Hesionidae (Hes)	55.6	698.2
	Polychaeta	Phyllodocida	Nephtyidae (Nep)	87.0	2,218.2
	Polychaeta	Phyllodocida	Nereididae (Ner)	98.2	8,507.3
	Polychaeta	Phyllodocida	Pilargiidae (Pil)	70.4	1,625.5
	Polychaeta	Phyllodocida	Pholoidae (Pho)	59.3	658.2
	Polychaeta	Spionida	Spionidae (Spi)	92.6	5,056.4
	Polychaeta	Terebellida	Terebellidae (Ter)	35.2	1,136.4
Arthropoda	Crustacea	Amphipoda	Aoridae (Aor)	59.3	2,421.8
	Crustacea	Amphipoda	Isaeidae (Isa)	87.0	6,900.0
	Crustacea	Amphipoda	Melitidae (Mel)	94.4	4,438.2
	Crustacea	Amphipoda	Oedicerotidae (Oed)	66.7	667.3
	Crustacea	Tanaidacea	Apseudidae (Aps)	90.7	40,083.6
	Crustacea	Tanaidacea	Pseudotanaidae (Pse)	37.0	4,265.5
	Crustacea	Isopoda	Anthuridae (Ant)	75.9	3,816.4
	Crustacea	Isopoda	Cirolanidae (Cir)	37.0	427.3
	Crustacea	Decapoda	Alpheidae (Alp)	40.7	98.2
Mollusca	Gastropoda	Neogastropoda	Marginellidae (Mar)	85.2	3,963.6
	Gastropoda	Cephalaspidea	Retusidae (Ret)	55.6	5,536.4
	Gastropoda	Mesogastropoda	Skeneopsidae (Ske)	38.9	956.4
	Gastropoda	Mesogastropoda	Stenothyridae (Ste)	35.2	581.8
	Bivalvia	Unidentified	Unidentified* (Uni)	44.4	338.2
	Bivalvia	Veneroida	Tellinidae (Tel)	81.5	17,134.5

Nereididae was the most commonly observed family with 98.2% occurrence whereas Terebellidae and Stenothyridae had lowest occurrence (35.2%). Apseudidae was the most abundant family with average density of 40,083.6 ind m^{-2}; on the other hand, the Alpheidae was least abundant with average density 98.2 ind m^{-2}.

FACTOR ANALYSIS

DO and TSS were omitted from the factor analysis due to high uniquenesses (0.975 and 0.848, respectively). The model provided an adequate fit using three factors (chi-squared = 20.23, 12 df, p-value = 0.063). Table 2 shows the loadings, with values less than 0.20 in magnitude suppressed. If only loadings greater in magnitude than 0.45 are considered, the three factors do not contain any overlapping variables.

Factor 1 encompasses salinity, containing positive loadings for Sal and wpH, and a negative loading for wDep as expected, with deeper water during the rainy season. Factor 2 represents the effect of sediment characteristics in the lake bed (sand-clay habitat), consisting of a positive loading for sand and a similar negative loading for clay. Factor 3 characterizes physical and chemical compositions in the lake, comprising positive loadings for TN, OC, and wTemp, and a negative loading for spH. Thus Factor 1 was defined as $-0.53 \times \text{wDep} + 0.70 \times \text{wpH} + 0.99 \times \text{Sal}$, Factor 2 as $0.94 \times \text{Sand} - 0.95 \times \text{Clay}$, and Factor 3 as $0.47 \times \text{OC} + 0.58 \times \text{TN} - 0.57 \times \text{spH} + 0.54 \times \text{wTemp}$. The three factors respectively accounted for 24.6%, 20.7%, and 13.7% of the variance in the environmental data, a total of 59.0%. The three factors were included in the regression model as predictors together with the two singleton variables omitted from the factor analysis, with each of these five predictor variables scaled to have mean 0 and standard deviation 1.

Table 2. Factor analysis (with loadings below 0.2 omitted)

Environmental variables	Factor 1	Factor 2	Factor 3
Organic carbon (OC)	-	-	**0.47**
Total nitrogen (TN)	0.34	-	**0.58**
Sediment pH (spH)	0.39	-	**– 0.57**
Water depth (wDep)	**– 0.53**	-	-
Water pH (wpH)	**0.70**	-	-
Salinity (Sal)	**0.99**	-	-
Water temperature (wTemp)	0.42	-	**0.54**
Sand	-	**0.94**	-
Clay	-	**– 0.95**	-
% Total variance	24.6	20.7	13.7
% Cumulative variance	24.6	45.3	59.0

REGRESSION ANALYSIS

The choice $c = 100$ in the transformation log $(1 + c \times \text{density})$ gave residuals satisfying the normality assumption. The left panel of Table 3 shows the corresponding individual regression coefficients and standard errors and r-squared values for each family after fitting the MMR model with all five environmental predictors included. The right panel shows the corresponding results for a reduced model containing only the two predictors that were statistically significant in the MANOVA (Table 4).

Table 3. Coefficients and standard errors (in parenthesis) from fitting multivariate multiple regression models with all five environmental predictors (left panel) and with only the two predictors found statistically significant in the MANOVA (right panel). Coefficients with *p*-values greater than 0.05 in both models are omitted; those adjudged not honestly statistically significant are shown in italics and those with *p*-values less than 0.01 are shown in bold.

Fam	5 predictors						2 predictors		
	Factor 1	Factor 2	Factor 3	–TSS	DO	r^2	Factor 1	Factor 2	r^2
Hes	**0.76** (0.28)	-	-	-	-	0.23	0.67 (0.25)	-	0.18
Uni	0.71 (0.27)	-	-	-	-	0.16	0.53 (0.25)	-	0.08
Spi	0.44 (0.21)	-	-	-	-	0.13	0.43 (0.19)	-	0.09
Gon	**1.01** (0.26)	-	–0.76 (0.31)	1.08 (0.45)	-	0.31	0.61 (0.25)	-	0.11
Pho	–0.61 (0.25)	0.43^{ns}(0.22)	-	-	-	0.33	**–0.79** (0.23)	0.47 (0.22)	0.28
Pse	–0.64 (0.28)	**0.81** (0.24)	-	-	-	0.38	–0.65 (0.25)	**0.85** (0.24)	0.32
Ant	-	**0.69** (0.24)	-	-	-	0.19	-	**0.73** (0.24)	0.17
Pil	-	–0.68 (0.25)	-	-	-	0.17	-	**–0.66** (0.24)	0.16
Aor	-	**0.87** (0.26)	-	*–1.11 (0.53)*	-	0.26	-	**0.75** (0.27)	0.14
Ter	-	**1.04** (0.20)	*0.60 (0.28)*	*0.92 (0.41)*	-	0.45	-	**1.08** (0.21)	0.35
Cir	0.42^{ns} (0.24)	0.51 (0.21)	0.67 (0.29)	-	-	0.30	**0.69** (0.23)	0.42^{ns} (0.21)	0.19
Ner	0.27^{ns} (0.15)	*0.30 (0.14)*	*0.42 (0.19)*	-	-	0.25	**0.41** (0.14)	0.27^{ns} (0.14)	0.17
Ste	0.35^{ns} (0.25)	-	*0.65 (0.30)*	-	-	0.25	0.52 (0.23)	-	0.17
Mel	0.20^{ns} (0.18)	-	0.58 (0.22)	-	-	0.24	0.42 (0.18)	-	0.11
Ret	-	-	-	**1.56** (0.58)	-	0.22	-	-	0.05
Cap	-	-	-	-	–0.86 (0.36)	0.15	-	-	0.07
Nep	-	–0.16^{ns} (0.24)	-	-	-	0.12	-	–0.42 (0.20)	0.08
Isa	-	0.46^{ns}(0.23)	-	-	-	0.16	-	0.47 (0.23)	0.09
Mar	-	0.42^{ns}(0.22)	-	-	-	0.16	-	*0.45 (0.21)*	0.15
Alp	-	-	-	-	-	0.09	-	-	0.04
Oed	-	-	-	-	-	0.08	-	-	0.04
Aps	-	-	-	-	-	0.05	-	-	0.04
Ske	-	-	-	-	-	0.02	-	-	0.01
Tel	-	-	-	-	-	0.10	-	-	0.08

The coefficients listed are the ones statistically significant at 5% and 1% (in bold). Since there are 120 regression coefficients in all and 5% of these would be expected to have *p*-values less than 0.05 even if all their corresponding population parameters were zero, the six largest *p*-values less than 0.05 are italicized to indicate failure to achieve "honest" significance. The additional coefficients (labeled *ns*) were not statistically significant in their fitted model, but achieved significance in the other model. Note that the coefficients for TSS are reversed in sign because most were negative, so this predictor is labeled –TSS.

Table 4. MANOVA decomposition for multivariate multiple regression model with five predictors.

Source of variance	Df	Pillai	approx F	Df (num)	Df (denom)	Prob (>F)
Intercept	1	0.990	104.092	24	25	< 0.0001
Factor 1	1	0.793	3.968	24	25	0.0005
Factor 2	1	0.778	3.641	24	25	0.0010
Factor 3	1	0.596	1.535	24	25	0.1468
□TSS	1	0.503	1.055	24	25	0.4466
DO	1	0.339	0.691	24	25	0.8157
Residuals	48					

CANONICAL CORRESPONDENCE ANALYSIS

Figure 4 shows biplots based on the CCA matching the two MMR analyses. The families are represented by dots whereas the environmental predictors are represented by arrows. Each arrow determines an axis in the plots, obtained by extending the arrows in both directions.

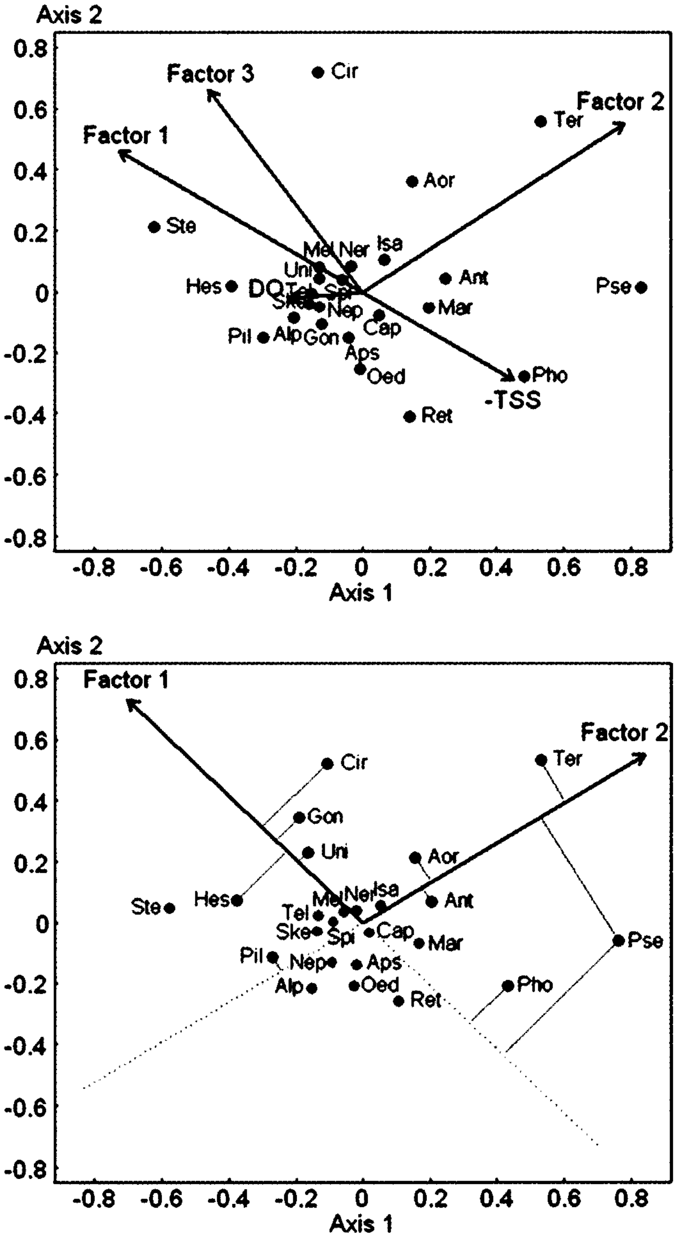

Figure 4. Upper panel: biplot of first two axes of CCA ordination diagram with 24 macrobenthic fauna families against five environmental predictors. Lower panel: similar biplot omitting the three environmental variables not statistically significant in the MANOVA, with families showing highly significant coefficients in the multivariate multiple regression model (p-values < 0.01) connected to the corresponding arrows representing the predictors

5. DISCUSSION

COMPARISON METHODS

This study aim was to identify potential environmental "key factors" accounting for the distribution of macrobenthic fauna communities, to improve understanding concerning benthic/abiotic interactions and ecosystem functioning. By replacing groups of correlated predictors by single variables, factor analysis was used to remove correlations between environmental parameters that mask their effects on the macrobenthos densities.

Multivariate multiple regression and canonical correspondence analysis were used to examine the relations between the macrobenthic fauna family densities and the reduced set of environmental predictors. Although each of these five predictors was found to be associated with at least one family, the corresponding MANOVA decomposition found only two of them to be statistically significant overall. The biplot produced by the canonical correspondence analysis is seen to be more informative when only these two predictors were included.

The MMR model containing all five predictors gives seven associations between a family density and an environmental determinant that are highly statistically significant (p-value < 0.01), and a further nine with p-value between 0.01 and 0.05. Ten families showed no evidence of an association with any of the five determinants. Most of these associations can also be seen in the CCA biplot.

The most noticeable difference between the results of the two methods is that Spionidae is found to be associated with the salinity factor in the MMR model but this association is not seen in the biplots. Since Spionidae is a marine benthos (Day and Blake [31]) with typical dominant species *Pseudopolydora kempi* and *Prionospio cirrifera* (Angsupanich et al. [22]), there is evidence supporting the MMR result in this case. In the biplot containing all five environmental predictors the arrows for Factor 1 and TSS have identical directions, but the correlation between these variables (0.30) is not high.

SCIENTIFIC FINDINGS

The results, both by MMR and CCA, clearly indicate that the salinity factor was positively associated with the densities of the Goniadidae, Hesionidae, and Spionidae and the unidentified families in the Bivalvia, and negatively associated with the densities of Pholoidae and Pseudotanaidae. This is in contrast with those analyzed based on the same data using BIOENV by Angsupanich et al. [22] which indicated no impact of salinity on benthos density. In general, salinity is an important factor affecting the distribution and structure composition of macrobenthic fauna in brackish water of coastal habitats (Mannino and Montagna [32], Ogunwenmo and Osuala [33], Nanami et al. [34]). Although Middle Songkhla Lake is not connected to the sea directly, this zone receives the effect of salinity from the saltwater inflow through the Lower Lake which is open to the Gulf of Thailand. Salinity is often regarded as a primary descriptor in estuarine ecosystems (Gaston [35], Lamptey and Armah [36]).

A sedimentary habitat contains information mirroring the functional biodiversity and activity patterns of macrobenthic fauna (Rosenberg et al. [37]). The main characteristics at the bottom of Middle Songkhla Lake are clay and silt (Angsupanich et al. [22]) except for station six, which is mainly sand (84.6%). We found that sand/clay excess was positively associated with the densities of Terebellidae, Aoridae, Pseudotanaidae, and Anthuridae, while a negative association was found with Pilargiidae. A typical genus *Sigambra* within Pilargiidae (Angsupanich et al. [22]) was found to be negatively related with sand-clay excess, a finding supported by a study in the southeastern Gulf of California reporting that the genus *Sigambra* was dominant in the areas of sand percentage of 1% or mud of 60-70% (Méndez [38]).

In addition, the genus *Marginella* within Marginellidae was also listed as being present in Middle Songkhla Lake (Angsupanich et al. [22]) thus showing a positive association with sand. This finding agrees with a study of invertebrate species identified in Fresh Creek, Bahamas where *Marginella* was listed as most commonly having the habitat type of sandflat (Layman and Silliman [39]).

Ten families (Nereididae, Stenothyridae, Nephtyidae, Isaeidae, Marginellidae, Alpheidae, Oedicerotidae, Apseudidae, Skeneopsidae, and Tellinidae) showed no evidence of association with any of the environmental variables. Although, Alpheidae was found to have the lowest density among the families included in our study, it is commonly found in the stomach contents of the dominant bottom feeding fish (*Osteogeneiosus militaris* and *Arius maculates*) in Middle

Songkhla Lake. Angsupanich et al. [39] implied that these catfish species feed opportunistically on a variety of prey in their environment coupled with preferential feeding. So the low occurrence of Alpheidae may have been due to its swift movement and consequent catching difficulty.

Nereididae is one of the most important polychaete due to its diversity and abundance, found not only in marine environments (Gonzalez-Escalante and Salazar-Vallejo [40]) but also in brackish water such as occurs in Middle Songkhla Lake. Fourteen species of Nereididae were reported in a former study (Angsupanich et al. [22]) and it seems that Nereididae is widespread in Middle Songkhla Lake where it was the highest species richness. No evidence of Nereididae variation with salinity was found, possibly due to species diversity within this family. Some species, such as *Ceratonereis hircinicola*, were widely spread in the high salinity areas (Angsupanich and Kuwabara [41]), whereas *Namalycastis indica* has been found to inhabit fresh to slightly brackish water in cisterns, pools and lagoons (Glasby [42]).

Songkhla Lake nowadays suffers from the use of coastal land and water resources for uncontrolled shrimp farming, the destruction of both mangrove areas and peat swamp forest, construction of intake and outfall structures, and the construction of a deep sea port (Chufamanee et al. [43]). The analytic methods we have used are designed to gain a better understanding of the environmental factors associated with macrobenthic fauna and can be used as an additional or alternative method for analysis of the relationships between environmental variables and the abundance of benthic organisms. This knowledge is useful for the natural resource management of estuarine environments.

ACKNOWLEDGEMENTS

We thank Professor Don McNeil for his advice on the statistical analysis. The data were collected using funds provided by the Biodiversity Research and Training Program (BRT), under the National Science and Technology Development Agency (NSTDA), together with the Thailand Research Fund (TRF) with the grant number 142016.

REFERENCES

[1] Weisberg, S.B., Dauer, D.M., Schaffner, L.C., and Frithsen, J.B., An estuarine benthic index of biotic integrity (B-BI) for Chesapeake Bay, *Estuaries*, 1997; **20**: 149-158.

[2] Borja, A., Franco, J., and Pérez, V., A marine biotic index to establish the ecological quality of soft-bottom benthos within European estuarine and coastal environments, *Mar. Pollut. Bull.*, 2000; **40**: 1100-1114.

[3] Ranasinghe, J.A., Thompson, B., Smith, R.W., Lowe, S., and Schiff, K.C., Evaluation of benthic assessment methodology in Southern California Bays and San Francisco Bay, Technical Report No. 432, Southern California Coastal Water Research Project, 2004.

[4] Gray, J.S., Aschan, M., Carr, M.R., Clarke, K.R., Rosenberg, R., and Warwick, R.M., Analysis of community attributes of the benthic macrofauna of Frierfjord/ Langesundfjord and in a mesocosm experiment, *Mar. Ecol. Prog. Ser.*, 1988; **46**: 151-165.

[5] Warwick, R.M., A new method for detecting pollution effects on marine macrobenthic communities, *Mar. Biol.*, **92**: 557-562.

[6] Clarke, K.R., and Warwick, R.M., *Change in Marine Communities: an Approach to Statistical Analysis and Interpretation*, Natural Environment Research Council, Plymouth, UK, 1994.

[7] Clarke, K.R., Somerfield, P.J., and Chapman, M.G., On resemblance measures for ecological studies, including taxonomic dissimilarities and a zero-adjusted Bray-Curtis coefficient for denuded assemblages, *J. Exp. Mar. Biol. Ecol.*, 2006; **330**: 55-80.

[8] Ter Braak, C.J.F., Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis, *Ecology*, 1986; **67**: 1167-1179.

[9] von Wehrden, H., Hanspach, J., Bruelheide, H., and Wesche, K., Pluralism and diversity-trends in the use and application of ordination methods 1990-2007, *J. Veg. Sci.*, 2009; **20**: 695-705.

[10] Rakocinski, C.F., Brown, S.S., Gaston, G.R., Heard, R.W., Walker, W.W., and Summers, J.K., Macrobenthic responses to natural and contaminant-related gradients in Northern Gulf of Mexico estuaries, *Ecol. Appl.*, 1997; **7**: 1278-1298.

[11] Hawkins, C.P., Norris, R.H., Hogue, J.N., and Feminella, J.W., Development and evaluation of predictive models for measuring the biological integrity of streams, *Ecol. Appl.*, 2000; **10**: 1456-1477.

[12] Joy, M.K., and Death, R.G., Development and application of a predictive model of riverine fish community assemblages in the Taranaki region of the North Island, New Zealand, *N.Z. J. Mar. Freshwat. Res.*, 2000; **34**: 241-252.

[13] Guerra-García, J.M., González-Villa, F.J., and García-Gómez, J.C., Aliphatic hydrocarbon pollution and macrobenthic assemblages in Ceuta harbour: a multivariate approach, *Mar. Ecol. Prog. Ser.*, 2003; **263**: 127-138.

[14] Hajisamae, S., and Chou, L.M., Do shallow water habitats of an impacted coastal strait serve as nursery grounds for fish?, *Estuar. Coast. Shelf Sci.*, 2003; **56**: 281-290.

[15] Morrisey, D.J., Turner, S.J., Mills, G.N., Williamson, R.B., and Wise, B.E., Factors affecting the distribution of benthic macrofauna in estuaries contaminated by urban runoff, *Mar. Environ. Res.*, 2003; **55**: 113-136.

[16] Ysebaert, T., Herman, P.M.J., Meire, P., Craeymeersch, J., Verbeek, H., and Heip, C.H.R., Large-scale spatial patterns in estuaries: estuarine macrobenthic communities in the Schelde estuary, NW Europe, *Estuar. Coast. Shelf Sci.*, 2003; **57**: 335-355.

[17] Quintino, V., Elliott, M., and Rodrigues, A.M., The derivation, performance and role of univariate and multivariate indicators of benthic change: Case studies at differing spatial scales, *J. Exp. Mar. Biol. Ecol.*, 2006; **330**: 368-382.

[18] Anderson, M.J., Animal-sediment relationships re-visited: Characterising species distributions along an environmental gradient using canonical analysis and quantile regression splines, *J. Exp. Mar. Biol. Ecol.*, 2008; **366**: 16-27.

[19] Glockzin, M., and Zettler, M.L., Spatial macrozoobenthic distribution patterns in relation to major environmental factors- A case study from the Pomeranian Bay (southern Baltic Sea), *J. Sea. Res.*, 2008; **59**: 144-161.

[20] Liang, S., Shieh, B., and Fu., Y., A structural equation model for physiochemical variables of water, benthic invertebrates, and feeding activity of waterbirds in the Sitsao wetlands of Southern Taiwan, *Zool. Stud.*, 2002; **41**: 441-451.

[21] Warton, D.I., and Hudson, H.M., A MANOVA statistic is just as powerful as distance-based statistics, for multivariate abundances, *Ecology*, 2004; **85**: 858-874.

[22] Angsupanich, S., Siripech, A., and Charoenpornthip, M., Macrobenthic fauna community in the Middle Songkhla Lake, Southern Thailand, *Songklanakarin J. Sci. Technol.*, 2005(a); **27**: 365-390.

[23] Ter Braak, C.J.F., and Šmilauer, P., CANOCO Reference manual and CanoDraw for Windows User's guide: Software for Canonical Community Ordination (version 4.5), Microcomputer Power, Ithaca, NY, USA, 2002.

[24] R Development Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, URL http://www.R-project.org, 2009.

[25] Browne, M.W., An overview analytic rotation in exploratory factor analysis, Multivariate. *Behav. Res.*, 2001; **36**: 111-150.

[26] Abdi, H., Factor rotations in factor analyses. *In*: Lewis-Beck, M., A. Bryman, T. Futing (Eds), *Encyclopedia of Social Sciences Research Methods*, Thousand Oaks, CA, Sage, 2003: 1-8.

[27] Mardia, K.V., Kent, J.T., and Bibby, J.M., *Multivariate Analysis*, Academic Press, USA, 1979.

[28] Venables, W.N., and Ripley, B.D., *Modern Applied Statistics with S*, 4th Edn., Springer Science + Business Media, New York, USA, 2002.

[29] Olson, C.L., On choosing a test statistic in multivariate analysis of variance, *Psychol. Bull.*, 1976; **83**: 579-586.

[30] Johnson, R.A., and Wichern., D.W., *Applied Multivariate Statistical Analysis*, 4th Edn., Prentice-Hall, USA, 1998.

[31] Day, R.L., and Blake, J.A., Reproduction and larval development of *Polydora Giardi* Mesnil (Polycheata: Spionidae), *Biol. Bull.*, 1979; **156**: 20-30.

[32] Mannino, A., and Montagna, P.A., Small-scale spatial variation of macrobenthic community structure, *Estuaries*, 1997; **20**: 159-173.

[33] Ogunwenmo, C.A., and Osuala., I.A., Physico-chemical parameters and macrobenthos of an estuarine creek and an artificial pond in Lagos, South-Western Nigeria, *Acta SATECH*, 2004; **1**: 128-132.

[34] Nanami, A., Saito, H., Akita, T., Motomatsu, K., and Kuwahara, H., Spatial distribution and assemblage structure of macrobenthic invertebrates in a brackish lake in relation to environmental variables, *Estuar. Coast. Shelf Sci.*, 2005; **63**: 167-176.

[35] Gaston, G.R., Lee, D.A.L., and Nasci, J.C., Estuarine macrobenthos in Calcasieu Lake, Louisiana: Community and trophic structure, *Estuaries and Coasts*, 1988; **11**: 192-200.

[36] Lamptey, E., and Armah, A.L., Factors affecting macrobenthic fauna in a tropical hypersaline coastal lagoon in Ghana, West Africa, *Estuaries and Coasts*, 2008; **31**: 1006-1019.

[37] Rosenberg, R., Magnusson, M., and Nilsson, H.C., Temporal and spatial changes in marine benthic habitats in relation to the EU Water Framework Directive: The use of sediment profile imagery, *Mar. Pollut. Bull.*, 2009 ; **58**: 565-572.

[38] Méndez, N., Relationships between deep-water polychaete fauna and environmental factors in the southeastern Gulf of California, Mexico, *Sci. Mar. (Barc.)*, 2007; **71**: 605-622.

[39] Angsupanich, S., Somsak, S., and Phrommoon, J., Stomach contents of the catfishes *Osteogeneiosus militaris* (Linnaeus, 1758) and *Arius maculatus* (Thunberg, 1792) in the Songkhla Lake, *Songklanakarin J. Sci. Technol.*, 2005(b); **27**: 391-402.

[39] Layman, C.A., and Silliman, B.R., Preliminary survey and diet analysis of juvenile fishes of an estuarine creek on Andros Island, Bahamas, *Bull. Mar. Sci.*, 2002; **70**: 199-210.

[40] Gonzalez-Escalante, L.E., and Salazar-Vallejo, S.I., A new estuarine species, *Nereis garwoodi* (Polychaeta: Nereididae), from Bahia Chetumal, Macican Caribbean coast, *Rev. Biol. Trop.*, 2003; **51**: 155-164.

[41] Angsupanich, S., and Kuwabara, R., Macrobenthic fauna in Thale Sap Songkla, a brackish lake in southern Thailand, *Lakes Reservoirs: Res. Manage.*, 1995; **1**: 115-125.

[42] Glasby, C.H., *The Namanereidinae (Polychaeta: Nereididae), Part 1, Taxonomy and phylogeny*, Records of the Australian Museum, 1999: 129 pp.

[43] Chufamanee, P., Boromthanarat, S., and Lønholdt, J., A case experience on integrated environment and water management towards people's livelihoods, Case story 1: Partnership policy in Songkhla Lake Basin, Workshop report. Songkhla workshop on linking management of catchments and coastal ecosystems, Royal Thai Ministry of Natural Resources and Environment, 2003: 15-32.

Appendix 3

Publication 3: Confidence intervals for adjusted proportions using logistic regression

Confidence Intervals for Adjusted Proportions Using Logistic Regression

Noodchanath Kongchouy (Corresponding author)
Department of Mathematics, Faculty of Science and Technology,
Prince of Songkla University, Hat Yai, Songkhla, 90112, Thailand
Tel: 66-84-063-1251 E-mail: nootchanath.k@psu.ac.th

Uraiwan Sampantarak
Pattani Inland Fisheries Research and Development Center,
Department of Fisheries, Ministry of Agricultural and Cooperatives, Pattani 94160, Thailand
Tel: 66-84-653-9391 E-mail: uraiwans@fisheries.go.th; uraiwan111@hotmail.com

This study was partially funded by the Graduate School, Prince of Songkla University.

Abstract

This paper presents confidence intervals for adjusted proportions using logistic regression with weighted sum contrasts. The methods are applied to data from two studies, (1) imposex percentages among female gastropods at different locations in the Gulf of Thailand adjusted for different species, and (2) complication-based neonatal morbidity risk for births at a major hospital adjusted for demographic factors.

Keywords: Confidence interval, Proportion, Logistic regression, Sum

1. Introduction

Odds ratios are conventionally used for assessing associations between binary outcomes and categorical risk factors. They are often preferred in scientific studies because they give valid confidence intervals for these associations for case-control studies as well as for cohort studies and cross-sectional studies (Fernandez et al., 1999; Lim & Tongkumchum, 2009; Peters et al., 2000). A further advantage is that methods such as Mantel-Haenszel adjustment and logistic regression are available for adjusting odds ratios for confounding bias arising from covariates associated with both the binary outcome and the risk factor of interest. These issues are discussed in detail in the biostatistical literature (see, for example, McNeil, 1996; Woodward, 1999).

For cohort and cross-sectional studies where the proportions or percentages of an adverse outcome are of primary interest, it is also important to give confidence intervals for comparing these proportions. If there are no covariates of interest, these confidence intervals can be computed, either directly from the observed proportions with some adjustment such as an arcsine transformation to ensure that the confidence intervals are between 0 and 1 (see, for example, Armitage & Berry, 1994) using a logistic regression model. However, the situation is more complicated when adjustments for covariates are required, and methods for constructing such confidence intervals are not routinely provided in statistical packages.

In this paper we describe a general method for computing confidence intervals for comparing several proportions after adjusting for categorical covariates. The method is illustrated using two recently published applications of scientific interest: imposex among female gastropods in the Gulf of Thailand (Swennen, Sampantarak, & Rattanadakul, 2009), and complication-based neonatal morbidity risk among babies at a major hospital in southern Thailand (Rachatapantanakorn & Tongkumchum, 2009).

2. Methods

2.1 Logistic Regression

Logistic regression (Hosmer & Lemeshow, 2000; Kleinbaum & Klein, 2002) is a statistical method widely used to model the association between a binary outcome probability - the probability of a specific outcome - and a set of fixed determinants. When the determinants are categorical factors, these factors can be structured as a multi-way contingency table of counts and the data for analysis comprise the proportions of adverse outcomes in the cells of this table. If, for example, the outcome variable Y takes values 0 and 1 (adverse outcome) and there are two factors with levels indexed by i and j, respectively, the model takes the form

$$\mathrm{Prob}[Y_{ij}=1]=1/\left(1+e^{-(\alpha_i+\beta_j)}\right). \tag{1}$$

Note that if r and c are the numbers of factor levels specified in the model by the sets of parameters $\{\alpha\}$ and $\{\beta\}$, the number of independent parameters is $r + c - 1$, so it is necessary to put a constraint on these parameters when fitting the model. This constraint is conventionally achieved by replacing $\alpha_i + \beta_j$ in the model by $\mu + \alpha_i + \beta_j$ where $\alpha_1 = 0$ and $\beta_1 = 0$, in which case the model is written as

$$\mathrm{Prob}[Y_{ij}=1]=1/\left(1+e^{-(\mu+\alpha_i+\beta_j)}\right). \tag{2}$$

This model provides estimates of odds ratios for comparing the outcome probabilities with respect to specified levels of each factor. Thus $\exp(\alpha_i)$ is the ratio of the odds of an adverse outcome for level i to that for level 1 for the first factor, whereas $\exp(\beta_i)$ is the ratio of the odds of an adverse outcome for level j to that for level 1 for the second factor. Thus each odds ratio in this model uses the first level of each factor as a baseline. To obtain odds ratios with respect to another baseline level, the data would need to be recoded so that the new baseline is constrained to take the value 0.

Under appropriate conditions on the pattern of zeros in the data (see, for example, Section 7.2 of Venables & Ripley, 2002), this logistic regression model is fitted using maximum likelihood and the results include estimates of the parameters and their standard errors, from which confidence intervals can be plotted.

2.2 Contrasts

When the constraints $\alpha_1 = 0$ and $\beta_1 = 0$ are used in the logistic regression model (2) the confidence intervals apply to the *differences* between each of the sets of parameters and the first parameter specified in each factor. These differences are known as *treatment contrasts.* In practice, it is often preferable not to single out a specific level of a factor as a basis for comparison, but rather to treat all factor levels in the same way. For linear models of the form

$$Y_{ij}=\alpha_i+\beta_j+\varepsilon_{ij}, \tag{3}$$

standard errors of these differences are obtained by using the standard sum contrasts available in commonly used software packages such as R program (R Development Core Team, 2007). However, as pointed out by Venables and Ripley (2002), these contrasts are not valid for unbalanced designs, which include logistic regression models. Thus it is necessary to construct specific contrasts for logistic regression, and this can be accomplished by using weighted sum contrasts rather than treatment contrasts (Tongkumchum & McNeil, 2009). These weighted sum contrasts provide standard errors for the differences between each factor level and their overall mean.

2.3 Adjustment for Covariates

The method is analogous to that commonly used by linear regression analysis for adjusting outcomes to reduce the effects of covariate factors, such as seasonal adjustment of unemployment rates. In this case, if $\{\beta\}$ represents the primary factor of interest and $\{\alpha\}$ the covariate factor, and ε_{ij} are independent errors with mean zero and common standard deviation, the model is given by Equation (3). The factors $\{\alpha\}$ and $\{\beta\}$ in this model are estimated as the row and column means of the data matrix y_{ij}, with a suitable constraint to ensure identifiability. To adjust for the covariate, the adjusted mean for level j of the primary factor is obtained by first removing the effect of the covariate from each observation by replacing y_{ij} by $y_{ij}-\hat{\alpha}_i$ and then adding a constant to ensure that the mean of the corrected observations remains the same as the mean of the original observations. As a result, the adjusted mean is $\overline{y}_j=\hat{\beta}_j+d$, where d is a constant chosen to ensure that the overall mean before and after the adjustment remains the same. It follows that $d=\overline{y}-\overline{\beta}$, where $\overline{\beta}$ is the mean of the estimated β parameters.

Similarly, the formula for adjusting the proportion for level j of a primary factor is

$$p_j^{*}=1/\left(1+e^{-(\beta_j+d)}\right). \tag{4}$$

Note that this result follows from the fact that the estimate given by logistic regression for the adjusted odds ratio for level j compared to level 1 is $\exp(\hat{\beta}_j)$ and this must equate to $\left(\frac{p_j^*}{1-p_j^*}\right)\Big/\left(\frac{p_1^*}{1-p_1^*}\right)$.

The constant d may be chosen to ensure that the overall proportion (or the total number N) of adverse outcomes before and after the adjustment is the same, that is

$$\sum_{i=1}^{r}\sum_{j=1}^{c} n_{ij}p_j^* = \sum_{i=1}^{r}\sum_{j=1}^{c} n_{ij}p_{ij} = N. \qquad (5)$$

Substituting Equation (4) into Equation (5), it follows that d must satisfy the equation

$$\sum_{j=1}^{c}\left(\sum_{i=1}^{r} n_{ij}\right)\Big/\left(1+e^{-(\beta_j+d)}\right) = N. \qquad (6)$$

Equation (6) is non-linear and cannot be solved explicitly to give an expression for the constant d. However, it can be solved using the Newton-Raphson iterative procedure with Marquardt damping to ensure convergence. Note that this method extends straightforwardly to any number of covariate factors.

3. Applications

3.1 Imposex among gastropods in the Gulf of Thailand

Female gastropods were collected from 56 sampling sites grouped into 13 areas around the Gulf of Thailand in 2006 (Swennen et al., 2009) and tested for imposex. Since different species have different sensitivities to be imposex and this variation could bias the estimation of imposex prevalence due to the fact that different species are found in different locations, it was necessary to take both factors (16 species groups and 13 areas) into account. The logistic regression model (Equation 2) was fitted to the data. Figure 1 shows plots of 95% confidence intervals for the proportions.

The overall imposex percentage was 25.2, indicated by the vertical lines in Figure 1. Although there were 208 (13 areas × 16 species groups) cells in the data table, 124 of these cells contained no data because relatively for species groups were found in each area. Thus the number of degrees of freedom after fitting the two-factor model was reduced from 71 (84 – 13) to 56 (84 – 13 – 15).

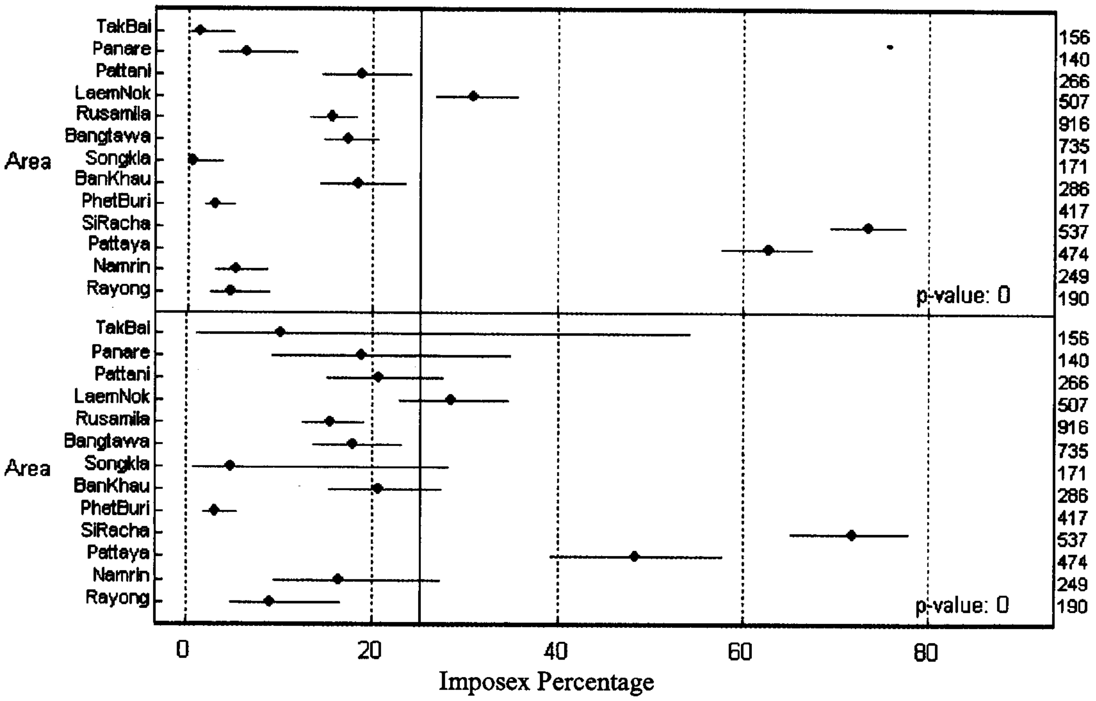

Figure 1. The 95% confidence intervals for percentages of female gastropods with imposex disease at various locations around the Gulf of Thailand; the upper and lower panels show crude and species-adjusted percentages, and the sample sizes are given on the right.

Figure 1 indicates that the crude percentages overestimated the true values in the Pattaya area and underestimated them in the Tak Bai, Pattani and Namrin areas. Due to the fact that 15 additional parameters were needed to take the species factor into account, the adjusted percentages have wider confidence intervals than the crude percentages. This effect is particularly notable for the Tak Bai area because the imposex was only found to occur among one of the three species groups observed there.

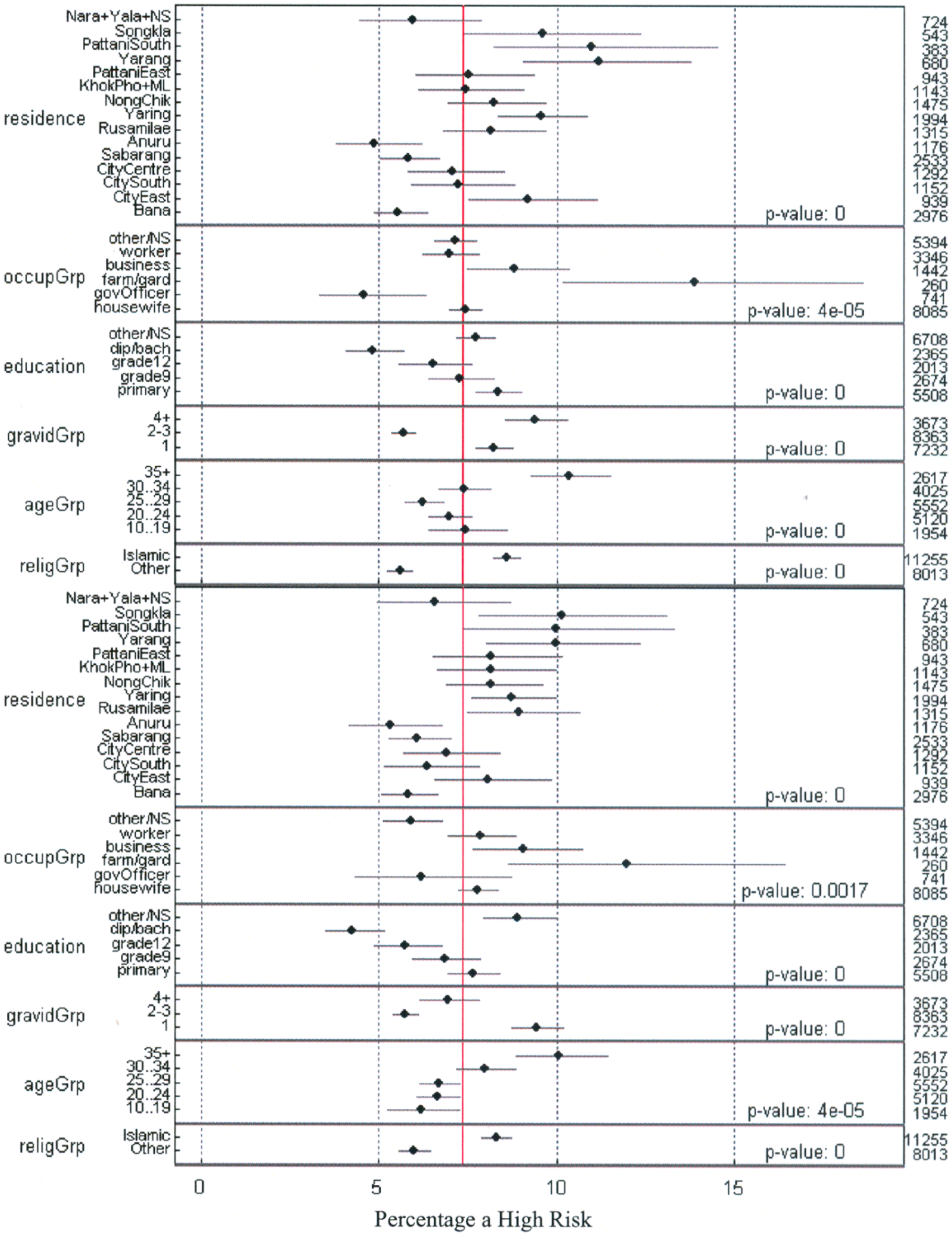

Figure 2. The 95% confidence intervals for complication-based risks among singleton deliveries to mothers with no previous Caesarean-section birth at Pattani Hospital, with respect to each of six demographic factors before (upper panel) and after (lower panel) adjustment for other factors.

3.2 Complication-based neonatal morbidity risk

Based on complications recorded for 19,268 singleton deliveries to mothers with no previous Caesarean-section birth at Pattani Hospital in Southern Thailand over a nine-year period from 1997 to 2005, Rachatapantanakorn and Tongkumchum (2009) classified babies as high or low risk, and used logistic regression to assess the effects of six demographic risk factors.

Figure 2 shows 95% confidence intervals of the crude and adjusted percentages for each determinant before and after adjusting for the other determinants. The main differences between the crude and adjusted risks occurred among three demographic factors: the number of pregnancies (gravid group), age-group and religion.

Mothers in gravid group 4 or more pregnancies had highest risk percentage before adjustment, but after adjusting for the other determinants the risk was highest for primigravid mothers. For the age-group effect, the crude percentage was lowest for mothers aged 25-29 years, whereas the adjusted percentage was lowest for mothers aged less than 20 years. It is also noteworthy that the quite large difference in risks between Muslim and non-Muslim mothers was reduced after adjusting for the other factors.

4. Discussion

In this paper we have described a simple method for adjusting proportions for categorical covariates based on a fitted logistic regression model that provides asymptotically valid confidence intervals for comparing proportions over different levels of a categorical risk factor. While this method is not entirely new (see, for example, related earlier work by Berthold et al., 2007; Graubard & Korn, 1999; Lane & Nelder, 1982), it is not widely used in scientific studies, particularly when comparing more than two proportions. A further advantage of the method is that by using appropriately weighted sum contrasts each proportion can be compared with the overall mean rather than with a specified reference group.

5. Acknowledgement

We are grateful to Professor Don McNeil for his helpful guidance.

References

Armitage, P., & Berry, G. (1994). *Statistical methods in medical research.* (3rd ed.). UK: Blackwell-science, Oxford.

Berthold, S. M., Wong, E. C., Schell, T. L., Marshall, G. N., Elliott, M. N., Takeuchi, D., & Hambarsoomians, K. (2007). U.S. Cambodian refugees' use of complementary and alternative medicine for mental health problems. *Psychiatric Services*, 58, 1212-1218.

Fernandez, E., Schiaffino, A., Rajmil, L., Badia, X., & Segura, A. (1999). Gender inequalities in health and health care services use in Catalonia (Spain). *Journal of Epidemiology and Community Health*, 53(4), 218-222.

Graubard, B. I., & Korn, E. L. (1999). Predictive margins with survey data, *Biometrics*, 55, 652-659.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression.* (2nd ed.). John Wiley and Sons.

Kleinbaum, D. G., & Klein, M. (2002). *Logistic Regression: a Self-Learning Text.* (2nd ed.). New York: Springer-Verlag.

Lane, P. W., & Nelder, J. A. (1982). Analysis of covariance and standardization as instances of prediction, *Biometrics*, 38(3), 613-621.

Lim, A., & Tongkumchum, P. (2009). Methods for analyzing hospital length of stay with application to inpatients dying in Southern Thailand. *Global Journal of Health Science*, 1(1), 27-38.

McNeil, D. (1996). *Epidemiological research methods.* Chichester: John Wiley & Sons.

Peters, A., Liu, E., Verrier, R. L., Schwartz, J., Gold, D. R., Mittleman, M., Baliff, J., Oh, A. J., Allen, G., Monahan, K., & Dockery, D. W. (2000). Air pollution and incidence of cardiac arrhythmia. *Epidemiology*, 11, 11-17.

R Development Core Team. (2007). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org.

Rachatapantanakorn, O., & Tongkumchum, P. (2009). Demographic determinants for caesarean delivery in Pattani hospital. *Southeast Asian Journal of Tropical Medicine and Public Health*, 40, 602-611.

Swennen, C., Sampantarak, U., & Rattanadakul, N. (2009). TBT-pollution in the Gulf of Thailand: a re-inspection of imposex incidence after 10 years. *Marine Pollution Bulletin*, 58, 526-532.

Tongkumchum, P., & McNeil, D. (2009). Confidence intervals using contrasts for regression model. *Songklanakarin Journal of Science and Technology*, 31, 151-156.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S.* (4th ed.). New York: Springer Science + Business Media, (Section 6.2).

Woodward, M. (1999). *Epidemiology: study design and data analysis.* USA: Chapman & Hall.

Appendix 4: Names of the study areas indicated by a letter and the name of a town or village on the adjacent coast, the positional coordinates, average depth of the sampling sites, the names of the species, number of males, number of females, and number of imposex in 2006

Study area	Site	Position	Depth (m)	Species	Males	Females	Imposex
A: Rayong	1	12° 35' 00" N 101° 29' 00" E	8-20	*Murex altispira*	8	31	0
				Murex trapa	0	4	0
				Thais lacera	2	3	3
				Babylonia areolata	11	20	0
				Pugilina cochlidium	75	52	2
				Hemifusus ternatanus	3	2	0
	2	12° 37' 47" N 101° 26' 46" E	0	*Semiricinula muricoides*	37	68	3
				Thais rufotincta	12	10	1
B: Namrin	3	12° 40' 39" N 101° 06' 47" E	0	*Semiricinula muricoides*	3	19	3
				Thais clavigera	9	15	9
				Morula musiva	84	87	0
	4	12° 40' 00" N 101° 02' 48" E	0	*Semiricinula muricoides*	15	33	0
				Thais clavigera	30	21	0
				Morula musiva	79	74	1
C: Pattaya	5	12° 58' 14" N 100° 53' 30" E	0	*Lataxiena blosvillei*	203	242	214
	6	12° 59' 45" N 100° 55' 22" E	0	*Semiricinula muricoides*	45	155	68
	7	13° 00' 18" N 100° 55' 24" E	0	*Pugilina cochlidium*	58	77	15
D: Si Racha	8	13° 07' 39" N 100° 53' 54" E	0	*Semiricinula muricoides*	32	126	109
	9	13° 09' 30" N 100° 54' 00" E	6	*Murex trapa*	2	3	3
				Pugilina cochlidium	9	6	4
				Hemifusus ternatanus	31	59	16
	10	13° 09' 00" N 100° 54' 27" E	0	*Semiricinula muricoides*	100	120	82
	11	13° 12' 16" N 100° 56' 00" E	0	*Lataxiena blosvillei*	152	124	111
				Pugilina cochlidium	0	1	1
	12	13° 10' 00" N 100° 50' 00" E	4	*Murex trapa*	15	7	5
				Hemifusus ternatanus	1	3	0

Appendix 4: (Cont.)

Study area	Site	Position	Depth (m)	Species	Males	Females	Imposex
	13	13° 10' 00" N 100° 49' 00" E	8	*Murex trapa*	35	37	27
	14	13° 10' 00" N 100° 47' 00" E	12-20	*Murex trapa*	51	50	36
				Nassarius livescens	1	0	0
				Hemifusus ternatanus	1	1	1
				Turricula javana	1	0	0
E: Phet Buri	15	13° 06' 00" N 100° 04' 00" E	5	*Nassarius stolatus*	86	78	0
	16	13° 06' 00" N 100° 07' 00" E	2-5	*Thais lacera*	52	62	1
				Babylonia areolata	1	0	0
				Pugilina cochlidium	4	4	0
				Hemifusus ternatanus	1	3	0
				Turricula javana	0	1	0
	17	13° 05' 00" N 100° 12' 00" E	6	*Thais lacera*	18	21	0
				Pugilina cochlidium	12	11	0
				Hemifusus ternatanus	1	0	0
	18	13° 03' 00" N 100° 16' 00" E	10-30	*Murex altispira*	0	2	0
				Murex trapa	146	232	12
				Pugilina cochlidium	1	0	0
				Hemifusus ternatanus	0	3	0
F: Ban Khau	19	12° 53' 00" N 100° 03' 00" E	6	*Murex trapa*	16	47	22
				Thais lacera	5	13	2
				Pugilina cochlidium	7	12	0
				Hemifusus ternatanus	44	61	0
	20	12° 53' 00" N 100° 06' 00" E	9-10	*Murex trapa*	47	91	17
				Thais lacera	4	6	0
				Pugilina cochlidium	20	22	0
	21	12° 53' 00" N 100° 10' 00" E	10-14	*Murex trapa*	2	13	7
				Thais lacera	19	15	5
	22	12° 10' 00" N 100° 10' 00" E	8-10	*Murex trapa*	2	5	0
				Thais lacera	2	0	0
				Turricula javana	0	1	0

Appendix 4: (Cont.)

Study area	Site	Position	Depth (m)	Species	Males	Females	Imposex
G: Songkla	23	07° 15' 00" N		*Murex altispira*	33	150	0
		100° 46' 00" E		*Rapana rapiformis*	30	19	1
				Hemifusus ternatanus	0	2	0
H: Bang Tawa	24	6° 52' 00" N	2	*Murex altispira*	0	1	0
		101° 04' 20" E		*Murex occa*	6	7	0
				Murex trapa	15	18	1
				Thais lacera	3	3	0
				Nassaria pusilla	12	44	1
				Nassarius siquijorensis	25	50	3
				Nassarius jacksonianus	0	3	0
				Hemifusus ternatanus	0	2	1
				Turricula javana	32	68	3
	25	6° 52' 00" N	2	*Murex occa*	1	1	0
		101° 07' 25" E		*Thais lacera*	1	3	1
				Nassarius siquijorensis	4	26	13
				Hemifusus ternatanus	0	5	2
				Turricula javana	0	6	2
	26	6° 51' 37" N 101° 09' 30" E	0	*Thais bitubercularis*	208	208	30
	27	6° 52' 00" N	2	*Murex occa*	2	2	0
		101° 09' 30" E		*Thais lacera*	3	3	0
				Babylonia areolata	0	4	0
				Nassaria pusilla	2	9	0
				Nassarius siquijorensis	14	103	40
				Nassarius jacksonianus	1	6	0
				Pugilina cochlidium	5	9	0
				Hemifusus ternatanus	1	11	4
				Turricula javana	22	101	26
	28	6° 58' 30" N	8-10	*Murex altispira*	2	3	0
		101° 10' 00" E		*Murex trapa*	2	0	0
				Rapana rapiformis	1	0	0
				Hemifusus ternatanus	1	0	0
	29	6° 52' 49" N	3-4	*Murex occa*	1	2	0
		101° 10' 00" E		*Murex trapa*	0	1	1

Appendix 4: (Cont.)

Study area	Site	Position	Depth (m)	Species	Males	Females	Imposex
				Pugilina cochlidium	9	4	0
				Hemifusus ternatanus	35	32	0
I: Rusamilae	30	6° 53' 48" N 101° 13' 55" E	3	*Murex occa*	16	19	2
				Thais lacera	12	6	0
				Nassarius siquijorensis	2	1	0
				Nassarius jacksonianus	19	42	5
				Pugilina cochlidium	0	1	0
				Hemifusus ternatanus	12	8	0
				Turricula javana	19	41	0
	31	6° 53' 20" N 101° 13' 00" E	3	*Murex occa*	20	22	3
				Nassarius jacksonianus	36	102	12
				Nassarius livescens	13	8	1
				Pugilina cochlidium	5	5	0
				Hemifusus ternatanus	7	6	0
				Turricula javana	33	54	1
	32	6° 53' 30" N 101° 13' 30" E	2	*Murex trapa*	125	182	28
				Thais lacera	25	29	2
				Hemifusus ternatanus	2	18	9
				Turricula javana	6	33	7
	33	6° 53' 30" N 101° 14' 00" E	3-5	*Murex occa*	130	48	19
	34	6° 55' 00" N 101° 13' 00" E	4	*Murex occa*	41	64	23
				Thais lacera	0	1	0
				Nassarius siquijorensis	7	7	0
				Nassarius stolatus	68	151	14
				Turricula javana	11	68	17
J: Laem Nok	35	6° 52' 40" N 101° 15' 30" E	2	*Murex altispira*	0	1	0
				Murex occa	67	43	13
				Murex trapa	0	2	0
				Thais bitubercularis	1	3	0
				Nassarius siquijorensis	5	14	2
				Nassarius jacksonianus	9	12	1
				Pugilina cochlidium	1	2	0
				Hemifusus ternatanus	1	3	1

Appendix 4: (Cont.)

Study area	Site	Position	Depth (m)	Species	Males	Females	Imposex
				Turricula javana	5	14	5
	36	6° 53' 00" N	1.5	*Murex occa*	4	6	4
		101° 18' 00" E		*Thais lacera*	59	21	5
				Pugilina cochlidium	78	69	3
	37	6° 53' 30" N	2	*Murex occa*	14	16	8
		101° 18' 30" E		*Thais lacera*	17	11	1
				Pugilina cochlidium	90	50	1
	38	6° 54' 00" N	1-3	*Murex occa*	0	1	1
		101° 19' 30" E		*Thais lacera*	22	22	4
	39	6° 54' 00" N	1.5	*Murex occa*	72	90	46
		101° 17' 00" E		*Murex trapa*	0	2	0
				Thais lacera	5	28	19
				Pugilina cochlidium	72	48	10
	40	6° 53' 30" N	3-4	*Murex occa*	45	35	29
		101° 17' 30" E		*Thais lacera*	27	14	4
K: Pattani Bay Mouth	41	6° 54' 30" N	4	*Murex occa*	4	8	7
		101° 15' 30" E		*Thais lacera*	1	6	5
				Pugilina cochlidium	5	2	1
	42	6° 55' 00" N	2-3	*Murex occa*	31	33	15
		101° 15' 00" E		*Thais lacera*	2	36	15
				Pugilina cochlidium	51	34	2
	43	6° 55' 16" N	5	*Murex occa*	12	13	0
		101° 14' 20" E		*Nassarius livescens*	14	13	0
				Pugilina cochlidium	5	1	0
				Hemifusus ternatanus	2	5	0
				Turricula javana	4	12	2
	44	6° 55' 24" N	5	*Murex occa*	5	7	0
		101° 14' 16" E		*Thais lacera*	2	2	0
				Nassarius jacksonianus	2	1	0
				Nassarius livescens	3	6	1
				Pugilina cochlidium	3	3	0
				Hemifusus ternatanus	0	1	1
				Turricula javana	2	6	0

Appendix 4: (Cont.)

Study area	Site	Position	Depth (m)	Species	Males	Females	Imposex
	45	6° 56' 05" N 101° 14' 55" E	0	*Nassarius stolatus*	73	77	1
L: Panare	46	7° 01' 00" N 101° 14' 00" E	10	*Murex altispira*	0	1	0
				Murex trapa	0	4	1
				Thais lacera	0	1	0
				Babylonia areolata	0	3	0
	47	7° 02' 00" N 101° 16' 00" E	10-15	*Nassarius livescens*	2	2	0
	48	7° 08' 00" N 101° 15' 00" E	20	*Phos senticosus*	1	1	0
				Nassarius siquijorensis	1	1	0
	49	6° 54' 45" N 101° 23' 00" E	8-10	*Murex altispira*	1	0	0
				Chicoreus banksii	0	2	0
				Rapana rapiformis	0	1	0
	50	7° 03' 00" N 101° 32' 00" E	20-30	*Murex altispira*	5	30	1
				Chicoreus banksii	7	8	1
	51	6° 54' 50" N 101° 28' 00" E	10	*Murex altispira*	6	12	0
				Rapana rapiformis	1	1	0
	52	6° 57' 30" N 101° 30' 00" E	20-25	*Murex altispira*	5	29	0
				Chicoreus banksii	7	24	4
				Rapana rapiformis	21	11	1
	53	6° 57' 00" N 101° 32' 00" E	20	*Murex altispira*	6	4	1
				Chicoreus banksii	5	5	0
M: Tak Bai	54	6° 16' 50" N 102° 03' 30" E	4-6	*Babylonia areolata*	144	131	2
	55	6° 14' 50" N 102° 05' 10" E	0	*Thais bituberculari*s	4	3	0
	56	6° 20' 00" N 102° 09' 00" E	15-20	*Murex altispira*	7	22	0
					3713	5044	1270

Vitae

Name: Mrs. Uraiwan Sampantarak

Student ID: 4848012

Educational Attainment:

Degree	**Name of institution**	**Year of Graduation**
B. Sc. (Fisheries)	Khon Kaen University	1993
M. Sc. (Agribusiness)	Prince of Songkla University	2003

Work-Position and Address:

Position: Fishery Biologist, Professional Level

Address: Pattani Inland Fisheries Research and Development Center,
Inland Fisheries Research and Development Bureau,
Department of Fisheries, Ministry of Agriculture and Cooperatives

List of Publication and Proceeding:

Publications: (1) Swennen, C., Sampantarak, U., and Ruttanadakul, N. 2009. TBT-pollution in the Gulf of Thailand: a re-inspection of imposex incidence after 10 years. *Marine Pollution Bulletin*, 58 (4): 526-532.

(2) Sampantarak, U., Kongchouy, N., and Angsupanich, S. 2009. Regression-based modeling of macrobenthic fauna density in Middle Songkhla Lake, Thailand. (Manuscript)

(3) Kongchouy, N., and Sampantarak, U. 2010. Confidence intervals for adjusted proportions using logistic regression. *Modern Applied Science*, 4 (6).

Proceeding: Sampantarak, U., Swennen, C., and Ruttanadakul, N. Imposex in neogastropods in the Gulf of Thailand. National Graduate Research Conference, 15-16 May 2009. Chiangmai. page 13.