Kasetsart J. (Nat. Sci.) 49 : 188 - 199 (2015)

Factor Analysis for Clustering and Estimating Fish Distribution Pattern in a Tropical Estuary in Southern Thailand

Niwadee Saheem[1], Sarawuth Chesoh[2,*] and Apiradee Lim[1]

ABSTRACT

This study investigated the pattern of the fish standing crop using data on weight by month and sampling locations in the Na Thap River inSouthern Thailand. Samples of 28 different estuarine fish species were collected monthly at four different locations on the river from June 2005 to December 2012. Factor analysis was used to group the number of species and hence produced four interpretable factors:factor 1 was represented by a group of marine organisms, factors 2 and 3 represented mainly euryhaline species and factor 4 was a group of mesohaline species. The results indicated that all four fish groups reached maximum levels during February and March. The fish standing crop by weight for each factor was significantly associated with the month of the year. The results suggest that the fish standing crop was associated with a season.

Keyword: multivariate analysis, linear regression, fish assemblage, Na Thap River, Southern Thailand

INTRODUCTION

Fish clustering is recognized as an essential indicator for aquatic ecosystem productivity and fisheries management strategies and the examination of fish distribution patterns is accepted as providing meaningful bio-indicators of diversity index, habitat disturbances and ecological health, especially in both temperate and tropical water bodies(Able, 2005; Hajisamae *et al*, 2006). Moreover, the estuarine region is often called the uterus of the sea and it is a unique and fragile ecosystem that must be managed carefully for the mutual benefit of all who enjoy and depend on it (Blaber, 2000; Moyle and Cech, 2004). Principally, fish habitat management should be implemented corresponding to sustainable fisheries management. Unfortunately, fisheries data are often complex and fish stock are poorly managed, so the relevant data must be cleaned and the distribution pattern manipulated before further analysis is possible.

Previous studies of fish catch data have been used to determine an appropriate statistical model to assess fisheries resources. However, when many variables are available, model fitting is complicated, so reducing the number of variables is necessary. Cluster analysis is commonly used to classify similar variables in fisheries (Callaway *et al*., 2002; Cartes *et al*., 2002; Jackson *et al*., 2010). Factor analysis can also be used to reduce the variables by removing redundancy or duplication

[1] Department of Mathematics and Computer Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, 94000, Thailand.

[2] Department of Science, Faculty of Science and Technology, Prince of Songkla University, Pattani Campus, 94000, Thailand.

* Corresponding author, e–mail: chesoh.s@hotmail.com

Received date : 02/04/14 Accepted date : 04/08/14

from a set of correlated variables (Zuur *et al*., 2003; Deka *et al*., 2005; Erzini *et al*., 2005; Chesoh and Choonpradub, 2011). This method also helps researchers to select a small group of variables that represent variables from a larger set and characterize correlations between variables or a set of observations.

The current study focused only on the fish standing crop by weight, which is a tool often used to assess fish abundances and their variations in the ecosystem (Fausch *et al.*, 1988; Saheem et al., 2014).The biological data from many studies has indicated that not only seasonal fluctuations affect organism distribution, but the development of fishery management also affects that variation (Angsupanich and Rakkheaw, 1997; Baisre, 2000; Hajisamae and Ibrahim, 2008). Seasonal and other factors which have an effect on the dependent variables in the model should be monitored and adjusted before fitting the model. The aims of this study were to determine the relationships between fish standing crop and species identity and to investigate the pattern of the fish standing crop by month and sampling location in the Na Thap River in southern Thailand from June 2005 to December 2012. The findings from this study will provide useful information for temporal and spatial variation in the fish standing crop in a tidal river.

MATERIALS AND METHODS

Sampling site and data collection

The Na Thap River is one of the main rivers in Songkhla province, Southern Thailand. The river originates from various streams in the highlands of the Thai-Malaysian border and flows eastward to the Gulf of Thailand at Ban Pak Bang Na Thap in Chana district for approximately 26.5 km. The river supports several agricultural activities and provides an important habitat for aquatic animals. Moreover, the river is surrounded by communities, fish farming and industrial development. The water body can be classified into three parts comprising freshwater, brackish water and the estuary ecosystem. This study included only the estuarine ecosystem where four measurement sites were established (site 1: Ban Khu Namrob, site 2: Ban Ma Ngon, site 3: Ban Khlongkha and site 4: Ban Pakbang Na Thap) as shown in Figure 1. Site 4 was located at the mouth of the river where it flows into the Gulf of Thailand. The fish standing crop data consisted of the total weights of all collected aquatic organisms from each sampling site using a 200 m^2 × 1 m depth purse seine net. The aquatic organisms from the samples were classified by species, according to Taki (1974) and Rainboth (1996) guidelines.

Statistical methods

The studied population comprised the monthly fish standing crops by weight of 28 estuarine fish species collected during June 2005–December 2012. For each species, there were 364 monthly observations for 8 yr from the four different sites. Spearman's rank correlation was used to identify and test the strengths of relationships between the fish species (Coti *et al.*, 2012). Factor analysis based on maximum likelihood was used to allocate the fish species into a smaller number of interpretable groups. The Promax rotation method (Johnson and Wichern, 2007) was applied and then data were reduced as shown in Equation 1. Thus, if y_{ij} is the fish standing crop weight in month i of species j, the factor model was formulated as

$$y_{ij} = \mu_j + \sum_{k=1}^{p} \lambda_j^{(k)} f_i^{(k)} \qquad (1)$$

where μ_j is the average fish standing crop weight in species j, the p column vectors $f^{(k)}$ in this model are called common factors and the p row vectors $\lambda^{(k)}$ are called their loadings.

A factor model also provides a separate uniqueness value for species which have high uniqueness and cannot be grouped together (Mardia, 1980).

After factor analysis, the factor scores were calculated using the average fish standing crop of each factor. The factor is represented by *k*. An additive linear model was used to investigate the effect of month and location on the average fish standing crop for each factor as shown in Equation 2:

$$f_{ij} = \mu + \alpha_i + \beta_j. \quad (2)$$

where f_{ij} is the average fish standing crop weight for factor *k* in each calendar month and location (1, 2, 3 and 4), μ is an overall mean, α_I is the coefficient for each month of the year and β_j is the coefficient for the collected site. All statistical and graphical analysis was carried out using the R program (R Development Core Team, 2010).

RESULTS

The fish standing crop during the period of study in the Na Thap River was comprised of 28 common species from 3phyla, 3classes, 8orders and 24 families as shown in Table 1.

Promax rotation after fitting four factors (groups of species) gave loadings as shown in Table 2. Factor 1 had nine species consisting of a mixture of four vertebrates and five invertebrates(Indian squid, octopus, mantis shrimp, painted stone crab, oceanic paddler crab, tongue sole, yellow stripe trevally, Indo-Pacific mackerel and the tiger-toothed conger). Most of the marine organisms in this group prefer to inhabit waters with a salinity range between 18 and 30 psu. Factor 2 was made up of 10 species with 9 vertebrates and 1 invertebrate (banana prawn). Factor 3 had six species with five decapoda and the greasy grouper. Most of the second and third groups of species (factor 2 and factor 3), for example the large-scaled goby, spotted scat, crescent grunter, streaked spinefoot, mullet, banana prawn, greasy grouper,

Figure 1 Study sites (1–4) in Na Thap River, Songkhla province, Southern Thailand.

mud crab, green tiger prawn, black tiger shrimp and blue swimming crab are euryhalines(capable of living in a wide range of salinity). Factor 4 consisted of three species of Metapenaeus (shrimp post larvae, greasy back shrimp and stork shrimp), which are mesohalines and spend most of their life in waters with a salinity range of 5–18 psu. The loading factors of some species were highest in more than one group; for example, green tiger prawn, acetes and greasy back shrimp, but they were categorized to the factor that yielded the highest loading scores. The last column in Table 2 represents the uniqueness of the 28 different estuarine species. Only one species, the yellow pike-conger, had a high uniqueness (0.849). Even though this species had high uniqueness, it was included in factor 2 because the species in this group were all vertebrates and the loading factor (0.496) was also high in factor 2. The factor analysis indicated that four factors were sufficient and could explain 54.2% of the total variance of the 28 estuarine species in the river.

The correlation matrix of fish standing crop weight in each species is shown in Figure 2 as a bubble plot, ordered by the four factors. The correlation matrix of the fish standing crop weight between each species before factor modeling (Figure 2a) and the correlation matrix of residuals reduced after factor modeling were used to see how well the factor model reduced these correlations.

Table 1 Selected species found in the Na Thap River from June 2005 to December 2012

No	Common name	Sciencetific name	Class	Order	Family
1	Greasy back shrimp	*Metapenaeus ensis*	Malacostraca	Decapoda	Aristeidae
2	Banana prawn	*Penaeus merguiensis*	Malacostraca	Decapoda	Penaeidae
3	Black tiger shrimp	*Penaeus monodon*	Malacostraca	Decapoda	Penaeidae
4	Green tiger prawn	*Penaeus semisulcatus*	Malacostraca	Decapoda	Penaeidae
5	Stork shrimp	*Metapenaeus tenuipes*	Malacostraca	Decapoda	Penaeoidea
6	Acetes	*Acetes sp.*	Malacostraca	Decapoda	Sergestidae
7	Shrimp post-larvae	–	Malacostraca	Decapoda	–
8	Oceanic paddler crab	*Neodorippe callida*	Malacostraca	Decapoda	Dorippidae
9	Painted stone crab	*Matuta planipes*	Malacostraca	Decapoda	Grapsidae
10	Blue swimming crab	*Portunus pelagicus*	Malacostraca	Decapoda	Portunidae
11	Mud crab	*Scylla serrata*	Malacostraca	Decapoda	Portunidae
12	Mantis shrimp	*Cloridopsis dubia*	Malacostraca	Stomatopoda	Squilloidea
13	Octopus	*Octopus sp.*	Cephalopoda	Octopoda	Octopodidae
14	Indian squid	*Photololigo duvauceli*	Cephalopoda	Teuthida	Ocypodoidea
15	Yellow pike-conger	*Congresox talabon*	Actinopterygii	Anguilliformes	Muraenesocidae
16	Mullet	*Liza sp.*	Actinopterygii	Mugiliformes	Mugilidae
17	Yellow stripe trevally	*Selaroides leptolepis*	Actinopterygii	Perciformes	Carangidae
18	Silver biddy	*Gerres filamentosus*	Actinopterygii	Perciformes	Gerreidae
19	Large-scaled goby	*Parapocryptes sp.*	Actinopterygii	Perciformes	Gobiidae
20	Johns snapper	*Lutjanus johnii*	Actinopterygii	Perciformes	Lutjanidae
21	Spotted scat	*Scatophagus argus*	Actinopterygii	Perciformes	Scatophagidae
22	Tiger-toothed croaker	*Otolithes ruber*	Actinopterygii	Perciformes	Sciaenidae
23	Indo-Pacific mackerel	*Rastrelliger brachysoma*	Actinopterygii	Perciformes	Scombridae
24	Greasy grouper	*Epinephelus sp.*	Actinopterygii	Perciformes	Serranidae
25	Streaked spinefoot	*Siganus javas*	Actinopterygii	Perciformes	Siganidae
26	Silver sillago	*Sillago sihama*	Actinopterygii	Perciformes	Silllaginidae
27	Crescent grunter	*Therapon jarbua*	Actinopterygii	Perciformes	Terapontidae
28	Tongue sole	*Cynoglossus sp.*	Actinopterygii	Pleuronectiformes	Cynoglossidae

There was substantial correlation remaining between each factor (Figure 2b).

After factor analysis, the data for each factor were explored against year elapsed during the study period, as shown in Figure 3. The fish standing crop weight for each factor was high in 2012.

Figure 4 shows the seasonal effect of standing crop for each factor. The seasonal factors were statistically significant for each factor. The fish standing crop weights in July to December were lower than average and the lowest was in December. On the other hand, the fish standing crop weights in January to June were higher than average, whereas, the highest was between February and March. As the data showed a seasonal effect on the fish standing crop for each factor, these data needed to be seasonally adjusted.

Figure 5 shows the fish standing crop data after adjusting for season by subtracting the monthly average for each site and then adding back the mean. The right panels show the average fish standing crop for each site by each factor with their confidence intervals. The results indicate that site has no effect for factor 2 and factor 3.

Table 2 Loadings greater than 0.1 and uniqueness of different fish species obtained from factor analysis,where shaded values show species in the same factor.

Fish species	Factor1	Factor2	Factor3	Factor4	Uniqueness
Indian squid	**0.981**		0.113	–0.172	0.164
Yellow stripe trevally	**0.972**			–0.193	0.264
Indo-Pacific mackerel	**0.910**		–0.264		0.429
Octopus	**0.795**				0.402
Mantis shrimp	**0.794**		0.112		0.297
Painted stone crab	**0.766**		0.164		0.333
Tongue sole	**0.731**			0.120	0.319
Oceanic paddler crab	**0.548**	0.137	0.242		0.371
Tiger-toothed croaker	**0.455**	0.266	0.169		0.396
Large-scaled goby	–0.245	**0.869**	0.106	–0.190	0.488
Spotted scat	–0.209	**0.743**	0.312		0.404
Crescent grunter	0.153	**0.631**			0.576
Silver biddy	–0.230	**0.611**	0.160	0.109	0.595
Streaked spinefoot		**0.610**		0.129	0.525
Johns snapper	0.187	**0.609**	–0.209		0.527
Mullet	0.170	**0.600**		–0.107	0.545
Silver sillago	0.143	**0.547**			0.537
Yellow pike-conger		**0.496**	–0.151		0.849
Banana prawn	0.211	**0.433**		0.199	0.445
Greasy grouper		0.149	**0.687**		0.291
Mud crab			**0.538**	0.215	0.390
Green tiger prawn	0.468		**0.522**	–0.174	0.410
Acetes	0.491	–0.139	**0.505**		0.287
Black tiger shrimp	0.239	0.280	**0.500**		0.310
Blue swimming crab	0.308		**0.370**	0.219	0.354
Shrimp post larvae	–0.206			**1.144**	0.005
Greasy back shrimp	0.344		0.166	**0.411**	0.385
Stork shrimp	0.296	0.250		**0.382**	0.373

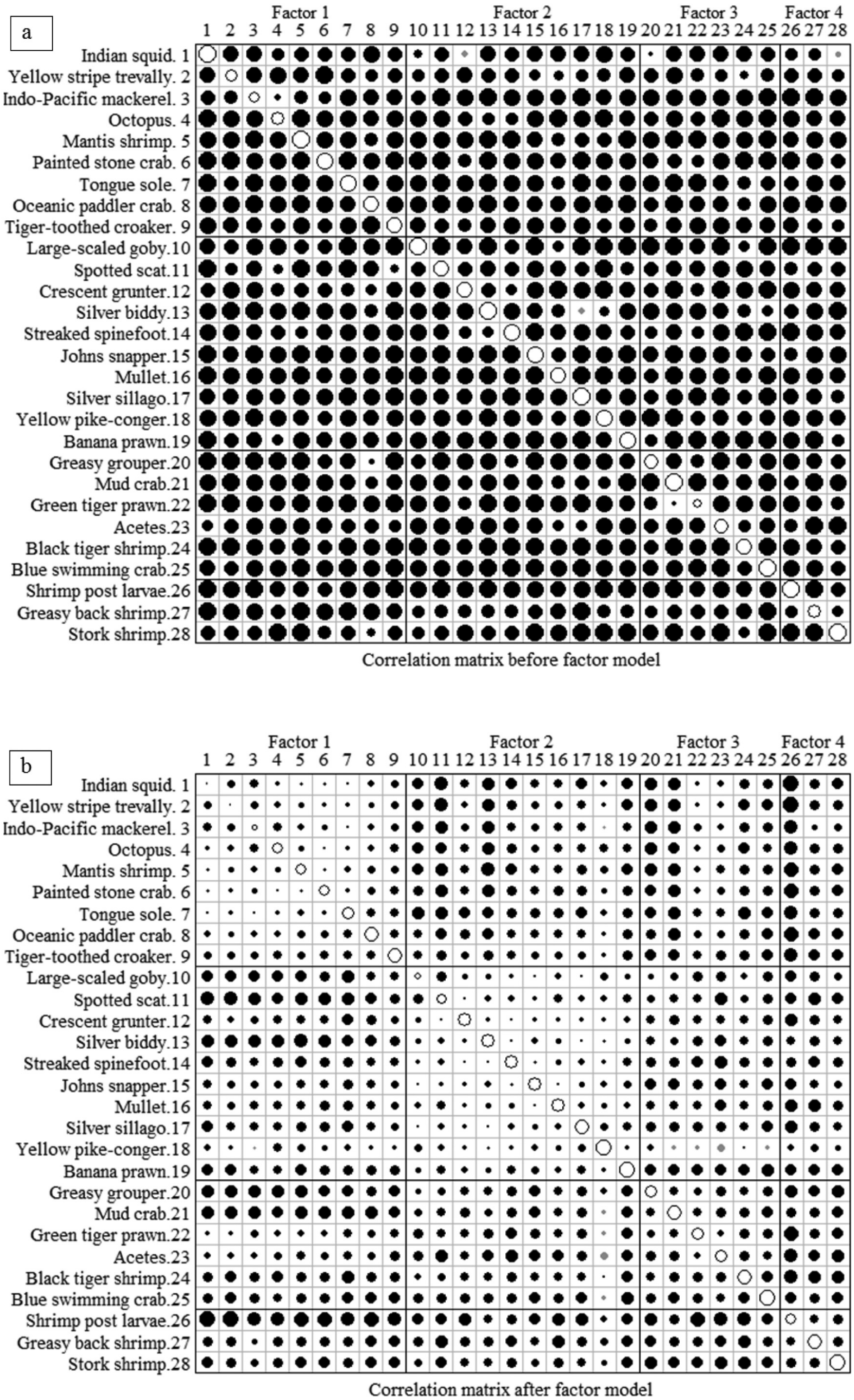

Figure 2 Bubble plots of correlations between fish standing crop in each species:(a) Before fitting the factor model;(b) After fitting the factor model. The size of each bubble indicates the magnitude of correlation for the matrix cell.

However, factor 1 and factor 4show a site effect with increasing proximity to the Gulf of Thailand. In the study, season explained 17–43% (based on the r-squared value in the right panel of Figure 4) of the standing crop weight variation, whereas only 0.5–12% (based on the r-squared value in the right panel of Figure 5) of standing crop variation can be explained by location.

DISCUSSION AND CONCLUSION

Four groups of species (factors) and their standing crop were associated with seasons. For all four factors, the average fish standing crop reached maximum levels in summer (February–April).The monsoon season and the salinity gradient of the intrusion of sea water are major natural phenomena affecting the water salinity in various locations in the Na Thap River estuary. In the rainy season, the river typically exhibits higher turbidity and violent water flow and is contaminated by nonpoint source pollution. In summer, the water is generally calm, clear and saline. Olukolajo and Oluwaseun (2008) also reported seasonal effects, with high percentages of estuarine fish species found in the dry season and freshwater fish found more in the rainy season in a tropical lagoon in Southwest Nigeria.

The current study revealed that the seasons have a greater effect on the fluctuations in the fish standing crop than do location influences. This result agreed with many studies which have reported that seasonal variation influences the number of individuals, species diversity and species distribution and moreover, fish distribution depends on several factors, including habitat,

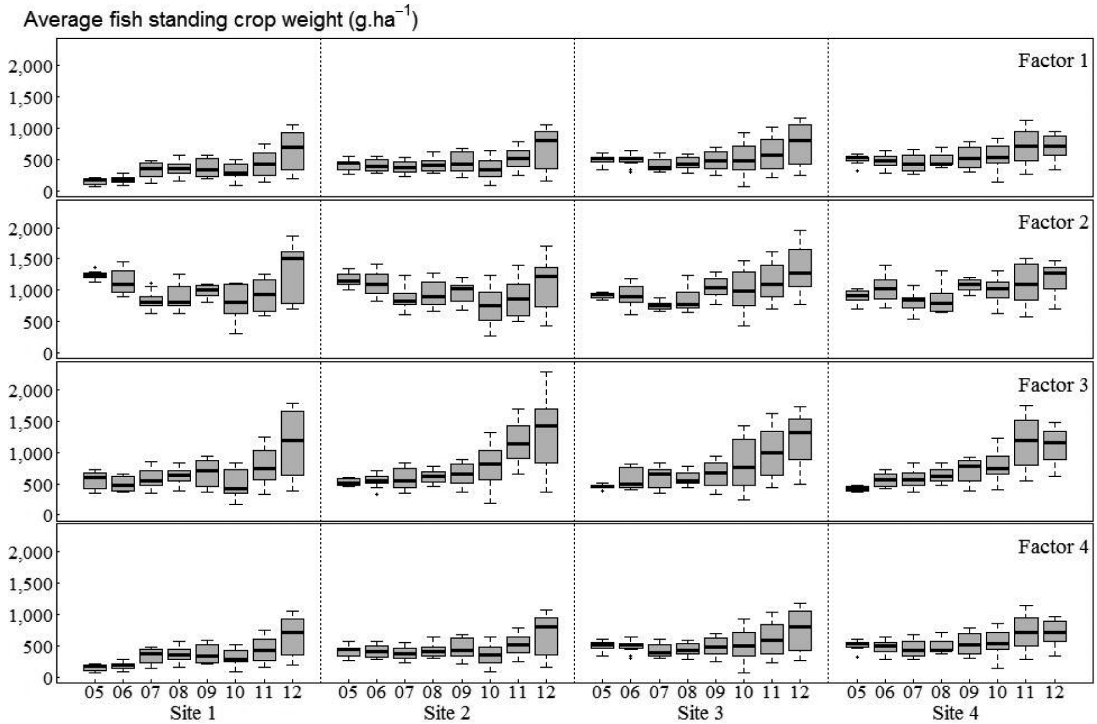

Figure 3 Box plots of average fish standing crop weights for four factors by site and year (2005–2012) for 28 species in the Na Thap river. (Each box represents the distribution of fish standing crop for 8 yr from the four different sites. The box size shows the interquartile range of data, the middle line in each box shows the median of the fish standing crop, the vertical bars show the range of data.)

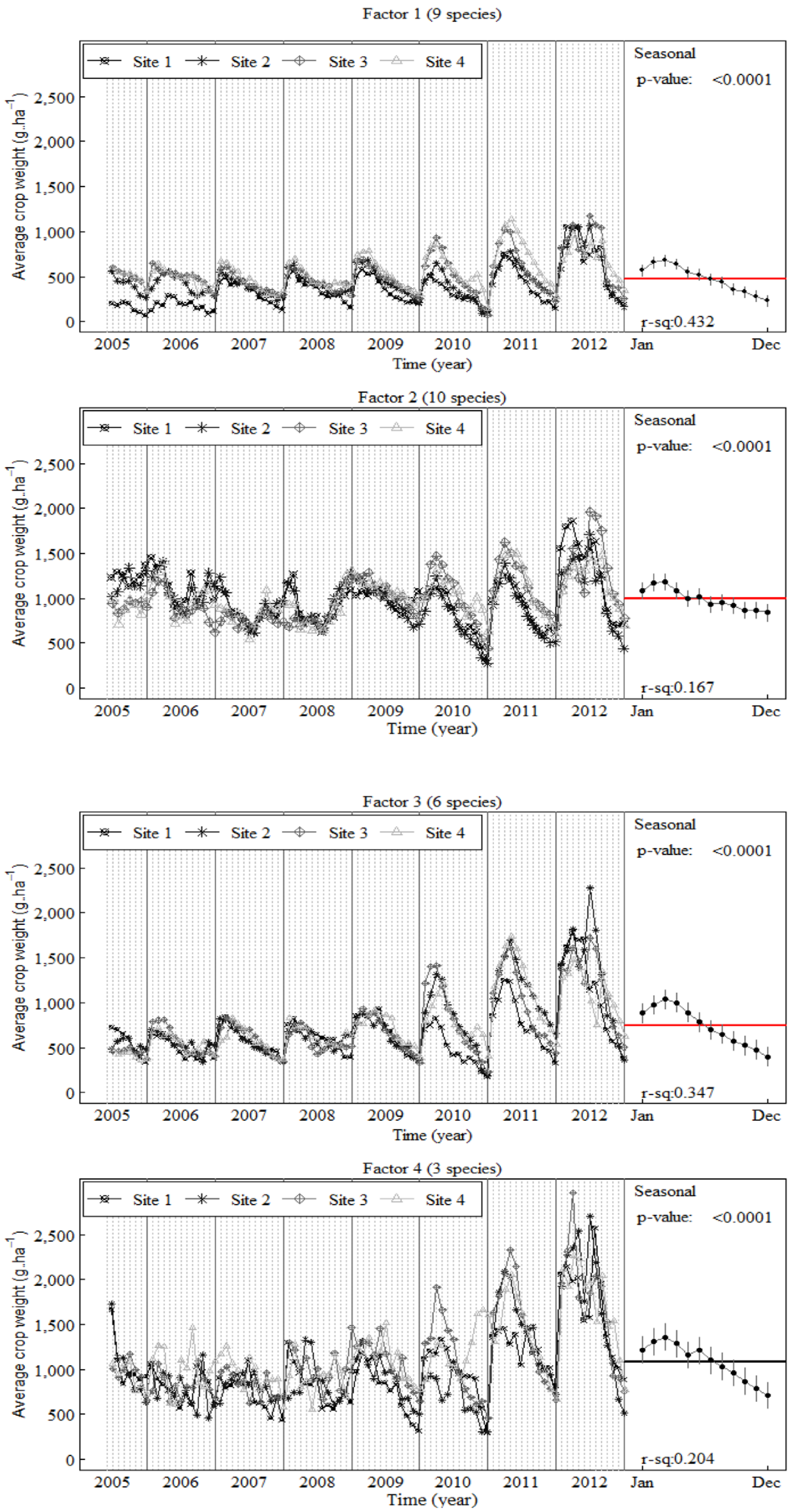

Figure 4 Average fish standing crop of each factor with 95% confidence interval bars of seasonal effect (r-sq. = Coefficient of determination; horizontal line in the right panels shows average fish standing crop weight of each factor during the study period).

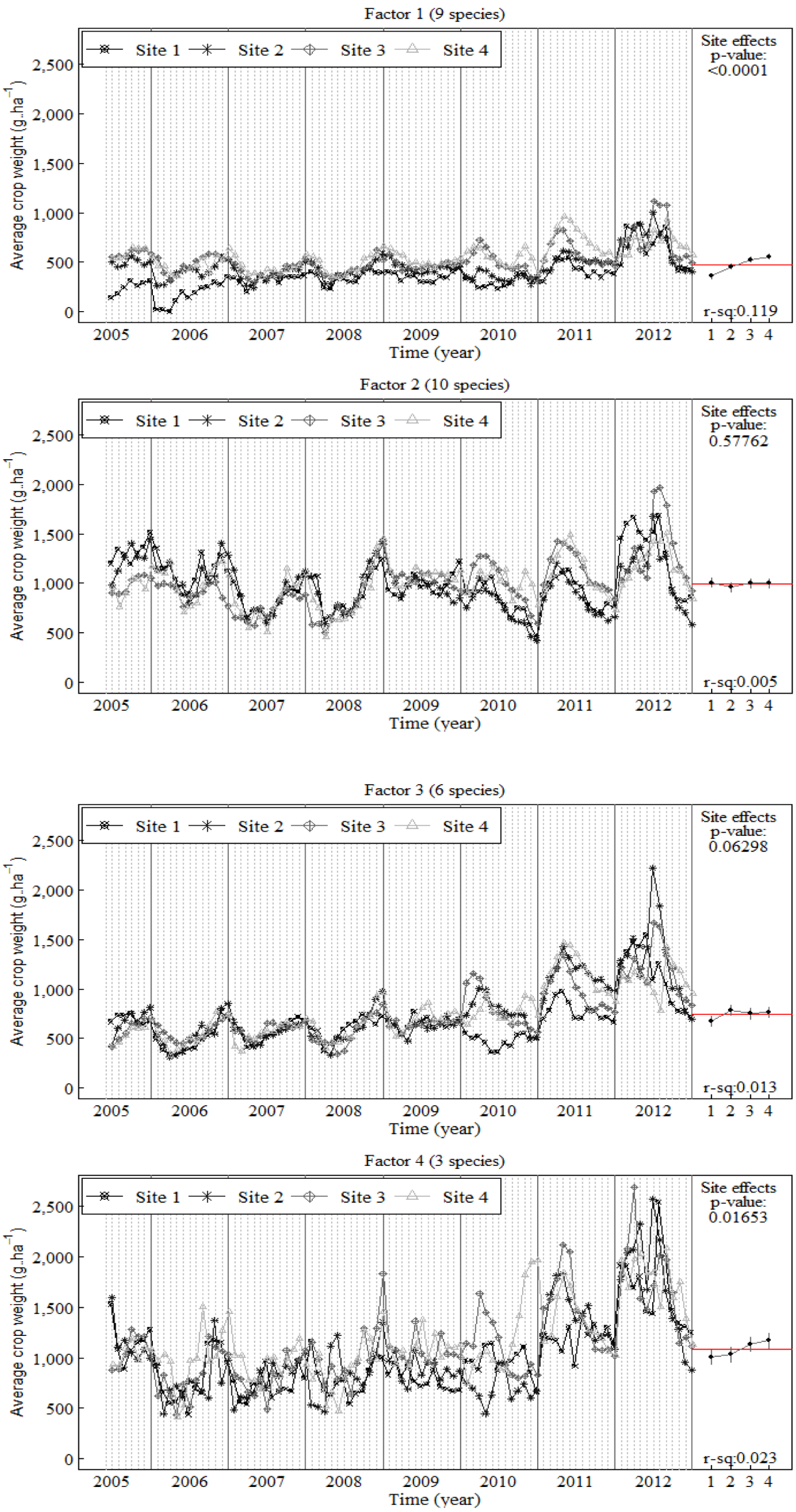

Figure 5 Average fish standing crop of each factor with 95% confidence intervals of site effect (r-sq. = Coefficient of determination; horizontal line in the right panels shows average fish standing crop weight of each factor during the study period).

nutrients, and water characteristics (Fischer and Eckmann, 1997; Pires *et al*., 1999; Barletta-Bergan *et al*, 2002; Hajisamae and Ibrahim, 2008; Ayoola and Kuton, 2009). The highest fish standing crop was recorded at the last sampling site at the mouth of the river for factor 1 and factor 4 because the estuarine fish species in those groups can only tolerate a narrow range of salinity concentration and they follow this salinity gradient throughout the Gulf of Thailand. This result is consistent with the reports of Barletta *et al*. (2005), Love *et al.* (2008) and Emmanuel and Chukwu (2010) who reported that saltwater intrusion impacted on fish diversity and fish distribution. In addition, the salinity increased from station 1 (upstream) to station 4 (downstream) in the range 14–38 psu. Most estuarine fish species spend much of their lives in the brackish water of the estuary ecosystem where salinity exceeds 12 psu (Blaber, 2000; Moyle and Cech, 2004).

Generally, fishing in the estuarine zone is based on catching targeted fish on their passage from feeding grounds to spawning and nursery grounds (Chesoh and Choonpradub, 2011). Knowledge of these patterns in the fish standing crop weight by month (time) and location (place) can encourage the regulatory authorities to better manage fisheries enhancement programs. Furthermore, the findings reinforce that these migratory fish species, particularly the marine organism group in factor 1,must be managed to sustain their diversity and abundance by banning fishing activity during migratory seasons. However, the study has some limitations, especially as fishing effort and the correlation of environmental parameters were not taken into account. It is recommended that the factor analysis method can be applied for grouping numerous fish species and linear modeling can be used to evaluate any variation. Based on estimates of the fish standing crop, good predictions of the distribution pattern are possible.

ACKNOWLEDGEMENTS

The authors express their deepest gratitude and appreciation to Prof. Don McNeil for making it possible to complete this report and also thank the Electricity Generation Authority of Thailand (EGAT) for providing valuable information for the study and the Prince of Songkla Scholarship for financial support.

LITERATURE CITED

Able, K.W. 2005. A re-examination of fish estuarine dependence: Evidence for connectivity between estuarine and ocean habitats. **Estuar. Coast. Shelf. Sci.** 64: 5–17.

Angsupanich, S. and S. Rakkheaw. 1997. Seasonal variation of phytoplankton community in Thale Sap Songkhla, a lagoonal lake in Southern Thailand. **Neth. J. Aquat. Ecol.** 30: 297–307.

Ayoola, S.O. and M.P. Kuton. 2009. Seasonal variation in fish abundance and physicochemical parameters of Lagos lagoon, Nigeria. **Afr. J. Environ. Sci. Tech.** 3: 149–158.

Baisre, J.A. 2000. **Chronicles of Cuban Marine Fisheries (1935–1995). Trend Analysis and Fisheries Potential.** Food and Agriculture Organization of the United Nations. Rome, Italy. 394 pp.

Barletta-Bergan, A., M. Barletta and U. Saint-Paul. 2002. Structure and seasonal dynamics of larval fish in the Caete´ River estuary in North Brazil. **Estuar. Coast. Shelf. Sci.** 54: 193–206.

Barletta, M., A. Barletta-Bergan, U. Saint-Paul and G. Hubold. 2005. The role of salinity in structuring the fish assemblages in a tropical estuary. **J. Fish. Biol.**66: 45–72.

Blaber, S.J.M. 2000. **Tropical Estuarine Fishes: Ecology, Exploitation and Conservation.** Blackwell. Oxford, UK. 384 pp.

Callaway, R., J. Alsvag, I. de Boois, J. Cotter, A. Ford, H. Hinz, S. Jennings, I. Kroncke, J.

Lancaster, G. Piet, P. Prince and S. Ehrich. 2002. Diversity and community structure of epibenthic invertebrates and fish in the North Sea. **ICES. J. Mar. Sci.** 59: 1199–1214.

Cartes, J.E., P. Abello, D. Lloris, A. Carbonell, P. Torres, F. Maynou and L.G. de Sola. 2002. Feeding guild of western Mediterranean demersal fish and crustaceans: an analysis based on a spring survey. **Sci. Mar.** 66: 209–220.

Chesoh, S. and C. Choonpradub. 2011. A model for clustering fish community structure with application to Songkhla Lake bi-monthly catches 2003–2006. **Turk. J. Fish. Aquat. Sci.** 11: 177–184.

Conti, L., G. Grenouillet, S. Lek and M. Scardi. 2012. Long-term changes and recurrent patterns in fisheries landings from large marine ecosystems (1950–2004). **Fish. Res.** 119–120: 1–12.

Deka, T.K., M.M. Goswami and M. Kakati. 2005. Causes of fish depletion – a factor analysis approach. **Naga.** 28: 37–42.

Emmanuel, B.E. and L.O. Chukwu. 2010. Spatial distribution of saline water and possible sources of intrusion into a tropical freshwater lagoon and the transitional effects on the lacustrine ichthyofaunal diversity. **Afr. J. Environ. Sci. Tech.** 4: 480–491.

Erzini, K., C.A.O. Inejih and K.A. Stobberup. 2005. An application of two techniques for the analysis of short, multivariate non-stationary time-series of Mauritanian trawl survey data. **ICES. J.Mar. Sci.** 62: 353–359.

Fausch, K.D., C.L. Hawkes and M.G. Parsons. 1988. **Models That Predict Standing Crop of Stream Fish from Habitat Variables: 1950–85**. Department of Agriculture, Forest Service, Pacific Northwest Research Station. La Grande, OR, USA. 52 pp.

Fischer, P. and R. Eckmann. 1997. Seasonal changes in fish abundance, biomass and species richness in the littoral zone of a large European lake, Lake Constance, Germany. **Arch. Hydrobiol.** 139: 433–448.

Hajisamae, S., P. Yeesin and S. Ibrahim. 2006. Feeding ecology of two sillaginid fishes and trophic interrelations with other co–existing species in the Southern part of South China sea. **Environ. Biol. Fish.** 76: 167–176.

Hajisamae, S. and S. Ibrahim. 2008. Seasonal and spatial variations of fish trophic guilds in a shallow, semi–enclosed tropical estuarine bay. **Environ. Biol. Fish.** 82: 251–264.

Jackson, D.A., S.C. Walker and M.S. Poos. 2010. Cluster analysis of fish community data: "New" tools for determining meaningful groupings of sites and species assemblages. **Am. Fish. Soc. Symp**.73: 503–527.

Johnson, R.A. and D.W. Wichern. 2007. **Applied Multivariate Statistical Analysis** (6th ed.). Prentice Hall. NJ, USA. 773 pp.

Love, J.W., J. Gill and J.J. Newhard. 2008. Saltwater intrusion impacts fish diversity and distribution in the Blackwater River drainage (Chesapeake Bay watershed). **Wetlands**28: 967–974.

Mardia, K.V., J.T. Kent and J.M. Bibby. 1980. **Multivariate Analysis.** Academic Press Inc. London, UK. 521 pp.

Moyle, P.B. and J.J. Cech. 2004. **Fishes, an Introduction to Ichthyology.** Prentice Hall. Upper Saddle River, NJ, USA. 726 pp.

Olukolajo, S.O. and K.A. Oluwaseun. 2008. Seasonal variation in the distribution and fish species diversity of a tropical lagoon in South–West Nigeria. **J. Fish. Aquat. Sci.** 3: 375–383.

Pires, A.M., I.G. Cowx and M.M. Coelho. 1999. Seasonal changes in fish community structure of intermittent streams in the middle reaches of the Guadiana basin, Portugal. **J. Fish. Biol.** 54: 235–249.

R Development Core Team. 2010. **R: A Language and Environment for Statistical Computing.** R Foundation for Statistical Computing. Vienna, Austria. 1706 pp.

Rainboth, W.J. 1996. **Fishes of the Cambodian**

Mekong. Food and Agriculture Organization of the United Nations. Rome, Italy. 321 pp.

Saheem, N., A. Lim and S. Chesoh. 2014. Predicting standing crop using lagged fingerling density of freshwater fish in the Na Thap River of Southern Thailand. **Songklanakarin J. Sci. Technol.** 36: 13–19.

Taki, Y. 1974. **Fishes of The Lao Mekong Basin.** U.S. Agency for International Development Mission to Laos. Vientiane, Laos. 232 pp.

Zuur, A.F., I.D. Tuck and N. Bailey. 2003. Dynamic factor analysis to estimate common trends in fisheries time series. **Can. J. Fish. Aquat. Sci.** 60: 542–552.