United Arab Emirates

Environment Agency Abu Dhabi (EAD)

Implementation of a National Fisheries Information System

(UAE-NFIS)

Training on advanced statistics

C. Stamatopoulos, PhD

Abu Dhabi, January 2018

"... I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "Law of Frequency of Error". The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement, amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshalled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along..."

Sir Francis Galton for the Central Limit Theorem
Natural Inheritance, 1889

Preface

A major task of most countries is the improvement of their data collection programmes relating to agricultural statistics, including fisheries and forestry. The advantage of sample-based fisheries statistical systems is that they can make large-scale data collection programmes more affordable when sharp limitations exist with regards to human resources and availability of operational funds. The Environment Agency of Abu Dhabi (EAD) has long recognized these needs and over the past period it has intended to make use of training materials and methodological/operational guidelines with the view of assisting its staff in their efforts to improve their performance in their field and office functions and responsibilities. This document was prepared as a principal training component of the UAE National Fisheries Information System (UAE-NFIS).

Data collection on catch, fishing effort, first-sale prices and average fish size is a key factor for basic fisheries statistical studies. This means that a statistical system that operates on a regular basis is not an end in itself but a valuable source of information and data that serves a wide variety of purposes. Consequently a regular fisheries statistical programme is judged with two criteria: (i) whether it operates in a cost-effective manner (and this concerns its developers and operators) and, (ii) whether its results are of good utility when diffused to their intended audience. This training course mainly concerns the second criterion and its objective is to provide system developers and operators with the theoretical basis for improving the methodological and operational aspects of their data analyses.

This document contains 10 chapters which will be dealt with in phases, following the schedule of the consultant's missions to the country. Some chapters have already appeared in earlier documents but it was thought that an integrated training document would facilitate the understanding of inter-related theoretical aspects. It is hoped that this manual will assist users to better explore the UAE-NFIS statistical database that has been implemented in 2018.

Constantine Stamatopoulos, PhD
Senior Advisor – Fisheries Statistics
Project for the Sustainable Management of Fisheries Resources in UAE

Abu Dhabi, January 2018

Contents

Chapter 1: Determining populations in sample-based fisheries surveys

Understanding the exact type and size of the population from which samples will be regularly collected is a prerequisite for a successful implementation of a fisheries statistical monitoring programme. There is one and only one way to achieving this and it simply consists of considering the population of answers that would result from a hypothetical census which would statistically cover all events in space and time. Such a census is always assumed to take place within a statistical context consisting of: (i) a month, (ii) a geographical stratum and, (iii) a specific boat-gear category.

1.1 Population of landings

To determine the population of landings $\mathbf{N_L}$ of a specific boat-gear category in a stratum and during a month, the following information is required:

(a) The calendar days $\mathbf{D}$;
(b) The average duration $\mathbf{d}$ of a fishing trip of the boat-gear in question;
(c) The number of boats-gears $\mathbf{B}$ that are operational during the month. A boat-gear is considered operational if it has made at least one trip during the month.

A boat-gear can on the average make $\mathbf{D}/\mathbf{d}$ landings during the $\mathbf{D}$ days, hence the number of landings will be:

$$\mathbf{N_L} = \mathbf{B} \times \mathbf{D}/\mathbf{d} \tag{1.1}$$

<u>Example 1</u>
Assuming the month of September 2015 and 250 operational speedboats with miscellaneous gear in the stratum (port) of Doha, then the following calculations apply.

- The number of calendar days is $\mathbf{D}$ = 30.
- The average trip duration is $\mathbf{d}$ = 1.
- The number of boats-gears is $\mathbf{B}$ = 250.

The population size of landings will thus be obtained from expression (1): $\mathbf{N_L}$ = 250 x 30/1= 7500 landings.

Example 2
Assuming the month of September 2015 and 100 operational launches with traps in the stratum (port) of Al Wakra. The average trip duration is 3 days. The following calculations apply.

- The number of calendar days is $\mathbf{D}$ = 30.
- The average trip duration is $\mathbf{d}$ = 3.
- The number of boats-gears is $\mathbf{B}$ = 100.

The population size of landings will thus be obtained from expression (1): $\mathbf{N_L}$ = 100 x 30/3= 1000 landings.

1.2 Population of boat-gear activities

In order to determine the Probability Boat Active (PBA) of a boat-gear category, the sampling scenario in use by Samaq Web is to interview fishermen on a weekly basis with regards to their fishing activities during last week. In some cases the question concerns the numbers of days at sea during past month and it is asked at the end of the reference month.

In the case of weekly samples the population of boat-gear activities $\mathbf{N_E}$ is determined as follows. Assuming that a census has been conducted on a weekly basis covering all operational boats. Consequently each week will produce $\mathbf{B}$ answers each referring to any of 0, 1, 2, 3, 4, 5, 6, or 7 days at sea. At the end of the month we will have collected 4 x B answers. Thus the size of the population of boat-gear activities $\mathbf{N_E}$ will be:

$$\mathbf{N_E} = 4 \times \mathbf{B} \qquad (1.2)$$

Likewise in the case of a monthly census each month will produce $\mathbf{B}$ answers each referring to 1, 2, … , 30 days at sea. Thus the size of the population of boat-gear activities $\mathbf{N_E}$ will be equal to **B**.

$$\mathbf{N_E} = \mathbf{B} \qquad (1.3)$$

Sections 1.1 and 1.2 are relevant to the spatial accuracy of landings and effort. All populations are convex, a fact that will simplify the calculation process. However, depending on the size of the populations there might be a need for a parallel computation of spatial accuracy using two completely independent methods that are referred to as Small Population Sampling Theory (SPST) and Large Population Sampling Theory (LPST) respectively.

1.3 Populations for temporal accuracy

Spatial accuracy measures the effectiveness of sampling operations in terms of amounts of samples collected over a reference period (i.e. a month). This however is not enough to ensure that all possible effort has been made to reduce bias in the estimates. For instance, if 32 landings have been collected on a single sampling day and none in the rest of the month, then the spatial accuracy would be over 90% but there would also be a high risk of bias if the fishing day in question has been particularly bad (negative bias) or particularly good (positive bias). This means that in order to ensure that the frequency of sampling is as adequate as the amount of samples collected, sampling operations ought to be distributed evenly over the reference period. The question then arises as to how many days should be needed.

Conceptually the problem is quite equivalent to that of the landings and boat-gear activities. Let us assume that a census has been conducted in space and time for landings and that on each day we have a complete picture of all landings that have occurred during this day. If the census had been conducted for November 2014 then we would have 30 different data sets covering all landings in the month. Thus if we intend to sample a number of days and estimate the monthly total the population size would be 30.

The temporal accuracy does not apply to effort samples since these are conducted at regular intervals and for the selected boats-gears they statistically cover the entire month.

Temporal accuracy always uses the Small Population Sampling Theory (SPST) methodology.

1.4 Closed seasons

So far we have examined the populations of landings and effort on the basis of a full calendar month with 28, 29, 30 or 31 days. In the case where a closed season begins or ends within a reference month, then the calendar days must be adjusted accordingly.

Chapter 2: Population variance and sampling error

Once a population has been identified it is possible to estimate its mean value and the variability of its elements. Generally the mean of a population is not known unless a census has been conducted. The only way to obtain an idea of its mean value is to collect samples and approximate the population mean using the sample mean.

Assuming a population of **N** elements and a random sample of size n consisting of elements ($x_1,..., x_n$), then the sample mean $\bar{x}$ is given by:

$$\bar{x} = \frac{1}{n}\sum x_i \qquad (2.1)$$

The sample mean calculated as in (2.1) is an *unbiased estimate* of the population mean. The term "unbiased estimate" implies that if it were possible to draw all possible samples of size n and each time calculating the respective sample mean, then the mean of the sample means would coincide with the population mean.

Likewise the sample variance:

$$s^2 = \frac{1}{n-1}\sum (x_i - \bar{x})^2 \qquad (2.2)$$

is also an unbiased estimate of the population variance.

Actually the main task is to obtain an idea of the variability of the sample mean and not of the population. For this purpose we use the sample variance to calculate the variance of the mean:

$$s_{\bar{x}}^2 = \frac{s^2}{n}(1 - \frac{n}{N}) \qquad (2.3)$$

It is recalled that n and N are the sizes of the sample and the population respectively. When the population is large the above expression is reduced to:

$$s_{\bar{x}}^2 = \frac{s^2}{n} \qquad (2.4)$$

Expressions (2.3) and (2.4) indicate that irrespective of the variability of the population, the variance of the sample mean tends to become zero as n increases. Generally, the sample mean is assumed to be found within a confidence interval whose lower and upper limits are calculated as follows:

$$s_{\bar{x}} = +\sqrt{s_{\bar{x}}^2} \quad \text{(sampling error)} \tag{2.5}$$

$$\text{Lower limit: } x_{-} = \bar{x} - 1.96 s_{\bar{x}} \qquad \text{Upper limit: } x_{+} = \bar{x} + 1.96 s_{\bar{x}} \tag{2.6}$$

The coefficient 1.96 is used to assure that the true population mean lies within the confidence interval determined by (2.6) with a probability of 95%. Using a lower probability it would make the interval narrower but riskier. Generally the confidence level of 95% is used in large-scale fisheries surveys.

The final step is to express the sampling error in relative rather than in actual terms. A typical indicator that expresses the relative error is the coefficient of variation (CV) which usually takes the following percentage form:

$$CV = 100 \frac{s_{\bar{x}}}{\bar{x}} \tag{2.7}$$

As a last note this section ought to pay tribute to the *Central Limit Theorem* which is literally "central" to all methodologies and operations relating to sample-based fisheries surveys and it is the basis for constructing confidence limits such as those described by expressions (2.3) – (2.7). Its implications are fundamental. Whatever the shape of the distribution of a population, the distribution of the sample means and of the sample variances are obeying the normal distribution, provided that many samples are drawn at random.

This theorem has an interesting history. Its first version was postulated by the mathematician de Moivre who in 1733 used the normal distribution to approximate the distribution of the number of heads resulting from many tosses of a fair coin. This finding was nearly forgotten until the French mathematician Pierre-Simon Laplace rescued it from obscurity in his monumental work *Théorie analytique des probabilités*, which was published in 1812. Laplace expanded De Moivre's finding by approximating the binomial distribution with the normal distribution. But as with De Moivre, Laplace's finding received little attention in his own time. It was not until the nineteenth century when, in 1901, the Russian mathematician Aleksandr Lyapunov defined it in general terms and proved precisely how it worked mathematically. Nowadays, the central limit theorem is considered to be the official sovereign of probability theory.

Sir Francis Galton, a principal founder of statistical surveys, described the Central Limit Theorem in the exhilarating terms shown in the preface.

Chapter 3: Estimation of Catch/Effort Variables

The objective of a sample-based fisheries statistical monitoring programme is to regularly collect catch and effort data on a monthly basis and produce estimates of total catch, catch by species, fishing effort by boat-gear, species prices and values and, average fish weight. To achieve this, a catch/effort assessment system collects data to populate the generic catch/effort equation:

ESTIMATED CATCH = (Sample overall CPUE) x Estimated Effort (3.1)

Equation 3.1 is repeated for each statistical entity that is a month, a port-stratum and a boat-gear category.

3.1 Estimation of overall CPUE

Assuming that several landings have been collected during a month in a port-stratum and for a specific boat-gear category. If in each sample the total landing of species is c and the duration of the related fishing trip is d, then the overall CPUE is estimated by:

$$\text{CPUE} = \frac{\sum q}{\sum d} \qquad (3.2)$$

For the species that are found in the landings over a month their proportions to the total sample landings will be:

$$p_i = \frac{\sum (\text{species}\_q)}{\sum q} \qquad (3.3)$$

3.2 Estimation of effort

The generic formula for estimating monthly fishing effort of a boat-gear at a port-stratum is the following:

Estimated Effort = (PBA) x (B) x (D) (3.4)

where:

- PBA is the Probability of a boat-gear being active on any day during the month;
- B is the total number of operational boats-gears enumerated at the end of a reference month;
- D is the number of calendar days in the month. Except for cases of closed seasons, this number is equal to the number of days in a month, i.e. 28, 29, 30 or 31.

3.2 Estimation of PBA

There are several ways of estimating the sample PBA. In UAE the two scenarios in use are:

(a) At the end of each week fishermen are asked to specify the number of days worked during past week. The possible answers are any of a_i=0, 1, … , 7 days. Assuming n samples of such answers, then the sample PBA s computed as:

$$\text{PBA} = \frac{\sum a_i}{7n} \qquad (3.5)$$

(b) At the end of each month (with calendar days D), fishermen are asked to specify the number of days worked during the month. The possible answers 1 , . . . , D days, where D is the number of days in the month. Note that 0 is not a possible answer because an operational boat must make at least one fishing trip during the month. Assuming n samples of such answers, then the sample PBA is computed as:

$$\text{PBA} = \frac{\sum a_i}{nD} \qquad (3.6)$$

Once the PBA has been estimated by either (3.5) or (3.6), the total effort is estimated from (3.4) and the total catch from (3.1).

3.3 Catch by species

Catch by species is estimated on the basis of sample species proportions given in (3.3). Thus estimated catch by species is computed as:

Species catch = (Species proportion) x (Estimated Total Catch) (3.7)
Species CPUE = (Species proportion) x (Overall Sample CPUE) (3.8)

3.4 Species prices

Species prices need not be collected at each landing. In theory one sample would be enough but the reliability of price estimates is proportional to the number of samples involving prices. Assuming that m samples contain collected prices, then the weighted average price of a species is calculated as:

$$p_i = \frac{\sum q_i p_i}{\sum q_i} \qquad (3.9)$$

where:

- q_i is a non-zero species quantity over a month;
- p_i is the collected price.

Gaps in the collection of prices cause inconsistent estimates of prices and values when monthly estimates a are aggregated in variable grouping schemes. For instance, average prices and values by boat-gear will not be 100% comparable to totals by species.

To avoid such discrepancies Samaq Web uses a sophisticated technique of price estimation in various stages of reliability. The result is that any grouping of estimates will result in comparable prices and values.

3.5 Average species weight

Along with catch, effort, and prices, Samaq Web collects information on the number of individuals appearing in the landings. There are three options available for the data collectors:

- For small fish they collect the number of individuals in one Kg.
- For fish of average size they collect the total number of individuals in the catch.
- For large fish they estimate the average weight of an individual.

Samaq Web analyzes the size data collected and for each statistical entity of a month, port-stratum and boat-gear category it furnishes the following two estimates:

1. Average weight of a species;
2. Numbers of individuals caught.

Chapter 4: Precision and Accuracy in Catch/Effort Estimates

4.1 Introduction

The *Precision* of an estimator (such as the sample CPUE or the sample PBA) expresses the degree of closeness among the results of a sampling process or, in dispersion terms, how the samples are placed with respect to their mean.

The *Accuracy* of an estimator expresses the closeness of the resulting estimates to the actual (and unknown) population means.

Although the terms Accuracy and Precision are more or less synonymous in colloquial terms, they are well distinct in a statistical context. For instance, a set of sampling results may be exact but not precise, or precise but not accurate, or neither precise nor accurate, or both accurate and precise (see Fig. 4.1). Furthermore, when a survey has an inherent systematic error (or bias), this error will not be resolved by increasing the sample size; the survey will continue to generate biased results. If the samples remain the same but the systematic error has been identified and resolved then this action will improve the accuracy but not the precision.

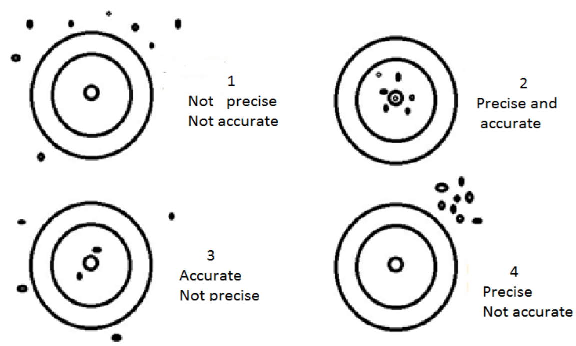

Figure 4.1 – Illustration of the difference between precision and accuracy: Case 1 -results that are neither precise nor accurate. Case 2 – results are both accurate and precise. Case 3 - exact results but somewhat dispersed (not precise). Case 4 - results are biased with good precision but completely inaccurate.

The index most commonly used to measure the precision of a sampling process is the *Coefficient of Variation (CV)*. This index is formulated on the basis of the standard error as discussed in Chapter 2. It is a very useful indicator and is always computed along with other statistics. Its advantages are: (i) simple to calculate and, (ii) simple to interpret. However, the CV is not applicable to populations of small size because the central limit theorem (refer to Chapter 2) is no longer in force. Most importantly, the CV is variable and cannot be used in advance in order to plan a sample-based survey.

For all these reasons sample-based programmes make use of an additional relative index (pessimistic accuracy) that will be discussed in detail in the coming sections.

4.2 Definition of accuracy A

Let us consider a population with N elements : $P = \{Y_1, Y_2, \ldots, Y_N\}$ and a population mean $\overline{Y}$.
Consider also a sample of size n : $\{X_1, X_2, \ldots, X_n\}$ and its mean $\overline{X}$ which is an unbiased estimate of $\overline{Y}$.

In relative terms *the accuracy* (i.e. proximity) of $\overline{X}$ with respect to $\overline{Y}$ is expressed by the relative accuracy index:

$$A = 1 - \frac{|\overline{Y} - \overline{X}|}{R} \qquad (4.2.1)$$

where R is the amplitude or range $Y_{max} - Y_{min}$ of the population values.

4.3 Normalization of populations

The term *normalization (or standardization)* is used to describe a specific type of numerical transformation that applies to a finite population and in a manner that each population element is uniquely mapped to a number between 0 and 1 inclusive.

Let us again consider a population of N elements, $P = \{Y_1, Y_2, \ldots, Y_N\}$, with a population mean $\overline{Y}$. We then apply the following linear transformation to each population element:

$$u_i = \frac{Y_i - Y_{min}}{Y_{max} - Y_{min}},\ i = 1, 2, \ldots, N \qquad (4.3.1)$$

For the expression above the following propositions are true:

1. The minimum element Y_{min} will be mapped to 0.
2. The maximum element Y_{max} will be mapped to 1.
3. All other elements will have values between and 1.
4. The range of the normalized population will be 1.
5. The mean $\overline{U}$ of the transformed population will be equal to: $\frac{\overline{Y} - Y_{min}}{Y_{max} - Y_{min}}$.

<u>Proof:</u>

$$\overline{U} = \frac{1}{N}\sum u_i = \frac{1}{NR}\sum (Y_i - Y_{min}) = \frac{N\overline{Y}}{NR} - \frac{NY_{min}}{NR} = \frac{\overline{Y} - Y_{min}}{Y_{max} - Y_{min}}$$

4.5 Equivalence of accuracy between original and normalized populations

Let us consider a population of N elements, $P = \{Y_1, Y_2, \dots, Y_N\}$, with a population mean $\overline{Y}$. Consider also a sample of size n: $\{X_1, X_2, \dots, X_n\}$ and a sample mean $\overline{X}$ that is an unbiased estimate of $\overline{Y}$.

Following that we examine the sample of n normalized elements: $\{u_1, u_2, \dots, u_n\}$ that corresponds to the original sample $\{X_1, X_2, \dots, X_n\}$. According to equation (4.2.1) the accuracy of $\overline{X}$ will be given as: $A = 1 - \frac{|\overline{Y} - \overline{X}|}{R}$.

Likewise the accuracy A_u of the normalized sample mean $\overline{u}$ of $\{u_1, u_2, \dots, u_n\}$ will be:

$$A_u = 1 - \frac{|\overline{U} - \overline{u}|}{1} = 1 - \frac{\left|\frac{\overline{Y} - Y_{min}}{R} - \frac{\overline{X} - Y_{min}}{R}\right|}{1} = 1 - \frac{|\overline{Y} - \overline{X}|}{R} = A$$

We have thus proved that the accuracy of a sample drawn from a finite population is equal to the accuracy of the corresponding sample taken from the normalized population.

The above property is very significant for the subsequent analysis of accuracy since it has eliminated the population range from the variables involved. As a result accuracy can now continually refer to normalized population with range $R = 1$. Expression (4.2.1) can thus take the simplified form:

$$A = 1 - |\overline{Y} - \overline{X}| \qquad (4.5.1)$$

where $\overline{X}$ and $\overline{Y}$ refer now to normalized elements between 0 and 1.

We are now going to examine again the equations concerning confidence limits (refer to Chapter 2). This can be summarized to take the form:

$$|\overline{Y} - \overline{X}| \le 1.96 s_{\overline{X}}$$

and it means that at 95% confidence level the error is equal to $1.96 s_{\overline{X}}$ and consequently the above expression takes the form:

$$A = 1 - 1.96 s_{\overline{X}} = 1 - 1.96 \frac{s}{\sqrt{n}} \sqrt{1 - \frac{n}{N}} \qquad (4.5.2)$$

where s is the standard deviation of a sample of normalized elements.

It can thus be concluded that the accuracy of an estimation process is a function of the population size N, the sample size n and the standard deviation of the sample.

Knowing the accuracy resulting from a sample is certainly useful but it should be far more useful if accuracy were known on an *a priori* basis which would permit determining the sample size required to achieve a pre-set accuracy level, for instance 95%. This is the most frequent question during the planning phase of a large-scale data collection programme. This problem will be discussed in the next session where the concept of "pessimistic accuracy" is introduced.

4.6 Probabilistic Pessimistic Accuracy

Let us again examine the normalized expression for accuracy:

$$A = 1 - 1.96 \frac{s}{\sqrt{n}} \sqrt{1 - \frac{n}{N}} \qquad (4.6.1)$$

Our new objective is to set-up an upper limit for the standard deviation s in samples drawn from populations that are either *convex* or *orthogonal* (see Figure 4.6.1). These two types of populations are the ones encountered in the fisheries survey programme of UAE. The third type concerns concave populations that are present only in sampling scenarios where the sample Probability Boat Active (PBA) is measured as the ratio (boat-gears active) / (boat-gears examined).

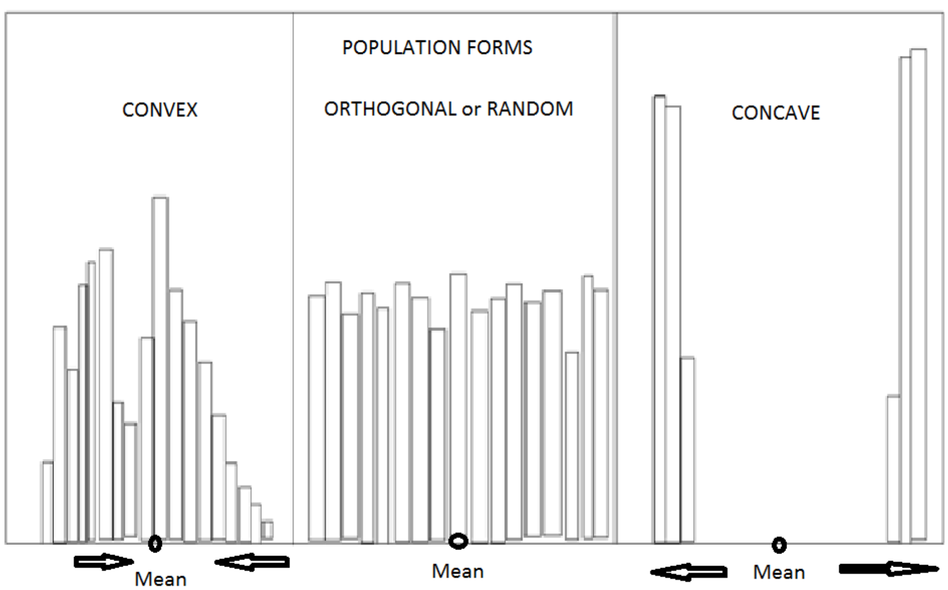

Figure 4.6.1 Convex, orthogonal (or random) and concave populations. In convex populations the values tend to concentrate around the mean. In orthogonal populations all values have about the same frequency. In concave populations values tend to move away from the mean.

According to statistical theory the standard deviation of a normalized population that is orthogonal (or random or uniform) is always higher than the standard deviation of a normalized convex population. This upper limit is given by:

$$\sigma_R = \sqrt{\frac{2N-1}{6(N-1)} - \frac{1}{4}} \qquad (4.6.2)$$

By substituting the standard deviation s in (4.6.1) by its maximum value given by (4.6.2) we obtain:

$$A \geq 1 - 1.96 \frac{\sigma_R}{\sqrt{n}} \sqrt{1 - \frac{n}{N}} \qquad (4.6.3)$$

The right term of the last relation determines a lower limit (i.e. a pessimistic accuracy) for all accuracies resulting from samples of size n that were drawn from a convex or random population. It specifies that all accuracies will be greater than this lower limit relation with a probability of 95% (equivalent to stating that only 5% of the accuracies might fall below that limit).

Concerning planning of surveys relation (4.6.3) provides exactly what a survey planner is looking for. Only the population size N is needed. For any pessimistic accuracy level (a practical starting level is 90%), we can determine the sample size that satisfies that level. To be noted that the actual accuracy will never be known *a priori*. All we will know is that the resulting accuracy will have a probability of 95% to be above the limit determined by (4.6.3).

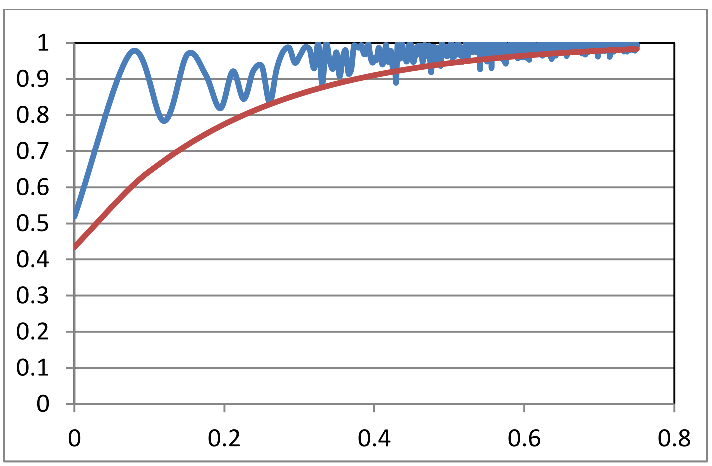

Figure 4.6.2 Fluctuation of sampling accuracy (blue line) when sample size varies between 1 and population size. The red line shows the pessimistic accuracy calculated by expression 4.6.3 for a convex or random population. Overall, all accuracy values lie above the pessimistic accuracy curve with few exceptions constituting about 5% of all values.

It is recalled that:

- The formulae discussed earlier are valid for convex or random populations;

- The exact accuracy is never known. What is known is its lower limit (pessimistic accuracy).
- Since the accuracy lower limit described in 4.6.3 can be computed at varying confidence levels (i.e. 90%, 95%, etc.), the pessimistic accuracy described above is a Probabilistic Accuracy.

Chapter 5: Non-Probabilistic Accuracy for small populations

5.1 Introduction

In the previous chapter the pessimistic probabilistic accuracy was defined on the basis of expression 4.6.3 which is repeated below:

$$A \geq 1 - 1.96\frac{\sigma_R}{\sqrt{n}}\sqrt{1-\frac{n}{N}}$$

It was stated that the right term of the equation constitutes a lower accuracy limit for all random or convex populations of size N. Thus, this lower limit (or pessimistic accuracy) can be illustrated as a known limit curve when sample size runs from 1 to N (Figure 4.6.2).

However as already stated this lower limit is based on the assumption that the central limit theorem (which ensures that the distribution of the sample means is normal) is in force. This is generally not true when the populations under study are small, i.e. they consist of 30, 50 or 100 elements. In such cases the pessimistic accuracy is excessively penalizing for the data collection since the required sample size for a desired accuracy level, say 90%, is too large.

This is illustrated in Figure 5.1. Here we have a population with size N=50. The dotted line represents the pessimistic accuracy curve drawn for all sample sizes 1-50. It is noted that the actual accuracy fluctuation stays well above the dotted line in most part of the sample size range. In practical terms this means that to achieve a given level of accuracy the probabilistic accuracy method will require more samples that are really required. In fact we should be looking for an improved accuracy curve that would look like the continuous curve of the same Figure 5.1.

To achieve such a curve the accuracy approach changes drastically and becomes more elaborate since it has to be based on algebraic and geometrical rather than probabilistic concepts.

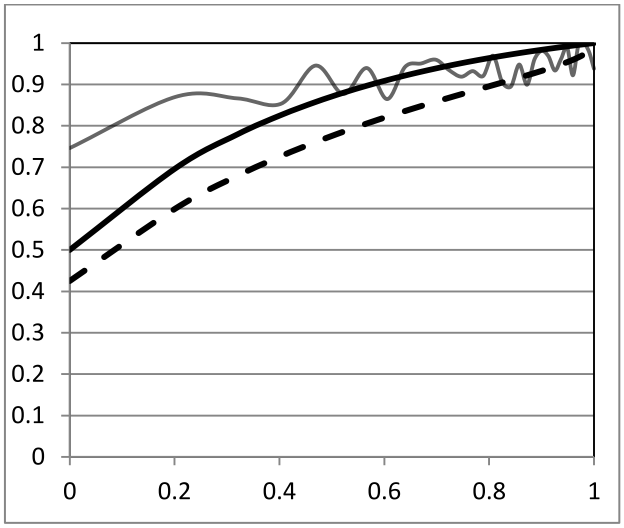

Figure 5.1 Pessimistic accuracy curve (dotted line) applied to a small population of size N=50. Notice that the fluctuation of accuracy stays well above the dotted line.

5.2 Variables and parameters of the non-probabilistic approach.

When the population size is limited (i.e. 15, 30, 50, 100, 500, 1000), it is essential to calculate accuracy using a non-probabilistic approach. The equation of accuracy for small populations is given by:

$$A = a_1 + a_2 N^{-kx} \qquad (5.2.1)$$

where the variable x is the ratio $\frac{\ln(n)}{\ln(N)}$ and the three parameters a_1, a_2, k are calculated by means of the following intermediate parameters:

- For convex populations (as in UAE) an intermediate parameter W is calculated as: $W = 0{,}75(1 - \frac{1}{N})$
- For concave populations (not applicable in UAE) the intermediate parameter W is calculated as: $W = 1 - \ln(1 + 0.5e^{\frac{1}{N}})$
- An intermediate parameter a is calculated as : $a = \frac{2WN^2}{(N-1)^2} - \frac{N+1}{N-1}$
- An intermediate parameter g is calculated as: $g = a + \frac{1-a}{N}$
- An intermediate parameter S is calculated as: $S = (1-a)(\frac{1}{\ln N} - \frac{1}{N \ln N} - \frac{1}{N})$
- The equation parameter k is then given by: $k = \frac{-2}{\ln N} \ln(\frac{S}{1-S-g})$

- The equation parameter a_2 is calculated as: $a_2 = \frac{(1 - S - g)^2}{2S + g - 1}$
- The equation parameter a_1 is calculated as: $a_1 = g - a_2$

The curve defined by (5.1) is not a limit-curve but an approximate curve. This means that it provides a more "economical" sample size than the one resulting from the probabilistic approach.

This approach is particularly useful in cases where sampling is done from a population of 28, 29, 30 or 31 days. In such a case the non-probabilistic approach provides more economical results than the probabilistic one. An example is provided in Table 5.2.1. Assuming a desired accuracy level of 90% (or 0.90) we notice that for the same sample size, for instance n=7, the non-probabilistic approach results in A=0.909, while the probabilistic one requires 17 samples to achieve the same level.

Table 5.2.1 The two accuracy approaches in a population of 30 days

Sample size	Non-probabilistic Accuracy A	Probabilistic Accuracy A
1	0,500	0,416
2	0,773	0,600
3	0,825	0,674
4	0,854	0,726
5	0,879	0,757
6	0,895	0,786
7	0,909	0,804
8	0,919	0,821
9	0,931	0,833
10	0,936	0,849
11	0,945	0,857
12	0,950	0,868
13	0,955	0,876
14	0,959	0,883
15	0,964	0,890
16	0,968	0,898
17	0,972	0,905
18	0,974	0,912
19	0,978	0,916
20	0,980	0,924
21	0,984	0,928
22	0,986	0,936
23	0,987	0,940
24	0,989	0,944
25	0,991	0,949
26	0,992	0,954
27	0,994	0,959
28	0,997	0,972
29	0,999	0,980
30	1.000	1.000

Chapter 6: Synthesis of accuracy indicators

6.1 Spatial and temporal accuracy

It has so far been shown that the accuracy of catch/effort estimates depends on the number of samples collected during a month for a specific boat-gear in a port-stratum. This quantitative accuracy, denoted by A_Q, is determined, as we have seen, on the basis of the population size and the sample size.

In parallel we also examine the temporal accuracy A_T which is determined by the number of sampling days during the month. This stems from the fact that the frequency of sampling (i.e. the number of days in a month on each sampling takes place) is as important as the number of samples collected. Let us for instance assume that 32 samples are enough to give an accuracy of at least 90%. If these samples are collected on a single day, the estimates run the risk of being biased if the day selected is either too “good” or too “bad”. It thus becomes evident that the 32 samples must be distributed over several days in order to avoid such a risk. The question then is: how many days are required to obtain a temporal accuracy A_T=90%.

The answer lies in considering the month (for instance September 2015) as a population of N=30 calendar days. If a hypothetical census has been implemented for the CPUE, we would then have 30 different actual (i.e. real) CPUE values. How many should be sampled to formulate a CPUE estimate with an accuracy A_T=90%?

Evidently the non-probabilistic approach for small populations should be used (refer to Chapter 5, Table 5.2.1). Using this approach the answer would be 7 days. In practice we use 8 days since in this manner we can plan the survey to sample twice a week, assuming four weeks in a month.

6.2 Overall Accuracy

The spatial and temporal accuracy indicators A_Q and A_T are always computed for the CPUE.

Concerning the PBA for effort estimation, the temporal accuracy is not needed because the sampling scenario in use by UAE-NFIS collects boat-gear activities using a sampling-in-space and census-in-time approach. In such a manner the temporal accuracy for the PBA is 100%.

The overall accuracy A for the monthly estimates is determined as the minimum value of:

- The spatial accuracy of the CPUE;

- The temporal accuracy of the CPUE;
- The spatial accuracy of the PBA.

The logic behind the above method is simple: if a census had been conducted for any two of the above three elements and a sample-based survey for the third, then the resulting accuracy would depend on the sample-based estimate. Hence (here again we use a pessimistic approach), the overall accuracy should be set to the lowest accuracy of the three.

A practical conclusion is that the spatial and temporal accuracy ought to have comparable values else there would be cases of wasted data collection effort. For instance if 128 landing samples have been taken aiming at a spatial accuracy of 95%, then the frequency of sampling should be increased to 12 sampling days and not stay at 8 (refer to Table 5.2.1). By limiting the sampling frequency to 8 days there would be no improvement to the overall accuracy since the system would still select 90% (being the lower of the two)

6.3 Sampling Uniformity Index (SUI)

In Section 6.2 describing the overall accuracy it was shown that the frequency of sampling is as important as the sample size and that the number of sampling days should be selected in a manner that the spatial and temporal accuracy are comparable.

The temporal aspect of sampling issue does not end here. Sufficient samples collected over a sufficient number of days are not enough if they are not uniformly distributed. For instance, we intuitively know that 32 samples taken uniformly over 8 days should be preferable to a scheme where the same 32 samples are distributed too irregularly (refer to Figure 6.3.1). How can we measure the uniformity aspect?

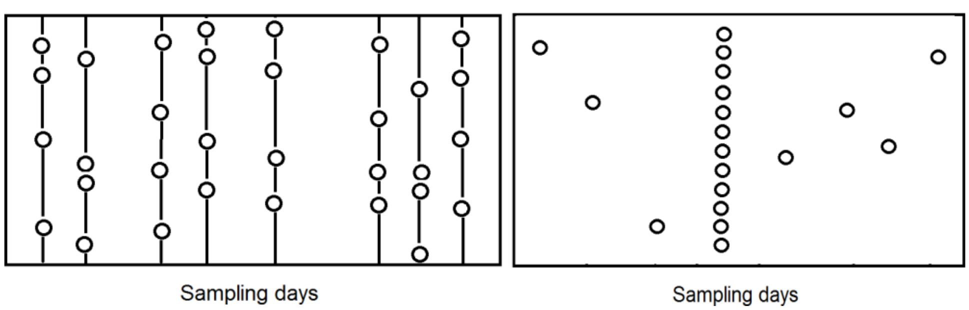

Figure 6.3.1 Illustration of a uniform and a non-uniform sampling scheme. Both use the same numbers of samples and sampling days. Both have the same spatial and temporal accuracy. However the first scheme is better because it distributes samples evenly over the sampling days.

The process is better presented using a numerical example.

Consider a large population and a sample of 32 elements collected over 8 days. The following data collection scheme was used:

Day	1	2	3	4	5	6	7	8
Samples	4	1	1	6	7	1	1	11

According to the theory already presented the temporal accuracy will be 90% and the spatial one will also be 90%.

We then analyze the uniformity of the samples over time as follows:

a) We calculate the average number of samples per day. In this case the average is equal to (32 samples) / (8 days) = 4 samples per day.
b) On each day we calculate the ratio: (n. samples)/average. If the result is >1 it is replaced by 1.

   Ratio 1: 4/4 = 1
   Ratio 2: 1/4 = 0.25
   Ratio 3: 1/4 = 0.25
   Ratio 4: 6/4 > 1 = 1
   Ratio 5: 7/4 > 1 = 1
   Ratio 6: 1/4 = 0.25
   Ratio 7: 1/4 = 0.25
   Ratio 8: 11/4 > 1 = 1

c) We next sum up the ratios and round up: (for instance 5.2 = 5, 5.7 = 6).

   Sum = 5

d) The resulting sum specifies the *number of virtual days* which theoretically penalizes the temporal accuracy A_T.
e) The ratio 5/8 = 0.625 is defined as Sampling Uniformity Index (SUI). The closer SUI is to 1 the more uniform the sampling scheme.

In theoretical terms SUI is calculated as follows:

(i) n is the total number of samples taken over the month.
(ii) d is the number of sampling days.
(iii) n_i (i=1,2,…,d) are the samples per day.
(iv) $\bar{n}$ is the arithmetic mean of n_i.
(v) v_i is the ratio $n_i / \bar{n}$. If v_i >1 it is set to 1.
(vi) v is the sum of all v_i rounded up.

SUI is then computed as the fraction: v / d.

Chapter 7: Estimation process and statistical diagnostics in UAE-NFIS

In this chapter a full UAE-NFIS example is presented which synthesizes all estimation approaches and statistical indicators so far discussed, namely:

Estimated variables
CPUE, PBA, prices, catch, effort and average fish size.

Statistical indicators
Coefficient of variation, spatial and temporal accuracy, SUI.

7.1 Estimation of fishing effort

Table 7.1 illustrates a full example of effort estimation for August 2015, the port-stratum of Al Khor and for launches with traps. Similar estimations have been computed for all statistical entities involved.

The estimated effort is 1179 days. It is easy to verify that this estimate is obtained by multiplying the number of boats-gears (67) by the sample PBA (0.568) and by the active days (31).

The spatial accuracy for PBA is 0.95 and the method of its calculation is SPST, the non-probabilistic approach used for small populations. It is recalled that UAE-NFIS also calculated accuracy with the probabilistic approach for medium-large populations. The SPST was selected since it apparently furnished a better accuracy value. However, this value (0.95) is not the best in the estimation process with the result that UAE-NFIS displayed an overall accuracy of 0.93 (top blue line).

Since in UAE-NFIS the sampling scenario for the PBA is sampling-in-space and census-in-time, the temporal accuracy and the associated SUI are equal to 1.

The coefficient of variation (CV) for the PBA (the CV indicates the dispersion of other hypothetical PBA sample means around the found mean of 0.568, is very good: only 4.2%. To be noted that a CV below 5% is empirically considered good precision, between 5% and 15% a good to average precision and above 15% a not good precision.

Table 7.1 Estimation of fishing effort

AL KHOR الخور (Minor stratum) + LB Traps لنش قرقور (Large boats) Accuracy = 0.93

Effort (WEEKLY)

Est.Effort :	1,179

Records :	**37**
Boats/Gears :	**67**
Active Days :	**31.0**
PBA :	**0.568**
Spatial Accur. :	0.95
Method for Accur. :	SPST
N.days :	-
Temp. Accur. :	1.00
Method for Accur. :	-
SUI :	1.00
CV (%) :	4.2 %

7.2 Estimation of overall CPUE, total catch, price and value

Table 7.2 illustrates a full example of catch estimation for August 2015, the port-stratum of Al Khor and for launches with traps. Similar estimations have been computed for all statistical entities involved.

The estimated overall CPUE is 191.87 Kg/boat-gear day. Total catch is obtained by multiplying this value by the estimated effort computed earlier (1,179 boat-gear days). This is equal to 226,214 Kg. The difference with the 226,182 Kg shown in the example is due to rounding errors.
The spatial accuracy for the CPUE PBA is 0.93 and was calculated using the SPST non-probabilistic approach. Likewise the temporal accuracy used the SPST method for small populations (here the population size is N=31) and resulted in 0.97. UAE-NFIS has selected the minimum of 0.95 (PBA), 0.93 and 0.97 to represent the overall accuracy of the estimation process.

The CV for the CPUE is rather high: 17%, but not unacceptable.

The SUI is 0.82 which indicates good uniformity of samples collected over the sampling days.
Lastly the average price (or unit value) of the total catch was estimated to be 17.04 QR/Kg. The total value of the catch is thus 17.04 x 226,182 = 3,854,141 QR. The small difference with the figure shown in Table 7.2 is due to rounding errors.

Table 7.2 Estimation of overall CPUE, total catch, price and value

Landings

Est.catch :	**226,182**
Records :	**37**
Sample catch :	**27,821.0**
Sample effort :	**145.0**
CPUE :	**191.87**
Aver.price :	**17.04**
Est.value :	**3,853,044**
Spatial Accur. :	0.93
Method for Accur. :	SPST
N.days :	17

Temp. Accur. :	0.97
Method for Accur. :	SPST
SUI :	0.82
CV (%) :	17.0 %

7.3 Estimation of catch by species, prices and average fish weight

Table 7.3 illustrates a full example of catch-by-species estimation for August 2015, the port-stratum of Al Khor and for launches with traps. Similar estimations have been computed for all statistical entities involved.

Due to lack of sufficient samples to estimate average fish size, the corresponding column is empty. However UAE-NFIS intends to streamline collection of fish weight in the coming months.

To be noted that summing up the species catch we find 226,182 Kg. Likewise the sum of species values is the same as the one shown in Table 7.2

Table 7.3 Estimation of catch by species, prices and average fish weight

By species	Aver.weight	N.fish in catch	Price	Value	CPUE	Est.catch
Baasi باسي Nemipterus bipunctatus (Threadfin bream)	-	-	19.684	54,731	2.36	2,780
Boukshina بوقشينة Lethrinus lentjan (Pink ear emperor)	-	-	10.964	401,114	31.03	36,585
Farsh فرش Diagramma pictum (Painted sweetlips)	-	-	7.233	28,580	3.35	3,951
Fasker فسكر Acanthopagrus bifasciatus (Twobar seabream)	-	-	9.908	17,400	1.49	1,756
Gane قين Scarus ghobban (Blue-barred parrotfish)	-	-	14.000	6,146	0.37	439
Hamour هامور Epinephelus coioides (Orange-spotted grouper)	-	-	58.433	556,769	8.08	9,528
Hamra حمره Lutjanus malabaricus (Malabar blood snapper)	-	-	27.155	123,188	3.85	4,536
Jash جش Carangoides bajad (Orangespotted trevally)	-	-	15.966	170,558	9.06	10,683
Jid جد Sphyraena flavicauda (Yellowtail barracuda)	-	-	13.680	10,010	0.62	732
Karari كراري Atule mate (Yellowtail scad)	-	-	17.148	57,716	2.86	3,366
Kobkob قبقب Portunus pelagicus (Blue swimmimg crab)	-	-	11.000	3,219	0.25	293
Koffar كوفر Argyrops spinifer (King soldier bream)	-	-	14.837	277,911	15.89	18,731
N'aimia نعيمية Pinjalo pinjalo (Pinjalo)	-	-	27.625	64,682	1.99	2,341
Naiser نيسر Lutjanus fulviflamma (Dory snapper)	-	-	7.439	35,926	4.10	4,829
Rebeeb ربيب Gnathanodon speciosus (Gold toothless treval)	-	-	24.075	112,739	3.97	4,683
Saafi صافي Siganus canaliculatus (Whit-spott spinefoot)	-	-	27.834	407,319	12.41	14,634
Saal صال Carangoides chrysophrys (Longnose trevally)	-	-	17.070	249,800	12.41	14,634
Laden لدن Epinephelus polylepis (Smallscaled grouper)	-	-	29.684	41,267	1.18	1,390
Semaan سمان Epinephelus bleekeri (Duskytail grouper)	-	-	33.349	209,849	5.34	6,293
Sh'ari شعري Lethrinus nebulosus (Sprankled emperor)	-	-	13.133	393,972	25.45	29,999
Siken سكن Rachycentron canadum (Cobia)	-	-	15.000	2,195	0.12	146
Sooli سولي Lethrinus microdon (Smalltooth emperor)	-	-	8.882	88,388	8.44	9,951
Sultan Ibrahim سلطان إبراهيم Parupeneus marga (Pearly goatfish)	-	-	28.100	4,112	0.12	146
Tabaan تبان Euthynnus affinis (Kawakawa)	-	-	9.219	132,217	12.17	14,341
Yanam ينم Plectorhinchus sordidus (Sordid rubberlip)	-	-	3.667	3,219	0.74	878
OTHER أنواع أخرى (Miscellaneous)	-	-	14.018	400,016	24.21	28,536

Chapter 8: Effort standardization in UAE-NFIS

8.1 Introduction

It is generally accepted that when working with a specific boat-gear category (for instance launches with traps) fishing mortality is proportional to the total fishing effort exerted by its fishing units. When it comes to measure the combined effect of fishing operations of the entire fleet to the exploitation of a fish stock, it becomes apparent that adding together effort exerted by different boat-gear categories is not always meaningful without first applying effort adjustment to increase its compatibility. There are various techniques for addressing such situations, the commonest of which is known as "standardization of fishing effort". In UAE the UAE-NFIS component of the National Fisheries Information System has recently incorporated effort standardization routines that combine elements of the simple (if not very recent) *normalized effort* (used by the North Sea Round Fish Working Group, ICES, 1980) and *relative fishing power* developed by Robson (1966). Although the existing literature offers a plethora of sophisticated methods (an excellent review is the one by Maunder, 2004), it was nevertheless considered preferable to first use approaches that (a) are only dependent on catch/effort data from commercial fisheries and, (b) can apply to data of (still) limited time coverage.

The need for effort standardization was first pointed out by the Steering Committee of the Sustainable Management of Fisheries Resources project and was followed up by the Fisheries Department of the Ministry of Environment. Thanks to the collective effort made by field staff and the national experts of the Fisheries Department the presented methodology was repeatedly tested using data of good quality, completeness and accuracy. It should be noted that the presented method is only the first step in introducing effort standardization as a regular operational component of UAE-NFIS; the approach in use will be further refined when catch/effort data involving more years have been made available.

8.2 Primary variables in effort standardization

In UAE-NFIS the fishing effort exerted by a fishing unit during a fishing trip is measured by the duration of the trip and is referred to as "boat-gear days". If there are m boat-gear categories and the statistical monitoring system produces 12 monthly catch/effort estimates per boat-gear category (as is the case with UAE-NFIS) then over a reference period of n years there will be $(m \times 12n)$ monthly effort estimates $E_{i,j}$, $i=1\ldots m$; $j=1\ldots 12n$.

Along with fishing effort the system estimates monthly catch $C_{i,j}$ and Catch-Per-Unit-Effort $CPUE_{i,j}$.

It should be noted here that UAE-NFIS treats combined CPUE's as weighted averages and not as simple arithmetic means of their components. For instance to combine monthly CPUE values of the same boat-gear category into a an annual CPUE, the standard UAE-NFIS procedure is to re-calculate the monthly catch and effort values involved according to the standard formula

$$CPUE = \sum(\text{Catch}) / \sum(\text{Effort}).$$

Table 8.2.1 illustrates an example of a full set of UAE-NFIS catch/effort estimates for 2014 which involves the three primary variables described above.

Table 8.2.1 UAE-NFIS catch/effort data for 2014 (all species) – Accuracy of estimates: 91.7%

Catch in Kg	**(01)**	**(02)**	**(03)**	**(04)**	**(05)**	**(06)**	**(07)**	**(08)**	**(09)**	**(10)**	**(11)**	**(12)**	**2014**
Launches with traps	697,000	690,000	810,000	1,099,000	1,009,000	892,000	672,000	674,000	791,000	740,000	728,000	800,000	**9,602,000**
Launches with kingfish net	221,000	212,000	226,000	314,000	249,000	190,000	126,000	224,000	221,000	166,000	189,000	318,000	**2,656,000**
Launches with misc. gear	6,000	4,000	9,000	4,000	4,000	5,000	2,000	5,000	8,000	13,000	11,000	6,000	**77,000**
Speedboats with misc. gear	357,000	283,000	351,000	459,000	296,000	229,000	170,000	214,000	516,000	341,000	267,000	384,000	**3,867,000**
Combined	**1,281,000**	**1,189,000**	**1,396,000**	**1,876,000**	**1,558,000**	**1,316,000**	**970,000**	**1,117,000**	**1,536,000**	**1,260,000**	**1,195,000**	**1,508,000**	**16,202,000**
Effort in boat-gear days	**(01)**	**(02)**	**(03)**	**(04)**	**(05)**	**(06)**	**(07)**	**(08)**	**(09)**	**(10)**	**(11)**	**(12)**	**2014**
Launches with traps	3,168	3,072	3,402	3,005	3,070	3,265	2,897	3,203	3,316	2,961	3,169	3,515	**38,043**
Launches with kingfish net	1,011	1,273	1,678	1,324	1,260	1,339	1,198	1,454	1,309	1,135	1,051	1,238	**15,270**
Launches with misc. gear	159	196	193	108	114	195	108	183	213	384	333	194	**2,380**
Speedboats with misc. gear	4,775	4,544	6,181	5,941	5,580	4,450	2,584	4,276	5,784	4,112	3,082	4,468	**55,777**
Combined	...	...	...	...	...	...	...	...	...	...	...	...	...
CPUE in kg / boat-gear day	**(01)**	**(02)**	**(03)**	**(04)**	**(05)**	**(06)**	**(07)**	**(08)**	**(09)**	**(10)**	**(11)**	**(12)**	**2014**
Launches with traps	220.0	224.6	238.1	365.7	328.7	273.2	232.0	210.4	238.5	249.9	229.7	227.6	**252.4**
Launches with kingfish net	218.6	166.5	134.7	237.2	197.6	141.9	105.2	154.1	168.8	146.3	179.8	256.9	**173.9**
Launches with misc. gear	37.7	20.4	46.6	37.0	35.1	25.6	18.5	27.3	37.6	33.9	33.0	30.9	**32.4**
Speedboats with misc. gear	74.8	62.3	56.8	77.3	53.0	51.5	65.8	50.0	89.2	82.9	86.6	85.9	**69.3**
Combined	...	...	...	...	...	...	...	...	...	...	...	...	...

8.3 Computational steps in effort standardization

The objective of the presented method is to achieve effort compatibility when different boat-gear categories are combined together. Specifically, its two tasks are:

Producing *total standardized effort* of combined boat-gear categories;
Computing *standardized CPUE*'s for combined boat-gear categories;

It should be noted here that the example given in this section and summarized in Tables 8.2.1 and 8.4.1 treats catch as a whole and without focusing on a specific fish stock; such a consideration is used only temporarily with the sole purpose of facilitating the presentation of the computational steps in effort standardization. In Section 8.5 that describes the results of the method readers will be presented with two case studies dealing with **Spangled emperor** (*Lethrinus nebulosus)* and **Narrow-barred Spanish mackerel** (*Scomberomorus commerson*) respectively; these are the two top species of the 2014 landings in UAE.

The method starts by considering the compatibility of CPUE's of different boat-gear categories. Since these involve incompatible effort values in the denominator they cannot be combined at monthly or annual levels (notice the absent values for effort and CPUE in the totals line in Table 8.2.1). This happens since they are viewed as weighted averages over a period of a month or a year.

On the other hand each of these CPUE's could be temporarily viewed as the representative catch by just *one* boat from each boat-gear category during *one* day.

Using this second concept for monthly CPUE's by boat-gear category and over 12n periods, a 2-dimensional array of *daily* yields $P_{i,j}$ can be formed where:

i=1…m (boat-gear categories);
j=1…12n (monthly estimates).

To be noted that the notation has changed from CPUE to P since a CPUE is expressed in Kg / boat-gear day while the newly assumed daily yields P are in Kg.

The method proceeds with the following notations and computations:

The sum of all daily yields is given by:

$$P = \sum_{i=1}^{m} \sum_{j=1}^{12n} P_{i,j} \qquad (8.3.1)$$

The arithmetic mean of all daily yields is given by:

$$\bar{P} = \frac{1}{(mx12n)} P \qquad (8.3.2)$$

Working with a boat-gear i it is found that its total daily yield is:

$$P_i = \sum_{j=1}^{12n} P_{i,j} \qquad (8.3.3)$$

and the arithmetic mean is:

$$\overline{P_i} = \frac{1}{12n} P_i \qquad (8.3.4)$$

The overall arithmetic mean $\overline{P}$ shown in (8.3.2) is now assumed to represent the overall daily yield of a new (and hypothetical) boat-gear category. To compare the overall performance of each actual boat-gear to the new hypothetical one the following ratio is used:

$$f_i = \frac{\overline{P_i}}{\overline{\overline{P}}} \qquad (8.3.5)$$

where $\overline{P_i}$ and $\overline{P}$ are obtained from (8.3.4) and (8.3.2) respectively.

This ratio is referred to as s*tandardization factor* since it is used for converting actual effort into a standardized one. Once calculated, the f_i is considered to remain constant across all periods.

Consequently each effort cell $E_{i,j}$ representing effort of boat-gear i in period j **c**an be converted to standardized effort using the expression:

$$E_{i,j}^{STD} = f_i E_{i,j} \qquad i=1\ldots m;\ \ j=1\ldots 12n. \qquad (8.3.6)$$

Adding up all m standardized (thus addable) monthly effort values for a period j will result in a monthly standardized effort E_j^{STD} which combines all boat-gear categories:

$$E_j^{STD} = \sum_{i=1}^{m} E_{i,j}^{STD} \qquad j=1\ldots 12n. \qquad (8.3.7)$$

The standardized CPUE's by boat-gear category are obtained by dividing each catch cell $C_{i,j}$ by the corresponding standardized effort $E_{i,j}^{STD}$ obtained from (8.3.6):

$$CPUE_{i,j}^{STD} = \frac{C_{i,j}}{E_{i,j}^{STD}} \qquad i=1\ldots m;\ \ j=1\ldots 12n. \qquad (8.3.8)$$

Lastly the combined standardized catch-per-unit-effort effort in a period j is calculated. Here the combined monthly catch of all boat-gear categories is divided by the combined monthly standardized effort obtained from (8.3.7).

Combined standardized Catch-Per-Unit-Effort: $CPUE_j^{STD} = \frac{\sum_{i=1}^{m} C_{i,j}}{E_j^{STD}} \qquad j=1\ldots 12n. \qquad (8.3.9)$

At this stage tasks (a) and (b) that were set-up at the beginning of this section have been achieved.

Two points arise now regarding: (a) consistency of standardized data and (b) their numerical treatment across different periods.

It is evident that the standardization factors formulated by the presented approach depend directly on the selection of a hypothetical boat-gear category to be used as standard. According to Robson (1966) the role of such a standard can also be played by any of the actual boat-gear categories; such a flexibility of choice would result in a different but equally valid set of standardization factors. Given that in studying the fluctuation and trend of standardized variables users require consistent sets of data, it becomes apparent that the standardization effort and CPUE values so far obtained need additional treatment in order to become independent of the initial selection of a boat-gear as standard.

One way of achieving this is to adopt the normalization approach used by the ICES North Sea Round Fish Working Group (1980). The approach consists of (i) calculating the arithmetic mean of a standardized variable across periods and, (ii) substituting each standardized value by its proportion to the mean. In such a manner the resulting normalized values are dimensionless and share a similar value scale.

It remains to be seen if such normalized values are independent of the choice of a boat-gear category as standard. This is rather easy to prove without performing tedious computations. Suffice to notice that all expressions involving standardized effort contain two factors: one which is the quotient $1/\overline{P}$ and another that is independent of $\overline{P}$ and depends only on the original data. Consider for instance expression (8.3.7) which computes the combined standardized effort for a given period j. By recalling that each $E_{i,j}^{STD} = f_i E_{i,j}$ and that $f_i = \frac{\overline{P_i}}{\overline{P}}$, this expression can also be written as:

$$E_j^{STD} = \sum_{i=1}^{m} E_{i,j}^{STD} = \sum_{i=1}^{m} f_i E_{i,j} = \frac{1}{\overline{P}} \sum_{i=1}^{m} \overline{P}_i E_{i,j} \qquad (8.3.10)$$

When the combined monthly standardized effort is summed across periods, its arithmetic mean will also contain $\frac{1}{\overline{P}}$. During normalization each standardized effort from (8.3.10) will be divided by the arithmetic mean thus canceling out $\frac{1}{\overline{P}}$ and making the obtained normalized effort independent of the initial choice of a boat-gear category as standard.

Working in a similar manner with the standardized CPUE's we find that their sums and arithmetic means contain an expression of $\overline{P}$ and other expressions that are independent of it. During the normalization process the expressions of $\overline{P}$ cancel out thus proving that the normalized CPUE's are independent of the initial choice of a boat-gear category as standard.

8.4 A numerical example

Table 8.4.1 shows the results of the standardization approach after it has applied to the UAE-NFIS catch/effort data of Table 8.2.1.

Here the standardization involves m=4 boat-gear categories and 12 catch/effort monthly estimates resulting a total of 48 CPUE's. It is recalled that during the standardization phase the notation of these CPUE's will temporarily be changed to P since they will be viewed as representing daily catches. Accordingly their units will be in Kg.

Calculation of standardization factors

First the sum of all 48 daily yields (12 yields for each of the 4 boat-gear categories) is calculated.

$$P = \sum_{i=1}^{m=4} \sum_{j=1}^{12} P_{i,j} = 6{,}366 \text{ Kg.}$$

The corresponding arithmetic mean $\overline{P}$ in (2) will be equal to 6,366/48 = 132.6 Kg.

Next step is the calculation of average daily yields for each boat-gear category using expressions (3) and (4).

Launches with traps:	$P_1 = 3{,}038.5$ Kg and $\overline{P}_1 = 253.2$ Kg.
Launches with kingfish net:	$P_2 = 2{,}107.5$ Kg and $\overline{P}_2 = 175.6$ Kg.
Launches with misc. gear:	$P_3 = 383.8$ Kg and $\overline{P}_3 = 32.0$ Kg.
Speedboats with misc. gear:	$P_4 = 836.2$ Kg and $\overline{P}_4 = 69.7$ Kg.

Calculation of standardization factors (STD) makes use of expression (8.3.5). Each of the above averages is divided by $\overline{P} = 132.6$ calculated earlier:

STD factor for launches with traps = 253.2/132.6 = 1.909.
STD factor for launches with kingfish net= 175.6/132.6 = 1.324.
STD factor for launches with misc. gear = 32.0/132.6 = 0.241.
STD factor for speedboats with misc. gear = 69.7/132.6 = 0.525.

These results are shown in the first block of Table 8.4.1. To be noted that once these factors have been calculated they apply to all 12 monthly columns of 2014.

Calculation of standardized effort

The second block of Table 8.4.1 illustrates standardized effort for each of the four boat-gear categories. All standardized effort figures by boat-gear category are resulting from the application of expression (8.3.6) to all effort cells in Table 8.2.1. For instance in January 2014 the actual effort of launches with traps is 3,168 boat-gear days. The standardization factor for this boat-gear category is 1.909. By multiplying the 3,168 actual boat-gear days by this factor we obtain a standardized effort of 6,048 boat-gear days (first cell of the second block in Table 8.4.1).

To be noted that since all standardized effort values are addable it is now possible to combine them vertically across boat-gear categories and then horizontally across months, thus obtaining a total effort figure for 2014 equal to 122,733 standardized boat-gear days.

Next line shows combined standardized effort in normalized form. The arithmetic mean of the 12 effort figures is 10,228 boat-gear days. The normalized value of the first entry is 9,934 / 10,228 = 0.971.The rest of the normalized effort values are calculated likewise.

Calculation of standardized CPUE's

The third block of Table 8.4.1 illustrates standardized CPUE's for each of the four boat-gear categories. All figures are resulting from the application of expression (8.3.8) to each CPUE cell in Table 1. For instance in January 2014 the standardized CPUE for launches with traps will be 697,000 Kg of catch (first cell in Table 8.2.1) divided by the corresponding standardized effort of 6,048 boat-gear days, which gives 115.2 Kg / boat-gear day.

A combined standardized CPUE is also computed using expression (8.3.9). Here the total catch for January 2014 is 1,281,000 Kg and the combined standardized effort is 9,934 boat-gear days, thus resulting a combined standardized CPUE of 128.9 Kg / boat-gear day.

Next line shows combined standardized CPUE in normalized form. The arithmetic mean of the 12 combined CPUE's figures is 131.7. The normalized value of the first entry is 128.9 / 131.7 = 0.979. The rest of the normalized effort values are calculated likewise.

To be noted that the notation for catch-per-unit-effort has returned back to CPUE since this variable is again calculated as a weighted average of catch divided by effort.

Figure 8.4.1 illustrates a plot of the normalized effort and CPUE contained in Table 8.4.1.

Table 8.4.1 Standardization of the UAE-NFIS catch/effort of Table 8.2.1

Standardization factors	(01)	(02)	(03)	(04)	(05)	(06)	(07)	(08)	(09)	(10)	(11)	(12)	2014
Launches with traps	1.909	1.909	1.909	1.909	1.909	1.909	1.909	1.909	1.909	1.909	1.909	1.909	**1.909**
Launches with kingfish net	1.324	1.324	1.324	1.324	1.324	1.324	1.324	1.324	1.324	1.324	1.324	1.324	**1.324**
Launches with misc. gear	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	0.241	**0.241**
Speedboats with misc. gear	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	0.525	**0.525**

Standardized effort	(01)	(02)	(03)	(04)	(05)	(06)	(07)	(08)	(09)	(10)	(11)	(12)	2014
Launches with traps	6,048	5,865	6,495	5,737	5,861	6,234	5,531	6,115	6,331	5,653	6,050	6,711	72,633
Launches with kingfish net	1,339	1,686	2,222	1,753	1,669	1,773	1,586	1,925	1,733	1,503	1,392	1,639	20,221
Launches with misc. gear	38	47	47	26	27	47	26	44	51	93	80	47	574
Speedboats with misc. gear	2,509	2,387	3,247	3,121	2,932	2,338	1,358	2,247	3,039	2,160	1,619	2,347	29,305
Combined	**9,934**	**9,986**	**12,011**	**10,638**	**10,489**	**10,392**	**8,501**	**10,331**	**11,155**	**9,409**	**9,142**	**10,745**	**122,733**
Normalized	**0.971**	**0.976**	**1.174**	**1.040**	**1.026**	**1.016**	**0.831**	**1.010**	**1.091**	**0.920**	**0.894**	**1.051**	

Standardized CPUE's	(01)	(02)	(03)	(04)	(05)	(06)	(07)	(08)	(09)	(10)	(11)	(12)	2014
Launches with traps	115.2	117.6	124.7	191.6	172.1	143.1	121.5	110.2	124.9	130.9	120.3	119.2	**132.2**
Launches with kingfish net	165.1	125.8	101.7	179.1	149.2	107.2	79.4	116.3	127.5	110.4	135.8	194.0	**131.3**
Launches with misc. gear	156.5	84.6	193.4	153.6	145.5	106.3	76.8	113.3	155.8	140.4	137.0	128.3	**134.2**
Speedboats with misc. gear	142.3	118.5	108.1	147.1	101.0	97.9	125.2	95.3	169.8	157.8	164.9	163.6	**132.0**
Combined	**128.9**	**119.1**	**116.2**	**176.4**	**148.5**	**126.6**	**114.1**	**108.1**	**137.7**	**133.9**	**130.7**	**140.3**	**132.0**
Normalized	**0.979**	**0.904**	**0.882**	**1.339**	**1.128**	**0.961**	**0.866**	**0.821**	**1.045**	**1.017**	**0.992**	**1.065**	

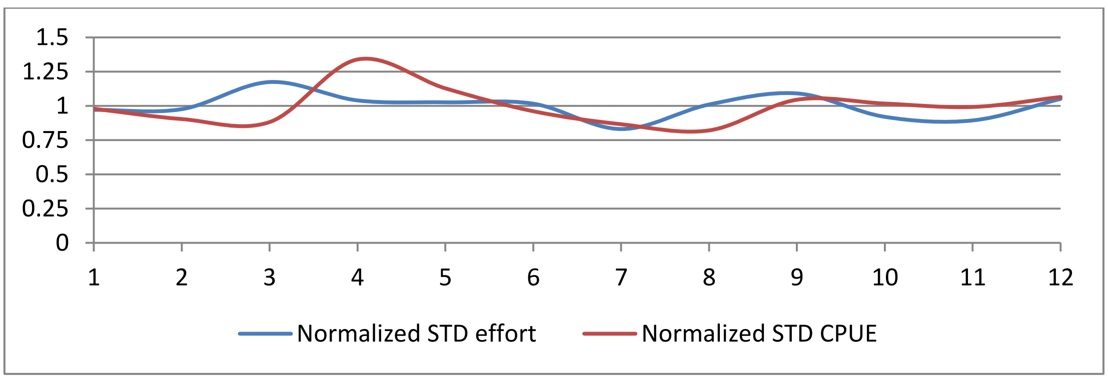

Figure 8.4.1 Plot of normalized effort and CPUE based on the 2014 UAE-NFIS catch/effort data

8.5 Results of the method

Application of effort standardization to the fishery of **Spangled emperor** (*Lethrinus nebulosus)*

As already mentioned in Introduction **Spangled emperor** (*Lethrinus nebulosus)* was the top species in 2014 with landings representing 16.2% of the total.

Figure 8.5.1 Spangled emperor (*Lethrinus nebulosus)*

This species is targeted by launches with traps and speedboats (tarads). Catches by the other two boat-gear categories are negligible and regarded as accidental. Consequently effort standardization focuses on the above two boat-gear categories. Launches with traps is the predominant boat-gear accounting for 76% of the species catches in 2013 and 71% in 2014.

Table 8.5.1 illustrates catch/effort data for 2013 and 2014. Since the effort exerted by the two boat-gear categories is not compatible no combined data are shown for effort and CPUE's in the last two columns.

Table 8.5.2 shows the results of the standardization process, including normalized values for effort and CPUE.

Figure 8.5.2 illustrates monthly plots of normalized effort and CPUE. There is a slight (but visible) rising trend for fishing effort and a declining one for the CPUE.

Table 8.5.1 Catch/effort data for Sh'ari شعري Lethrinus nebulosus (Spangled emperor) (2013 – 2014)
Accuracy of estimates: 90.6%

	Launches with traps			Speedboats			Combined		
Period	**Catch**	**Effort**	**CPUE**	**Catch**	**Effort**	**CPUE**	**Catch**	**Effort**	**CPUE**
1	111,130	4,561	24.37	61,683	2,594	23.78	172,813	...	...
2	98,669	3,911	25.23	69,040	2,485	27.79	167,709	...	...
3	118,539	2,625	45.17	68,689	2,313	29.70	187,228	...	...
4	382,464	3,475	110.06	57,079	3,120	18.29	439,543	...	...
5	272,046	2,790	97.50	62,590	4,086	15.32	334,636	...	...
6	190,795	2,524	75.60	39,851	1,705	23.38	230,646	...	...
7	150,694	2,812	53.59	24,752	1,143	21.65	175,446	...	...
8	129,136	3,010	42.90	12,556	816	15.39	141,692	...	...
9	89,015	2,812	31.66	33,806	2,771	12.20	122,821	...	...
10	115,355	3,267	35.31	53,946	2,769	19.48	169,301	...	...
11	119,403	3,228	36.99	47,075	6,417	7.34	166,478	...	...
12	128,556	2,907	44.22	59,225	3,733	15.87	187,781	...	...
2013	**1,905,802**	**37,922**	**50.26**	**590,294**	**33,952**	**17.39**	**2,496,096**	...	...
1	133,786	3,168	42.23	104,538	2,911	35.92	238,324	...	...
2	148,305	3,072	48.27	82,761	2,842	29.12	231,066	...	...
3	193,881	3,402	56.99	47,560	5,484	8.67	241,441	...	...
4	298,826	3,005	99.44	111,581	5,941	18.78	410,407	...	...
5	239,525	3,070	78.02	81,146	5,580	14.54	320,671	...	...
6	159,508	3,265	48.86	48,273	4,371	11.04	207,781	...	...
7	89,878	2,897	31.02	31,842	2,584	12.32	121,720	...	...
8	111,765	3,203	34.90	39,273	1,869	21.01	151,038	...	...
9	107,153	3,316	32.31	32,439	3,009	10.78	139,592	...	...
10	106,558	2,961	35.98	40,819	4,112	9.93	147,377	...	...
11	124,312	3,169	39.23	32,808	2,064	15.89	157,120	...	...
12	130,134	3,515	37.03	88,952	4,468	19.91	219,086	...	...
2014	**1,843,630**	**38,044**	**48.46**	**741,993**	**45,236**	**16.40**	**2,585,623**	...	...

Table 8.5.2 Standardized effort and CPUE for Sh'ari شعري Lethrinus nebulosus (Spangled emperor) (2013 – 2014)

	Launches with traps			Speedboats			Combined			
Period	**STD factor**	**STD effort**	**STD CPUE**	**STD factor**	**STD effort**	**STD CPUE**	**STD effort**	**Norma-lized**	**STD CPUE**	**Norma-lized**
1	1.467	6,691	16.61	0.533	1,383	44.61	8,073	1.261	21.41	0.644
2	1.467	5,737	17.19	0.533	1,324	52.13	7,061	1.103	23.75	0.714
3	1.467	3,850	30.78	0.533	1,233	55.72	5,083	0.794	36.83	1.107
4	1.467	5,098	75.01	0.533	1,663	34.32	6,761	1.056	65.01	1.955
5	1.467	4,093	66.45	0.533	2,178	28.74	6,271	0.980	53.36	1.604
6	1.467	3,703	51.52	0.533	909	43.86	4,611	0.720	50.02	1.504
7	1.467	4,125	36.52	0.533	609	40.62	4,734	0.739	37.06	1.114
8	1.467	4,416	29.23	0.533	435	28.87	4,851	0.758	29.21	0.878
9	1.467	4,125	21.57	0.533	1,477	22.89	5,602	0.875	21.92	0.659
10	1.467	4,793	24.06	0.533	1,476	36.55	6,269	0.979	27.01	0.812
11	1.467	4,735	25.21	0.533	3,420	13.76	8,156	1.274	20.41	0.614
12	1.467	4,265	30.13	0.533	1,990	29.77	6,255	0.977	30.02	0.903
2013	**1.467**	**55,631**	**34.25**	**0.533**	**18,097**	**32.62**	**73,728**		**33.86**	
1	1.467	4,648	28.78	0.533	1,551	67.38	6,199	0.968	38.44	1.156
2	1.467	4,507	32.90	0.533	1,515	54.64	6,022	0.941	38.37	1.154
3	1.467	4,991	38.84	0.533	2,923	16.27	7,914	1.236	30.51	0.917
4	1.467	4,408	67.77	0.533	3,167	35.24	7,575	1.183	54.18	1.629
5	1.467	4,504	53.17	0.533	2,974	27.28	7,478	1.168	42.88	1.289
6	1.467	4,789	33.30	0.533	2,330	20.72	7,119	1.112	29.19	0.878
7	1.467	4,250	21.14	0.533	1,377	23.12	5,627	0.879	21.63	0.650
8	1.467	4,699	23.78	0.533	996	39.42	5,695	0.890	26.52	0.797
9	1.467	4,865	22.02	0.533	1,604	20.23	6,469	1.010	21.58	0.649
10	1.467	4,344	24.52	0.533	2,192	18.62	6,536	1.021	22.55	0.678
11	1.467	4,649	26.73	0.533	1,100	29.82	5,749	0.898	27.33	0.822
12	1.467	5,156	25.23	0.533	2,382	37.35	7,538	1.177	29.06	0.874
2014	**1.467**	**55,811**	**33.03**	**0.533**	**24,111**	**30.77**	**79,922**		**32.35**	

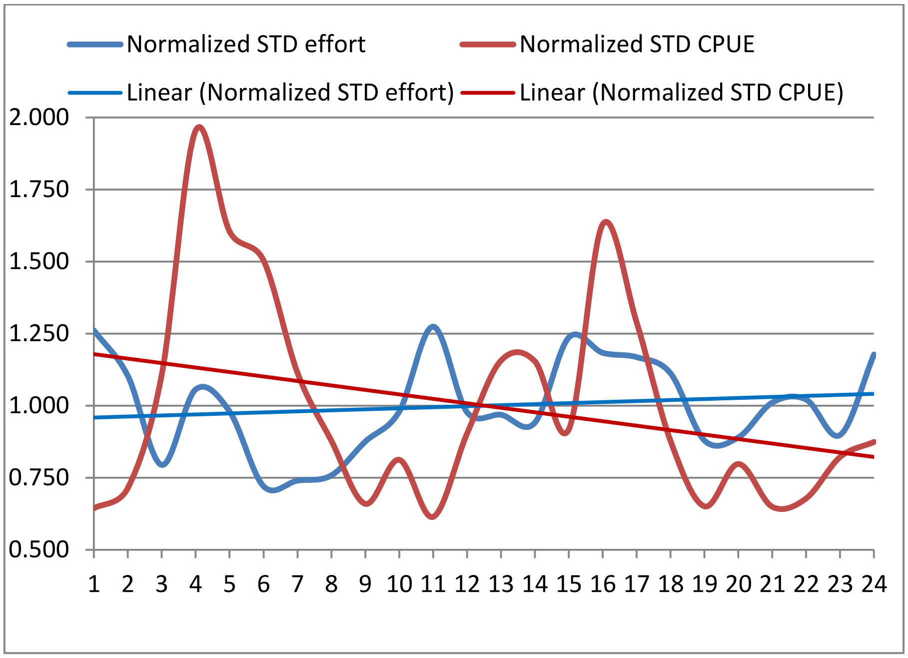

Figure 8.5.2 Monthly plots of normalized effort and CPUE for Sh'ari شعري Lethrinus nebulosus (Spangled emperor) (2013 – 2014). There is a slight (but visible) rising trend for fishing effort and a declining one for the CPUE.

Application of effort standardization to the fishery of **Narrow-barred Spanish mackerel** (*Scomberomorus commerson)*

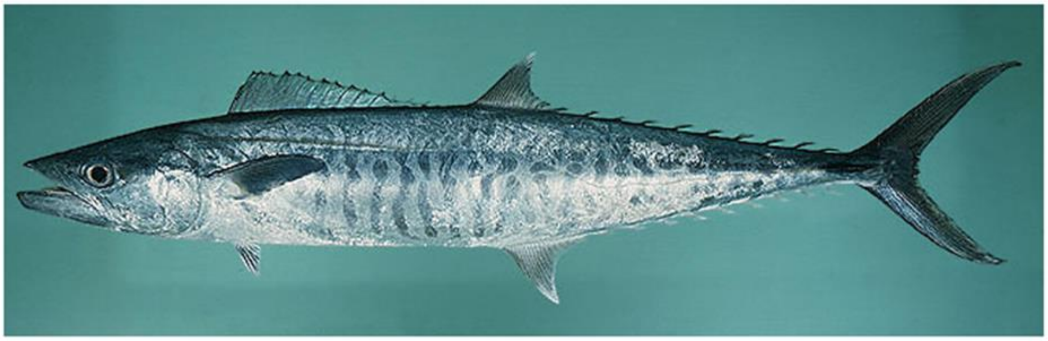

Figure 8.5.3 Narrow-barred Spanish mackerel (*Scomberomorus commerson*)

This important species (second in the 2014 ranked landings and representing 10.5% of the total) is targeted by launches with kingfish net and speedboats (tarads). Catches by launches with miscellaneous gear are negligible and are not included in the case study. Launches with kingfish net is by far the predominant boat-gear accounting for 90% of the species catches in 2013 and 95% in 2014.

Table 8.5.3 illustrates catch/effort data for 2013 and 2014. Since the effort exerted by the two boat-gear categories is not compatible no combined data are shown for effort and CPUE's in the last two columns.

Table 8.5.4 shows the results of the standardization process, including normalized values for effort and CPUE.

Figure 8.5.4 illustrates monthly plots of normalized effort and CPUE. There is a slight (but visible) declining trend for both fishing effort and the CPUE.

Table 8.5.3 Catch/effort data for Narrow-barred Spanish mackerel (*Scomberomorus commerson*) (2013 – 2014). Accuracy of estimates: 88.3%

	Launches with kingfish net			Speedboats			Combined		
Period	**Catch**	**Effort**	**CPUE**	**Catch**	**Effort**	**CPUE**	**Catch**	**Effort**	**CPUE**
1	258,119	1,874	137.77	10,157	508	20.00	268,276	...	...
2	133,628	2,213	60.39	0	0	0.00	133,628	...	...
3	199,616	2,138	93.36	10,770	950	11.34	210,386	...	...
4	118,478	1,261	93.93	90,642	1,784	50.81	209,120	...	...
5	130,612	1,285	101.68	48,423	823	58.81	179,035	...	...
6	94,031	948	99.24	8,809	896	9.83	102,840	...	...
7	127,899	1,532	83.49	1,903	257	7.40	129,802	...	...
8	175,797	1,644	106.93	0	0	0.00	175,797	...	...
9	188,768	1,972	95.73	2,536	1,700	1.49	191,304	...	...
10	287,010	2,015	142.40	324	1,573	0.21	287,334	...	...
11	114,090	1,565	72.92	30,691	6,579	4.67	144,781	...	...
12	112,053	1,237	90.56	22,743	2,368	9.61	134,796	...	...
2013	**1,940,103**	**19,683**	**98.57**	**226,998**	**17,438**	**13.02**	**2,167,101**	...	...
1	97,925	1,011	96.85	8,168	2,911	2.81	106,093	...	...
2	117,490	1,273	92.30	0	0	0.00	117,490	...	...
3	125,140	1,678	74.56	6,281	2,250	2.79	131,421	...	...
4	255,711	1,324	193.14	11,835	5,032	2.35	267,546	...	...
5	147,988	1,260	117.44	5,351	4,786	1.12	153,339	...	...
6	102,938	1,339	76.90	1,044	1,299	0.80	103,982	...	...
7	64,857	1,198	54.16	699	68	10.25	65,556	...	...
8	112,760	1,454	77.54	1,277	1,240	1.03	114,037	...	...
9	131,961	1,309	100.82	9,974	2,285	4.37	141,935	...	...
10	113,214	1,135	99.75	20,350	2,379	8.55	133,564	...	...
11	144,751	1,051	137.69	10,800	2,064	5.23	155,551	...	...
12	200,665	1,238	162.07	193	2,010	0.10	200,858	...	...
2014	**1,615,400**	**15,270**	**105.79**	**75,972**	**26,324**	**2.89**	**1,691,372**	...	...

Table 8.5.4 Standardized effort and CPUE for Narrow-barred Spanish mackerel (*Scomberomorus commerson*) (2013 – 2014)

	Launches with kingfish net			Speedboats			Combined			
Period	**STD factor**	**STD effort**	**STD CPUE**	**STD factor**	**STD effort**	**STD CPUE**	**STD effort**	**Norma-lized**	**STD CPUE**	**Norma-lized**
1	1.840	3,448	74.86	0.160	81	125.27	3,529	1.188	76.02	1.398
2	1.840	4,072	32.81	0.160	0	0.00	4,072	1.370	32.81	0.603
3	1.840	3,935	50.73	0.160	152	70.99	4,087	1.375	51.48	0.947
4	1.840	2,321	51.04	0.160	285	318.27	2,606	0.877	80.24	1.476
5	1.840	2,364	55.25	0.160	131	368.37	2,495	0.840	71.74	1.319
6	1.840	1,744	53.93	0.160	143	61.55	1,887	0.635	54.50	1.002
7	1.840	2,819	45.37	0.160	41	46.35	2,860	0.963	45.38	0.835
8	1.840	3,026	58.10	0.160	0	0.00	3,026	1.018	58.10	1.068
9	1.840	3,629	52.02	0.160	271	9.34	3,900	1.313	49.05	0.902
10	1.840	3,709	77.38	0.160	251	1.29	3,960	1.333	72.55	1.334
11	1.840	2,879	39.62	0.160	1,050	29.22	3,930	1.323	36.84	0.677
12	1.840	2,277	49.21	0.160	378	60.17	2,655	0.894	50.77	0.934
2013	**1.840**	**36,224**	**53.56**	**0.160**	**2,784**	**81.53**	**39,008**		**55.56**	
1	1.840	1,861	52.62	0.160	465	17.57	2,326	0.783	45.62	0.839
2	1.840	2,343	50.15	0.160	0	0.00	2,343	0.788	50.15	0.922
3	1.840	3,089	40.51	0.160	359	17.49	3,448	1.160	38.12	0.701
4	1.840	2,437	104.95	0.160	803	14.73	3,240	1.090	82.58	1.518
5	1.840	2,319	63.81	0.160	764	7.00	3,083	1.038	49.73	0.915
6	1.840	2,464	41.78	0.160	207	5.04	2,671	0.899	38.93	0.716
7	1.840	2,204	29.43	0.160	11	64.20	2,215	0.745	29.60	0.544
8	1.840	2,676	42.13	0.160	198	6.45	2,874	0.967	39.67	0.730
9	1.840	2,409	54.78	0.160	365	27.34	2,774	0.933	51.17	0.941
10	1.840	2,089	54.20	0.160	380	53.57	2,469	0.831	54.10	0.995
11	1.840	1,935	74.82	0.160	330	32.77	2,264	0.762	68.70	1.263
12	1.840	2,279	88.07	0.160	321	0.60	2,600	0.875	77.27	1.421
2014	**1.840**	**28,103**	**57.48**	**0.160**	**4,203**	**18.08**	**32,306**		**52.36**	

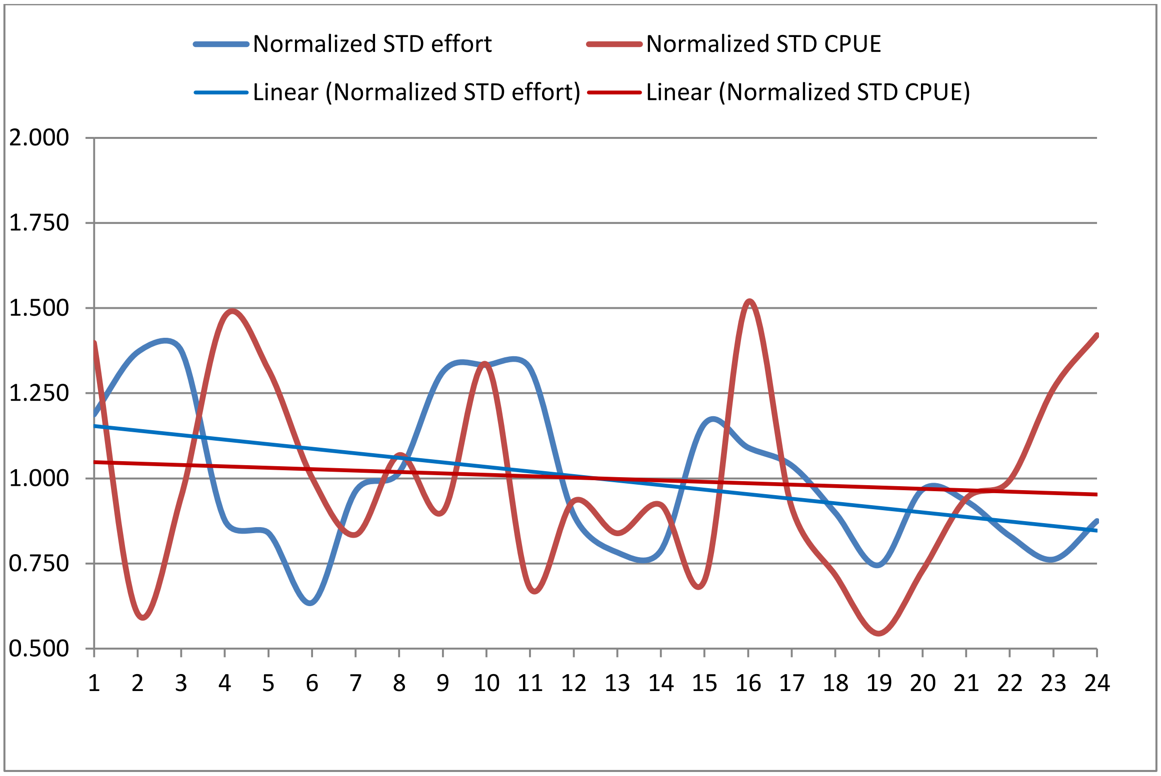

Figure 8.5.4 Monthly plots of normalized effort and CPUE for Narrow-barred Spanish mackerel (*Scomberomorus commerson*) (2013 – 2014). There is a slight (but visible) declining trend for both fishing effort and the CPUE.

8.6 Observations

8.6.1 Comparison to other methods

As mentioned in the Introduction UAE-NFIS has recently adopted the presented approach that combines elements of the *normalized relative effort* (used by the North Sea Round Fish Working Group, ICES, 1980) and the *relative fishing power* developed by Robson (1966). It was also mentioned that although the existing literature offers a plethora of other more recent and more sophisticated methods it was nevertheless considered preferable to first try out approaches that (a) depend only on catch/effort data from commercial fisheries and, (b) are applicable to situations of limited time coverage.

The Robson basic concept of relative fishing power was adopted in formulating effort standardization factors. The presented method uses a variation to the Robson concept; instead of arbitrarily selecting an existing CPUE to use as standard it uses for this purpose a mean daily yield of one fishing unit of a hypothetical boat-gear category. This variation does not constitute a real difference since Robson states that in choosing a CPUE standard "any boat-gear is as good as another". It is the authors' view, however, that involving all boat-gear categories in the source data makes the selection of the CPUE standard less arbitrary.

On the other hand the fact remains that users should be free to use any standard that would be appropriate or convenient for their work. This means that several standardized datasets, all equally valid but different from each other, might be resulting from the same source data.

To overcome this problem the presented method further processes the standardized data with the objective of making them consistent irrespective of the initial choice of a CPUE as standard. It was shown that such an objective can be achieved by means of a normalization process such as the one adopted by the North Sea Round Fish Working Group, ICES (1980). In this commonly used statistical technique each standardized value (whether referring to effort or to CPUE) is replaced by its proportion to the arithmetic mean of all values. In such a manner the resulting normalized values are dimensionless and share a similar value scale. Moreover the authors explicitly demonstrated that normalized values obtained in this manner are independent of the choice of the boat-gear standard initially selected.

Lastly the presented method follows the same concept of dynamic standardization shown in both ICES and Robson approaches. Monthly and annual standardization factors (and hence normalized effort and CPUE's) vary when the source data cover different numbers of years. For instance, launches with kingfish net have a standardization factor of 1.840 over the period January 2013 – December 2014. This value will be different when the source data will extend to December 2015, December 2016, etc. Such a consideration is essential in order for the standardized variables to be compatible across all periods, a criterion that would not be met if standardization was to apply for each year separately.

8.6.2 Equivalent approaches for the formulation of standardization factors

Expression (8.3.5) specifies that the standardization factor for a specific boat-gear category is directly defined as the ratio of its overall CPUE (viewed temporarily as the average daily yield $\overline{P_i}$ of a single fishing unit over all periods) to the average daily yield $\overline{P}$ of a hypothetical boat-gear. It is recalled that $\overline{P_i}$ is obtained from expression (8.3.4) and $\overline{P}$ from expressions (8.3.1) and (8.3.2).

Viewing CPUE's as overall daily yields of single fishing units was considered both convenient and practical in the present study. The chosen approach however does not preclude the adoption of other hypotheses which can produce the same results by means of different interpretations of the CPUE's. For instance an alternative approach is to formulate standardization factors on the basis of *days needed catching the same arbitrary quantity Q*. Under such a scheme the days needed for each boat-gear to catch Q will be $Q/\overline{P_i}$. Next a hypothetical boat-gear category with catch-per-unit-effort equal to $\overline{P}$ is considered. Here the number of days needed to catch Q is equal to $Q/\overline{P}$. Since the number of days needed is in reverse proportion to the relative importance of a boat-gear (i.e. higher performance implies fewer days to catch a given quantity Q) we divide the second ratio by the first, thus obtaining the same standardization factor $f_i = \frac{\overline{P_i}}{\overline{P}}$.

8.6.3 The problem of data gaps

Maunder (2004) has stressed the importance of paying due attention to situations in which there are data gaps in the datasets. The remedies are not always simple and in some cases they become quite elaborate.

Data gaps are more frequent in monthly data. For instance a species may disappear temporarily from the landings as a result of the seasonality of the fishery. Or a new and important boat-gear category may enter the fishery at a certain point thus creating gaps in the effort and catch of the earlier periods. Likewise a boat-gear might disappear at a certain point thus creating gaps in the effort and catch of the periods that come after.

It is the authors’ view that the problem of data gaps does not affect the presented method since the standardization factors are calculated on the basis of cumulative daily yields covering the entire reference period. It was shown that the standardization process applies to a matrix of source data in which cells may as well contain zeroes (for instance the speedboats in February and August 2013 and in February 2014). In mathematical terms the only condition for a boat-gear category to participate in the process is to have at least one non-zero entry in the matrix. In practice, however, boat-gear categories showing small and scattered quantities of accidental catch are not included in the process as was for instance the case or launches with miscellaneous gear catching kingfish.

8.6.4 Reliability of catch/effort data

Another point worth addressing is the reliability of catch/effort estimates that constitute the data source for the standardization process.

In UAE-NFIS catch/effort estimates go through a gauntlet of several quality checks before they are reported. Data reliability is rigorously monitored by means of sampling schedules at the beginning of each operational month and work progress reports that are being consulted by the system administrator at any instance. Sampling schedules indicate the number of sampling days to be used and the number of samples to be collected on each sampling day; all such norms apply to all ports and all boat-gear categories. Samples of landings and boat-gear activity are collected in a parallel manner, using different sampling norms and independently of each other. The aim of such planning is to achieve an overall accuracy of catch/effort estimates that stays above 90%; this has been consistently achieved from 2014 onwards. This threshold of 90% is rather empirical but it is generally accepted as satisfactory in large-scale statistical monitoring systems. In more exact terms the real accuracy achieved is not known; the approaches used by UAE-NFIS make use of the “pessimistic” accuracy concept in which the resulting accuracy stays above a pre-set lower limit (Stamatopoulos, 2003). It is also a composite index incorporating a spatial accuracy (a function of sample size) and a temporal accuracy (that depends on sampling frequency). In addition to the above two relative indices of accuracy the Sampling Uniformity Index (SUI) monitors the uniformity of samples over the sampling days and it penalizes the temporal accuracy in cases of uneven concentrations of samples favouring certain sampling days.

Chapter 9: Multi-variate ranking of catch/effort variables

9.1 Introduction

Ranking the values of a single variable is a commonly used technique in the analysis of catch/effort estimates. In such a ranking the most or least important element appears on top, followed by all other elements in descending or ascending order of importance. In the case of descending order (i.e. from maximum to minimum values) ranking is usually accompanied by two types of percentages: the first indicating the proportion of a ranked value to the total and the second the cumulative proportion to that point. This type of percentage permits a quick understanding of the most important elements which, when combined, account for a given proportion to the grand total.

Table 9.1.1 provides an example of such a ranking. It refers to catch by species in 2014 and displays the top 12 species the combined catch of which accounts for about 75% of the 2014 production. The first column indicates the species, the second the catch in 1000 Kg, the third indicates the percentage to the total and the fourth the cumulative percentage.

Table 9.1.1 Ranking and cumulative percentages for species catch (in '000 Kg) for 2014

2014 annual totals : Ranking and cumulative percentages			
Sh'ari شعري Lethrinus nebulosus (Spangled emperor)	2,625.3	16.2 %	**16.2 %**
Kanaad كنعد Scomberomorus commerson (Spanish mackerel)	1,702.7	10.5 %	**26.7 %**
Saafi صافي Siganus canaliculatus (Whit-spott spinefoot)	1,344.5	8.3 %	**35.0 %**
OTHER أنواع أخرى (Miscellaneous)	1,117.9	6.9 %	**41.9 %**
Boukshina بوقشينة Lethrinus lentjan (Pink ear emperor)	1,064.1	6.6 %	**48.4 %**
Hamour هامور Epinephelus coioides (Orange-spotted grouper)	968.0	6.0 %	**54.4 %**
Qurqufan قرقفان Rhabdosargus haffara (Haffara seabream)	702.4	4.3 %	**58.7 %**
Koffar كوفر Argyrops spinifer (King soldier bream)	592.8	3.7 %	**62.4 %**
Farsh فرش Diagramma pictum (Painted sweetlips)	574.4	3.5 %	**65.9 %**
Rebeeb ربيب Gnathanodon speciosus (Gold toothless treval)	557.2	3.4 %	**69.4 %**
Jash جش Carangoides bajad (Orangespotted trevally)	510.7	3.1 %	**72.5 %**
Tabaan تبان Euthynnus affinis (Kawakawa)	355.6	2.2 %	**74.7 %**

Table 9.1.2 provides a second example showing the top 8 species accounting for about 75% of the total value of the 2014 production.

Table 9.1.2 Top 8 species accounting for about 75% of the total value (in '000 QR) of the 2014 production

2014 annual totals : Ranking and cumulative percentages			
Kanaad كنعد Scomberomorus commerson (Spanish mackerel)	44,765	18.2 %	**18.2 %**
Hamour هامور Epinephelus coioides (Orange-spotted grouper)	41,323	16.8 %	**35.0 %**
Saafi صافي Siganus canaliculatus (Whit-spott spinefoot)	31,158	12.7 %	**47.7 %**
Sh'ari شعري Lethrinus nebulosus (Spangled emperor)	23,955	9.7 %	**57.4 %**
OTHER أنواع أخرى (Miscellaneous)	15,342	6.2 %	**63.7 %**
Rebeeb ربيب Gnathanodon speciosus (Gold toothless treval)	11,111	4.5 %	**68.2 %**
Boukshina بوقشينة Lethrinus lentjan (Pink ear emperor)	7,805	3.2 %	**71.4 %**
Jash جش Carangoides bajad (Orangespotted trevally)	7,118	2.9 %	**74.3 %**

In the first ranking of species catches the top species is Sh'ari *(Lethrinus nebulosus)*. In the second example the top species in terms of value is Kanaad (*Scomberomorus commerson*).

9.2 Cases requiring multi-variate ranking

By examining the two ranking examples provided in the previous section the question arises as to which species is the most important according to a combined criterion involving both catch and value. In general there can exist several criteria, equally important, for ranking elements (the shown cases referred to species but the method to be presented can deal with any type of element, such as boat-gear types, strata, etc.).

If an element appears on the top in all single ranking lists then there is no doubt that it is the most important. But generally elements are placed differently in these lists and it is often difficult to measure their overall importance by visual examination. It is in this respect that a multi-variate ranking comes in as a simple and useful tool.

9.3 Normalization of quantitative criteria

Let us assume a list of m elements denoted by: e_i, i=1…m.
We also assume that each element e_i is associated to an array of n values v_{ij}, j=1…n, which constitute numerical criteria of importance amongst the elements.
Since the columns (i.e. criteria values) of the matrix v_{ij} have different value ranges, the first step is to normalize these columns by mapping them to a common numerical scale. Generally a convenient scale is that of 0-1. Such a mapping is achieved by the following computations:

Determining the minimum and maximum values $\min_j$ and $\max_j$ for each column j.
All values v_{ij} will be mapped to normalized values u_{ij} using the formula:

$$u_{ij} = \frac{v_{ij} - \min_j}{\max_j - \min_j} \qquad (9.3.1)$$

Expression 9.3.1 indicates that each column will have values between 0 and 1 inclusive.

9.4 Distances from “ideal point”

Next we define the "ideal point" as the n-dimensional point with coordinates ($b_i, b_2, \ldots, b_n$). In other words for each of the n criteria we select the "best" value to be represented in the ideal point. A b_j will be 1 if the criterion favours high values or 0 if otherwise.

Next each element e_i will be assigned a distance d_i from the ideal point computed as:

$$d_i = \sqrt{\sum_{j=1}^{n} (u_{ij} - b_j)^2} \qquad (9.4.1)$$

During this process we also compute the maximum distance found D.

9.5 The rank indicator R

For each element e_i the following ranking indicator is computed:

$$R_i = 100 \text{ x } \frac{d_i}{D} \qquad (9.5.1)$$

This indicator expresses in percentage form the relative importance of an element using the combined numerical criteria described by the matrix v_{ij} described in Section 9.3.

9.6 Numerical example

The UAE-NFIS utility used for multi-variate ranking made use of two tables relating to species catches and species values respectively. Fifty five species participated in the process. The general procedure described in the previous sections furnished the results shown in Table 9.6.1. *Scomberomorus commerson* (Spanish mackerel) came up as the most important species in considering both catch and value. It is recalled that in terms of catch this species was second in the ranking list. *Lethrinus nebulosus* (Spangled emperor) which was top species in the catch list appears second in the combined ranking.

Table 9.6.1 Application of multi-variate ranking to species catch and values

Rank	Ranked elements	Relative importance (in %)
1	Kanaad كنعد Scomberomorus commerson (Spanish mackerel)	75.15
2	Sh'ari شعري Lethrinus nebulosus (Spangled emperor)	67.12
3	Saafi صافي Siganus canaliculatus (Whit-spott spinefoot)	59.35
4	Hamour هامور Epinephelus coioides (Orange-spotted grouper)	55.03
5	Boukshina بوقشينة Lethrinus lentjan (Pink ear emperor)	28.04
6	Rebeeb ربيب Gnathanodon speciosus (Gold toothless treval)	22.99
7	Qurqufan قرقفان Rhabdosargus haffara (Haffara seabream)	18.63
8	Jash جش Carangoides bajad (Orangespotted trevally)	17.65
9	Koffar كوفر Argyrops spinifer (King soldier bream)	17.64
10	Farsh فرش Diagramma pictum (Painted sweetlips)	14.29
11	Bedha بدحة Gerres longirostris (silver-biddy)	12.33
12	Karari كراري Atule mate (Yellowtail scad)	9.87
13	Kobkob قبقب Portunus pelagicus (Blue swimmimg crab)	9.42

14	Tabaan تبان Euthynnus affinis (Kawakawa)	8.88
15	Jid جد Sphyraena flavicauda (Yellowtail barracuda)	8.61
16	Semaan سمان Epinephelus bleekeri (Duskytail grouper)	8.59
17	Zubaidi زبيدي Carangoides malabaricus (Malabar trevally)	7.63
18	Saal صال Carangoides chrysophrys (Longnose trevally)	7.11
19	Dal'ah ضلعه Scomberoides commersonnianus (Queenfish)	6.05
20	Sooli سولي Lethrinus microdon (Smalltooth emperor)	5.94
21	Yanam ينم Plectorhinchus sordidus (Sordid rubberlip)	5.18
22	Naiser نيسر Lutjanus fulviflamma (Dory snapper)	4.82
23	Fasker فسكر Acanthopagrus bifasciatus (Twobar seabream)	4.68
24	Shiem شم Arius thalassinus (Giant sea catfish)	4.57
25	Hamra حمره Lutjanus malabaricus (Malabar blood snapper)	4.43
26	Khathaag خثاق Sepia pharaonis (Cuttle fish)	3.53
27	Gargor جرجور Carcharhinus dussumieri (White cheek shark)	2.97
28	Hamaam حمام Carangoides gymnostethus (Bludger)	2.81
29	Sultan Ibrahim سلطان إبراهيم Parupeneus marga (Pearly goatfish)	2.64
30	Laden لدن Epinephelus polylepis (Smallscaled grouper)	2.46
31	N'aimia نعيمية Pinjalo pinjalo (Pinjalo)	2.18
32	Helali هلالي Plectorhinchus gaterinus (Blackspotted rubbe)	1.81
33	Gane قين Scarus ghobban (Blue-barred parrotfish)	1.78
34	Siken سكن Rachycentron canadum (Cobia)	1.31
35	Sh'aam شعم Acanthopagrus latus (Yellowfin seabream)	1.28
36	Biyah بياح Moolgarda seheli (Bluespot mullet)	1.22
37	Shinainuwa شنينوة Cephalopholis hemistiktos (Yellowfin hind)	0.99
38	Shaqra شقره Lutjanus argentimaculatus (Mangrove red snap)	0.95
39	Ebzeimi إبزيمي Scolopsis taeniata (Black-streaked monocle b)	0.86
40	Ywaf يواف Anodontostoma chacunda (Gizzard shad)	0.53
41	Baasi باسي Nemipterus bipunctatus (Threadfin bream)	0.48
42	Wahra وحرة Platycephalus indicus (Bartail flathead)	0.46
43	Subaity سبيطي Sparidentex hasta (Sobaity seabream)	0.44
44	Umm El Rubian أم الروبيان Thenus orientalis (Falt head locust lobster)	0.35
45	Safi sneifi صافي صنيفي Siganus luridus (Dusky spinefoot)	0.35
46	Bertamah برطامة Epinephelus multinotatus (White-blotched gro)	0.23
47	Emaad عماد Platax orbicularis (Orbicular batfish)	0.19
48	Hagool حاقول Tylosurus crocodilus crocodilu (Hound needle)	0.19
49	Anfooz عنفور Pomacanthus maculosus (Yellowbar angel fish)	0.17
50	Battana بطانة Crenidens crenidens (Monocle bream)	0.11
51	Zieb ذيب Terapon jarbua (Jarbua terapon)	0.08
52	Lisan لسان Brachirus orientalis (Oriental sole)	0.06
53	Umm El Laban أم اللبن Scolopsis taeniata (monocle bream)	0.04
54	Hef حف Chirocentrus dorab (Dorab wolf-herring)	0.01

55	Kasour كاسور Saurida tumbil (Greater lizard fish)	0.00

9.7 Compound trend of variables

In examining trends of basic fisheries variables such as catch, effort, values, etc. we often encounter the same question as in ranking: the individual trends are clear but there is no easy indication as to what is the general direction of the overall fisheries. For instance we may have a rising catch trend accompanied with stable values and declining effort. The question arises as to whether the general situation could be described by one trend line.

The approach used here resembles the one used for multi-variate ranking. First the data are normalized (using however a slightly different scheme) and then geometrical distances are calculated in order to plot a single curve.

9.8 Normalization of values

Let us assume a list of m elements representing periods: e_i , i=1…m. Such periods are usually months covering a year span, e.g. January 2012 – December 2015. In such a case m=4 years x 12 months = 48 periods.

We also assume that each period e_i is associated to an array of n values v_{ij}, j=1…n for n fisheries variables that are examined over the years and months.

Since the columns of the matrix v_{ij} have different value ranges (imagine for instance catch values in Kg, effort values in boat-gear days, etc.), the first step is to normalize all values of each variable across periods.

This time a normalization scheme is used by means of which each value v_{ij} is transformed to the ratio:

$$u_{ij} = \frac{v_{ij}}{\overline{V_j}} \qquad (9.8.1)$$

where $\overline{V_j}$ is the arithmetic mean of a variable j over all periods i:

$$\overline{V_j} = \frac{1}{m}\sum_{i=1}^{m} v_{ij} \qquad (9.8.2)$$

The resulting normalized values u_{ij} are dimensionless and share similar value scales. Next step involves the definition of a “low point” on the basis of which a single trend will be drawn. This n-dimensional point will be formed by selecting the minimum values of the variables of positive impact (e.g. catch) and the maximum ones for those with negative impact (such as effort). In other words each element b_j of point ($b_j, b_j, \ldots, b_n$) will be defined as:

$$b_j = \mathrm{Min}(v_{ij}) \text{ for variables of positive impact} \qquad (9.8.3)$$

$$b_j = \mathrm{Max}(v_{ij}) \text{ for variables of negative impact} \qquad (9.8.4)$$

9.9 Distances from low point

For each period i=1…m the following distance from low point is computed:

$$d_i = \sqrt{\sum_{j=1}^{n}(u_{ij} - b_j)^2} \qquad (9.9.1)$$

Next the average distance $\overline{D}$ is defined as:

$$\overline{D} = \frac{1}{m}\sum_{i=1}^{m} d_i \qquad (9.9.2)$$

and is used to normalize distances as follows:

$$d_i^* = \frac{d_i}{\overline{D}} \qquad (9.9.3)$$

The values of normalized distances constitute the single trend aimed by this method.

Figure 9.9.1 illustrates an example of multi-variate trend. In this example the variables to be combined are total catch, total value and total fishing effort. In formulating the low point described by expressions 9.8.3 and 9.8.4 the first expression was used for catch and values, considering that the impact of these two variables is positive for high values. Instead fishing effort is used as a direct index of fishing pressure and as an indirect index of running costs; this means that low effort values have a positive impact while high ones have a negative impact. Hence for effort expression 9.8.4 was used.

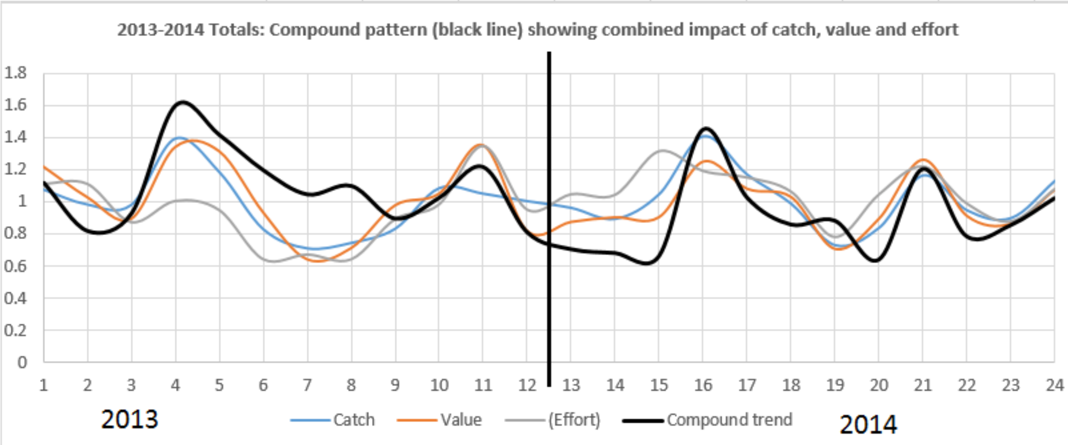

Figure 9.9.1 – Illustrating the single trend (black line) of total catch, total value and total effort for the period January 2013 – December 2014.

Chapter 10: Clusters of fishing patterns

10.1 Introduction

Clustering[1] of variables using multiple quantitative criteria is a commonly used statistical technique in fisheries applications. There is a wide variety of such applications ranging from "static" cases (e.g. clustering of boat-gears types according to constant characteristics) to "dynamic" ones where the criteria vary with time. Irrespective of the type of application, however, the theoretical concepts and approaches remain the same and these will be reviewed in the present chapter.

10.2 Elements, multiple criteria and normalization

Let us assume a list of m elements denoted by e_i, i=1…m. We also assume that each element e_i is associated to an array of n values v_{ij}, j=1…n, which constitute numerical criteria of similarity amongst the elements. In this manner the elements e_i with coordinates v_{ij} represent m points in the n-dimensional space.
Since the columns (i.e. criteria values) of the matrix v_{ij} have different value ranges, the first step is to normalize these columns by mapping them to a common numerical scale. Generally a convenient scale is that of 0-1. Such a mapping is achieved by the following computations:

- Determining the minimum and maximum values $\min_j$ and $\max_j$ for each column j.
- All values v_{ij} will be mapped to normalized values u_{ij} using the formula:

$$u_{ij} = \frac{v_{ij} - \min_j}{\max_j - \min_j} \qquad (10.2.1)$$

- Expression 10.2.1 indicates that each column will have values between 0 and 1 inclusive.

[1] In this document the Wroclaw Taxonomy technique is used. Originally, the method was developed for biological analyses but it was found quite suitable for other applications using quantitative clustering criteria.

10.3 Distance matrix, cluster pairs and clustering

For every two elements e_k and e_l with $k \neq l$ we use their normalized coordinates to calculate their geometrical distance d_{kl}:

$$d_{kl} = \sqrt{\sum_{j=1}^{n} (u_{kj} - u_{lj})^2} \qquad (10.3.1)$$

This new matrix of distances has the following properties:

- It contains m(m-1) distances (since zeroes are not included).
- It is symmetrical, i.e. $d_{kl} = d_{lk}$.

Next we search for pairs of points e_k and e_l such that e_l is the closest to e_k and at the same time e_k is the closest to e_l. Such pairs of points are called *cluster pairs* because they constitute the cores around which other points will be clustered.

Assuming that N such pairs have been identified, then the method will result in N clusters, each of which will be formed as follows.

- The mid point of each cluster pair (e_k, e_l) is formed with coordinates:

  $\frac{u_{kj} + u_{lj}}{2}$, j=1,2, … , n
- Each point e_i that is not a cluster pair, will be clustered around the pair for which its mid point is the closest to e_i.

In the specific case of fishing patterns along periods, the method also calculates degrees of resemblance between points in percentage form (refer to Tables 10.4 and 10.5). These indexes result from the proportion of months over 12 months that the normalized values of two points stat in the same value range of either 0-0.25, 0.26-0.5, 0.51-0.75, 0.76-1.

Numerical Examples

10.4 Clustering of species CPUE's (2014 data).

	Clustered elements	Element with closest resemblance	Association strength (%)
Cluster 1	Baasi باسي Nemipterus bipunctatus (Threadfin bream)	Kasour كاسور Saurida tumbil (Greater lizard fish)	**91.7**
	Kasour كاسور Saurida tumbil (Greater lizard fish)		**91.7**
Cluster 2	Biyah بياح Moolgarda seheli (Bluespot mullet)	Wahra وحرة Platycephalus indicus (Bartail flathead)	**100.0**
	Wahra وحرة Platycephalus indicus (Bartail flathead)		**100.0**
Cluster 3	Dal'ah ضلعه Scomberoides commersonnianus (Queenfish)	Naiser نيسر Lutjanus fulviflamma (Dory snapper)	**83.3**
	Naiser نيسر Lutjanus fulviflamma (Dory snapper)		**83.3**
	Ebzeimi إبزيمي Scolopsis taeniata (Black-streaked monocle b)		**66.7**
	Karari كراري Atule mate (Yellowtail scad)		**58.3**
	Ywaf يواف Anodontostoma chacunda (Gizzard shad)		**41.7**
	Zubaidi زبيدي Carangoides malabaricus (Malabar trevally)		**75.0**
Cluster 4	Gane قين Scarus ghobban (Blue-barred parrotfish)	Jid جد Sphyraena flavicauda (Yellowtail barracuda)	**91.7**
	Jid جد Sphyraena flavicauda (Yellowtail barracuda)		**91.7**
	Hef حف Chirocentrus dorab (Dorab wolf-herring)		**75.0**
	Jash جش Carangoides bajad (Orangespotted trevally)		**66.7**
	Sooli سولي Lethrinus microdon (Smalltooth emperor)		**75.0**
Cluster 5	Gargor جرجور Carcharhinus dussumieri (White cheek shark)	Safi sneifi صافي صنيفي Siganus luridus (Dusky spinefoot)	**91.7**
	Safi sneifi صافي صنيفي Siganus luridus (Dusky spinefoot)		**91.7**
Cluster 6	Hagool حاقول Tylosurus crocodilus crocodilu (Hound needle)	Umm El Laban أم اللبن Scolopsis taeniata (monocle bream)	**91.7**
	Umm El Laban أم اللبن Scolopsis taeniata (monocle bream)		**91.7**
	Battana بطانة Crenidens crenidens (Monocle bream)		**66.7**
	Rebeeb ربيب Gnathanodon speciosus (Gold toothless treval)		**66.7**
Cluster 7	Hamour هامور Epinephelus coioides (Orange-spotted grouper)	Khathaag خثاق Sepia pharaonis (Cuttle fish)	**100.0**
	Khathaag خثاق Sepia pharaonis (Cuttle fish)		**100.0**
	Bedha بدحة Gerres longirostris (silver-biddy)		**66.7**
	Fasker فسكر Acanthopagrus bifasciatus (Twobar seabream)		**91.7**
	Helali هلالي Plectorhinchus gaterinus (Blackspotted rubbe)		**66.7**

	Sh'aam شعم Acanthopagrus latus (Yellowfin seabream)		75.0
	Shiem شم Arius thalassinus (Giant sea catfish)		83.3
Cluster 8	Hamra حمره Lutjanus malabaricus (Malabar blood snapper)	N'aimia نعيمية Pinjalo pinjalo (Pinjalo)	91.7
	N'aimia نعيمية Pinjalo pinjalo (Pinjalo)		91.7
	Anfooz عنفوز Pomacanthus maculosus (Yellowbar angel fish)		66.7
	Emaad عماد Platax orbicularis (Orbicular batfish)		75.0
	Semaan سمان Epinephelus bleekeri (Duskytail grouper)		75.0
Cluster 9	Kobkob قبقب Portunus pelagicus (Blue swimmimg crab)	Siken سكن Rachycentron canadum (Cobia)	66.7
	Siken سكن Rachycentron canadum (Cobia)		66.7
	Hamaam حمام Carangoides gymnostethus (Bludger)		58.3
Cluster 10	Koffar كوفر Argyrops spinifer (King soldier bream)	Yanam ينم Plectorhinchus sordidus (Sordid rubberlip)	75.0
	Yanam ينم Plectorhinchus sordidus (Sordid rubberlip)		75.0
	Kanaad كنعد Scomberomorus commerson (Spanish mackerel)		75.0
	Zieb ذيب Terapon jarbua (Jarbua terapon)		75.0
Cluster 11	Laden لدن Epinephelus polylepis (Smallscaled grouper)	OTHER أنواع أخرى (Miscellaneous)	100.0
	OTHER أنواع أخرى (Miscellaneous)		100.0
	Boukshina بوقشينة Lethrinus lentjan (Pink ear emperor)		66.7
Cluster 12	Lisan لسان Brachirus orientalis (Oriental sole)	Subaity سبيطي Sparidentex hasta (Sobaity seabream)	100.0
	Subaity سبيطي Sparidentex hasta (Sobaity seabream)		100.0
	Tabaan تبان Euthynnus affinis (Kawakawa)		83.3
Cluster 13	Qurqufan قرقفان Rhabdosargus haffara (Haffara seabream)	TOTALS	66.7
	TOTALS		66.7
	Farsh فرش Diagramma pictum (Painted sweetlips)		41.7
	Sh'ari شعري Lethrinus nebulosus (Sprankled emperor)		50.0
Cluster 14	Saafi صافي Siganus canaliculatus (Whit-spott spinefoot)	Sultan Ibrahim سلطان إبراهيم Parupeneus marga (Pearly goatfish)	75.0
	Sultan Ibrahim سلطان إبراهيم Parupeneus marga (Pearly goatfish)		75.0
	Shaqra شقره Lutjanus argentimaculatus (Mangrove red snap)		50.0
	Shinainuwa شنينوة Cephalopholis hemistiktos (Yellowfin hind)		41.7
Cluster 15	Saal صال Carangoides chrysophrys (Longnose trevally)	Umm El Rubian أم الروبيان Thenus orientalis (Falt head locust lobster)	91.7

Umm El Rubian أم الروبيان Thenus orientalis (Falt head locust lobster)	**91.7**
Bertamah برطامة Epinephelus multinotatus (White-blotched gro)	**83.3**

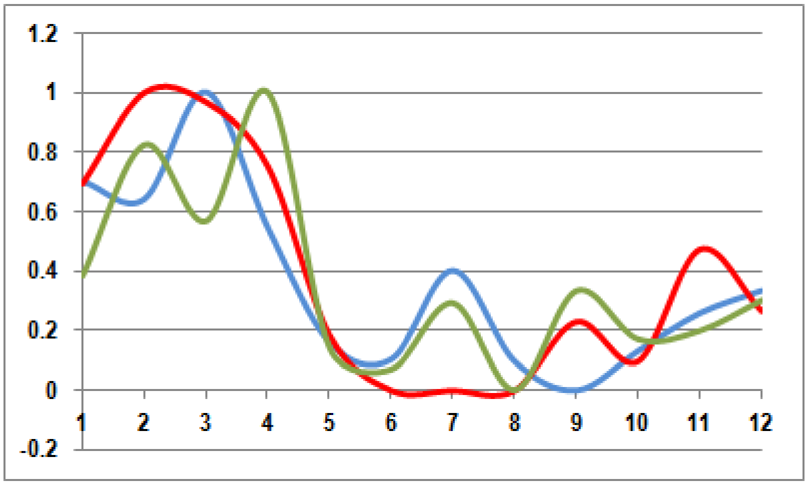

Figure 10.4.1 Graphical example of the resemblance pattern of Cluster 7, depicting Orange-spotted grouper (blue line), Sepia pharaonis (red) and Twobar seabream (green).

10.5 Clustering of species prices (2014 data)

	Clustered elements	Element with closest resemblance	Association strength (%)
Cluster 1	Anfooz عنفوز Pomacanthus maculosus (Yellowbar angel fish)	Wahra وحرة Platycephalus indicus (Bartail flathead)	**100.0**
	Wahra وحرة Platycephalus indicus (Bartail flathead)		**100.0**
	Ebzeimi إبزيمي Scolopsis taeniata (Black-streaked monocle b)		**91.7**
	Emaad عماد Platax orbicularis (Orbicular batfish)		**75.0**
Cluster 2	Battana بطانة Crenidens crenidens (Monocle bream)	Khathaag خثاق Sepia pharaonis (Cuttle fish)	**83.3**
	Khathaag خثاق Sepia pharaonis (Cuttle fish)		**83.3**
Cluster 3	Bedha بدحة Gerres longirostris (silver-biddy)	Farsh فرش Diagramma pictum (Painted sweetlips)	**66.7**
	Farsh فرش Diagramma pictum (Painted sweetlips)		**66.7**
	Gane قين Scarus ghobban (Blue-barred parrotfish)		**58.3**
	Jash جش Carangoides bajad (Orangespotted trevally)		**41.7**
	Kobkob قبقب Portunus pelagicus (Blue swimmimg crab)		**58.3**
Cluster 4	Boukshina بوقشينة Lethrinus lentjan (Pink ear emperor)	N'aimia نعيمية Pinjalo pinjalo (Pinjalo)	**100.0**
	N'aimia نعيمية Pinjalo pinjalo (Pinjalo)		**100.0**
	Bertamah برطامة Epinephelus multinotatus (White-blotched gro)		**91.7**
	Hamaam حمام Carangoides gymnostethus (Bludger)		**75.0**
	Lisan لسان Brachirus orientalis (Oriental sole)		**66.7**
	Saal صال Carangoides chrysophrys (Longnose trevally)		**58.3**
	Semaan سمان Epinephelus bleekeri (Duskytail grouper)		**100.0**
Cluster 5	Fasker فسكر Acanthopagrus bifasciatus (Twobar seabream)	Koffar كوفر Argyrops spinifer (King soldier bream)	**91.7**
	Koffar كوفر Argyrops spinifer (King soldier bream)		**91.7**
	Hamour هامور Epinephelus coioides (Orange-spotted grouper)		**66.7**
	Hef حف Chirocentrus dorab (Dorab wolf-herring)		**83.3**
	Kasour كاسور Saurida tumbil (Greater lizard fish)		**75.0**
	Qurqufan قرقفان Rhabdosargus haffara (Haffara seabream)		**83.3**
	Sh'aam شعم Acanthopagrus latus (Yellowfin seabream)		**66.7**
	Sh'ari شعري Lethrinus nebulosus (Sprankled emperor)		**75.0**
	TOTALS		**66.7**
Cluster 6	Gargor جرجور Carcharhinus dussumieri (White cheek shark)	Kanaad كنعد Scomberomorus commerson (Spanish mackerel)	**83.3**
	Kanaad كنعد Scomberomorus commerson (Spanish mackerel)		**83.3**
	Safi sneifi صافي صنيفي Siganus luridus (Dusky spinefoot)		**66.7**

	Shaqra شقره Lutjanus argentimaculatus (Mangrove red snap)		**66.7**
	Siken سكن Rachycentron canadum (Cobia)		**83.3**
Cluster 7	Hagool حاقول Tylosurus crocodilus crocodilu (Hound needle)	Jid جد Sphyraena flavicauda (Yellowtail barracuda)	**83.3**
	Jid جد Sphyraena flavicauda (Yellowtail barracuda)		**83.3**
	Baasi باسي Nemipterus bipunctatus (Threadfin bream)		**58.3**
	Biyah بياح Moolgarda seheli (Bluespot mullet)		**83.3**
	Dal'ah ضلعه Scomberoides commersonnianus (Queenfish)		**41.7**
	Helali هلالي Plectorhinchus gaterinus (Blackspotted rubbe)		**50.0**
	Umm El Rubian أم الروبيان Thenus orientalis (Falt head locust lobster)		**50.0**
	Zubaidi زبيدي Carangoides malabaricus (Malabar trevally)		**50.0**
Cluster 8	Hamra حمره Lutjanus malabaricus (Malabar blood snapper)	Yanam ينم Plectorhinchus sordidus (Sordid rubberlip)	**75.0**
	Yanam ينم Plectorhinchus sordidus (Sordid rubberlip)		**75.0**
	Subaity سبيطي Sparidentex hasta (Sobaity seabream)		**66.7**
	OTHER أنواع أخرى (Miscellaneous)		**41.7**
Cluster 9	Laden لدن Epinephelus polylepis (Smallscaled grouper)	Sooli سولي Lethrinus microdon (Smalltooth emperor)	**100.0**
	Sooli سولي Lethrinus microdon (Smalltooth emperor)		**100.0**
	Shinainuwa شنينوة Cephalopholis hemistiktos (Yellowfin hind)		**91.7**
	Tabaan تبان Euthynnus affinis (Kawakawa)		**66.7**
	Ywaf يواف Anodontostoma chacunda (Gizzard shad)		**50.0**
Cluster 10	Rebeeb ربيب Gnathanodon speciosus (Gold toothless treval)	Sultan Ibrahim سلطان إبراهيم Parupeneus marga (Pearly goatfish)	**91.7**
	Sultan Ibrahim سلطان إبراهيم Parupeneus marga (Pearly goatfish)		**91.7**
	Karari كراري Atule mate (Yellowtail scad)		**83.3**
	Naiser نيسر Lutjanus fulviflamma (Dory snapper)		**58.3**
	Saafi صافي Siganus canaliculatus (Whit-spott spinefoot)		**75.0**
	Shiem شم Arius thalassinus (Giant sea catfish)		**83.3**
	Umm El Laban أم اللبن Scolopsis taeniata (monocle bream)		**91.7**
	Zieb ذيب Terapon jarbua (Jarbua terapon)		**83.3**

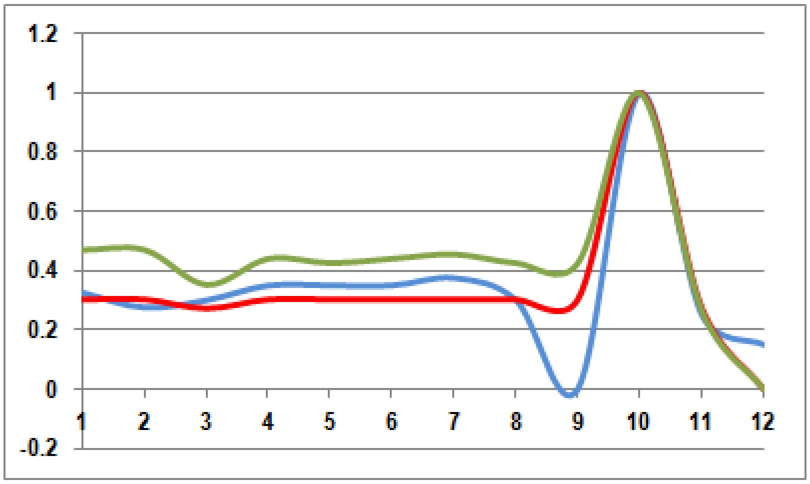

Figure 10.5.1 Graphical example of the resemblance pattern of Cluster 4, depicting Pink ear emperor (blue line), Pinjalo (red) and Duskytail grouper (green).