Calculation of Calibration Factors from the Comparative Fishing Trial between FRS *Africana* and RV *Dr Fridtjof Nansen*

LUYANDA LENNOX ANTONY

Submitted in fulfilment of the requirements for the

Degree of Magister Scientiae

in the

Faculty of Science

Department of Statistical Sciences

at the

University of Cape Town

February 2008

Supervisors: Prof T. Dunne (University of Cape Town, South Africa)

Dr RW. Leslie (Department of Environment and Tourism, South Africa)

DECLARATION

"I, LUYANDA LENNOX ANTONY, know the meaning of plagiarism and declare that all of the work in the thesis, save for that which is properly acknowledged, is my own".

Luyanda Lennox Antony

ACKNOWLEDGMENTS

I would like to send my gratitude to the people mentioned below for all their advice, support and assistance in the completion of this study. Firstly, to my supervisors, Prof Tim Dunne of UCT and Dr Rob Leslie, thank you for your patience, good guidance and constructive feedback which you have given me from the beginning to the completion of this dissertation. Thank you Prof Tim Dunne for making yourself available to supervise this project when we approached you. Thank you.

To Marine and Coastal Management, especially Research and Development, thank you for assisting me in collection of the data at sea and its extraction of the database. Thank you for providing funds to me so this project can be conducted. Thank you for being good friends and for all your encouraging words during the difficult times of this project.

To my wife and daughter, mother, father and sister and brothers, I thank you for your patient understanding, love and faith through all the hard times, especially my sister, who has always pushed me even when I thought of giving up the project. Thank you.

To my family and friends, I thank you for all the support you have given me during this study.

SUMMARY

Between 2001 and 2004 four sets of intercalibration trials were conducted by South Africa's Marine and Coastal Management between South African fisheries research ship (FRS) *Africana*, using the old and new trawl gear, and the Norwegian research vessel (RV) *Dr Fridtjof Nansen*. The aim of the experiment was to calculate conversion factors to convert RV *Dr Fridtjof Nansen's* catch as if it were FRS *Africana*'s catch, and to convert the historical catches of FRS *Africana* (old trawl gear) into FRS *Africana* (new trawl gear) catches.

Conversion factors and their confidence intervals were calculated using various analytical methods for the two important South African hake species: *Merluccius capensis* (Shallow-Water Cape Hake) and *Merluccius paradoxus* (Deep-Water Cape Hake), and also the following important demersal species *Genypterus capensis* (Kingklip), *Cynoglossus zanzibarensis* (Red-spotted tonguefish), *Lophius vomerinus* (Monkfish), *Trachurus capensis* (Horse Mackerel), *Chelidonichthys capensis* (Cape Gurnard), *Chelidonichthys queketti* (Lesser Gurnard), *Zeus capensis* (Cape Dory), and *Helicolenus dactylopterus* (Jacopever).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES

LIST OF TABLES

LIST OF BOXES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1. Background Information

The provision of biological advice for the management of marine demersal resources requires estimates of the current abundance of exploited age-groups and the size of recruiting year-classes (Doubleday 1981). Some information on the age structure of the resource and its current status can be obtained by methods such as cohort analysis of estimates of the age composition of commercial catches, especially if accurate calibration of terminal fishing mortality using catch rates is possible. However, data on commercial fishing frequently have shortcomings in terms of accuracy and usefulness as an indicator of stock abundance and seldom provide useful indices of the sizes of recruiting year-classes (Doubleday 1981).

In finding solutions to these difficulties, scientists have turned increasingly to the use of research vessel survey indices of abundance (Doubleday 1981). These indices of abundance have the advantage of consistent methodology from year to year, thereby avoiding the problems of effort creep inherent in indices derived from commercial fisheries. According to Newby *et al.* 2004, effort creep is the term applied to the continual increase in catching power that occurs in fisheries as a result of technical innovation or the uptake of unregulated fishing inputs (allocative efficiency) and improvements in technical efficiency.

In addition, 'surveys are better able to forecast recruitment through the deployment of smaller-meshed nets than are permitted in the commercial fisheries. The accumulation of extended data series from surveys which can be intercalibrated with estimates of abundance from cohort analyses has increased

confidence in abundance estimates obtained from such surveys' (Doubleday, 1981).

Demersal trawl surveys have therefore assumed a key role in the provision of scientific advice for fishery management (Doubleday 1981). These surveys also generate valuable data on distribution of demersal species and on biological parameters such as growth rates, feeding behaviour, and incidence of parasites (Doubleday 1981).

Bottom trawl surveys are widely used for monitoring demersal stocks when only an index of abundance is required (Vazquez 2002, von Szalay and Brown 2001, Warren 1996, Walsh and Orr 1998). These surveys provide biomass and annual yield estimates of major commercial species (and non-commercial by-catch species for which few or no data are available from the commercial fishery) and monitor changes in abundance over time (von Szalay and Brown 2001). The estimation of total biomass from catch per unit effort (CPUE), however, involves several crucial assumptions (such as CPUE is proportional to the total biomass, constant catchability, and that such biological parameters as natural mortality coefficient and growth coefficient are small in magnitude), rendering such estimates rather imprecise.

Scientific demersal surveys are an important source of information for estimating the abundance of fisheries populations in fisheries science (Pelletier 1998). Many fisheries stock assessments depend mainly on the time series of abundance indices obtained from annual fisheries surveys (Pelletier 1998). Pelletier (1998) reports that 'the temporal continuity of such time series may be compromised by a change in survey vessel or in fishing equipment'. Catch resulting from any fishing operation can be considered dependent on three factors: (i) vessel and fishing characteristics, e.g. fishing gear, technological equipment, crew, etc; (ii) characteristics of sampled populations, e.g. abundance level, spatial distribution,

availability to fishing; and (iii) environmental conditions, e.g. weather, hydrology, depth, and substrate type.

Pelletier (1998) reports some of the recommendations which were given in a workshop held in Canada in 1980. The scientists who gathered at this workshop noticed that proportion of catches differed from vessel to vessel both quantitatively and qualitatively, even though the vessels fished in the same area. The workshop recommended that (i) pairwise / parallel trawling should be incorporated in the allocation of vessels by area for multi-vessels surveys, permitting comparison of fishing power as data sets are accumulated over the years; (ii) whenever possible, the same vessel should be retained over time (Pelletier 1998).

Before old vessels are replaced by new ones, calibration studies should be undertaken. The need for intercalibration experiments between research vessels was thus acknowledged in multi-vessel survey programs and when survey vessels or trawl gears are replaced over time.

In the 1980s some intercalibration studies were undertaken in the southeast Atlantic under the auspices of the International Commission for Southeast Atlantic Fisheries (ICSEAF), but they were all conducted in Namibian waters (then known as South West Africa). At that stage surveys were conducted off South Africa using the FRS *Africana*, off southern Namibia (between the Cunene and Walvis Bay) by the Spanish using the *Chicha Touza* (The Chicha was a commercial vessel chartered by the Spanish research organisation in Barcelona), and off northern Namibia by the Russians using 2 contracted commercial vessels. Comparative trawls were conducted between the FRS *Africana* and *Chicha Touza*, and between *Chicha Touza* and the Russian vessels. More recently the work between the FRS *Africana* and the RV *Dr Fridtjof Nansen* was conducted and this work is reported in this thesis.

1.2. Vessels and Trawl Gears

Two research vessels, South African FRS *Africana* (Figure 1.1) and the Norwegian RV *Dr Fridtjof Nansen* (Figure 1.2), have been used during intercalibration experiments in South African waters between 2001 and 2004. During the intercalibration trials in 2001 and 2002, FRS *Africana* was fitted with the "Old" 2-pannel 180 German trawl gear, and was then fitted with "New" high lift 4-pannel 180 German trawl gear during the 2004 intercalibration trials. RV *Dr Fridtjof Nansen* was using Reketral "Gisund Supper" trawl gear during both periods of the intercalibration trials.

1.2.1. Research Vessels

FRS *Africana*

Figure 1.1: FRS *Africana*

FRS *Africana* is the flagship of the Marine and Coastal Management fleet of oceanographic research ships. Her main role is a platform for research and monitoring undertaken to inform and guide the management of South Africa's offshore fisheries. The ship, built in Durban and commissioned in 1982, has a

raft-mounted machinery system enclosed in an acoustic hood to reduce noise interference with her extensive suite of acoustic equipment (unpublished FRS Africana survey report, 2000). There are eight laboratories providing comprehensive facilities for various research disciplines and three containerized laboratories for specialized duties. The vessel is based in Cape Town, South Africa, and is managed for the department by Smit Pentow Marine. Below are the vessel particulars (unpublished FRS Africana survey report, 2000):

Type:	Steel Hulled Fisheries Research
Ship Class:	Lloyds Register 100 A1 +
	Ice Class II Keel Laid: 1979 (Commissioned in March 1982)
Builders:	Dorbyl Marine, Durban
Gross Tonnage:	2471.11 tons
Nett Tonnage:	617.78 tons
Displacement:	3337 tons
Length:	77.85m
Breath:	15.25m
Moulded Draft:	5.7m
Horsepower:	1790 kW (2400 BHP)
Cruising Speed:	12 knots
Minimum Speed:	1 knot
Range:	20000 nautical miles
Endurance:	45 days
Complement:	52 (33 crew and 19 scientific/other staff)

RV *Dr Fridtjof Nansen*

Figure 1.2: RV *Dr Fridtjof Nansen*

1.2.2. Trawl Gears

There are conflicting requirements for the trawl gear used for groundfish surveys. It is desirable to use similar gear to that used by the commercial trawl industry at the start of the survey series, but large commercial-size catches are undesirable for surveys. The potential size of the mean catch per station can be reduced by using smaller trawl gear, or by shortening the trawl duration.

However, reducing the trawl gear size or trawl duration introduces errors. As the stern of the vessel rises and falls on the swells (stern surge), the net is jerked. The magnitude of this stern surge is dependent on the size of the vessel. The effect of the stern surge is reduced by the drag (inertia) of the net. The larger the net, the greater the inertia, which reduces the stern surge, but makes it more difficult for the vessel to tow the net. Therefore the size of the net should be matched to the size of the vessel to obtain maximum efficiency. Thus reducing the size of the net in order to reduce the average size of the catch for research

purposes can result in a mismatch between the size of the vessel and the net which could have severe effect on the fishing efficiency of the net through factors such as stern surge.

Trawl duration is taken as the time that the net is in contact with the seabed and fishing efficiently. However, it is difficult to precisely determine the start and end times for the trawls. Estimation of the precise time that the net settled on the seabed (the "start" of the trawl) depends on the experience and expertise of the Fishing Master. Also, the end of the trawl is commonly taken as the time that hauling commences, however, the trawl does not leave the seabed immediately as the slack in the trawl warps must first be taken up, and this delay increases with the depth of the water.

The error in estimating the precise start and end of the trawl is independent of the duration of the trawl, therefore, the relative magnitude of error introduced increases as total trawl duration decreases, e.g. if one can determine the start time to an accuracy of say 5 minutes, then the relative error for a 3 hour tow is only $\pm 3\%$ (of 175-185 minutes), whereas it is $\pm 50\%$ (of 5-15 minutes) for a 10 minute tow. In addition, as the trawl net is permanently open, there is some catch taken in the water column both before and after this period.

Consequently, in groundfish surveys around the world it is general practice to use nets not much smaller than the ideal size for the power of the research vessel (to limit the effect of stern surge) and to limit tow duration to 30 minutes (amounting to a tolerance of $\pm 17\%$, i.e. assuming that start time can be determined to an accuracy of say 5 minutes as in previous paragraph).

The recent development of sophisticated net monitoring equipment has enabled the start and end of the effective time (the period that the net is in contact with the seabed) to be accurately measured thereby reducing the error introduced by

shortening the trawl duration. However, it is still desirable to use trawl duration in excess of 20 minutes.

1.2.2.1. FRS *Africana* Trawl Gear

When the FRS *Africana* was commissioned in 1982, the trawl nets used by the older, smaller RV *Africana II* were simply transferred to her. However, the FRS *Africana* was fitted with 32mm main trawl wires, and therefore needed larger, heavier trawl doors (otter boards) than the RV *Africana II*.

In the mid 1990s the FRS *Africana* was fitted with SCANMAR net monitoring equipment. This monitoring showed that the trawl net was over-spread i.e. the large trawl doors pulled the wings further apart than the ideal net geometry. This over-spreading of the trawl net in turn pulled the headline down so that the vertical mouth opening was only 2m (instead of the expected 4m). In addition, the light footrope (a rope-wrapped chain) was often lifted clear of the sea bed. It was immediately obvious that the trawl system would have to be replaced with a system that was balanced to the size and power of the FRS *Africana* to address these problems.

But any change to the trawl gear would break the time series, therefore it was important that all changes be done at the same time (so that there was only a single break in the time series) and that the new gear be calibrated against the old gear. A decision was taken to fit the FRS *Africana* with lighter, smaller diameter trawl warps (28mm diameter), thus enabling the use of smaller trawl doors. It was therefore decided that the new gear should be introduced when the 32mm trawl warps (then in use) were replaced.

A standard method to calibrate between two sets of trawl gears on the same vessel is to trawl with one net, switch to the other net and then to trawl parallel to, and in close proximately to, the previous trawl track. In this way a large number

of "trawl pairs" can be generated to enable a comparison of the fishing efficiency of the two trawl nets. However, it would be impossible to switch between the trawl warps in this manner; therefore it was decided to calibrate the old trawl gear against a trawl gear in a second vessel.

Once the new trawl gear was fitted to the first vessel, the new trawl gear could also be calibrated against the second vessel, thereby using the second vessel as a "bridge" between the two time-series of biomass estimates. The vessel selected for this purpose was the Norwegian research vessel RV *Dr Fridtjof Nansen*, as she had been used to conduct trawl surveys off Namibia, and the calculated calibration factors could then also be used to standardize surveys off South Africa and Namibia.

For the ideal experimental design the "bridge" vessel should use the same net selected as the new configuration. Thus the first phase would compare FRS *Africana* using the old gear with RV *Dr Fridtjof Nansen* using the new gear thus yielding a calibration factor comprising a gear effect and a vessel effect. For the second phase both vessels would be using the same gear for which the calibration factor would be a vessel effect only. However, the compromise to simultaneously bridge the old and new time series for the FRS *Africana* and determine factors to standardize between surveys conducted off South Africa and Namibia meant that the RV *Dr Fridtjof Nansen* had to use her standard trawl gear. Thus both phases of the actual experimental design have vessel and gear effects

Old "180 German" trawl gear on FRS Africana

The old trawl gear (Figure 1.3) on FRS *Africana* is equipped with 32mm diameter trawl warps, with a 50tons breaking strain. Trawl doors are of type $7m^2$ WV-doors, weighing 1500kg. This trawl gear has a headline of 35.86m long, footrope 57m long, and a 75mm mesh size in the codend with 35mm mesh liner. The

sweeps were 50m. The estimated door spread is 121m and no strappings are used.

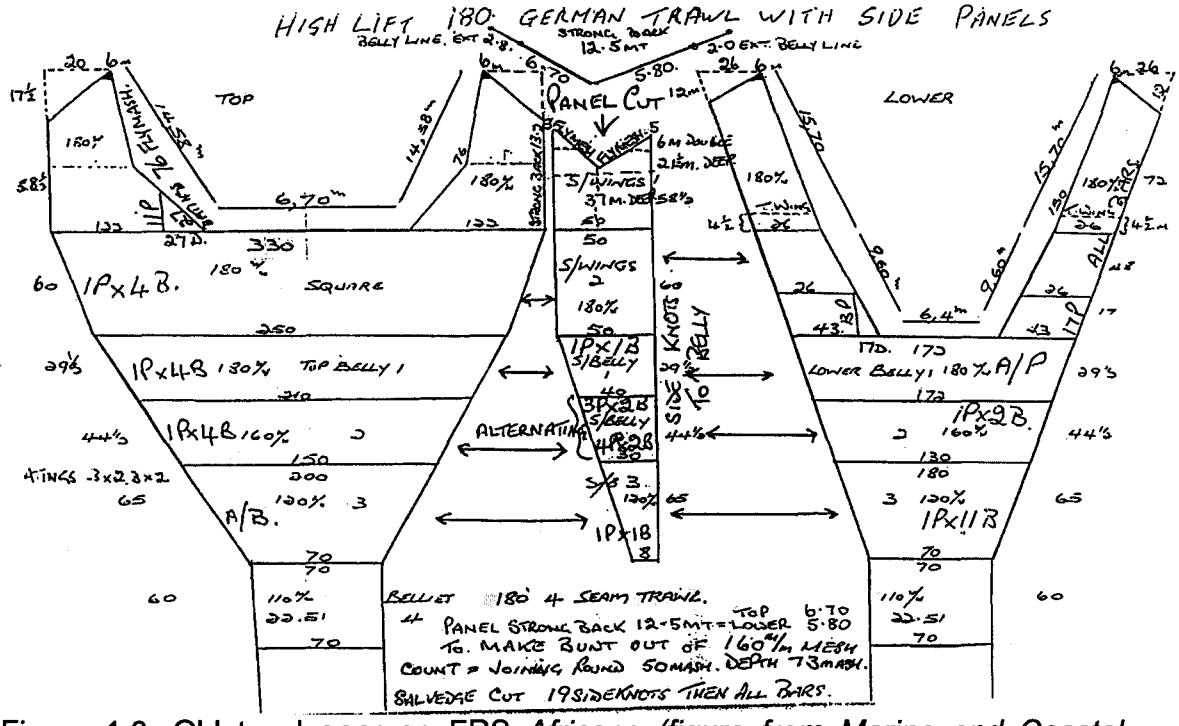


Figure 1.3: Old trawl gear on FRS *Africana (figure from Marine and Coastal Management cruise reports)*. Note that the side panels were not used.

New "High Lift 180 German" trawl gears on FRS Africana

The new trawl gear (Figure 1.4) on FRS *Africana* is equipped with 28mm diameter Bridon steel core trawl warps, with a 50tons breaking strain. Trawl doors are of type 5.5m^2 multi-purpose Morgere doors, weighing 1500kg. This trawl gear has a headline of 36.30m long with 330kg of lifts; and a footrope 54.7m long. The footrope has a modified rockhopper and hoppers of 8" diameter and spacers of 6" diameter threaded on a chain. Therefore the hoppers are only marginally bigger than the rest of the footrope, thereby preventing escapement under the footrope. The sweeps were reduced to 9m to minimise herding. The door spread varies according to sea depth. It is estimated at 65m in shallow waters and 80m in deeper waters.

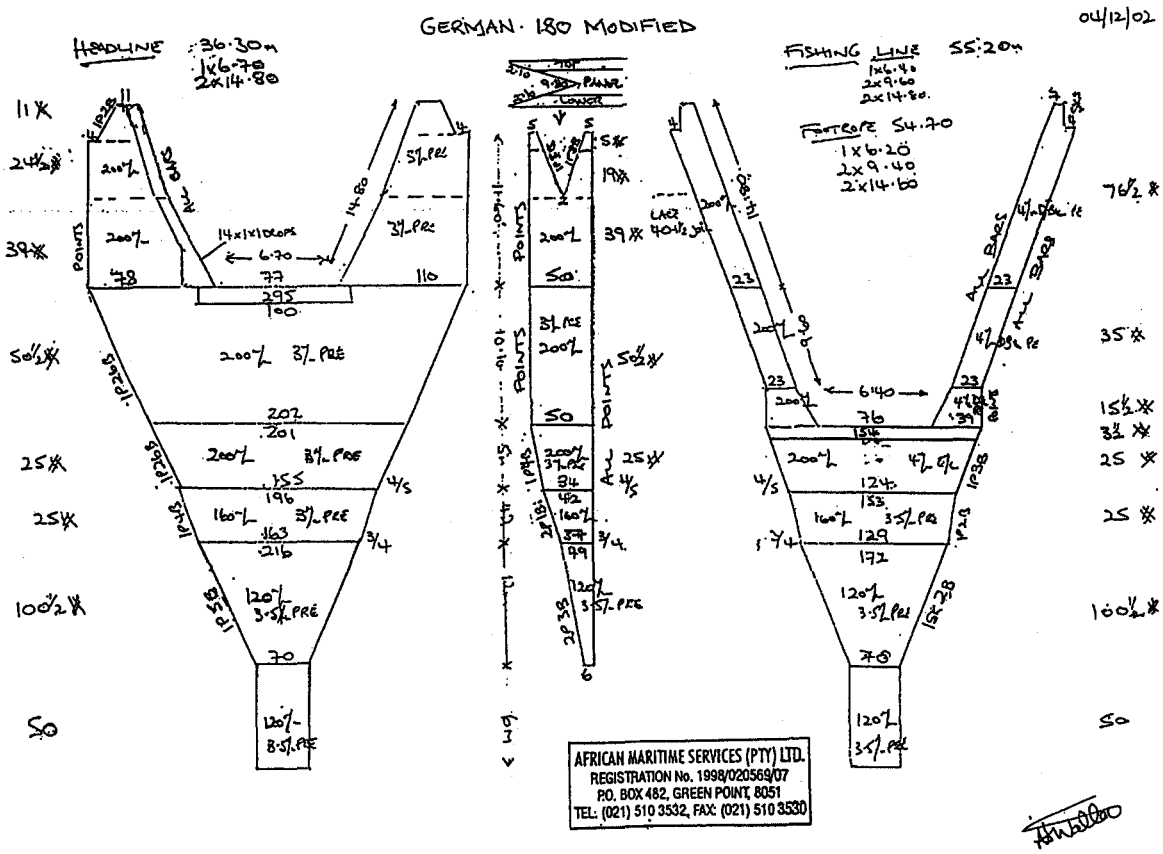


Figure 1.4: New trawl gear on FRS *Africana (figure from Marine and Coastal Management)*

1.2.2.2. RV *Dr Fridtjof Nansen* Trawl Gear

The trawl gear used by RV Dr *Fridtjof Nansen* is called Reketral "Gisund Supper" bottom trawl (Figure 1.5). This trawl gear has a headline of 31m, footrope 47m and a 20 mm mesh size in the codend with an inner net of 10mm mesh size. The estimated opening is 6m. A restraining rope is fitted between the trawl warps 130m ahead of the trawl doors to limit the door spread and maintain a constant distance between the doors. The distance between wings during towing is about 18m. The sweeps are 40m long.

The footrope is "Rockhopper-type" 12 inch rubber bobbins gear. The doors are of 'Thyboron' combi type, 7.81m^2, 1670kg. Their distance when trawling is about 45-55m in average, depending on the depth (least distance on the low depths). At depths greater than 300m the trawl gear was equipped with a tickler chain, which is suppose to improve the catchability of bottom living species.

The SCANMAR system is used on all trawl hauls. This equipment consists of sensors, a hydrophone, a receiver, a display unit and a battery charger. Communication between sensors and ship is based on acoustic transmission. The doors are fitted with sensors to provide information on their distance. A height sensor is fitted to the bottom trawl to measure the opening and provide information on clearance and bottom contact.

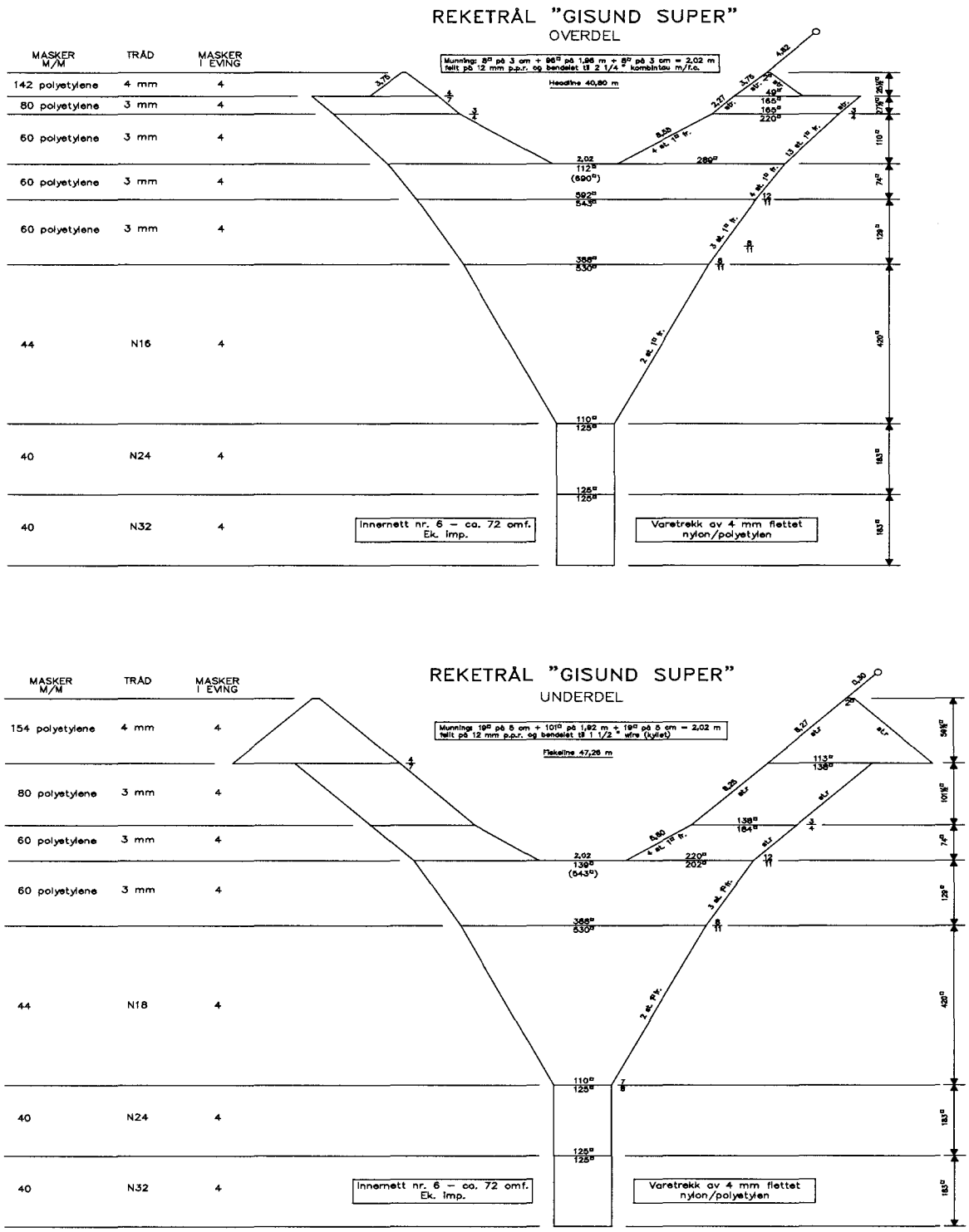


Figure 1.5: Trawl gear on RV *Dr Fridtjof Nansen*

1.3. Problem context and definition

The time-series of relative biomass indices collected by the FRS *Africana* includes both a winter (1986-1990) and a summer (1985-2002) series. Figure 1.6 presents the hake time series up to 2002. Due to technical problems with the ship, there were no direct biomass surveys by the FRS *Africana* in 1998, 2000 or 2001. The Norwegian survey vessel RV *Dr Fridtjof Nansen* conducted summer surveys for 2000 and 2001; however the relative biomass indices from the two vessels are not directly comparable due to differences in trawl gear and vessel characteristics. Further, the trawl configuration on FRS *Africana* was changed in May 2003, thereby breaking its time series of survey biomass indices. It is imperative that a calibration factor be determined to link future survey indices with the historic time series.

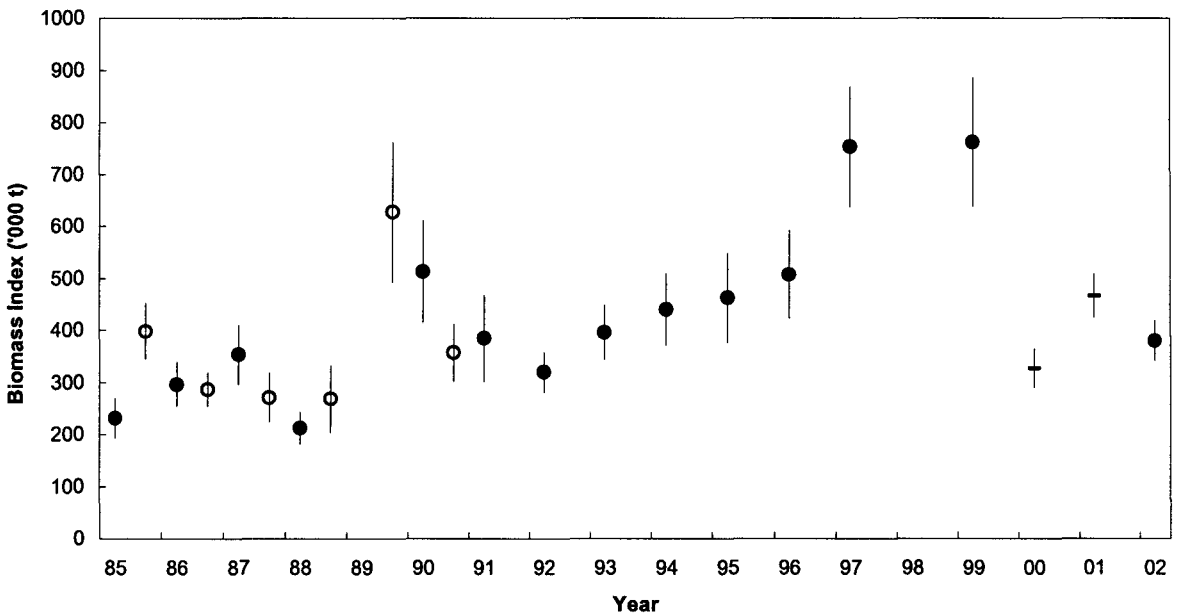


Figure 1.6: Direct survey biomass estimates for Cape hake (0-500m) from summer (solid circles) and winter (open circles) surveys of the West Coast by the FRS *Africana* and summer surveys by the RV *Dr Fridtjof Nansen* (horizontal bars). Error bars represent one standard error.

This study seeks to explore the problem and also aims to address the following questions:

1. To calculate conversion factors to relate biomass estimates determined by FRS *Africana* (with old trawl gear) with biomass estimates obtained by RV *Dr Fridtjof Nansen* in 2000 and 2001 for the following species:

    - *Merluccius capensis* (Shallow-Water Cape Hake)
    - *Merluccius paradoxus* (Deep-Water Cape Hake)
    - *Genypterus capensis* (Kingklip)
    - *Cynoglossus zanzibarensis* (Red-spotted tonguefish)
    - *Lophius vomerinus* (Monkfish)

    And where extent of data permits:

    - *Trachurus capensis* (Horse Mackerel)
    - *Chelidonichthys capensis* (Cape Gurnard)
    - *Chelidonichthys queketti* (Lesser Gurnard)
    - *Zeus capensis* (Cape Dory)
    - *Helicolenus dactylopterus* (Jacopever)

2. To calculate conversion factors to relate biomass estimates determined by FRS *Africana* (with new trawl gear) since 2003 with biomass estimates obtained by RV *Dr Fridtjof Nansen* for the suite of species listed in 1.

3. To combine the old gear and new gear conversion factors to relate FRS *Africana* (with old trawl gear) to FRS *Africana* (with new trawl gear) for the relevant suite of species.

4. To determine whether or not the conversion factors are independent of:
    - Size of the catch
    - Mean length of fish in catch (hake only)

- Fishing depth

5. To determine whether or not the calibration factors are independent of various environmental parameters such as water temperature, oxygen content, etc. This aspect of the study is of lower priority and is dependent on the suitability of the available data and time constraints.

1.4. Species description

The study attempts to calculate conversion factors for 10 demersal species that are found in the South African waters. These 10 fish species are the most important demersal fish species in the South African fish industry and are the targeted fish species during the demersal cruises conducted by the Marine and Coastal Management. This section introduces the 10 fish species and provides a brief background on each of them.

Merluccius capensis (Shallow-Water Cape Hake)

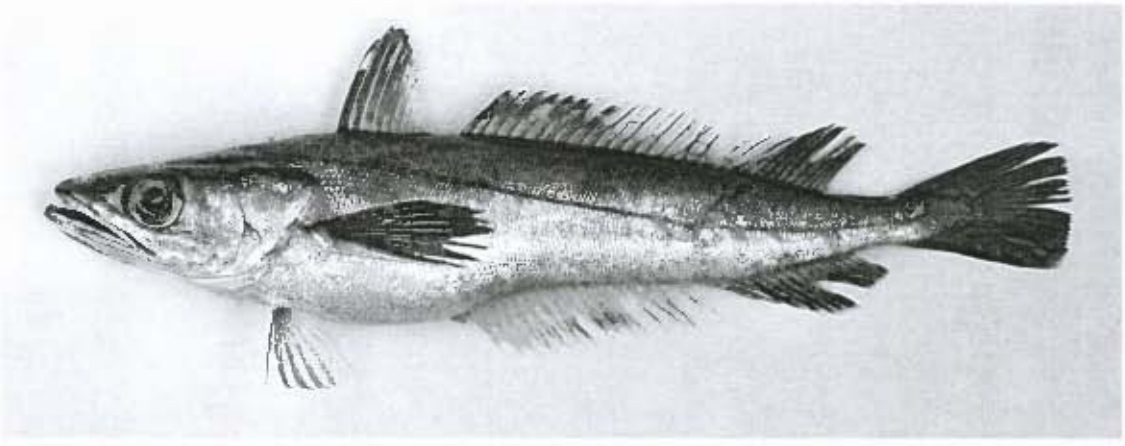


Figure 1.7: *Merluccius capensis*. [1]

Merluccius capensis is found mostly in the Southern Angola to KwaZulu Natal in South Africa, also on Valdivia Bank off Namibia. The fish are found near the

bottom during the day, and move into midwater at night. During the day, adults rest near the bottom on the continental shelf and upper slope, to 550m (Heemstra and Heemstra 2004). They are light brown dorsally, silvery to white ventrally, lack pigment spots on the gill tubercles and have 49-53 vertebrae (Smith and Heemstra 1986).

Females grow faster than males and mature at 45-60cm (Punt and Leslie, 1991). They grow to a maximum length of 140cm. Breeding occurs throughout the year, and peaks of reproductive activity occur mostly in August and September.

Juveniles feed on small crustaceans and small deep-sea fishes such as lantern fish, whereas larger individuals feed chiefly on small hake (mainly *M. paradoxus*) and jack mackerel. Cannibalism is also common in larger fish. Together with the deep-sea hake, they are abundant in trawl fishery, but catches have fluctuated greatly from year to year due to overfishing and vagaries of recruitment (Heemstra and Heemstra 2004).

***Merluccius paradoxus* (Deep-Water Cape Hake)**

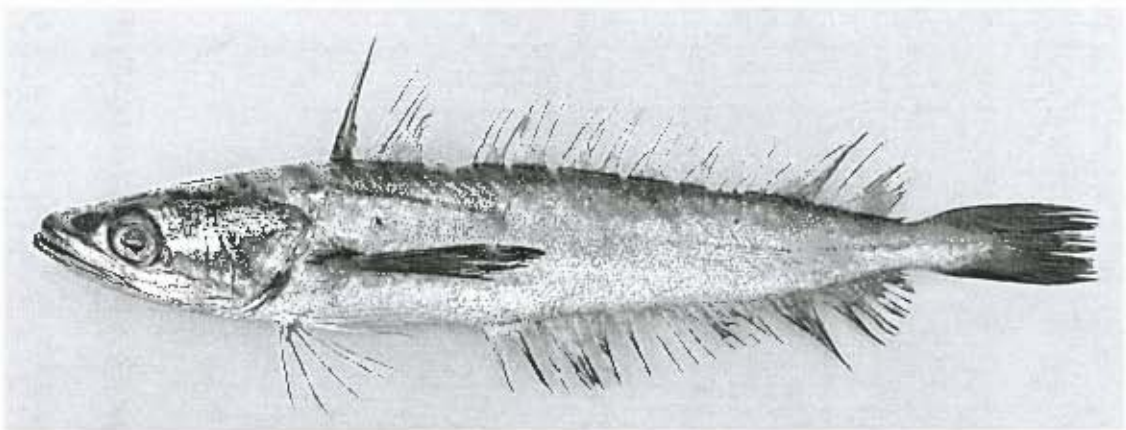


Figure 1.8: *Merluccius paradoxus.*

Merluccius paradoxus can be given a similar description to the *M. capensis* in terms of structure, they are also dark brown dorsally, silvery to white ventrally

(Smith and Heemstra 1986). They differ from *M. capensis* in possessing prominent pigment spots on the gill tubercles and in a higher vertebral count (54-58 vertebrae). An older fish can grow up to a maximum of 115 cm in length, which is slightly smaller than maximum length of *Merluccius capensis*. Distribution is Southeast Atlantic and Southwest Indian: Cape Frio, Namibia and south to the Agulhas Bank and east to East London in South Africa. They have been recorded off KwaZulu-Natal (Marine and Coastal Management unpublished records).

They are found near the sea bottom depths. They feed on fish, mysids, euphausiids and squids. The young feed mainly on euphausiids, but the diet becomes polyphagous with growth. Cannibalism has been observed in larger individuals

***Genypterus capensis* (Kingklip)**

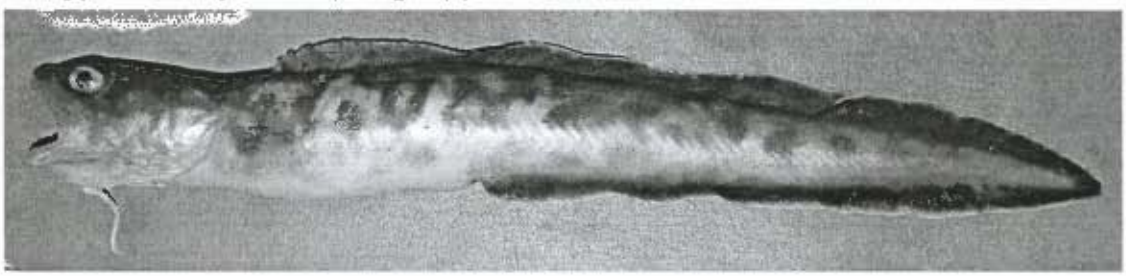


Figure 1.9: *Genypterus capensis*. [1]

Genypterus capensis, also known as Kingklip, is the most sought after species of the commercially important species in the South African fishery. They are found mostly in the Eastern Atlantic: Walvis Bay, Namibia to Algoa Bay, South Africa. Kingklip occurs in rocky areas of the shelf and upper continental slope in 50 – 550m (Heemstra and Heemstra 2004). They mature at 4 to 6 years when females are about 62cm and males about 48cm; and after maturity females grow faster than males (Heemstra and Heemstra 2004). Though they mature at a smaller size length, these fishes can grow up to a maximum of 180 cm in length.

Juveniles occur mostly in shallow waters (< 200m) and feed on a variety of benthic fish, crustaceans and squid, while adults eat mainly fish (Heemstra and Heemstra 2004). The longline fishery catches mainly large adults of this fish and juveniles are mostly caught by hake trawlers.

***Cynoglossus zanzibarensis* (Red-spotted tonguefish)**

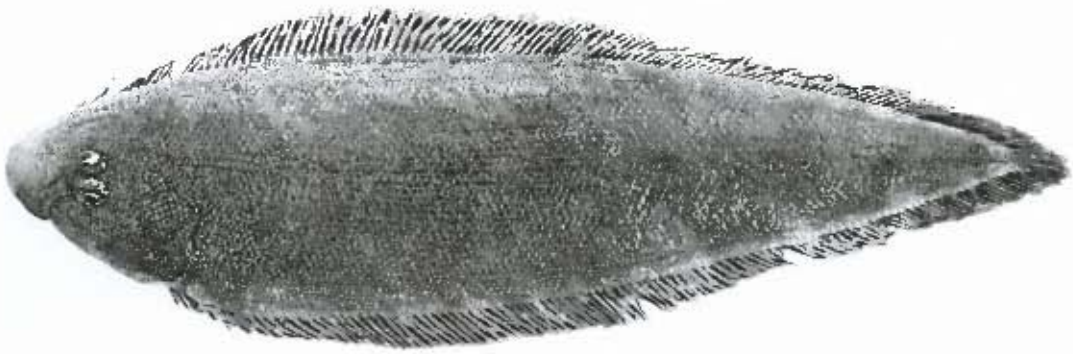


Figure 1.10: *Cynoglossus zanzibarensis.* [1]

Cynoglossus zanzibarensis, known as red-spotted tonguefish, is found mostly in the Eastern Atlantic and Western Indian Ocean; from southern Namibia to South Africa and east to Kenya (Heemstra and Heemstra 2004). They are a shallow-water species, mainly found on sand bottoms in 10 – 430m, and juveniles usually tide in pools. It feeds on benthic invertebrates and fish (Heemstra and Heemstra 2004).

Cynoglossus zanzibarensis females mature at 28cm and also spawn small pelagic eggs throughout the year (Heemstra and Heemstra 2004). They have a maximum length of only 32cm. *Cynoglossus zanzibarensis* are eaten by many piscivores (sharks, bony fish and man). They are common along South African south coast and one of South Africa's best eating 'soles' (Heemstra and Heemstra 2004). It is marketed as 'lemon sole' but affectionately known as 'sandrat' by fishermen. (Heemstra and Heemstra 2004). It is caught mainly as a by-catch in the hake and sole trawl fisheries.

***Lophius vomerinus* (Monkfish)**

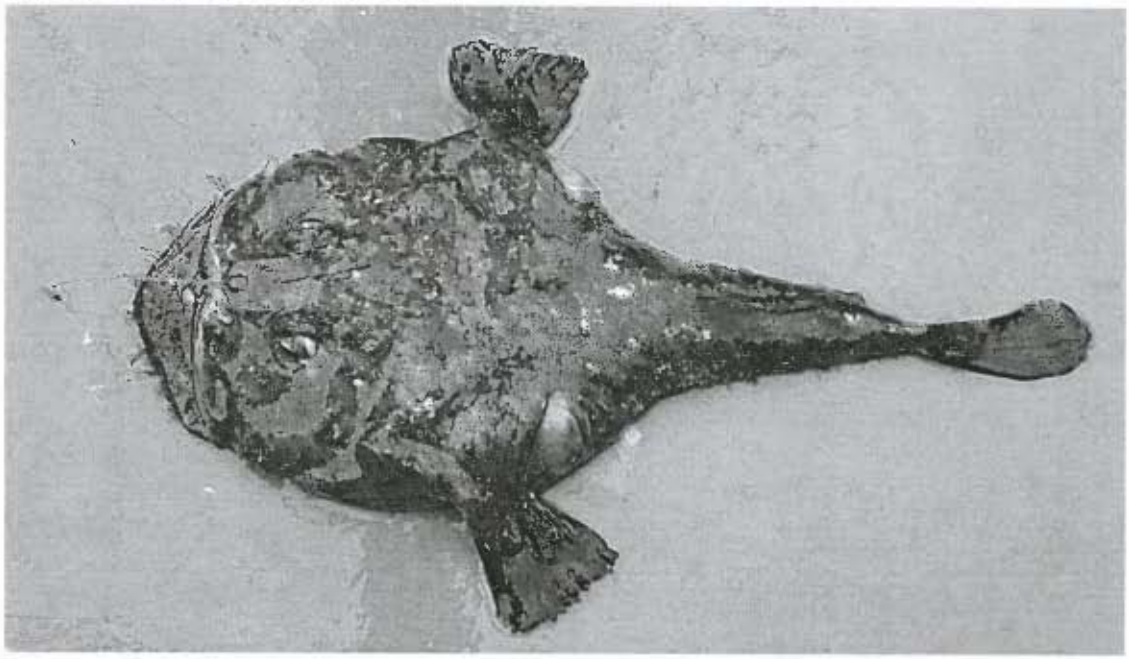


Figure 1.11: *Lophius vomerinus*. [1]

Lophius vomerinus, known as Monkfish, is found mostly in the Southeast Atlantic: Namibia to East London, South Africa. It occurs on the deeper continental shelf and upper slope. They mature at 40cm and the spawning is usually in summer with the eggs embedded in ribbon-like veils that float at surface (Heemstra and Heemstra 2004) and can grow up to a maximum size of 100cm.

They feed mainly on bottom-living fishes, often on pilchard, round herring and horse mackerel. They are also good eating, but it is usually the tails (rear half of the body) that appear in the markets (Heemstra and Heemstra 2004).

***Trachurus capensis* (Horse Mackerel)**

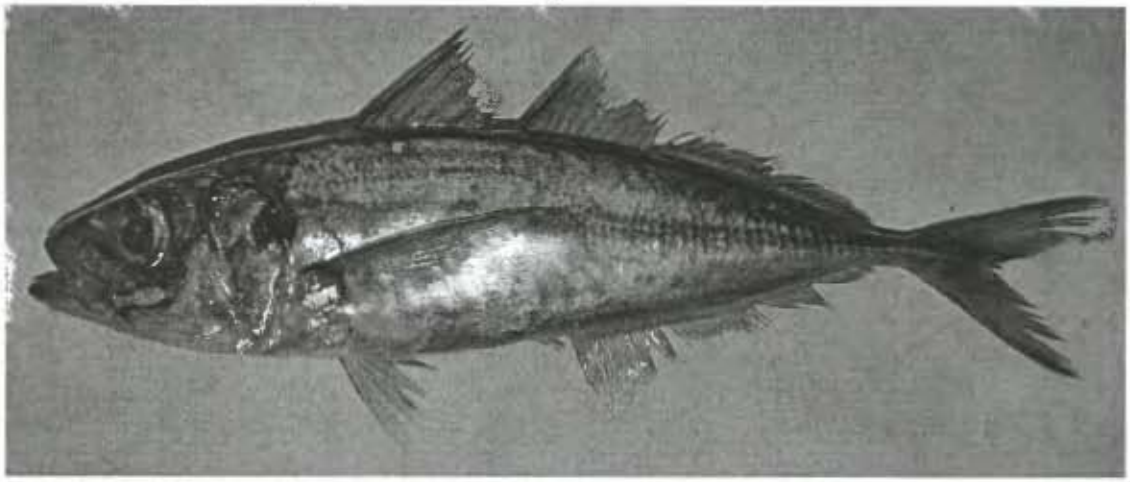


Figure 1.12: *Trachurus capensis*. [1]

Trachurus capensis, also known as the Horse Mackerel, is found in Eastern Atlantic and Southern Indian Oceans: southern Angola to Port Alfred, South Africa. They occur mainly over the continental shelf, often over sand bottoms. They rise to feed in surface waters at night and are found close to the bottom during the day. They have a maximum size of 60cm. Juveniles feed mainly on copepods while adults prey on fish and a wide range of invertebrates.

***Chelidonichthys capensis* (Cape Gurnard)**

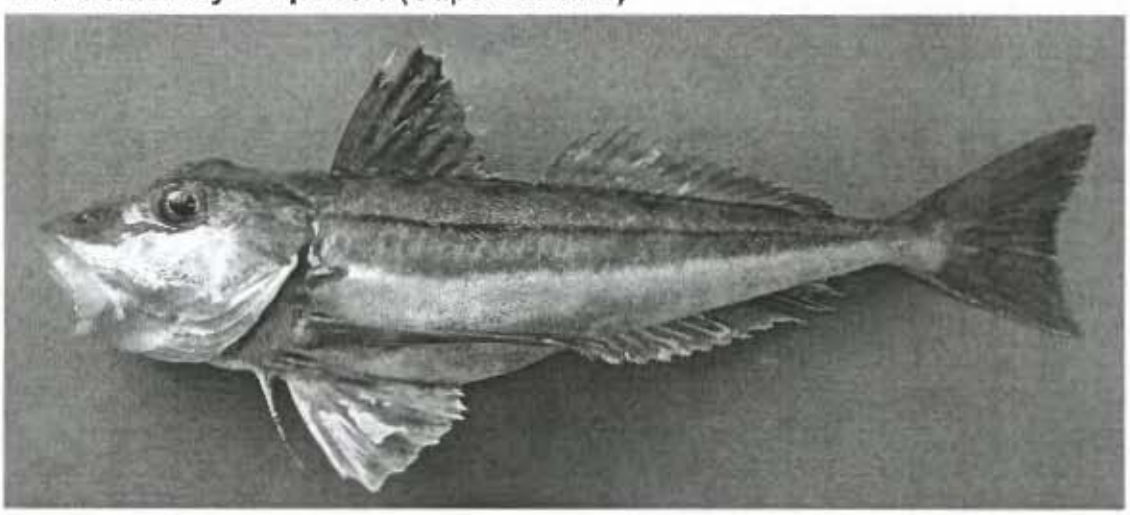


Figure 1.13: *Chelidonichthys capensis*. [1]

Chelidonichthys capensis, also known as the Cape Gurnard, is found in the Southeast Atlantic and Western Indian Ocean: Cape Fria, Namibia to Maputo, Mozambique. It occurs over sandy and muddy bottoms in coastal areas. Females mature at 35cm while males mature at 37cm (Heemstra and Heemstra 2004). *Chelidonichthys capensis* can grow to a maximum of 75 cm in size. Spawning occurs in summer with pelagic eggs (Heemstra and Heemstra 2004). They feed on fishes and crustaceans. They are excellent food fish and an important commercial species caught mainly by trawlers (Heemstra and Heemstra 2004).

***Chelidonichthys queketti* (Lesser Gurnard)**

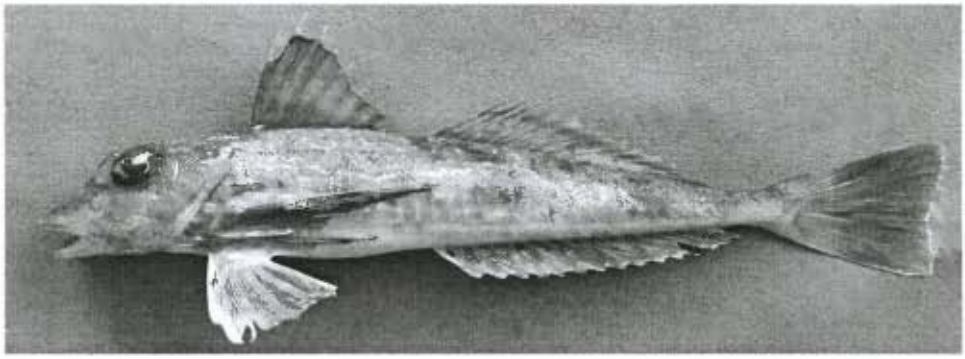


Figure 1.14: *Chelidonihthys queketti*. [1]

Chelidonihthys queketti, also known as the Lesser Gurnard, is found in the Western Indian Ocean: Southern Mozambique southward to Table Bay, South Africa. They occur in shallow waters to 150 m depth. They grow to a maximum length of 35cm.

***Zeus capensis* (Cape Dory)**

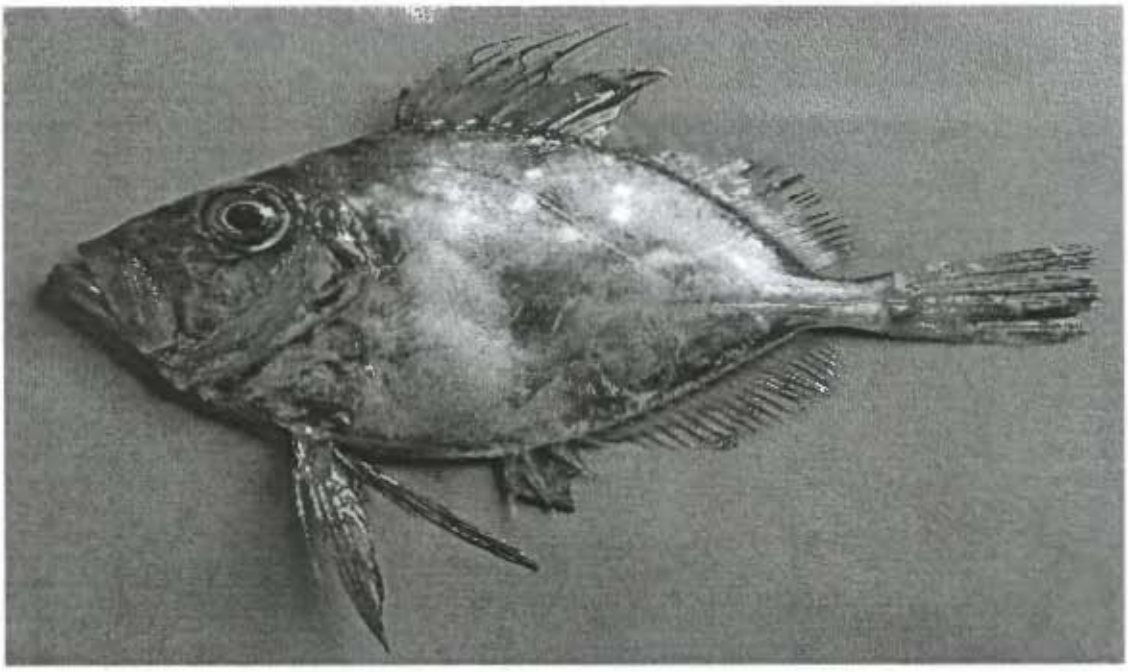


Figure 1.15: *Zeus capensis*. [1]

Zeus capensis, also known as the Cape Dory, is found in the Western Indian Ocean: Mozambique around the Cape to St. Helena Bay, South Africa. They occur near the bottom of the sea or in midwater. They feed on a variety of fishes, cephalopods and crustaceans and they are an excellent food fish.

***Helicolenus dactylopterus* (Jacopever)**

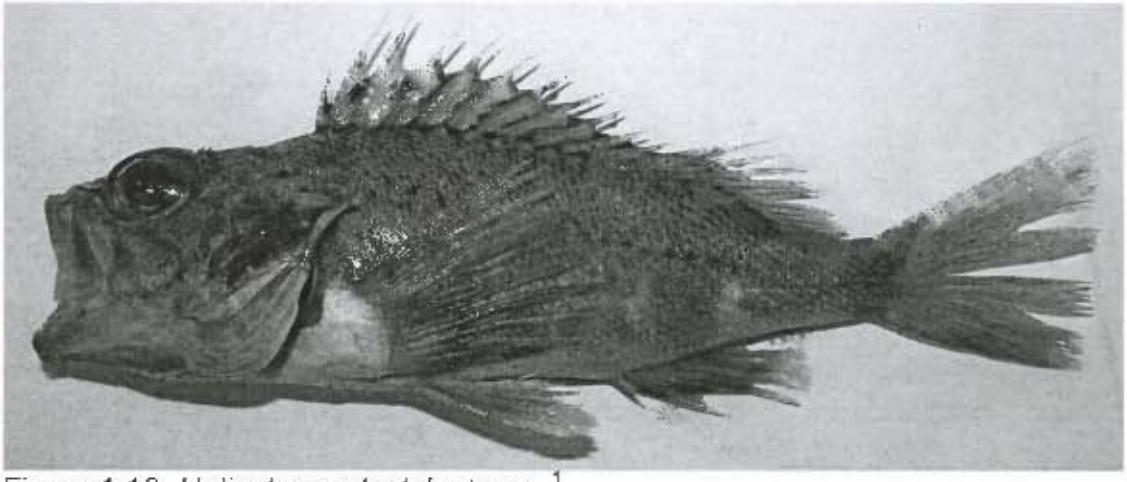


Figure 1.16: *Helicolenus dactylopterus*. [1]

Helicolenus dactylopterus, also known as the Jacopever, is found in the Western Atlantic: Nova Scotia, Canada to Venezuela. Eastern Atlantic: Iceland and Norway to the Mediterranean and the Gulf of Guinea, including Madeira, the Azores, and the Canary Islands; also Walvis Bay, Namibia to Natal, South Africa. They occur in soft bottom areas of the continental shelf and upper slope at depths 55 – 550m (Heemstra and Heemstra 2004). Adults are usually found at depths below 250m. The maximum size found on these fishes is 47cm.

Jacopever feeds on both benthic and pelagic organisms (crustaceans, fishes, cephalopods, and echinoderms). Fertilization occurs internally, and the eggs are released before they hatch; larvae are pelagic (Heemstra and Heemstra 2004). Spawning is in winter, and juveniles are prey sixgill sharks and hake. Jacopever is of some commercial importance as a by-catch of trawlers (Heemstra and Heemstra 2004).

[1] The photos of species in section 1.4 (except for *M. paradoxus*) were taken by Dr R.W. Leslie, a co-supervisor for this thesis. *M. paradoxus* photograph was taken by Oddgier Alvheim, IMR, Bergen.

1.5. Conclusion

We have now been introduced to the problem the study is trying to address. We've also been introduced to the two vessels that were used to collect data during the comparative trials, and to the trawl gears they were using during those trials. We've also learned about the 10 fish species of interest that were caught during the trials. These 10 fish species have a great value to the South African fishery in terms of both food consumption and economy.

In the next Chapter we will learn about the previous studies conducted in the field of comparative trials. We will also learn more about comparative trials and what one needs to do to make sure they are successful. And lastly, and definitely not least, the study areas where the study was conducted.

CHAPTER 2

INTERCALIBRATION PROCESS

2.1. Introduction

There have been studies on the comparative (matched) trials conducted in Europe during the 1990s (Pelletier 1998, Vazquez 2002, Gavaris and Brodie 1984, Warren 1996, Walsh and Orr 1998). Comparative trials have to be conducted every time there is a break in the time series of data due to gear changes or vessel changes. In this chapter, we first discuss what is needed for successful comparative trials, and then give a literature review on some aspects of comparative trials.

2.2. Comparative trials

Before comparative trials are conducted, their objectives must be clearly defined and adequate research plans developed by the participating biologists. Most important is a detailed knowledge of the characteristics of the participating research vessels and the gears used by these vessels for survey work.

It must also be ensured that the overall conditions on each vessel during the comparative trials, with respect to the standard gear and its handling by the normal crew are kept as far as possible the same as those which pertain during ordinary survey work. There should be no changes in the gear used, including the net, warps, otter boards, etc., or in operating procedures such as pull of the winch, trawling speed, duration of trawling, etc.

According to literature (Doubleday 1981), there are two different ways of comparing the catch rates of survey vessels, the direct and indirect methods. Doubleday (1981) defined the direct method as having two vessels fish side by

side on the same fishing grounds under the similar conditions (sea depth, temperature, oxygen content, salinity, etc), and the indirect method as comparing quantitative catch data of survey vessels which fished the stations rather independently within a specific area and time period, under more or less similar conditions. Doubleday (1981) then noted that in both cases, comparative fishing results are characterized by large variation due to the various factors influencing catch rates (factors such as vessel trawling speed, vessel horse power, trawl gear geometry, etc), many of which cannot be controlled by design, but simultaneous matching will be superior because it guarantees a profile of nearly equivalent spatial and temporal conditions rather than an assumption of comparable conditions on aggregate.

In consequence, there will always be some degree of uncertainty as to the exact differences or ratios in the catching power of the survey vessels. These differences or ratios are also not necessarily constant for different fish species, depending on the special type of gear used on each participating vessel.

Doubleday (1981) further mentions that the selection of a suitable fishing ground and season largely determines the success of a comparative fishing experiment. "The experiment should be carried out during a period of favourable weather conditions to avoid undue and costly loss of time and to improve comparability of results. It is also necessary not only to select an area where trawling can be undertaken without difficulty but also where fish concentrations are dense enough for good catches to allow meaningful comparisons. Direct comparison of catches by vessels fishing side by side is based on the assumption that the number of fish in the path of the trawl is more or less the same for each vessel, at any time. Planning of such experiments should also take into account some fish species behavioural aspects like diurnal migration" (Doubleday 1981).

Both vessels participating in the experiment must have the same trawl duration (usually the standard trawl duration is 30 minutes). The trawl duration begins

when the net is in the water and the vessel starts fishing until it stops fishing, and this is measured by what is called a net-sonde (Doubleday 1981). On a side trawler, the duration of trawl may be counted from the time when the warps are blocked up until they are released and haul-back begins. On a stern trawler, trawling usually begins when the appropriate warp length has been played out and the declination of the warps has stabilized and ends when haul-back begins (Doubleday 1981).

According to Doubleday (1981), after every trawl, the trawl gear should be checked carefully for damage or evidence of improper operation, and every effort should be made to maintain a constant trawling speed for each trawl, and to the extent possible, fishing should be carried out at more or less similar depths for each pair of trawls. During the experiment, all events should be carefully recorded in a standard way agreed upon prior to the commencement of operations, and it is essential that procedures for regular communication between the participating vessels be established under the leadership of the cruise leader (Doubleday 1981).

Doubleday (1981) stated that for each vessel, the treatment of catches on board will depend on the facilities and manpower available and also on the size of catches and the number of species to be investigated. For large catches (in terms of numbers caught), random sampling may be required with subsequent proportional adjustment of the results to total catches (in cases of length measurements). For the demersal surveys conducted by Marine and Coastal Management, the length frequencies within the sampled fish of each trawl are adjusted by multiplying the frequencies of the length class by the ratio of the species catch weight to the species sample weight. The reason for this adjustment of length frequencies is to estimate the approximate total number of all fish that were caught in the trawl (see Section 4.1).

As a minimum requirement the record for each trawl should show the weight (kg) of total catch, weight (kg) of sampled catch, number and weight (kg) of each major species studied, length frequencies of each major species, and weight (kg) of by-catch (other fish, invertebrates, organic and inorganic material). The method and instruments of measurement for each species should be the same on each vessel. If time and manpower permit, more extensive biological sampling and evaluation would be desirable (Doubleday 1981).

The geographical range of validity of the experiment should be as wide as possible. The number of trawls actually required to provide meaningful results depends on the variability between trawls. Because it is difficult to predict the minimum number of trawls needed, changes in the program may have to be decided during the execution of the experiment. However, because of various factors which cannot be controlled, a high level of accuracy should not be expected.

It must be considered that not all trawl pairs may be suitable for comparison. Therefore, enough paired data should be collected to allow for the rejection of doubtful cases. It is possible that one vessel or both catches nothing for a specific major fish species, leading to zero values in the data. Some vessels use manual measuring boards to take the lengths of the fish, and this practice can lead to various outliers in the data where for example, the length is recorded as 160cm instead of 16.0cm.

Adjustments and rejection of such cases should be made only on an objective basis after careful analysis of the data and application of statistical methods. Some statistical methods may be log based which might reduce the effect of the outliers but will give rise to adjustment issues in the presence of zero values. The safest way would be to compare the results of all trawls with the results obtained after rejection of doubtful cases and to evaluate any emerging contrasts by means of differences or ratios. Upon completion of the experiment, a detailed

statistical analysis of all valid data is required to elaborate the conversion factors for the catches of the various species relative to the different vessels.

2.3. Literature review

Literature review on studies related to comparative trawls is discussed below. These studies were mainly from the 1990's and very few studies were conducted recently. Most of the studies that were conducted recently focuses on subjects such as herding, escapement, temperature, fish behaviour. These factors, herding, escapement, fish behaviour to the trawl gears used in the study, are such an important aspect when it comes to calculation of conversion factors, and such studies should be undertaken in future especially for these demersal species.

Von Szalay and Brown (2001) analyzed catch per unit effort data collected during a parallel trawl comparison of two vessels (NMFS and ADF&G) in 1997. Both vessels had substantially different trawl gear. They used the Kappenman's method to estimate fishing power correction (FPC) for 4 species and a mean squared error-based decision rule to determine whether the use of fishing power correction factors was warranted.

Pelletier (1998) gives a review of intercalibration of fisheries research survey vessels. He describes an experiment conducted between two French research vessels and also reports some methods he used to achieve the objectives of intercalibration. He reports that intercalibration experiments should be designed to minimize sources of variability in the catch. He contrasts two approaches to sampling design, one where two vessels trawled independently in small areas that were assumed to be homogenous with respect to fish abundance and environmental conditions. He concluded that such designs (designs based on independent trawls) could not account for all of the sources of variability, and a substantial fraction of spatial heterogeneities can possibly remain as residual

variability, inflating catch variance and decreasing both the precision of conversion coefficient estimates and the power of statistical tests.

Pelletier (1998) contrasts the above design with one based on paired trawls, and he found that the paired trawls considerably reduced the consequences of spatio-temporal variabilities because the two vessels trawled simultaneously at the same speed and as closely as possible to one another with the distance generally ranging from 1/4 to 1 nautical mile (nm)[2]. Pelletier (1998) estimated conversion coefficients using a quasi-likelihood method. Quasi-likelihood approaches may be considered as a generalization of likelihood approaches in that they do not require a full specification of the distribution of the observation (Pelletier 1998).

A number of comparative trials have been conducted in the past by the Northwest Atlantic Fisheries Organization (NAFO). Annual NAFO Scientific Council Meetings discuss many issues involving comparative trials. They publish reports which are available on their website http://www.nafo.ca/. They conduct these comparative trials for different suites of species, with the aim of calculating conversion factors between one trawl gear and the other.

The sampling strategy used is similar to the strategy in this study, generally paired trawls. They use both length data and catch data to calculate conversion factors.

However, some of their comparative trials were not conducted with two vessels. For example, Vazquez (2002) reports a comparative trial between the Lofoten survey gear and a Campelen 1800 shrimp trawl gear carried out on Flemish Cap in July 1999-2001. The method followed for the comparison was to make alternative trawls with both gears in the same geographical position. For an example, trawls made with one trawl gear were repeated the next day with the

[2] 1 nautical mile = 1,852 kilometers

alternative. This sampling design eliminates vessel to vessel sources of variation but introduces contrasts of time at each geographical position.

In his report Vazquez (2002) calculated length distribution using paired tows for which species catch ratio Campelen / Lotofen or its inverse was greater than 10 (cod, *A. plaice* and *S. fasciatus*), 20 (*G. halibut* ad grenadier), 30 (*S. mentella*), 50 (juvenile redfish and *S. marinus*) or 60 times (shrimp). This procedure was designed to exclude data from paired trawls where one of the gears caught nothing or almost nothing, so as to make the sample more homogenous. The ratios of Campelen / Lotofen frequency by length class were calculated by fit to a bilinear function:

$$r = \frac{(al+b)}{l+c} \tag{2.3.1}$$

where r = frequency ratio

l = length of fish

a,b,c = parameters

The minimum square fit was then weighted by the total joint frequency at each length group, and in cases where the Lotofen frequency was zero, frequencies were grouped including five original length classes in each group (Vazquez. 2002).

Wilderbuer, Kappenman and Gunderson (1998) applied four analytical techniques to comparative trawl data to obtain fishing power correction (FPC) factors for 12 major commercial species that were caught by two resource assessment trawls used by the National Marine Fisheries Services (Alaska Fisheries Science Center). The aim was to convert catches of the polyethylene trawl gear to the nylon trawl gear. These techniques included ratios of catch per

unit effort (CPUE), randomized block analysis of variance, standard least squares regression, and the Kappenman (1992) method.

The first method used by Wilderbuer *et al.* (1998), which was developed by Cochran (1977), calculated the FPC as the ratio of the mean CPUE of the nylon trawl to the mean CPUE of the polyethylene trawl as follows:

$$\hat{R} = \frac{\frac{\sum_{j=1}^{n} Y_j}{n}}{\frac{\sum_{i=1}^{m} X_i}{m}}. \qquad (2.3.2)$$

where: Y_j and X_i are the catch per unit effort of the nylon trawl and the polyethylene trawl respectively, and

n and m are the number of trawled stations by the nylon trawl and the polyethylene trawl respectively

The second method used by Wilderbuer *et al.* (1998), is the ratio (nylon:polyethylene) of the estimated CPUEs calculated from the randomized block ANOVA model via maximum likelihood estimation procedures. They log-transformed CPUE data as $\log_e$(CPUE + 1) so that when one haul of a trawl pair was zero for a given species, the antilog would be greater than zero. In other words, if one vessel has zero catch, the $\log_e$(CPUE + 1) = 0, and the antilog of this value is 1. The estimated $\log_e$(CPUE + 1) of net i (= 1, 2) during haul pair j was the grand mean from both nets (μ) plus the estimated effect of the net under examination ($\pm t_i$) and the estimated haul pair (randomized block) effect (h_j). These parameters (μ, t_i , and h_j) were estimated by fitting the following model to log transformed observed values of CPUEs:

$$\log_e\left(CPUE_{ij} + 1\right) = \mu + t_i + h_j + \epsilon_{ij} \qquad (2.3.3)$$

Wilderbuer *et al.* (1998) then estimated the FPC as follows:

$$\hat{FPC} = \frac{\hat{CPUE}_{nylon}}{\hat{CPUE}_{poly}} = e^{2t\left(1+0.5s^2\right)} \qquad (2.3.4)$$

Where s^2 is the grand variance from the two nets.

A third analytical method used by Wilderbuer *et al.* (1998) to estimate the FPC invokes the multiplicative model

$$CPUE_{ij} = e^{\mu+t_i+\epsilon_{ij}} \qquad (2.3.5)$$

According to the model in (2.3.5), the data are transformed the same way as in the randomized block ANOVA model as $\log_e$(CPUE + 1), which allows this model to be estimated by ordinary least-squares regression of antilogged values. In other words, CPUE data were first transformed into $\log_e$(CPUE + 1), then antilogged into original scale. This transformation was done so that when one haul of a trawl pair was zero for a given species, the antilog would be greater than zero. Wilderbuer *et al.* (1998) then obtained the FPC as the ratio of the backtransformed estimated regression equations:

$$CPUE_i = e^{\mu+t\left(1+0.5s^2\right)} \qquad (2.3.6)$$

Therefore, the FPC is estimated as:

$$\hat{FPC} = \frac{\hat{CPUE}_{nylon}}{\hat{CPUE}_{poly}} \qquad (2.3.7)$$

The final FPC estimator used by Wilderbuer *et al.* (1998) was the method developed by Kappenman (1992). This method is discussed in detail in Chapter 3 of this study.

Some of the methods discussed above will be adopted in this study. Most of the studies conducted as comparative trials are not recent; they date back as far as the early 1990s. There were no perceived problems with overfishing at that time

as there are today, and the methods of fishing have improved, but not changed much.

Brandão, Rademeyer and Butterworth (2004) used a Generalized Linear Model assuming a negative binomial distribution to calculate a multiplicative bias calibration factor, for each Hake species, using same data as in this study.

Before Brandão *et al* (2004) is explored further, the concept of Generalized Linear Models is first discussed so as to provide an idea of how their results came about. The purpose of this section on negative binomial model is to discuss in detail and possibly critique the unpublished paper by Brandão *et al* (2004) who calculated a multiplicative bias calibration factor on the same hake data used in this study by using a Generalized Linear Model assuming a negative binomial distribution of the hake catches.

Generalized Linear Modelling

The Generalized Linear Model (GLM) can be used to predict responses for dependent variables with discrete or continuous distributions as well as dependent variables which are non-linearly related to the predictor variables.

Consider a set of independent random variables $Y_1,\ldots,Y_N$ each with a distribution from the exponential family with the properties:

(a) The distribution of each Y_i is of canonical form and depends on a single natural parameter.

The density is of

$$f(y_i;\zeta_i) = s_i(y_i)t_i(\zeta_i)\exp(a_i(y_i)b_i(\zeta_i)) \qquad (2.3.8)$$

with a single natural parameter ζ_i where *a, b, s, t* are known functions.

Equation (2.3.8) can be rewritten in the form:

$$f(y_i;\zeta_i) = \exp\left[a_i(y_i)b_i(\zeta_i) + \log t_i(\zeta_i) + \log s_i(y_i)\right] \tag{2.3.9}$$

and if $a(y_i) = y_i$, equation (2.3.9) is said to be in the canonical form and $b(\zeta_i)$ is called the natural parameter of the distribution as opposed to ζ_i, a parameter in (2.3.8).

Therefore, the canonical form of (2.3.8) is given by:

$$f(y_i;\zeta_i) = \exp\left[y_i b_i(\zeta_i) + \log t_i(\zeta_i) + \log s_i(y_i)\right] \tag{2.3.10}$$

(b) The distributions of the Y_i's are of the same canonical form so that subscripts on b, s, and t are not required.

We have from (2.3.10):

$$f(y_i;\zeta_i) = \exp\left[y_i b(\zeta_i) + \log t(\zeta_i) + \log s(y_i)\right] \tag{2.3.11}$$

The joint probability density function of $Y_1,\ldots,Y_N$ is given by

$$f(y_1,\ldots,y_N;\zeta_1,\ldots,\zeta_N) = \prod_{i=1}^{N} \exp\left[y_i b(\zeta_i) + \log t(\zeta_i) + \log s(y_i)\right] \tag{2.3.11}$$

$$\therefore \quad f(y_1,\ldots,y_N;\zeta_1,\ldots,\zeta_N) = \exp\left[\sum_{i=1}^{N} y_i b(\zeta_i) + \sum_{i=1}^{N} \log t(\zeta_i) + \sum_{i=1}^{N} \log s(y_i)\right] \tag{2.3.12}$$

In contrast to the standard linear model where a linear combination $X\underline{\beta}$ of some explanatory variables X given by $\underline{\beta}$ allows $\mu_i = X_i^T \underline{\beta}$ and $y_i = \mu_i + \varepsilon_i$, we consider a GLM model with expected value μ_i of Y_i:

$$g(\mu_i) = X_i^T \beta \tag{2.3.13}$$

where X is a (N x k) matrix, β is a (k x 1) matrix and g in (2.3.13) is a monotone differentiable function called the **link function**.

In summary, the GLM model has the following three components:

(a) Response variables $Y_1,\ldots,Y_N$ which are assumed to share the same distribution from the exponential family;

(b) A set of k explanatory variables X and unknown parameters β.

(c) A monotone link function g.

In obtaining the maximum likelihood estimators of the parameters β for the GLM model described in the above paragraphs, the log-likelihood function for independent responses $Y_1,\ldots,Y_N$ from an exponential family, with a common canonical form, is given by:

$$\ln L(\zeta; y) = \sum_{i=1}^{N} y_i b(\zeta_i) + \sum_{i=1}^{N} t(\zeta_i) + \sum_{i=1}^{N} s(y_i) \tag{2.3.14}$$

where

$$E(Y_i) = \mu_i = -\frac{t'(\zeta_i)}{b'(\zeta_i)} \tag{2.3.15}$$

and

$$g(\mu_i) = x_i^T \beta = \eta_i \tag{2.3.16}$$

where, as defined previously, g is some monotone and differentiable function.

The estimator $\hat{\zeta}$ (of the observed values y) can be obtained by differentiating the log-likelihood function with respect to each element ζ_j of ζ and solving the equations simultaneously.

i.e. $$\frac{\partial \ln L(\zeta; y)}{\partial \zeta_j} = 0 \qquad \text{for } j = 1,\ldots,n \tag{2.3.17}$$

and similarly for β:

$$\frac{\partial \ln L(\zeta; y)}{\partial \beta} = 0 \tag{2.3.18}$$

It can be shown from Dobson (1990) that

$$\frac{\partial \ln L(\zeta; y)}{\partial \zeta_j} = U_j = \sum_{i=1}^{N} \frac{(y_i - \mu_i) x_{ij}}{\mathrm{var}(Y_i)} \left(\frac{\partial \mu_i}{\partial \eta_i} \right) \tag{2.3.19}$$

where x_{ij} is the j^{th} element of $\mathbf{x}_i^T$.

In general, the equations $U_j = 0 \ (j = 1,\ldots,k)$ are non-linear and can be solved by numerical iteration methods such as **Newton-Raphson method** or the **Fisher method of scoring**[3].

Brandão *et al.* (2004) defined their GLM for the Hake density as:

$$Y = C_{sp} = E \exp\left(\mu + \alpha_q + \beta_{pair}\right) + \varepsilon \tag{2.3.20}$$

where:

- C_{sp}: is the total catch of either species (*M. capensis* or *M. paradoxus*).
- E: is an offset which represents the effort extended by a trawl measured here as the swept-area trawled.
- μ: the intercept.

[3] Newton-Raphson method and Fischer method of scoring are used to solve non-linear equations.

q: a factor with 3 levels associated with the survey vessel-gear combination ("old *Africana*", "new *Africana*" or "*Nansen*"). Also known as the "catchability coefficient".

α_q effects of the factor levels associated with factor q.

β_{pair}: a factor with 205 levels associated with trawl pairs between the old Africana and the Nansen and between the new Africana and the Nansen survey vessels.

ϵ: error term assumed to be negative binomial in distribution.

The exponential component is C_{sp}.

The logarithmic link function was assumed so that the expected value of hake catches is given by:

$$\mu_y = E\left(C_{sp}\right) \tag{2.3.21}$$

where the monotone link function g is

$$g\left(\mu_y\right) = \ln\left(E\right) + \mu + \alpha_q + \beta_{pair} \tag{2.3.22}$$

According to the theory, [4]the negative binomial distribution can be used to describe the distribution arising from an experiment consisting of a sequence of independent trials (r), subject to several constraints. Firstly each trial results in success or failure, the probability of success for each trial, p, is constant across the experiment and finally the experiment continues until a fixed number of successes has been achieved. Its probability density function is given as follows:

$$f\left(y_i \mid r_i, p\right) = \binom{r_i + y_i - 1}{y_i} p^r \left(1-p\right)^y \tag{2.3.23}$$

In its canonical form:

$$f\left(y_i \mid r_i, p\right) = \exp\left[y_i \ln\left(1-p\right) + r_i \ln p + \ln\binom{r_i + y_i - 1}{y_i} \right] \tag{2.3.24}$$

has canonical parameters

[4] Definition from wilkipedia: http://en.wikipedia.org

$b(\zeta) = \ln(1-p)$ where $\zeta = 1-p$, and

$t(\zeta) = r_i \ln p$

Using equation (2.3.15) to calculate the mean function of $a(y_i)$, we obtain

$$E\left[a(y_i)\right] = r_i \frac{1-p}{p} \tag{2.3.25}$$

And the variance of $a(y_i)$ is given by:

$$Var\left[a(y_i)\right] = \frac{t'(\zeta_i)b''(\zeta_i) - t''(\zeta_i)b(\zeta_i)}{\left[b'(\zeta_i)\right]^3}$$

$$Var\left[a(y_i)\right] = r_i \frac{1-p}{p^2} \tag{2.3.26}$$

In Brandão *et al* (2004), catch (C_{sp}) is taken as proportional to a negative binomial count. The purpose of the negative binomial model in this instance is not the theoretical or classical description of successes until failure, but the property that the variance is larger than the mean in contrast to the Poison model where there is equal mean and variance. Therefore, in this instance, there is a log likelihood formed by 205 negative binomial terms each with their own r_i, y_i, and p_i and then the maximum likelihood estimation and other score function properties allow estimation of parameters. They determined that the distribution of hake catches has a quadratic variance function of the form: $mean + {mean^2}/{k}$, where k is the overdispersion parameter.

This form of variance follows from the assumption that hake catches have a negative binomial distribution.

Brandão *et al* (2004) therefore calculated conversion factors from catchability coefficient q from the following 2 equations:

$$\ln q_{new}^{capensis} = \ln q_{old}^{capensis} - 0.494 \tag{2.3.27}$$

and

$$\ln q_{new}^{paradoxus} = \ln q_{old}^{paradoxus} - 0.053 \qquad (2.3.28)$$

Brandão *et al* (2004) calculated conversion factors for new Africana vs old Africana, using Nansen as the bridge vessel, for *Merluccius capensis* as $e^{-0.494} = 0.61$ (from equation 2.3.27), and for *Merluccius paradoxus* as $e^{-0.053} = 0.95$ (from equation 2.3.28). It will be interesting to compare their results with the results of this study discussed in Chapter 5.

CHAPTER 3

RESEARCH METHODOLOGY

3.1. Introduction

This chapter focuses on how the data were collected and describes the methods that will be used for the analysis of that data. We begin by explaining the experimental method used in collection of the data. In other words, we describe the intercalibration experiment used between RV *Dr Fridtjof Nansen* and FRS *Africana*, and also describe in detail what was observed during the experiment. Study areas where the data were collected are also provided in this chapter. The second part of this chapter explains the methods used in analyzing the data. We explain the conditions to be met for the methods to be appropriate and how can they be used.

3.2. Experimental Method

The FRS *Africana* and RV *Dr Fridtjof Nansen* trawled parallel, as close together as possible. Each pair of trawls was of 30min duration and only took place during full daylight hours. After every second trawl, the vessels switched relative positions i.e. neither vessel was offshore of the other vessel for more than two stations in sequence. The FRS *Africana* fishes at a faster trawling speed (3.5 knots[5]) than the RV *Dr Fridtjof Nansen* (3.0 knots); therefore, on average, she travels 1.75nm during the 30 min trawling period whereas RV *Dr Fridtjof Nansen* travels 1.5nm. To account partially for this difference in speed, trawl pairs were started with RV *Dr Fridtjof Nansen* ahead of FRS *Africana* so that the two vessels were alongside approximately at the midpoint of the trawl track.

[5] 1 knot = 1.852km/h

The majority of stations were completed on the clear trawl grounds north of Cape Columbine, but also included some stations on rougher grounds, and from a range of depths below 600m.

For each trawl pair, the whole catch, or a sub-sample if the catch was too large, was sorted into species and weighed to determine the species composition by mass. A sub-sample of the catch of each species was counted to determine species catch composition by number. Within those counts, length frequencies were determined for shallow-water hake *Merluccius capensis*, deep-water hake *Merluccius paradoxus*, horse mackerel, west coast sole *Austroglossus microlepis*, sand sole, monk, squid, spiny dogfish *Squalus megalops*, kingklip, jacopever *Helicolenus dactylopterus* and some linefish species such as snoek, shark, and yellowfin, *inter alia*.

Work usually started in the morning and an average of 7 trawl-pairs was completed per day. RV *Dr Fridtjof Nansen* uses a restraining strap to maintain a door spread of 50m during normal survey fishing operations. The mouth opening of the FRS *Africana*'s trawl is about 30m as opposed to 20m for the RV *Dr Fridtjof Nansen*; therefore one would expect the FRS *Africana*'s catch to involve a factor for mouth opening of about 1.5 times that of the RV *Dr Fridtjof Nansen*. Also, one expects an additional factor for trawling speed since FRS *Africana* covers ±1.75nm whereas RV *Dr Fridtjof Nansen* covers ±1.5nm. On the other hand, RV *Dr Fridtjof Nansen* uses a small meshed codend which could increase (or decrease) her catch relative to FRS *Africana*, at least in respect of smaller size classes of fish.

In general, catches (Figure 3.1) by the FRS *Africana* were more than 1.5 times the mass of those for the RV *Dr Fridtjof Nansen*, especially in stations completed at midday and those in shallow waters. A similar pattern was observed with the "old *Africana* gear", and was attributed to increased herding at midday associated with improved visibility through higher light intensity at midday.

is evidence for contrasting relative performance of the gear in shallow and deeper water, then it may be necessary to consider employing a restraining strap on the FRS *Africana* gear in future.

3.3. Existing Data

The data (trawl station data) collected during the intercalibration trials that were used for the analysis in this thesis consists of the following:

- Station number and trawl number
- Catch composition (mass (kg) and number per species)
- Start and end positions (latitude and longitude) of the stations
- Starting time and duration (minutes) of trawling
- Fishing depth (meters)
- Total mass (kg) of all species caught and total mass of each species caught
- Length frequency data for both hake species
- A CTD dip (supplying a profile of salinity, temperature and oxygen content with depth).

There were 110 paired trawls in the intercalibration trials conducted between FRS *Africana* (with the old trawl gear) and RV *Dr Fridtjof Nansen* during 7 days in October 2001 and 9 days in January 2002. Also, there were 98 paired trawls of intercalibration trials conducted between FRS *Africana* (with the new trawl gear) and RV *Dr Fridtjof Nansen* during 6 days in February 2004 and 8 days in September 2004.

Table 3.1 and 3.2 report the number of matched pairs of non-zero catch data from a common location for each species collected during the two sets of intercalibration trials.

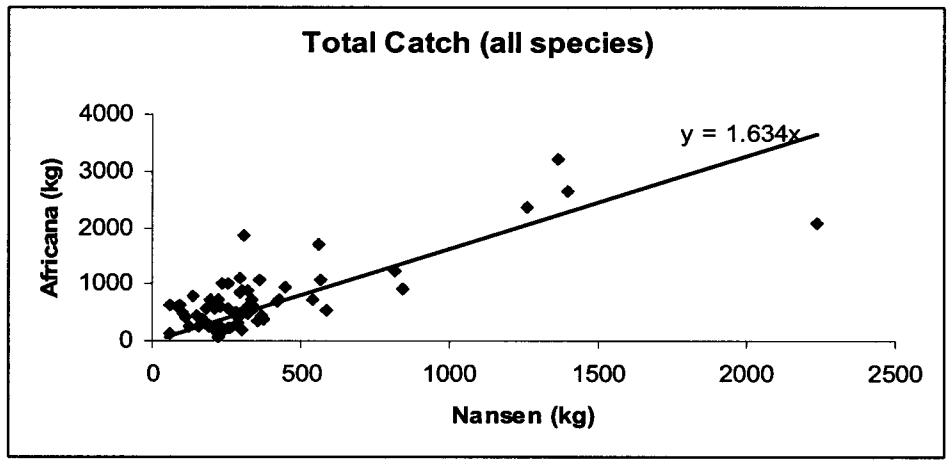


Figure 3.1: Total catch (kg) of FRS *Africana* against total catch (kg) of RV *Dr Fridtjof Nansen*

The average door spread for the new gear used by the FRS *Africana* is only 70m (reduced from 121m). It was hoped that this smaller spread would eliminate or substantially reduce the herding effect. The higher midday catches observed in this experiment (Figure 3.2) may indicate that the problem has not been successfully resolved.

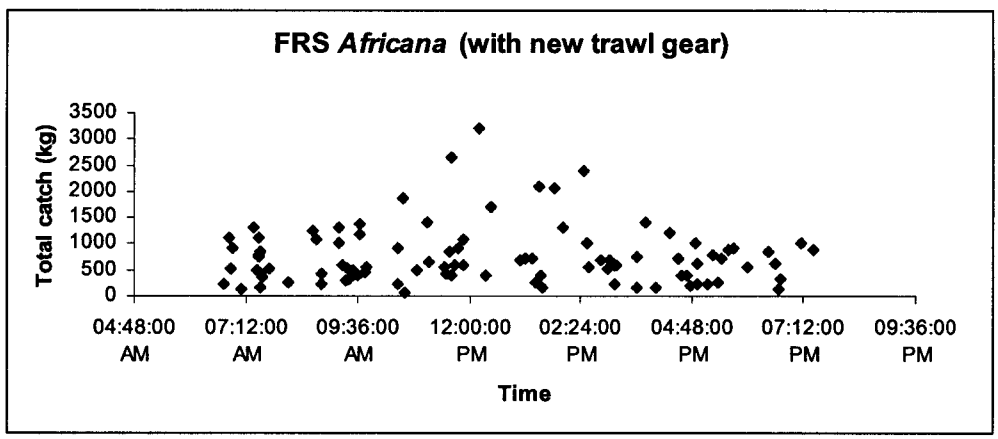


Figure 3.2: Total catch (kg) of FRS *Africana* (with new trawl gear) against time of day.

The door spread on FRS *Africana* increases with increasing depth, being about 65m in shallow waters and up to 80m in deep waters. In the deep waters, the increased door spread (and wingspread) would result in a lower vertical mouth opening. This change could account for the observation that relative catch volumes for the two vessels were more similar in deep waters than in shallow waters. If there

for other species. The contrast can be explained by the fact that the two hake species were the target species, and the surveys were designed to target larger catches of hake. The problem of zero catches can be explored either by ignoring pairs with zero catches by an analysis of differences, or by adding a small constant to allow for calculation of ratios. In both tables, *Merluccius paradoxus* was caught in every station. The two species, *Lophius vomenirus* and *Helicolenus dactylopterus*, were also caught in almost every station.

Hake distribution (and abundance) is affected by environmental conditions such as depth and water temperature. Studies have shown that catch rates are also affected by oxygen content. We do not expect these environmental factors to affect the conversion factors between the two vessels, as they were fishing in similar areas which are similar in conditions.

3.4. Study areas

Demersal surveys took place on the soft trawling grounds west of Cape Town from Cape Agulhas (20^0 E) to the international border of South Africa and Namibia just beyond Port Nolloth, and comparative trials were conducted during these surveys in the areas shown in Figure 2.1. All the trawls for comparative trials were at depths below 600m. The map in Figure 2.1 depicts the areas where the study trawls were undertaken.

Species	No. of pairs (total)	Pairs with zeros	% zeros
Merluccius capensis	94	11	11.7
Merluccius paradoxus	109	5	4.59
Genypterus capensis	99	16	16.16
Cynoglossus zanzibarensis	38	13	34.21
Lophius vomerinus	107	15	14.02
Trachurus capensis	70	12	17.14
Chelidonichthys capensis	42	9	21.43
* *Chelidonichthys queketti*	-	-	-
Zeus capensis	74	22	29.73
Helicolenus dactylopterus	109	9	8.26

Table 3.1: (109 trawls in 2001 and 2002) FRS *Africana* (old trawl gear) and RV *Dr Fridtjof Nansen.*

Species	No. of pairs (total)	Pairs with zeros	% zeros
Merluccius capensis	78	4	5.13
Merluccius paradoxus	95	2	2.11
Genypterus capensis	68	27	39.71
* *Cynoglossus zanzibarensis*	-	-	-
Lophius vomerinus	95	43	45.26
Trachurus capensis	55	6	10.91
Chelidonichthys capensis	39	10	25.64
* *Chelidonichthys queketti*	-	-	-
Zeus capensis	64	22	34.38
Helicolenus dactylopterus	94	1	1.06

Table 3.2: (95 trawls in 2004) FRS *Africana* (new trawl gear) and RV *Dr Fridtjof Nansen.*

** very few or no data available*

From tables 3.1 and 3.2 it is apparent that for each species except *Merluccius paradoxus* there were some paired trawls for which both vessels had zero catch of that species. It also appears that the percentage of zero catches on one of the vessels is lower for the 2 hake species and *Helicolenus dactylopterus* but higher

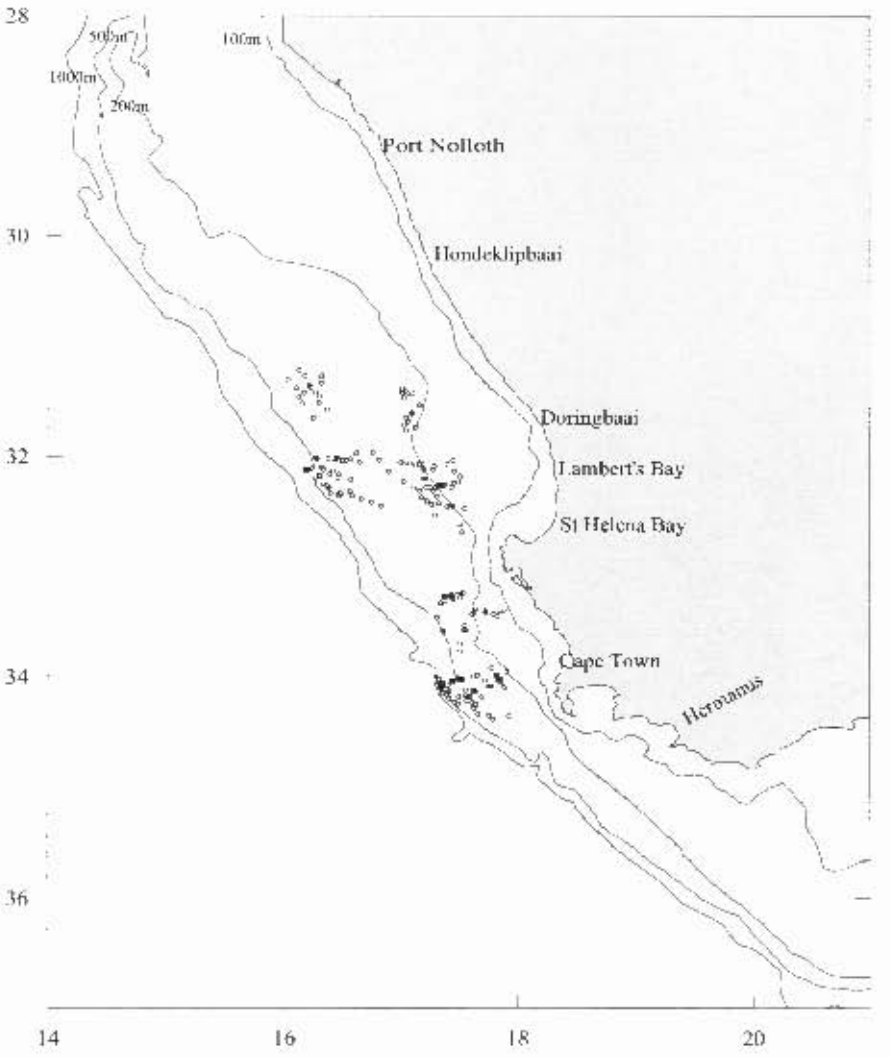


Figure 3.3. Map of the west coast of South Africa showing trawled stations during the intercalibration trials (2001 and 2002).

3.5. Analytical Methods

In examining bivariate data we may observe that the three usual assumptions underlying the linear regression model are not met. For each X, the corresponding Y's may not be distributed as Gaussians; their variances may be unequal; or their means may lie on a broken line or a curve rather than on a straight line. It is for these reasons that some data transformation methods were

explored in this study, as devices to better match data to assumptions of analysis.

Figures 3.3 and 3.4 presents the distribution of the marginal univariate total catch data (of all species) for both RV *Dr Fridtjof Nansen* and FRS *Africana (old and new trawl gear)*. Clearly, we can see from the histograms that the two catch variables are positively skewed and violate Gaussian assumptions. Our concern is to establish transformed data for which a conversion factor β for a model $\underline{y} = \beta x + \varepsilon$ for the bivariate data, where $\underline{\epsilon}$ is the vector of residuals, will admit suitable regression methods.

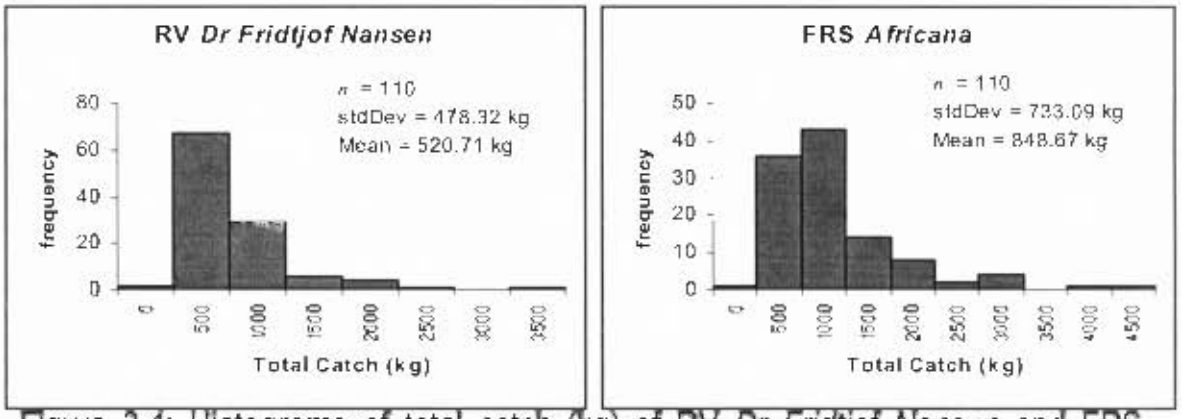


Figure 3.4: Histograms of total catch (kg) of RV *Dr Fridtjof Nansen* and FRS *Africana* (with old trawl gear)

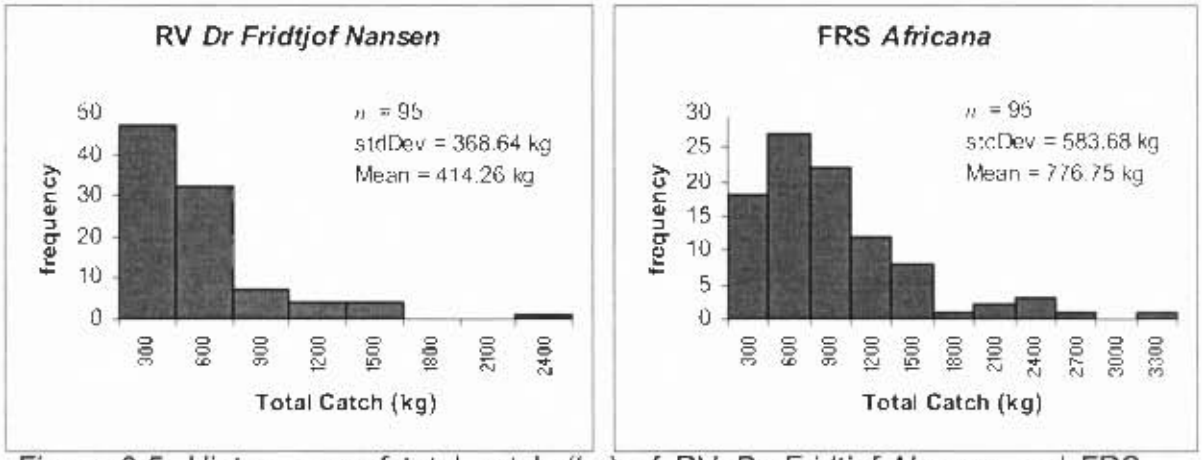


Figure 3.5: Histograms of total catch (kg) of RV *Dr Fridtjof Nansen* and FRS *Africana* (with new trawl gear)

We discuss below various transformation methods which then led to the calculation of the corresponding conversion factors.

3.5.1. Data Transformation

Transformations are the application of a mathematical modification to the values of a variable. Transformations are commonly used tools that can serve many functions in quantitative analysis of data. There are atleast five aims of data transformations for statistical analysis, especially for linear models (Quinn and Keough 2002):

- to make the data and the model error terms closer to a normal distribution (i.e. to make the distribution of the data symmetrical),
- to reduce any relationship between the mean and the variance (i.e. to improve homogeneity of variances), often as a result of improving normality,
- to reduce the influence of outliers, especially when they are at one end of a distribution,
- to improve linearity in regression analyses, and
- to make effects that are multiplicative on the raw scale additive on a transformed scale, i.e. to reduce the size of interaction effects.

Data transformations can be univariate:

$$X \Rightarrow \ln X$$
$$Y \Rightarrow \ln Y$$

Or bivariate:

$$(X,Y) \Rightarrow (X,R) \ for \ R = \frac{Y}{X}$$

The purpose of transformations is to obtain a form of data that satisfies the assumptions of analysis, e.g., common variance for all $i(and\ x_i)$, or a specific

distributional form. The most common use of transformations in biology is to help the data meet the distributional and variance assumptions required for linear models (Quinn and Keough 2002).

When a transformation of data is made in order to perform some calculations under suitable criteria, it is necessary to know how to change ("invert" or "back-transform") the transformation back into the original units in order to interpret any final estimation values. With a back-transformation of a model coefficient, the standard error will no longer be symmetric, so a better way to represent variability in the original units is obtained by back-transforming the end points of confidence intervals from the transformed data.

We begin our discussion by introducing the weighted least squares transformation methods, followed by the log transformation methods, then the Kappenman method and lastly, but not least, the Box-Cox model. Under the weighted least squares criterion, two conversion multiplier approaches are derived, namely, the WLS regression through origin (1) and the WLS regression through origin (2).

Under the log transformation criterion, the power approaches are derived, namely, inversion of the linear regression model and the inversion of the linear regression through origin. The Kappenman and the Box-Cox models are also power transformation approaches.

3.5.1.1. Weighted Least Squares

Weighted least squares (WLS) are useful for estimating the values of model parameters when the response values have differing degrees of variability over the combinations of the predictor values. Unlike least squares, however, each data term under the weighted least squares criterion involves an additional

weight, w_i, that determines how much each observation in the data set influences the final parameter estimates.

General principles of weights of data cases involve weights w_i for case i.

Ordinary least squares (OLS)

Here the model

$Y = \beta X + \varepsilon$, where the variance of the residuals is given by $\text{var}(\varepsilon_i) = \sigma^2$, under an OLS criterion, leads to an estimate of the slope of the regression line through the origin (i.e. the ratio of Y to X) given by:

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} = \frac{x'y}{\underline{x}'\underline{x}}$$

Unweighted data have all $w_i = \frac{1}{n}$ and $\sum w_i = 1$. However, estimates for weighted data will generally have $w_i > 0$, $\sum w_i \neq 1$ and weighted estimates of a common location parameter μ involved in all random variables X_i with possibly different variance σ_i^2 will have $\hat{\mu} = \sum w_i X_i$ and $\text{var}(\hat{\mu}) = \sum w_i^2 \sigma_i^2$. For $\hat{\mu}$ to be unbiased we require $\sum w_i = 1$. Given this constraint, and variances σ_i^2, a minimum variance unbiased estimator $\hat{\mu}$ is obtained when $w_i \sigma_i^2 = 1$ for all i.

Choice of unequal weights w_i may be informed by other information, particularly about the array of variances σ_i^2. If $X_i \sim \textit{random variable}\left(\mu, v_i \sigma^2\right)$ but independently over i, then we may choose: weights $w_i \propto v_i^{-1}$, i.e., $\left(w_i = \frac{v_i^{-1}}{\sum\left(\frac{1}{v_i}\right)} \right)$,

which yields an unbiased estimator $\hat{\mu}_w = \sum w_i X_i$ for mean μ, and estimator variance $\text{var}(\hat{\mu}_w) = \sum w_i^2 v_i \sigma^2 = \left(\sum \frac{1}{v_i} \right)^{-1} \sigma^2$.

In contrast, weights $w_i \propto v_i^{-\frac{1}{2}}$, i.e. $\left(w_i = \frac{v_i^{-\frac{1}{2}}}{\sum \left(v_i^{-\frac{1}{2}} \right)} \right)$, yields an unbiased estimator $\hat{\mu}_w = \sum w_i X_i$ for mean μ, and estimator variance $\text{var}(\hat{\mu}_w) = \sum w_i^2 v_i \sigma^2 = \left(\sum v^{-\frac{1}{2}} \right)^{-2} \sigma^2$.

If $\sigma_i^2 = \sigma^2$ then $\text{var}(\hat{\mu}) = \left(\sum w_i^2 \right) \sigma^2$ and minimum variance unbiased estimator $\hat{\mu}$ is attained when $w_i = \frac{1}{n}$.

Below we discuss two types of weighted least squares, adapting the common mean to a ratio model criterion with reference to the ordinary least squares. The modelling process explores linear models with fixed zero intercept.

WLS regression line through the origin (1)

If $y_i = \beta x_i + \varepsilon_i$ where $var(\varepsilon_i) = x_i \sigma^2$,

then weighting $w_i = \frac{\frac{1}{\sqrt{x_i}}}{\sum \left(\frac{1}{\sqrt{x_i}} \right)}$ gives a model with independent common variance residuals:

$$\left(\frac{y_i}{\sqrt{x_i}} \right) = \beta_{01} \sqrt{x_i} + e_i \quad \text{where} \quad \text{var}(e_i) = \sigma^2$$

Then the OLS calculation on the transformed data (with common variance) for the ratio parameter β, we obtain:

$$\hat{\beta}_{01} = \frac{\sum \frac{y_i}{\sqrt{x_i}} \cdot \sqrt{x_i}}{\sum \left(\sqrt{x_i}\right)^2} = \frac{\sum y_i}{\sum x_i} = \frac{\bar{y}}{\bar{x}}$$

where the ratio parameter estimate $\hat{\beta}_{01}$ has the variance:

$$\text{var}\left(\hat{\beta}_{01}\right) = \frac{\sigma^2}{\sum x_i} = \frac{\sigma^2}{n\bar{x}}$$

Then $$\hat{\text{var}}\left(\hat{\beta}_{01}\right) = \frac{\hat{\sigma}_1^2}{\sum x_i} = \frac{\hat{\sigma}_1^2}{n\bar{x}}$$

for $$\hat{\sigma}_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(\frac{y_i - \hat{\beta}_{01}x_i}{\sqrt{x_i}}\right)^2 = \frac{1}{n-1}\left(\sum \frac{y_i^2}{x_i} - \hat{\beta}_{01}^2\left(\sum x_i\right)\right).$$

Thus, under the assumed conditions, a 95% CI for β is

$$\hat{\beta}_{01} \pm 1.96 * s.e\left(\hat{\beta}_{01}\right)$$

and $$E(Y / X = x_0) = x_0\hat{\beta}_{01}$$

This model may be appropriate when X_i is interpreted as a proxy for the number of elements in Y_i. The estimator is simply the quotient of the totals (or equivalently the means).

WLS regression line through the origin (2)

Consider $y_i = \beta_{02}x_i + \varepsilon_i$ where $\text{var}(\varepsilon_i) = \sigma^2 x_i^2$. Then weighting $w_i = \frac{\frac{1}{x_i}}{\sum\left(\frac{1}{x_i}\right)}$ gives a model $R_i = \frac{y_i}{x_i} = \beta_{02} + e_i$ with transformed values that are identically and

independently distributed and we may use OLS in the transformed model to find an unbiased estimator for β_{02} as:

$$\hat{\beta}_{02} = \frac{1}{n}\sum R_i = \frac{1}{n}\sum \frac{y_i}{x_i}$$

The variance of this estimator is then given by:

$$\text{var}\left(\hat{\beta}_{02}\right) = \frac{\text{var}\left(R_i\right)}{n} = \frac{\sigma^2}{n}$$

Then $$\hat{\text{var}}\left(\hat{\beta}_{02}\right) = \frac{1}{n}\frac{\sum\left(R_i - \bar{R}\right)^2}{n-1}$$

$$\hat{\text{var}}\left(\hat{\beta}_{02}\right) = \frac{1}{n}\frac{\sum R_i^2 - \frac{1}{n}\left(\sum R_i\right)^2}{n-1}.$$

Thus a 95% CI for β is:

$$\left[\bar{R} \pm 1.96 * s.e.\left(\bar{R}\right)\right]$$

and $$E\left(Y / X = x_0\right) = x_0\bar{R}$$

This type of model is appropriate when X_i is interpreted as a proxy for a multiplicative factor. The estimator is the mean of the ratios $\left(\bar{R}\right)$ in contrast to the ratio of the means $\left(\hat{\beta}_{01} = \frac{\bar{y}}{\bar{x}}\right)$.

Results are presented in Section 5.3.

3.5.1.2. Log transformation

In this study we seek to manage the effects upon analysis and prediction of catch variances that appear to increase as catch volumes increases. For our objective, we will wish to have the logarithmic transformation yield transformed data which follow a Gaussian distribution. We say:

$X \sim \log Normal(\mu, \sigma^2)$ if $\log X \sim N(\mu_{\ln X}, \sigma^2_{\ln X})$ and
$Y \sim \log Normal(\mu, \sigma^2)$ if $\log Y \sim N(\mu_{\ln Y}, \sigma^2_{\ln Y})$

where X and Y are random variables in our applications representing non-zero catch per unit area (CPUA) of RV *Dr Fridtjof Nansen* and FRS *Africana* respectively.

This structure is used for each species in the studied suite of species. If there is evidence of a linear relationship between the two log-transformed variables for a species, then the following methods can be applied:

Inversion of the linear regression equation

For the multiplicative model $Y = e^{\alpha} X^{\beta}(E)$, with multiplicative error E, transforming (Y, X) to $(\log Y, \log X)$ over the model, we have the derived linear regression model:

$$\log Y = \alpha + \beta \log X + \varepsilon .$$

Thus, $\log y_i = \alpha + \beta \log x_i + \varepsilon_i$ where $i = 1, 2, \cdots, n$ is a model for the log-transformed data.

Scatterplot techniques may lead to adoption of the usual regression assumptions and hence to linear regression on the transformed data.

Therefore, having estimates $\hat{\alpha}, \hat{\beta}$ from the log-transformed data, provided that $\hat{\beta} \cong 1$, the conversion factor $k = e^{\alpha}$ describes the relationship between the paired trawls.

The 95% CI for $k = e^{\alpha}$ is given by $\exp(\hat{\alpha} \pm 1.96 \times \text{s.e}(\hat{\alpha}))$. It is therefore admissible to make an inference about $\alpha = 0$, *inter alia.*

We note that identical regression estimates $\hat{\beta}$ and standard errors for $\hat{\beta}$ are obtained regardless of the choice of logarithmic base.

The reverse transformation of the 95% confidence intervals for parameters $\hat{\beta} \pm 1.96 * s.e.(\hat{\beta})$ and for fitted values $x_0\hat{\beta} \pm 1.96 * s.e.(x_0\hat{\beta})$ yield corresponding interval estimates for $E(y / X = \exp(x_0))$ in the original scale.

The correlation coefficient is given by:

$$p = \frac{\left[(\underline{\log x})'(\underline{\log y})\right]}{\sqrt{\left[(\underline{\log x})'(\underline{\log x})\right]\left[(\underline{\log y})'(\underline{\log y})\right]}}$$

Inversion of the linear regression through origin

For $\alpha = 0$ in the multiplicative model $Y = e^{\alpha} X^{\beta}(E)$, we obtain the special case:

$$Y = X^{\beta} * E$$

which can be written as a linear model in a log-log space: $\log Y = \beta \log X + \log E$

Thus, $y_i = \beta x_i + \varepsilon_i \qquad for\ i = 1,2,\cdots,n$

If after examining the scatterplot of the transformed values it seems reasonable to assume that $\varepsilon_i = \ln E_i$ are zero mean, common variance, independent, and collectively sufficiently close to Gaussian (normal) distribution, we may fit (by OLS) the transformed model to obtain:

$$\hat{\beta} = \frac{\sum(\log Y * \log X)}{\sum(\log X)^2} = \frac{\underline{y}'\underline{x}}{\underline{x}'\underline{x}}$$

The standard error is given by:

$$s.e\left(\hat{\beta}\right)=\frac{\hat{\sigma}}{\sqrt{\sum\left(\log X * \log X\right)}}=\frac{\hat{\sigma}}{\sqrt{\underline{x}'\underline{x}}}$$

where $$\hat{\sigma}^2=\frac{1}{n-1}\left[\underline{y}'\underline{y}-\frac{\left(\underline{y}'\underline{x}\right)^2}{\underline{x}'\underline{x}}\right]=\frac{1}{n-1}\left[\underline{y}'\underline{y}-\hat{\beta}^2\left(\underline{x}'\underline{x}\right)\right]$$

Then, $\hat{y}\left|x=x_0\right.$ is estimated by

$$E\left(\hat{y}\right)=x_0\beta$$

and $$s.e.\left(\hat{y}\right)=\frac{x_0\hat{\sigma}}{\sqrt{\underline{x}'\underline{x}}}$$

so that $\hat{Y}\left|X=\exp\left(x_0\right)\right.$ has an approximate mean $E\left(\hat{Y}\right)=X^{\beta}$

Results are presented in Section 5.3.

3.5.1.3. Kappenman's Method

Kappenman (1992)'s model assumes that two catch per unit area (CPUA) positive random variables have unknown but identical distributions, except possibly for the values of the scale parameters of the distributions. The assumption of a common distribution for X and Y amounts to an assumption that the two variables each measure the same phenomenon, on possibly different scales.

The probability distribution functions (p.d.f)'s of X and Y are related. For $X=b_xW$ and $Y=b_yW$ so that $X=\frac{b_x}{b_y}Y$ and $Y=\frac{b_y}{b_x}X$, we have

$$\Rightarrow \quad f_x\left(w\right)\rightarrow\frac{1}{b_x}f\left(\frac{x}{b_x}\right) \text{ and } f_y\left(w\right)\rightarrow\frac{1}{b_y}f\left(\frac{y}{b_y}\right)$$

Kappenman (1992) envisaged independent X and Y samples as data, to which to apply this common distribution assumption. The bivariate data structure does not apply so that regression-like methods do not apply to his data.

The assumption of transforming identical distributions is reasonable for comparable vessels and gear, but is rendered even more plausible when the two random variables are each sampled simultaneously at the same randomly chosen set of geographical locations.

Here X = random variable representing CPUA for RV *Dr Fridtjof Nansen*
and Y = random variable representing CPUA for FRS *Africana*

The original method of Kappenman (1992), if applied to the bivariate trawl data, ignores the explicit pairing of the design. The Kappenman (1992) assumption of scale transformations to equivalent distributions may preserve aspects of the pairing when $m = n$. However, in allowing m, n to be distinct sample sizes, it is clear that the pairings are suppressed.

However, if Kappenman's method is applied to paired trawls as in this design, the effect of any ratio differential effort per unit area between the vessels will be confounded into the distributions of the masses per trawl.

Kappenman (1992) envisaged m data points $x_1 \ldots x_m$ and n data points $y_1 \ldots y_n$. He seeks a power transformation which gives Gaussian (normal) marginal distribution.

The first step in the method is to find a value d so that X^d and $r^d Y^d$ have approximately the same Gaussian distribution. Kappenman (1992) estimates the conversion factor as the ratio, $\hat{r}$, of the scale parameters $\frac{b_x}{b_y}$ at which the two marginals are equivalent. The transformation parameter d reduces to a power

transformation of the two variables (X, rY) to move from a common non-Gaussian distribution to a common Gaussian.

From Kappenman (1992), the estimator $\hat{r}$ is the value that minimizes the function:

$$g(r) = \left(\frac{n+m}{n+m-1}\right)^2 \left[\sum\left(x_i^d - \frac{\sum x_i^d + r^d \sum y_j^d}{n+m}\right)^2 + r^{2d}\sum\left(y_j^d - \frac{\sum y_j^d + \frac{\sum x_i^d}{r^d}}{n+m}\right)^2\right]$$

$$-\left(\frac{n}{n-1}\right)^2\left[\sum\left(x_i^d - \frac{\sum x_i^d}{n}\right)^2\right] - \left(\frac{m}{m-1}\right)^2\left[\sum\left(y_j^d - \frac{\sum y_j^d}{m}\right)^2\right] r^{2d}$$

where *d* satisfies the following highly non-linear system of 7 equations:

$$\frac{n+m}{d} + u - \frac{1}{w}\left[\sum x_i^{2d} \ln x_i - v\sum x_i^d \ln x_i + q^2 \sum y_j^{2d} \ln y_j - vq\sum y_j^d \ln y_j\right] = 0$$

for:

$$u = \sum \ln x_i + \sum \ln y_j \; ; \qquad v = \frac{\sum x_i^d + q\sum y_j^d}{n+m};$$

$$w = \frac{\sum\left(x_i^d - v\right)^2 + \sum\left(qy_j^d - v\right)^2}{n+m} \qquad q = \frac{-s + \sqrt{s^2 - 4pt}}{2p};$$

$$p = \frac{n}{n+m}\sum y_j^{2d} - \frac{n}{(n+m)^2}\left(\sum y_j^d\right)^2; \qquad s = \frac{m-n}{(n+m)^2}\sum x_i^d \sum y_j^d;$$

and $\quad t = \frac{m}{(n+m)^2}\left(\sum x_i^d\right)^2 - \frac{m}{n+m}\sum x_i^{2d}$.

The values of $\hat{r}$ and d can be found by using Microsoft Excel Solver routine with the Newtonian option. The Solver algorithm for the above non-linear equations is as follows:

Solve for: $\hat{r}$ and d

by minimising the objective function:

$$g(r)=\left(\frac{n+m}{n+m-1}\right)^2\left[\sum\left(x_i^d-\frac{\sum x_i^d+r^d\sum y_j^d}{n+m}\right)^2+r^{2d}\sum\left(y_j^d-\frac{\sum y_j^d+\frac{\sum x_i^d}{r^d}}{n+m}\right)^2\right]$$

$$-\left(\frac{n}{n-1}\right)^2\left[\sum\left(x_i^d-\frac{\sum x_i^d}{n}\right)^2\right]-\left(\frac{m}{m-1}\right)^2\left[\sum\left(y_j^d-\frac{\sum y_j^d}{m}\right)^2\right]r^{2d}$$

subject to the constraint:

$$\frac{n+m}{d}+u-\frac{1}{w}\left[\sum x_i^{2d}\ln x_i-v\sum x_i^d\ln x_i+q^2\sum y_j^{2d}\ln y_j-vq\sum y_j^d\ln y_j\right]=0$$

Options: use automatic scaling and assume Newtonian search.

The transformed variables X^d and r^dY^d can then be tested for normality. If the two variables X^d and r^dY^d show evidence of normality, then the solver tool or the programming tool used was able to obtain the necessary values of *d* and *r*. Noting that if variables X and rY have the same distribution, so too do $\frac{1}{r}X$ and Y.

The conversion factor for relating Y to X can be determined by $\frac{1}{\hat{r}}$.

Methods to derive standard errors and confidence intervals for the Kappenman method are yet to be developed. However, in this study, the Kappenman conversion factor was iteratively bootstrapped 1000 times using a resampling method with replacement. The bootstrapping was performed by recording and

running Macros in Excel. The Visual Basic code which was used to run the Macros is presented in the Appendix at the end of the thesis.

Therefore, from these 1000 bootstrap estimates of the Kappenman conversion factors, 95% confidence intervals were calculated by using a percentile method. The construction of these confidence intervals using the percentile method implies that 95 values (95% of the bootstrap estimates) would fall between the lower and upper values of the interval.

Results are presented in Section 5.3.

3.5.1.4. Box-Cox Extended Method

Another method of imposing equal marginal Gaussian probability distribution functions (p.d.f)'s for transformed bivariate variables is an extended Box and Cox (1964) transformation method. In this study, the Box and Cox (1964) was extended further to transform both the response and explanatory variables to normality. The use of this method was also exhibited in the three options mentioned below:

a) univariate transformation of x and y,
b) a common transformation parameter λ for both x and y, and
c) different transformation parameters λ_1, λ_2 for x and y respectively.

Let x = random variable representing CPUA RV *Dr Fridtjof Nansen*
y = random variable representing FRS *Africana*

Option 1: Univariate transformation of x and y

This option transforms variables x and y individually, to v and w. Therefore, for x, this approach seeks to obtain the following likelihood function:

$$L_v = \frac{1}{\sigma_v\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\left(\frac{v-\mu_v}{\sigma_v}\right)^2\right]\right\}$$

and for y,

$$L_w = \frac{1}{\sigma_w\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\left(\frac{w-\mu_w}{\sigma_w}\right)^2\right]\right\}$$

where:

$$v = \frac{x^{\lambda_1}-1}{\lambda_1 \dot{x}^{(\lambda_1-1)}} \quad \text{and} \quad w = \frac{y^{\lambda_2}-1}{\lambda_2 \dot{y}^{(\lambda_2-1)}} \qquad \text{for } \lambda_v \text{ and } \lambda_w \neq 0$$

and:

v and w are transformed variables of x and y respectively.

λ_1 and λ_2 are the transformation parameters for v and w respectively.

$\dot{x}$ and $\dot{y}$ are geometric means of x and y variables respectively.

The transformation parameters λ_v and λ_w are chosen as the values that maximises the likelihood functions L_v and L_w. They can be calculated using the Solver tool in Excel using the following Solver algorithm:

for v:

Solve for λ_v

by maximising the objective function: $LnL_v = -\ln\sigma_v - \ln\pi - \frac{1}{\sigma_v^2}\sum_{i=1}^{n}(v-\mu_v)^2$

subject to the constraint:

$$\frac{d\ln L_v}{d\lambda_v} = -\frac{1}{\sigma_v}\sum_{i=1}^{n}[v-\mu_v]\left[\frac{\lambda_v x^{\lambda_v}\ln x_i - \left(x_i^{\lambda_v}-1\right)\left(1+\lambda_v \ln \dot{x}\right)}{\lambda_v^2\left(\dot{x}^{\lambda_v-1}\right)}\right] = 0$$

and using an automatic scaling option and assuming a Newtonian search from Solver.

and for w:

Solve for λ_w

by maximising the objective function: $LnL_w = -\ln\sigma_w - \ln\pi - \frac{1}{\sigma_w^2}\sum_{i=1}^{n}\left(w-\mu_w\right)^2$

subject to the constraint:

$$\frac{d\ln L_w}{d\lambda_w} = -\frac{1}{\sigma_w}\sum_{i=1}^{n}[w-\mu_w]\left[\frac{\lambda_w x^{\lambda_w}\ln y_i - \left(y_i^{\lambda_w}-1\right)\left(1+\lambda_w \ln \dot{y}\right)}{\lambda_w^2\left(\dot{y}^{\lambda_w-1}\right)}\right] = 0$$

and using an automatic scaling option and assuming a Newtonian search from Solver.

If the transformation parameters are known, the estimates of parameters σ_v^2, σ_w^2, μ_v, and μ_w are found as usual. In this approach, the effect of the mixed product term in the joint pdf is ignored.

Since x and y are both transformed independently to v and w, the ratio of the variances of the v and w can be calculated and explored further to see if the transformation stabilised the variation of the transformed variables to similar. Thereafter a regression line through the origin must be fitted, implying that if v = 0, then $\overline{w} = \overline{v}$. If a straight line regression appears to be a satisfactory fit, the relation $w_i = \hat{\beta}v_i + \varepsilon_i$ is found, where, in the simplest situations, the residuals ε follows $N\left(0,\sigma^2\right)$. The least squares estimate of the slope parameter, β, is:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} v_i w_i}{\sum_{i=1}^{n} v_i^2} = conversion\ factor = \hat{k}$$

The residual mean square is given by:

$$s_{w.v}^2 = \frac{\sum_{i=1}^{n} w_i^2 - \frac{\left(\sum_{i=1}^{n} v_i w_i\right)^2}{\sum_{i=1}^{n} v_i^2}}{n-1}$$

with $(n-1)$ d.f. Confidence limits for β are

$$\hat{\beta} \pm t_{n-1;\alpha/2} s_b$$

where t is read from the t-table with (n-1) degrees of freedom, and

$$s_b = \frac{s_{w.v}}{\sqrt{\sum_{i=1}^{n} v_i^2}}.$$

But before the above is performed, one needs to first test the null hypothesis that the regression line, which is assumed straight, goes through the origin. This hypothesis test can be achieved by performing the following t-test:

$$t_{calc} = \frac{\overline{w} - \hat{\beta}\overline{v}}{s_{w.v}\sqrt{\frac{1}{n} + \frac{\overline{v}^2}{\sum v_i^2}}}$$

with (n-2) degrees of freedom.

Option 2: Transformation using common transformation parameter (λ)

After transforming x and y to v and w respectively, using a common transformation parameter, one seeks to obtain:

$$L = \frac{1}{2\pi\sigma_v\sigma_w\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{v-\mu_v}{\sigma_v}\right)^2 - 2\rho\left(\frac{v-\mu_v}{\sigma_v}\right)\left(\frac{w-\mu_w}{\sigma_w}\right) + \left(\frac{w-\mu_w}{\sigma_w}\right)^2\right]\right\}$$

where:

$$v = \frac{x^{\lambda} - 1}{\lambda \dot{x}^{(\lambda-1)}} \quad \text{and} \quad w = \frac{y^{\lambda} - 1}{\lambda \dot{y}^{(\lambda-1)}} \qquad \text{for } \lambda \neq 0$$

and:

v *and* w are transformed variables of x and y respectively.

λ is the common transformation parameter.

$\dot{x}$ and $\dot{y}$ are geometric means of x and y variables respectively.

The common transformation parameter λ is chosen as the value that maximises the likelihood function $L_{v,w}$. This value can be calculated using the Solver tool in Excel using the following Solver algorithm:

Solve for λ

by maximising the objective function:

$$LnL_{v,w} = -\ln(2\pi) - \ln\sigma_v - \ln\sigma_w - \frac{1}{2}\ln\left(1-\rho^2\right) - \frac{1}{2\sigma_v^2\left(1-\rho^2\right)}\sum_{i=1}^{n}(v-\mu_v)^2 +$$

$$\frac{\rho}{\sigma_v\sigma_w\left(1-\rho^2\right)}\sum_{i=1}^{n}(v-\mu_v)(w-\mu_w) - \frac{1}{2\sigma_w^2\left(1-\rho^2\right)}\sum_{i=1}^{n}(w-\mu_w)^2$$

subject to the constraint:

$$\frac{d\ln L_{v,w}}{d\lambda} = -\frac{1}{\sigma_v^2\left(1-\rho^2\right)}\sum_{i=1}^{n}(v-\mu_v)\frac{dv}{d\lambda} + \frac{\rho}{\sigma_v\sigma_w\left(1-\rho^2\right)}\left[\sum_{i=1}^{n}(w-\mu_v)\frac{dv}{d\lambda} + \sum_{i=1}^{n}(v-\mu)\frac{dw}{d\lambda}\right] -$$

$$\frac{1}{\sigma_w^2\left(1-\rho^2\right)}\sum_{i=1}^{n}(w-\mu_w)\frac{dw}{d\lambda} = 0$$

where $\dfrac{dv}{d\lambda} = \dfrac{\lambda x_i^{\lambda}\ln x_i - \left[x_i^{\lambda} - 1\right]\left[1 + \lambda\ln\dot{x}\right]}{\lambda^2\dot{x}^{\lambda-1}}$

and $\dfrac{dw}{d\lambda} = \dfrac{\lambda y_i^{\lambda}\ln y_i - \left[y_i^{\lambda} - 1\right]\left[1 + \lambda\ln\dot{y}\right]}{\lambda^2\dot{y}^{\lambda-1}}$

and using an automatic scaling option and assuming a Newtonian search from Solver.

If the transformation parameters are known, the estimates of parameters σ_v^2, σ_w^2, μ_v, μ_w, and ρ are found as usual.

We note that the univariate marginals being Gaussian is not a sufficient condition for bivariate normality.

Once *v and w* are found, the regression line through the origin must be fitted. In this case, the regression line through the origin implies that if $v = 0$, then $\overline{w} = 0$. If a straight line regression appears to be a satisfactory fit, the relation

$$w_i = \hat{\beta} v_i + \varepsilon_i$$

is found, where, in the simplest situations, the residuals ε follows $N\left(0, \sigma^2\right)$. The least squares estimate of the slope parameter, β, is:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} v_i w_i}{\sum_{i=1}^{n} v_i^2} = conversion\ factor = \hat{k}$$

The residual mean square is given by:

$$s_{w.v}^2 = \frac{\sum_{i=1}^{n} w_i^2 - \frac{\left(\sum_{i=1}^{n} v_i w_i\right)^2}{\sum_{i=1}^{n} v_i^2}}{n-1}$$

with $(n - 1)$ d.f. Confidence limits for β are

$$\hat{\beta} \pm t_{n-1;\alpha/2} s_b$$

where t is read from the t-table with $(n$-1) degrees of freedom, and

$$s_b = \frac{s_{w.v}}{\sqrt{\sum_{i=1}^{n} v_i^2}}.$$

But before the above is performed, one needs to first test the null hypothesis that the regression line, which is assumed straight, goes through the origin. This hypothesis test can be achieved by performing the following t-test:

$$t_{calc} = \frac{\overline{w} - \hat{\beta}\overline{v}}{s_{w.v}\sqrt{\frac{1}{n} + \frac{\overline{v}^2}{\sum v_i^2}}}$$

with (n-2) degrees of freedom.

Option 3: Transformation using different transformation parameters (λ_1 and λ_2)

After transforming x and y to v and w using different transformation parameters, we seek to obtain:

$$L = \frac{1}{2\pi\sigma_v\sigma_w\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{v-\mu_v}{\sigma_v}\right)^2 - 2\rho\left(\frac{v-\mu_v}{\sigma_v}\right)\left(\frac{w-\mu_w}{\sigma_w}\right) + \left(\frac{w-\mu_w}{\sigma_w}\right)^2\right]\right\}$$

where:

$$v = \frac{x_i^{\lambda_v} - 1}{\lambda_v \dot{x}^{(\lambda_v - 1)}} \quad \text{and} \quad w = \frac{y_i^{\lambda_w} - 1}{\lambda_w \dot{y}^{(\lambda_w - 1)}} \qquad \text{for } \lambda_v \text{ and } \lambda_w \neq 0$$

and:

v *and* w are transformed variables of x and y respectively.

λ_v and λ_w are the distinct transformation parameters for x and y respectively.

$\dot{x}$ and $\dot{y}$ are geometric means of x and y variables respectively.

The common transformation parameter λ is chosen as the value that maximises the likelihood function $L_{v,w}$. The value of λ can be calculated by the Solver tool in Excel using the following Solver algorithm:

Solve for λ_v and λ_w

by maximising the objective function:

$$LnL_{v,w} = -\ln(2\pi) - \ln\sigma_v - \ln\sigma_w - \frac{1}{2}\ln(1-\rho^2) - \frac{1}{2\sigma_v^2(1-\rho^2)}\sum_{i=1}^{n}(v-\mu_v)^2 +$$

$$\frac{\rho}{\sigma_v\sigma_w(1-\rho^2)}\sum_{i=1}^{n}(v-\mu_v)(w-\mu_w) - \frac{1}{2\sigma_w^2(1-\rho^2)}\sum_{i=1}^{n}(w-\mu_w)^2$$

subject to the two constraints:

(1) $$\frac{d\ln L_{v,w}}{d\lambda_v} = -\frac{1}{\sigma_v^2(1-\rho^2)}\sum_{i=1}^{n}(v-\mu_v)\frac{dv}{d\lambda} + \frac{\rho}{\sigma_v\sigma_w(1-\rho^2)}\sum_{i=1}^{n}(w-\mu_v)\frac{dv}{d\lambda} = 0$$ and

(2) $$\frac{d\ln L_{v,w}}{d\lambda_w} = \frac{\rho}{\sigma_v\sigma_w(1-\rho^2)}\sum_{i=1}^{n}(v-\mu)\frac{dw}{d\lambda} - \frac{1}{\sigma_w^2(1-\rho^2)}\sum_{i=1}^{n}(w-\mu_w)\frac{dw}{d\lambda} = 0$$

where $$\frac{dv}{d\lambda_v} = \frac{\lambda_v x_i^{\lambda_v}\ln x_i - \left[x_i^{\lambda_v} - 1\right]\left[1 + \lambda_v \ln \dot{x}\right]}{\lambda_v^2 \dot{x}^{\lambda_v - 1}}$$

and $$\frac{dw}{d\lambda_w} = \frac{\lambda_w y_i^{\lambda_w}\ln y_i - \left[y_i^{\lambda_w} - 1\right]\left[1 + \lambda_w \ln \dot{y}\right]}{\lambda_w^2 \dot{y}^{\lambda_w - 1}}$$

and using an automatic scaling option and assuming a Newtonian search from Solver.

Again, marginal univariate Gaussian distributions alone do not guarantee the desired likelihood. Once the transformation parameters λ_v and λ_w are known, the estimates of parameters σ_v^2, σ_w^2, μ_v, μ_w, and ρ can be found as usual.

Once *v and w* are found, the regression line through the origin must be fitted. In this case, the regression line through origin implies that if v = 0, then $\overline{w} = 0$. If a straight line regression appears to be a satisfactory fit, the relation

$$w_i = \hat{\beta} v_i + \varepsilon_i$$

is found, where, in the simplest situations, the residuals ε follows $N(0, \sigma^2)$. The least squares estimate of the slope parameter, β, is:

$$\hat{\beta} = \frac{\sum_{i=1}^{n} v_i w_i}{\sum_{i=1}^{n} v_i^2} = conversion\ factor = \hat{k}$$

The residual mean square is given by:

$$s_{w.v}^2 = \frac{\sum_{i=1}^{n} w_i^2 - \frac{\left(\sum_{i=1}^{n} v_i w_i\right)^2}{\sum_{i=1}^{n} v_i^2}}{n-1}$$

with $(n-1)$ d.f. Confidence limits for β are

$$\hat{\beta} \pm t_{n-1;\alpha/2} s_b$$

where t is read from the t-table with (n-1) degrees of freedom, and

$$s_b = \frac{s_{w.v}}{\sqrt{\sum_{i=1}^{n} v_i^2}}.$$

But before the conversion factor (which is the slope parameter of the regression line through origin) is calculated, one needs to first test the null hypothesis that the regression line, which is assumed straight due to Gaussian marginals, goes through the origin. This hypothesis test can be achieved by performing the following t-test:

$$t_{calc} = \frac{\overline{w} - \hat{\beta}\overline{v}}{s_{w.v}\sqrt{\frac{1}{n} + \frac{\overline{v}^2}{\sum v_i^2}}}$$

with (n-2) degrees of freedom.

The results of the three methods are presented in Section 5.3.

3.5.2. Summary of the analytical methods

<table>
<tr><th colspan="3">WEIGHTED LEAST SQUARES: Section 3.4.1.1</th></tr>
<tr><th>OLS</th><th>WLS1</th><th>WLS2</th></tr>
<tr><td>$y_i = \beta x_i + \varepsilon_i$</td><td>$\left(\frac{y_i}{\sqrt{x_i}}\right) = \beta\sqrt{x_i} + \varepsilon_i$</td><td>$\frac{y_i}{x_i} = \beta_{02} + \varepsilon_i$</td></tr>
<tr><td>$\beta_0 = \frac{x'y}{\underline{x}'\underline{x}}$</td><td>$\hat{\beta}_{01} = \frac{\bar{y}}{\bar{x}}$</td><td>$\hat{\beta}_{02} = \frac{1}{n}\sum R_i = \frac{1}{n}\sum \frac{y_i}{x_i}$</td></tr>
<tr><th colspan="3">LOG TRANSFORMATIONS: Section 3.4.1.2</th></tr>
<tr><th colspan="2">Inversion of LS after transf</th><th>Inversion of LS (α=0) after transf</th></tr>
<tr><td colspan="2">$Y = e^{\alpha} X^{\beta}(E)$
$\therefore \log Y = \alpha + \beta \log X + \log E$
$\therefore y_i = \alpha + \beta x_i + \varepsilon_i$</td><td>$Y = X^{\beta}(E)$
$\therefore \log Y = \beta \log X + \log E$
$\therefore y_i = \beta x_i + \varepsilon_i$</td></tr>
<tr><td colspan="2">With $\hat{\beta} \cong 1$, we use $\hat{k} = e^{\alpha}$</td><td>$\hat{\beta} = \frac{\underline{y}'\underline{x}}{\underline{x}'\underline{x}}$</td></tr>
<tr><th colspan="3">KAPPENMAN METHOD: Section 3.4.1.3</th></tr>
<tr><th colspan="3">BOX-COX EXTENDED METHOD: Section 3.4.1.4</th></tr>
<tr><th>Option 1</th><th>Option 2</th><th>Option 3</th></tr>
<tr><td>Univariate transf of x and y</td><td>Using a common transf parameter for both x and y</td><td>Using different transf parameters for both x and y</td></tr>
</table>

Table 3.3: Summary of the analytical methods used in the study

3.6 Conversion between FRS *Africana* (old trawl gear) and FRS *Africana* (new trawl gear)

The methods discussed in Section 3.4 are only applicable when one needs to calculate conversion factors between FRS *Africana* (both old and new trawl gear) against RV *Dr Fridtjof Nansen*, as there are data from the comparative trials conducted between the two vessels for the years 2001, 2002, and 2004. These methods cannot be applied to data where comparative trials were not conducted.

The third question to be addressed by the study deals with finding conversion factors to convert catch data of FRS *Africana* (old trawl gear) to the catch data of FRS *Africana* (new trawl gear). There were no comparative trials conducted between the two trawl gears of FRS Africana, therefore a method to calculate conversion factors between the two gears needed to be devised. It was therefore decided by the scientists at Marine and Coastal Management that RV *Dr Fridtjof Nansen* should be used as the bridge vessel to convert data between FRS *Africana* (old trawl gear) to FRS *Africana* (new trawl gear).

Using RV *Dr Fridtjof Nansen* as the bridge vessel implies that the conversion factors calculated from the comparative trials between FRS *Africana* (old and new trawl gear) against RV *Dr Fridtjof Nansen* can be used to estimate conversion factors to convert historical catch data from FRS *Africana* (old trawl data) as if they were catch data from FRS *Africana* (new trawl gear) as follows:

As seen from Section 3.4 for some analytical methods, the concept of conversion factor implies a multiplicative model for the conversion of average catches, for an example, the conversion of the RV *Dr Fridtjof Nansen* catches to FRS *Africana catches*. The conversion factor from the comparative trials between FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* was calculated as follows:

$$Afr_{oldGear} = \hat{k}_1 Nan \qquad (3.5.1)$$

where:

$Afr_{oldGear}$ = FRS *Africana* (old trawl gear) data;

$\hat{k}_1$ = conversion factor; and

Nan = RV *Dr Fridtjof Nansen* data.

And the conversion factor from the comparative trials between FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen* was calculated as follows:

$$Afr_{newGear} = \hat{k}_2 Nan \qquad (3.5.2)$$

where:

$Afr_{newGear}$ = FRS *Africana* (new trawl gear) data

$\hat{k}_2$ = conversion factor

Nan = RV *Dr Fridtjof Nansen* data

Therefore, using (3.5.1) and (3.5.2) to find conversion factors to convert FRS *Africana* (old trawl gear) historical data to FRS *Africana* (new trawl gear), the following method applies:

$$Afr_{newGear} = \frac{\hat{k}_2}{\hat{k}_1} Afr_{oldGear} \qquad (3.5.3)$$

3.7 Conclusion

The main aim of this chapter was to introduce and describe the analytical methods used to analyse the intercalibration data collected by FRS *Africana* (using old and new trawl gear) and RV *Dr Fridtjof Nansen*. And the conditions to be met for the methods to be appropriate and how can they be used were also discussed in detail. The methods were based on regression analysis, with the exception of Kappenman (1992) method and Box-Cox Extended method. The data collection methods were also described at the beginning of this chapter. The chapter went further by discussing how the conversion factors between FRS *Africana* (with old trawl gear) and FRS *Africana* (with new trawl gear) were calculated.

The next chapter discusses the length frequency data collected for the two commercially important South African Hake species, *Merluccius capensis* (Shallow-water hake) and *Merluccius paradoxus* (Deep-water hake).

CHAPTER 4

HAKE STUDY

4.1. Introduction

There are two types of hake species that are caught in the South African waters, namely, *Merluccius capensis*, known as shallow-water Cape Hake, and *Merluccius paradoxus*, known as the deep-water Cape Hake. These two species constitute a major component of the South African trawl fishery, dominating the demersal catches. The two species were discussed in detail in Chapter 1.

Since the study is targeted at both hake species, it is imperative that before conversion factors are calculated for these two species, their size classes (in terms of length in cm) should be explored in order to determine whether or not different conversion factors should be calculated for different size classes. In other words, we are trying to ascertain if there are any size effects in the hake data which will warrant calculation of size based conversion factors. Also, we will explore the sea depth data to see if different conversion factors should be calculated at different depths. This chapter focuses on length class as a proxy for size and the data for both the FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* data and FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen* will be examined.

The length frequency data collected by both vessels consists of the following information:

- Station number
- Trawl number
- Fishing depth (m)
- Date and time trawling took place
- Duration of the trawl (usually 30minutes)

- Catch weight (kg) – for each measured and dumped species of fish
- Sample weight (kg) – for each measured species of fish

Not all the fish caught from trawling are necessarily measured for length. For instance, if the catch of *Merluccius capensis* is relatively small, the Catch Weight would be the same as the Sample Weight. In other words, all the *Merluccius capensis* is weighed and length frequencies taken. However, if the catch of the *Merluccius capensis* is large, a sample that is intended to be representative (both small and large fish) is taken, and although non-random, then weighed and length frequencies are recorded and their total mass is referred to as the Sample Weight. The remainder of the catch is weighed and dumped to the sea. The weight of the dumped catch of *Merluccius capensis* is then added to the weight of the Sample weight of the same species to give what is referred to as the Catch weight.

The length frequencies within the sampled fish of each trawl are then adjusted using the formulae $\left(CatchWt/SampleWt\right) * frequencies$ to estimate for an approximate total number of all fish that were caught in that trawl.

4.2. Exploratory analysis of the hake length frequency data

Merluccius capensis

A 5x2x4 table reporting the 5 length frequencies of *M.capensis* from both vessels across 7 different depth classes is given in Table 4.1. The numbers in the table represents number of fish (*Merluccius capensis*) for the different length classes at each depth class and for each research vessel. It appears that both vessels are successful at catching the small *fish* (length between 21 and 40cm) at depths less than 200m. In shallower waters, it appears that RV *Dr Fridtjof Nansen* was

slightly more efficient than FRS *Africana*, but only for small fish. FRS *Africana* was dominant in catching bigger fish, especially at deeper waters.

			DEPTH (m)				
		VESSEL	[100-200]	[201-300]	[301-400]	[401-500]	TOTAL
LENGTH (cm)	[1-20]	Nan	140	3	0	0	**143**
		Afr	57	1	0	0	**58**
	[21-40]	Nan	3098	1322	6	3	**4429**
		Afr	3017	874	4	4	**3899**
	[41-60]	Nan	145	1052	536	240	**1973**
		Afr	442	2177	1050	275	**3944**
	[61-80]	Nan	8	102	332	193	**635**
		Afr	17	249	840	250	**1356**
	[81-100]	Nan	2	2	27	20	**51**
		Afr	2	9	74	14	**99**
	TOTAL		**6928**	**5791**	**2869**	**999**	**16587**

Table 4.1: Adjusted *M. capensis* length class frequencies from FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* at different depth intervals

The fact that RV *Dr Fridtjof Nansen* was catching lesser number of *M. capensis* than FRS *Africana* at depths greater than 300m suggests that at these depths there may be differences due to the gear of the two vessels.

We also examine some graphical versions of the same table to seek any kind of pattern from which we can make inferences. Figure 4.1 and 4.2 show that in this study both vessels caught higher numbers of smaller *M. capensis* in shallower waters (depth <= 200), but that number caught decreases for larger *M. capensis* in deeper waters.

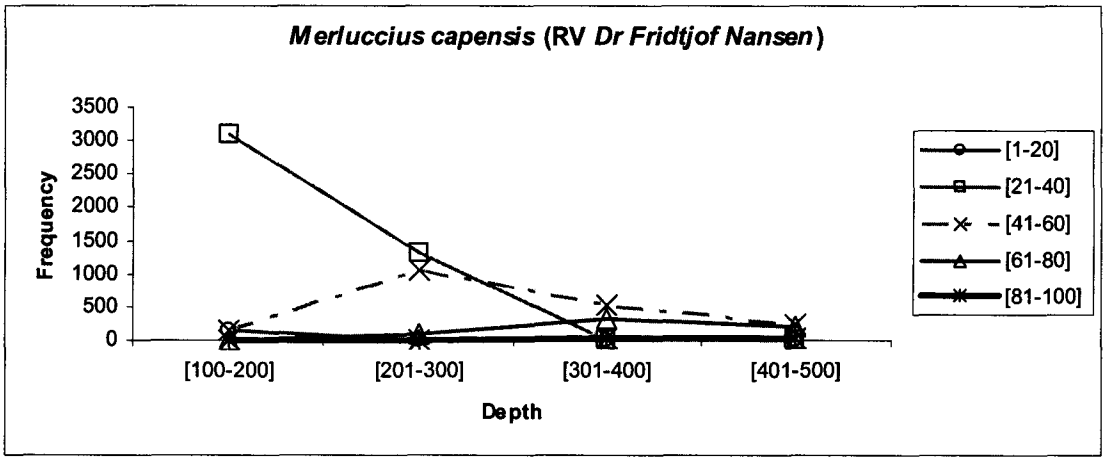


Figure 4.1: Adjusted *M. capensis* frequencies of size classes at each depth class for RV *Dr Fridtjof Nansen*.

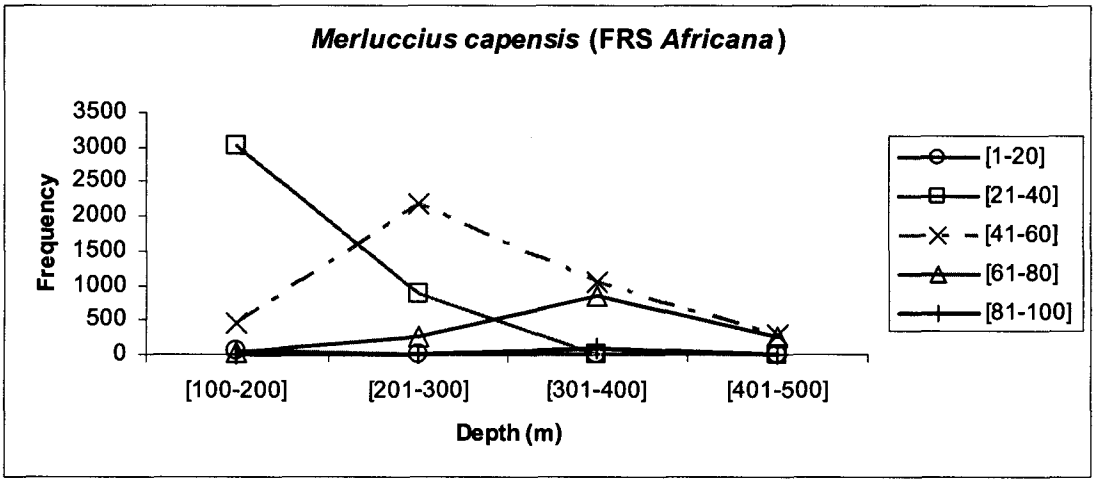


Figure 4.2 Adjusted *M. capensis* frequencies of size classes at each depth class for FRS *Africana*.

Figure 4.3 and Figure 4.4 present ratios of the frequencies for FRS *Africana* to the frequencies for RV *Dr Fridtjof Nansen* and logs of those frequency ratios respectively. Throughout all depths, one can observe that the frequency ratios of the smaller fish (lengths <20cm) was less than one. This outcome confirms the fact that the RV *Dr Fridtjof Nansen* appears to catch the smallest fish more efficiently than the other vessel. This pattern can also be observed in Figure 4.4. The same can be said for lengths between 20 and 40cm, except in the deeper waters, where FRS *Africana* was dominant. All other length sizes had log ratios

greater than or equal to zero, reflecting that FRS *Africana* was more effective at all depths for fish in those catch classes.

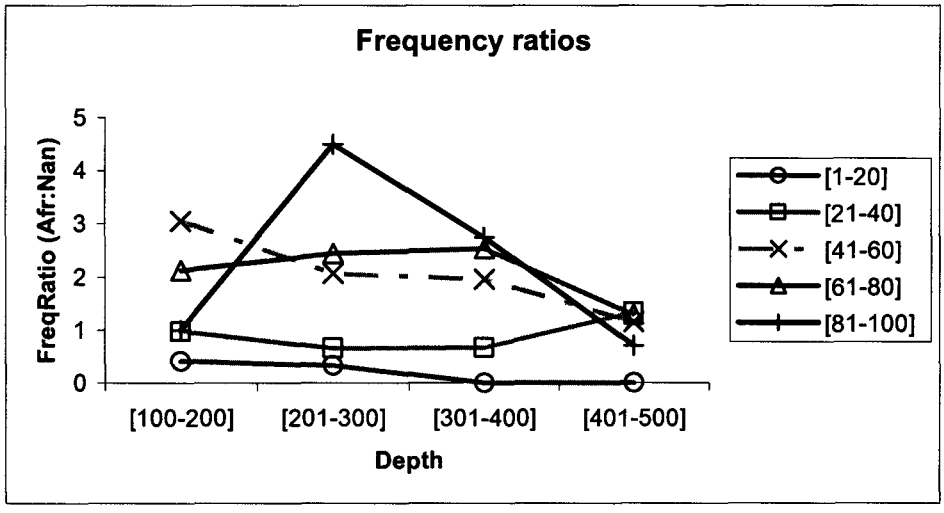


Figure 4.3: Frequency ratios (FRS *Africana* : RV *Dr Fridtjof Nansen*) for *M. capensis*.

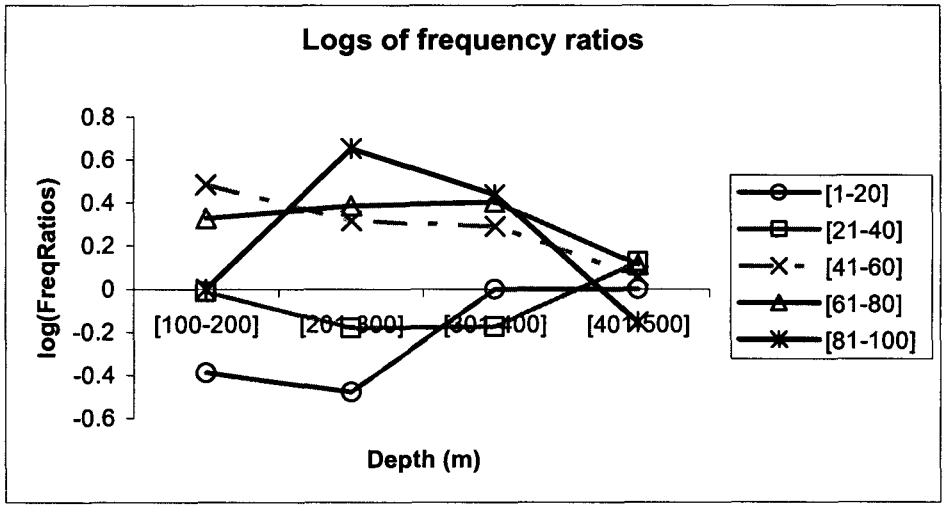


Figure 4.4: Log frequency ratios (FRS *Africana* : RV *Dr Fridtjof Nansen*) for *M. capensis*.

Merluccius paradoxus

A 5x2x5 table reporting length class frequencies of *Merluccius paradoxus* from both vessels across different depths is given in Table 4.2. The numbers in the Table 4.2 represents number of fish (*Merluccius paradoxus*) for the different length classes at each depth class and for each research vessel. A similar approach to that used for *Merluccius capensis* is adopted.

			DEPTH (m)					
	VESSEL		[101-200]	[201-300]	[301-400]	[401-500]	>500	TOTAL
LENGTH (cm)	[1 - 20]	Nan	6136	17442	922	588	42	**25130**
		Afr	9766	26714	1109	77	0	**37666**
	[21 - 40]	Nan	3289	33982	27746	33434	848	**99299**
		Afr	12449	71968	35859	17164	112	**137552**
	[41 - 60]	Nan	5	501	2836	12512	359	**16213**
		Afr	52	3232	5635	10698	228	**19845**
	[61 - 80]	Nan	0	0	193	1689	58	**1940**
		Afr	0	42	450	2583	27	**3102**
	[81 - 100]	Nan	0	0	0	30	3	**33**
		Afr	0	0	18	44	1	**63**
	TOTAL		**31697**	**153881**	**74768**	**78819**	**1678**	**340843**

Table 4.2: Adjusted *M. paradoxus* length class frequencies from FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* at different depth intervals

Figures 4.5 and 4.6 present the same data graphically to exhibit vessel contrasts. Both vessels caught few *M. paradoxus* less than 20cm in length at depths more than 400m. The bigger fish (above 50cm) were caught mainly in deeper waters, and smaller fish in shallow waters.

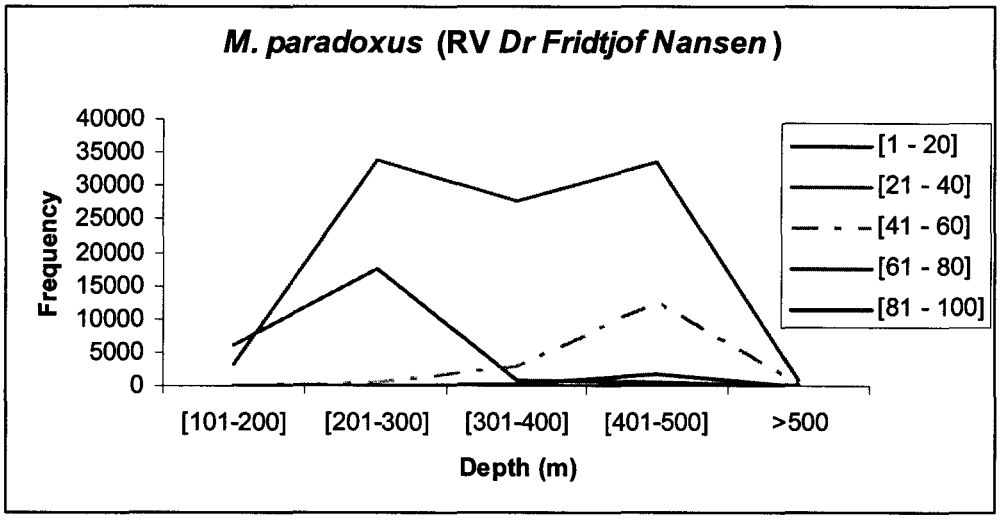


Figure 4.5: Adjusted *M. paradoxus* frequencies of size classes at each depth class for RV *Dr Fridtjof Nansen*.

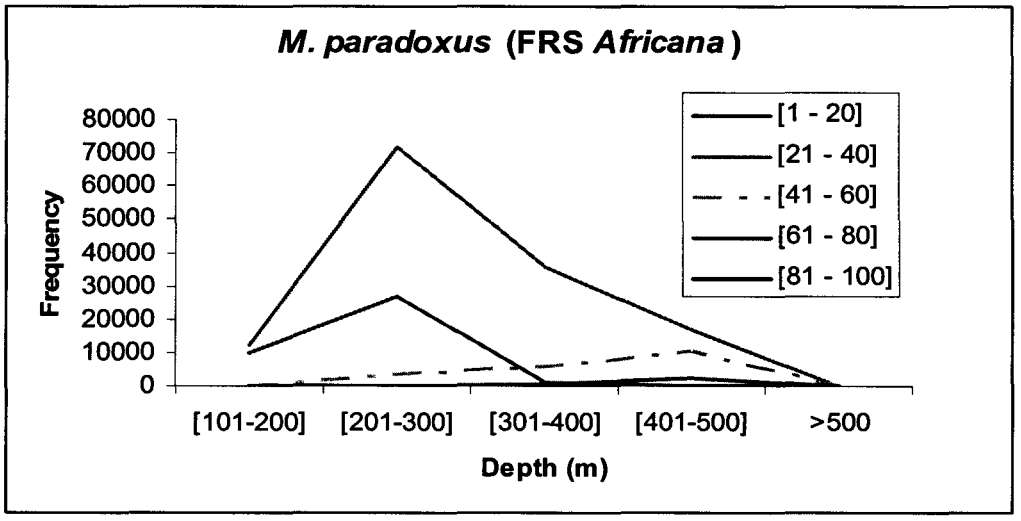


Figure 4.6: Adjusted *M. paradoxus* frequencies of size classes at each depth class for FRS *Africana*

Figure 4.7 presents frequency ratios of FRS *Africana* to RV *Dr Fridtjof Nansen*. FRS *Africana* was more efficient for fish of sizes 41cm to 60cm for depths lesser than 300m. Figure 4.8 reflects the fact that RV *Dr Fridtjof Nansen* was more efficient than FRS *Africana* at catching young fish of sizes less than 40cm at deeper waters.

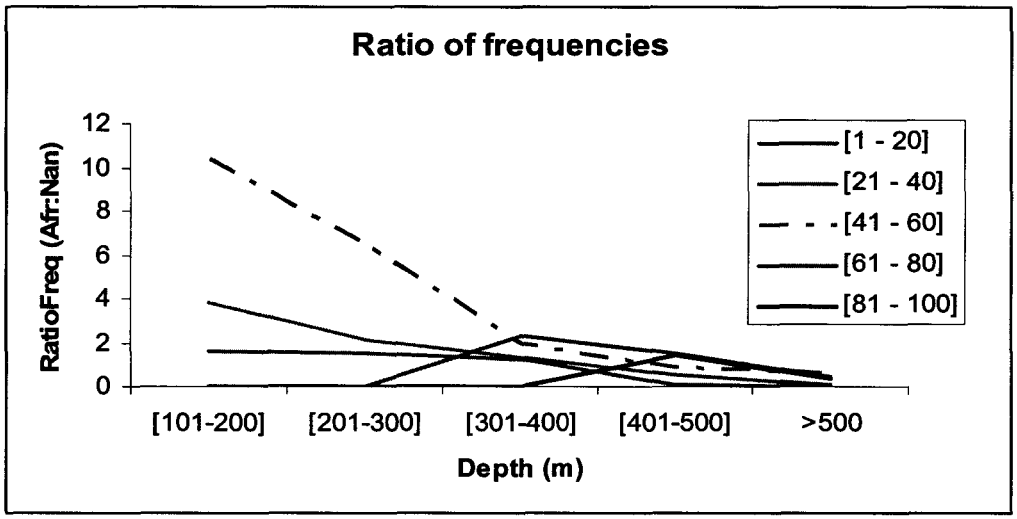


Figure 4.7: Frequency ratios (FRS *Africana* : RV *Dr Fridtjof Nansen*) for *M. paradoxus*.

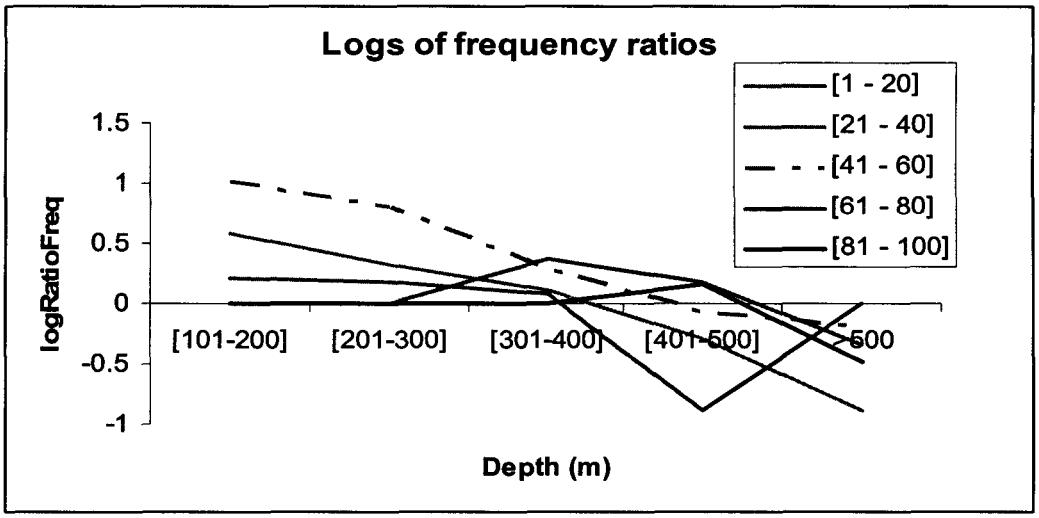


Figure 4.8: Logs of frequency ratios (FRS *Africana* : RV *Dr Fridtjof Nansen*) for *M. paradoxus*.

4.3. Hake Data Modelling

The length frequency data of both vessels were therefore submitted to Statistica 6.1 for log linear analysis. Log linear models treat the cell frequencies as counts distributed as a Poisson random variable. We used log-linear modelling to

determine if there are any relationships existing between the three variables, depth, vessel, and length. For both species, the log-linear model tested is:

$$\log f_{ijk} = \text{constant} + \lambda_i^{\text{depth}} + \lambda_j^{\text{vessel}} + \lambda_k^{\text{length}} + \lambda_{ij}^{\text{depth x vessel}} + \lambda_{ik}^{\text{depth x length}} + \lambda_{jk}^{\text{vessel x length}} \quad (4.3.1)$$

where: i refers to depth class (there are 5 classes); j refers to vessel class (there are two classes), k refers to length class (there are 5 classes), and log f_{ijk} are the logs of the frequencies between the three variables.

The test results of the fit of this model to the frequencies in the Tables 4.1 and 4.2 are as follows:

Model results for *Merluccius capensis*

```
  Table to be analyzed:
                          (1)          (2)          (3)
                        DEPTH       VESSEL       LENGTH
                          4      x     2      x     5

  Minimum cell frequency:       0.   Maximum:     3098.   Sum:
16597.

  Model to be tested: 23,31,33

  Delta: .5000 ;   Maximum iterations:  50 ;  Conv. criterion:
.0100
  Convergence reached after # of iterations:  23
                                              df         p
  Maximum Likelihood Chi-square: 12.816        12    .38259
            Pearson Chi-square: 14.193         12    .28854
```

Box 4.3: Log-linear modelling results for *Merluccius capensis*

The results in Box 4.3 are extracted from the Statistica software for the log linear analysis of frequencies. The Box provides the three variables (depth, vessel, and length) together with the number of classes each variable has. We can observe from Box 4.3 that the model in (4.3.1) fits the data for *M. capensis* (p-value =

with depth, and this finding affects the catching efficiency of the two trawl gears from the two vessels.

Figure 4.9 presents plots of unstandardised residuals (from regression of CPUA FRS *Africana* against CPUA RV *Dr Fridtjof Nansen*) against depth. For *M. capensis*, high values of residuals are observed in the depths between 300m and 400m for old trawl gear plot, and only one high residual point observed in the new trawl gear plot. Consistent residual values are observed in the *M. paradoxus* plot, except for a few high residual points in the depths above 400m. These residual plots clearly indicate that depth has an effect on the catch data, but the evidence is not strong enough to warrant conversion factors for different depths.

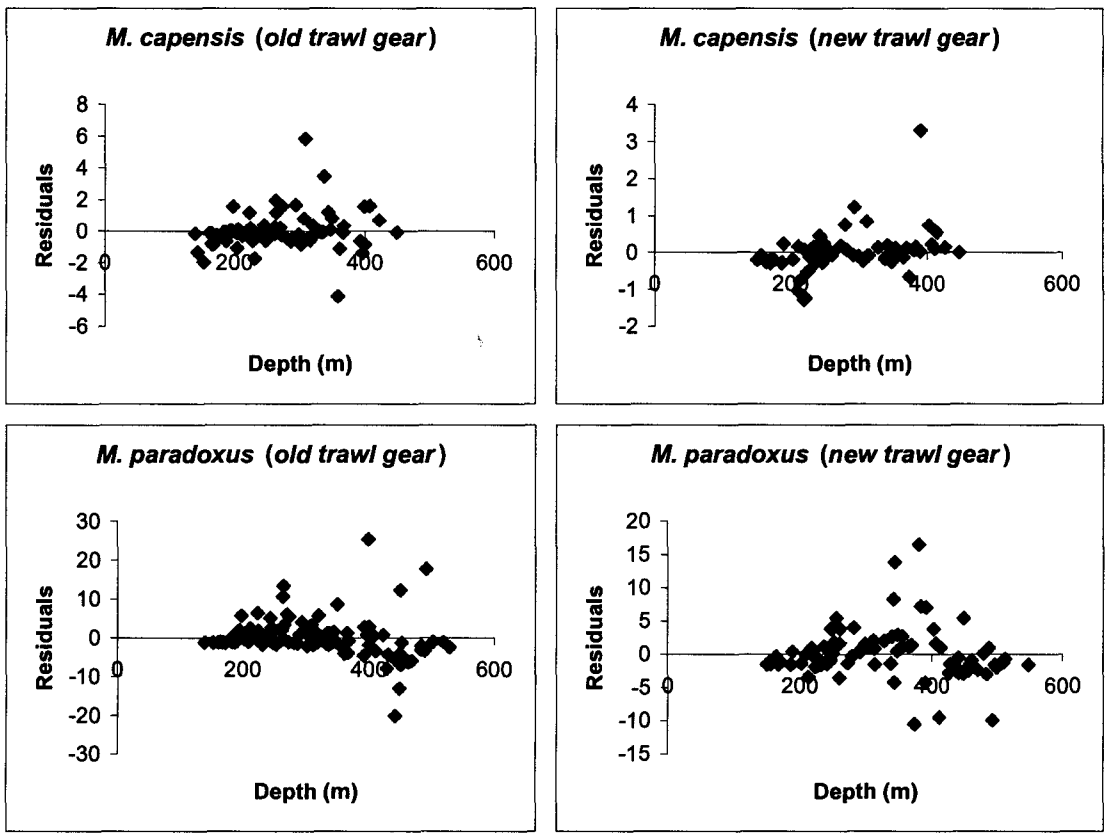


Figure 4.9: Plot of unstandardised residuals against depth for the two hake species, *M. capensis* and *M. paradoxus*. Residuals are from the regression analysis for CPUA FRS *Africana* (old and new trawl gear) and CPUA RV *Dr Fridtjof Nansen*.

0.383), meaning that the categorical model is sufficient to explain the frequencies in Table 4.1.

Model results for *Merluccius paradoxus*

```
Table to be analyzed:
                          (1)          (2)          (3)
                        DEPTH       VESSEL      LENGTH
                          5      x     2      x     5

  Minimum cell frequency:        0.   Maximum:    71968.   Sum:
340843.

  Model to be tested: Constant

  Delta: .5000 ;   Maximum iterations:  50 ;  Conv. criterion:
.0100
  Convergence reached after # of iterations:    0
                                                 df         p
  Maximum Likelihood Chi-square: 8789E2          49    0.0000
            Pearson Chi-square: 1339E3          49    0.0000
```

Box 4.4: Log-linear modelling results for *Merluccius paradoxus*

The results for *M. paradoxus* presented in Box 4.4 shows that the model in (4.3.1) did not fit this species data. This contrast (different results for *M. capensis* and *M. paradoxus*) implies that there is evidence of an interaction between the three variables for *M. capensis*, and none for *M. paradoxus*. In the next section, Conclusion, we will then examine the CPUA residuals against depth to see if any patterns can be observed.

4.4. Conclusion

The objective of this chapter was to uncover relationships that may exist between depth, length, and vessels before any conversion factors are calculated for the two hake species. From the tables and graphs of frequencies and frequency ratios presented in Section 4.2, it is evident that the size of the hake changes

The catch (kg) data from the comparative trawls was standardised to catch per unit area (CPUA) to compensate for the differences in trawling speed of the two research vessels and the fact that FRS *Africana* covers more area than RV *Dr Fridtjof Nansen* (see Section 3.2 of Chapter 3). The CPUA data has units g/m^2, which are the units for residuals in Figure 4.9 above.

Table 4.4 presents rank correlation between the residuals and depth for the two hake species data from FRS *Africana* (old and new trawl gears) against RV *Dr Fridtjof Nansen*. The non-significant p-value of 0.0830 for *M. capensis* (old trawl gear) and 0.4889 for *M. paradoxus* (new trawl gear) implies lack of relationship between the two variables. These findings strengthen the fact that depth does not have strong effect in the catch data.

The significant p-values of 0.000 for *M. capensis* (new trawl gear) and 0.005 for *M. paradoxus* (old trawl gear) implies evidence of a relationship between the residuals and rank, but the low values of Spearman R implies that the relationship, while non zero, is not very strong. This finding therefore is evidence that depth has an effect in the catches, but not a strong effect.

Pair of variables	N	Spearman R	p-value
Merluccius capensis			
Rank Depth & Rank Residuals (Old trawl gear)	83	0.191	0.083
Rank Depth & Rank Residuals (new trawl gear)	74	0.467	0.000
Merluccius paradoxus			
Rank Depth & Rank Residuals (Old trawl gear)	104	-0.276	0.005
Rank Depth & Rank Residuals (new trawl gear)	93	-0.073	0.489

Table 4.3: Rank correlations between residuals and depth

CHAPTER 5

DATA ANALYSIS

5.1. Introduction

Before any kind of analysis, one must explore the descriptive statistics and graphical representation of the data so that an appropriate method of analysis can be chosen. In this chapter, we present both graphical description of the data and some descriptive statistics. We will provide summaries for each of 8 species, beginning from the comparative trials between FRS *Africana* with the old trawl gear and RV *Dr Fridtjof Nansen,* and then the comparative trials between FRS *Africana* with the new trawl gear and RV *Dr Fridtjof Nansen.* We will then perform numerical data analysis using the methods of the previous chapter to explore for any consistency of influences and any noteworthy contrasts.

5.2. Exploratory Data Analysis

5.2.1. FRS *Africana* (old trawl gear) and RV *Dr Fridtjof Nansen*

Figure 5.1 presents the scatterplots of catches standardized to catch density (CPUA in g/m^2) from single trawls, paired by location and time, for all species under study. The raw catches (kg) from the trawls were standardized to density (CPUA) to compensate for the difference in cross section of water trawled by the gear of the two research vessels.

FRS *Africana* trawls at a faster speed than RV *Dr Fridtjof Nansen,* hence the somewhat higher CPUA values from FRS *Africana.* This factor has an effect in the ratio of catches of the two vessels, if they trawl for the same periods of time.

For each species, the scatterplot on the left presents all the data, including pairs where there were zero catches for one vessel for the given species, and the scatter plot on right presents all the data excluding the pairs where one vessel had zero catch for the given species. This strategy is to determine what effect excluding pairs with zero catches for either vessel has on the estimates of the regression line or of a conversion factor.

Where there is evidence of a linear relationship, the slope parameter is not affected, but the intercept does rise above zero after exclusion of these cases where either vessel had a zero catch. In all the scatterplots where the value of r provides a good fit, except for *Helicolenus dactylopterus* and *Zeus capensis*, the value of the coefficient of determination $\left(r^2\right)$ decreases after removal of zero catches. This decrease in r^2 implies that removal of zero catches does not necessarily improve the model fit.

Some scatterplots do not have a regression line, implying that there was no strong evidence for a linear relationship between the catch densities of the two vessels for that species. Regression coefficients and the coefficient of determinations are both reported on the graphs. In almost all the scatter plots, it can be easily concluded that FRS *Africana* clearly had greater catch densities than RV *Dr Fridtjof Nansen*.

From the graph of the CPUA of *Trachurus capensis*, two points can be considered outliers, representing two occasions on which catch of this species was very high on one vessel and low on the other. Removal of these outliers can make a considerable difference to the scatterplot and to the estimates, but because they are correctly recorded values, we do not discard them. In short, we will present the results with their inclusion.

Some of the methods discussed in Chapter 4 require the data from the two vessels to be linearly related. This assumption means that since there are

weaker linear relationships between the two research vessels for the CPUA data of *Genypterus capensis*, *Lophius vomerinus*, and *Trachurus capensis*, these methods cannot be applied for these species. Other methods which do not require data to be linearly related, like the Kappenman (1992)'s method, will then be used for these species.

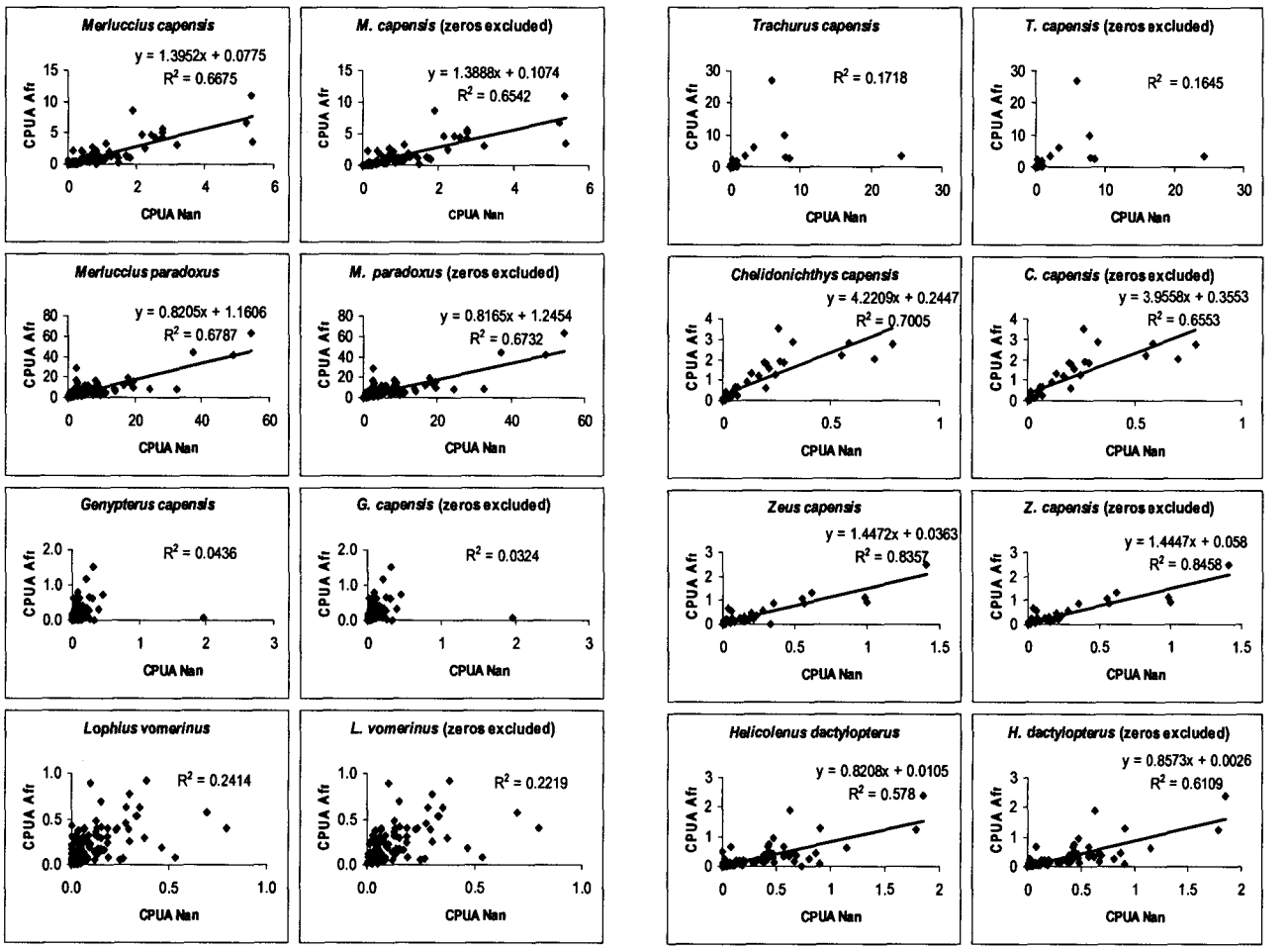


Figure 5.1: Scatterplots of CPUA for FRS *Africana* (old trawl gear) against RV *Dr. Fridtjof Nansen* for all species before and after zero catches were excluded.

Figure 5.2 comprises the histograms of catch density (CPUA in g/m^2) from single trawls of all species under study from both research vessels. Few species, especially from RV *Dr Fridtjof Nansen*, had high percentage of zero catches. A decision was taken to drop the pairs with any zero catches. The exclusion of zero catches, where one vessel caught a particular species and the other vessel did

not catch any of that particular species, could be seen from Figure 5.1 as not having a great impact on the analysis of the data. The R^2 values when zero catches are excluded are not varying much compared to when zero catches are included. These zero catches are (in some cases) a result of one vessel being unable to fish that particular station because of a broken net, and in such instances inclusion of the pairs may introduce bias.

Visually, all the species appear to have CPUA values that are heavily skewed to the right. Low values were considerably more frequent than high values.

The number of paired trawls is reported in the histograms. The skewness in the CPUA data of both research vessels makes it difficult to apply statistical techniques that rely on symmetry or Gaussian density.

Chelidonichthys capensis, also known as the Cape Gurnard, was caught in only 42 trawls (Table 3.1 in Chapter 3), of which 9 trawls had paired zero catch densities. Again, FRS *Africana* had high CPUA values compared to RV *Dr Fridtjof Nansen*, with the maximum CPUA of 3.6 g/m^2 compared to 0.8 g/m^2 for RV *Dr Fridtjof Nansen*. This contrast means that FRS *Africana* was catching more of the bigger Cape Gurnard than its counterpart. This pattern was evident in almost all the other species.

Descriptive statistics of all the species will be discussed later in this chapter. Observing the histograms, one can expect high variation in the data and high skewness and kurtosis. These features necessitate that the data be transformed to permit the use of Gaussian density methods discussed in Chapter 3, which require data to show some evidence of normal distribution and regular variance across the observation range.

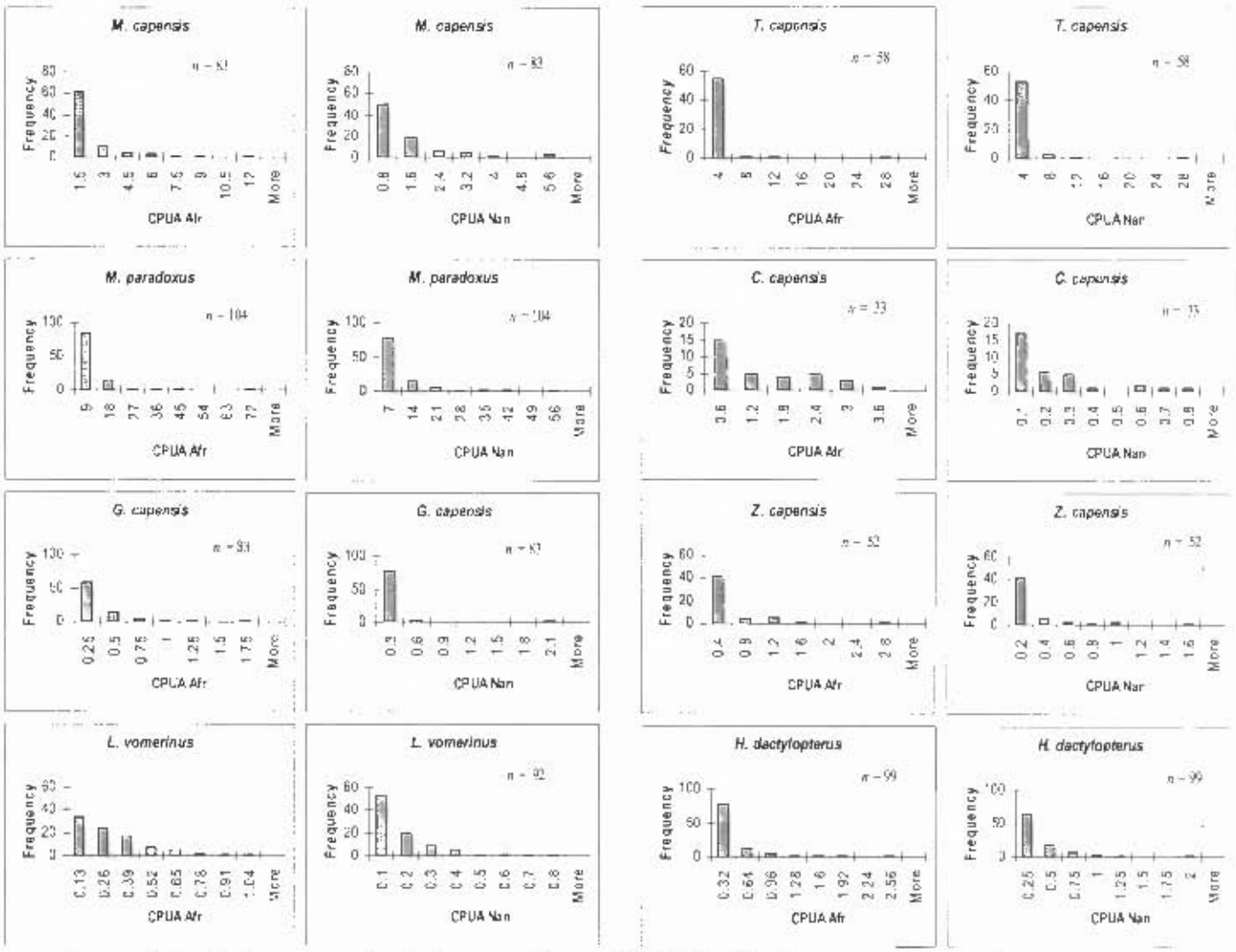


Figure 5.2: Histograms of catch per unit area (CPUA) of both research vessels, for each species

Figure 5.3 presents plots of the difference and ratio of catch per unit area (CPUA) from the two vessels against trawl depths for each species under study. These plots provide a clear picture of the range of depths at which one vessel was dominant over the other for that particular species in terms of CPUA. The negative difference in CPUA reflects the fact that RV *Dr Fridtjof Nansen* had higher catches than FRS *Africana*.

Also, by presenting these plots one seeks to explore if there is a depth effect in the data. In other words, is there evidence to suggest that the conversion factors which must be calculated to convert the data of RV *Dr Fridtjof Nansen* to FRS *Africana* (old trawl gear) should be calculated at different depths?

Merluccius capensis (shallow-water hake) and *Merluccius paradoxus* (deep-water hake) were discussed in detail in the previous chapter, where it was concluded that depth has an effect on the conversion factor, but it is not certain how large that effect is. The CPUA differences between the two vessels were not very high for most of the species, but there were some peaks both positive and negative in deeper waters.

The CPUA ratios clearly show the dominance FRS *Africana* had on RV *Dr Fridtjof Nansen*, especially on deeper waters. RV *Dr Fridtjof Nansen*'s catch for some of the species was very small, resulting in the higher ratios for FRS *Africana*. This contrast is due to the fact that FRS *Africana* trawls at a higher speed than RV *Dr Fridtjof Nansen*. This factor has an effect in the ratio of catches of the two vessels, if they trawl for the same periods of time.

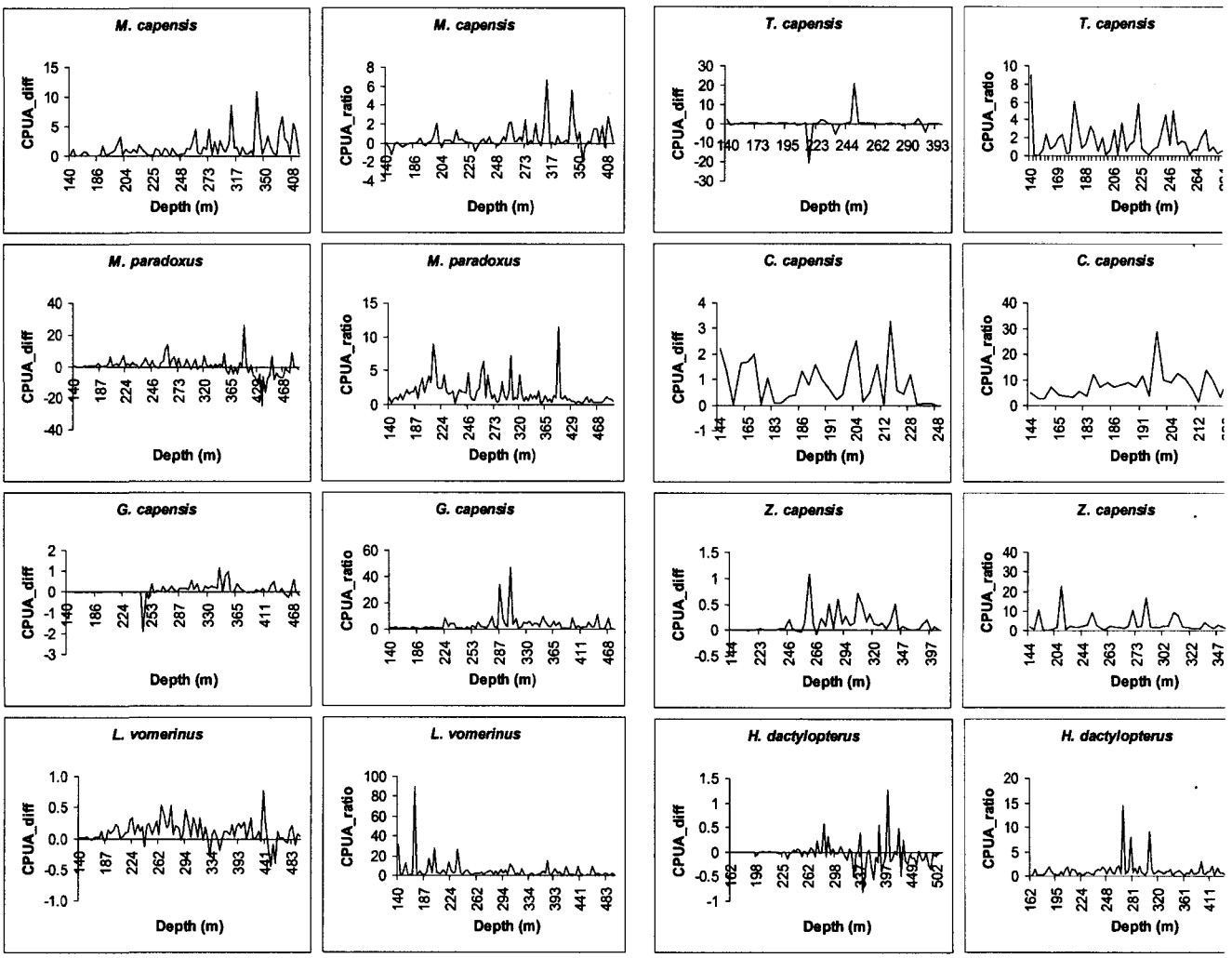


Figure 5.3: Plots of difference (left) and ratio (right) of vessel CPUA against trawl depth, for each species

Table 5.1 and 5.2 presents descriptive statistics of catch per unit area (CPUA) data of both RV *Dr Fridtjof Nansen* and FRS *Africana* (old trawl gear) respectively, for each species. We have seen how the data are distributed; now we need to see what kind of variation we have in the data. Firstly, from both tables, the average CPUA and the median CPUA are far from each other, which might suggest lack of symmetry.

The variance of the CPUA values is smaller for some species, but bigger for others. This smaller variance might be due to the fact that catches (kg) were scientifically standardized to density.

In conclusion, we may clearly infer that the data are not normally distributed and would need to be transformed in order to be able to use the methods presented in Chapter 3.

RV *Dr Fridtjof Nansen*								
	M.cap	***M.par***	***G.cap***	***L.vom***	***T.cap***	***C.cap***	***Z.cap***	***H.dact***
Mean	0.98	5.95	0.11	0.13	1.19	0.17	0.17	0.25
Std Error	0.12	0.91	0.03	0.02	0.48	0.04	0.04	0.03
Median	0.65	2.68	0.05	0.08	0.09	0.07	0.05	0.10
Std dev	1.13	9.23	0.23	0.15	3.64	0.21	0.28	0.34
Variance	1.28	85.20	0.05	0.02	13.22	0.04	0.08	0.12
Range	5.38	54.51	1.96	0.79	24.29	0.78	1.41	1.85
Min	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Max	5.38	54.52	1.96	0.79	24.29	0.78	1.42	1.85
Count	83	104	83	92	58	33	52	99

Table 5.1: Descriptive statistics of CPUA data from RV *Dr Fridtjof Nansen*, for each species.

FRS *Africana*								
	M.cap	***M.par***	***G.cap***	***L.vom***	***T.cap***	***C.cap***	***Z.cap***	***H.dact***
Mean	1.47	6.10	0.19	0.24	1.13	1.04	0.30	0.22
Std Error	0.21	0.90	0.03	0.02	0.50	0.18	0.06	0.04
Median	0.86	3.72	0.06	0.18	0.10	0.64	0.15	0.08
Std dev	1.94	9.19	0.27	0.20	3.81	1.02	0.45	0.37
Variance	3.77	84.36	0.07	0.04	14.55	1.03	0.20	0.14
Range	11.00	63.57	1.51	0.91	26.81	3.52	2.49	2.40
Min	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00
Max	11.00	63.58	1.51	0.92	26.81	3.53	2.49	2.40
Count	83	104	83	92	58	33	52	99

Table 5.2: Descriptive statistics of CPUA data from FRS *Africana* (old trawl gear), for each species.

NB: *M.cap – Merluccius capensis;*
M.par – Merluccius paradoxus;
G.cap – Genypterus capensis;
L.vom – Lophius vomerinus;
T.cap – Trachurus capensis;
C.cap – Chelidonichthys capensis;
Z.cap – Zeus capensis;
H.dact – Helicolenus dactylopterus

5.2.2. FRS *Africana* (new trawl gear) and RV *Dr Fridtjof Nansen*

As in Section 5.2.1, a similar approach will be used to explore the data from comparative trials between FRS *Africana* (new trawl gear) and RV *Dr Fridtjof Nansen*. Figure 5.4 presents scatterplots (some with regression lines) of catch per unit area (g/m^2) for FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen* for each species before and after zero catches were excluded.

As in Section 5.2.1, we observe that where there is evidence of a linear relationship, the slope parameter is not affected, but the intercept does rise above zero after exclusion of zero values. We can also observe that the value of r^2 does not improve after removal of the zero cases, which have low linear regression residuals.

Only *Genypterus capensis* and *Lophius vomerinus* did not exhibit a linear relationship between the CPUA data of the two vessels. For these two species, FRS *Africana* was very dominant and RV *Dr Fridtjof Nansen* had many zero catches for both species. This lack of linearity in the CPUA data of two vessels for the two species can be attributed to the trawl gear difference the two vessels were employing, and is consistent with the results obtained in the previous section for old trawl gear on FRS *Africana* with the exception of only *Trachurus capensis*.

The *Genypterus capensis* scatterplot was influenced by an outlier value, which is very evident in Figure 5.4. Removal of this outlier greatly changes the scatterplot as in Figure 5.5 resulting in the two vessels becoming linearly related.

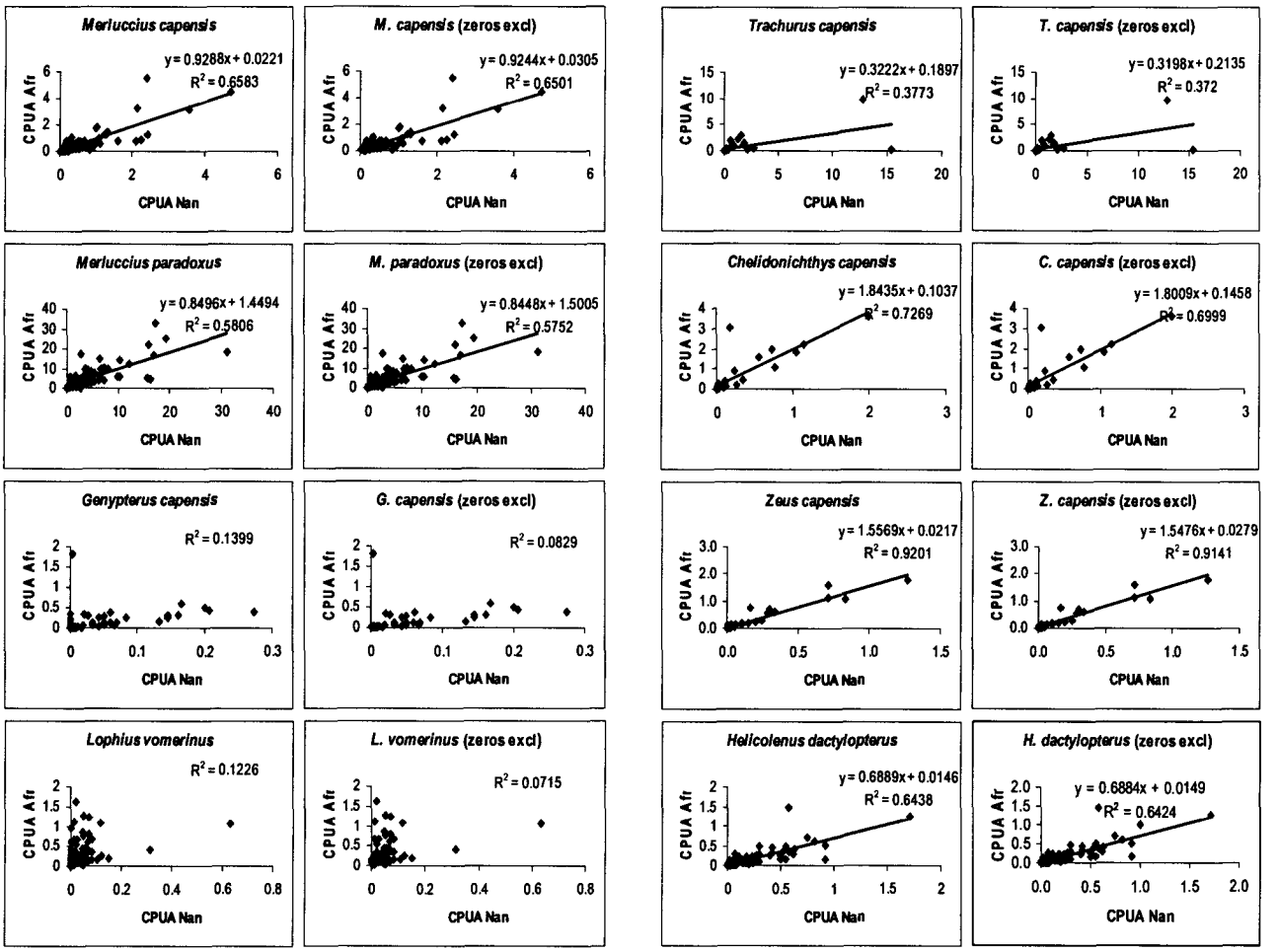


Figure 5.4: Scatterplots of CPUA for FRS *Africana* (new trawl gear) against RV *Dr. Fridtjof Nansen* for all species before and after zero catches were excluded

Because the outlier point (high against the y-axis) in the *Genypterus capensis* scatterplot of Figure 5.4 was a correctly captured data point, it cannot be simply deleted. Its removal as seen in Figure 5.5 confirms the fact that deletion of outliers does indeed have a significant effect on the scatterplot of the two variables.

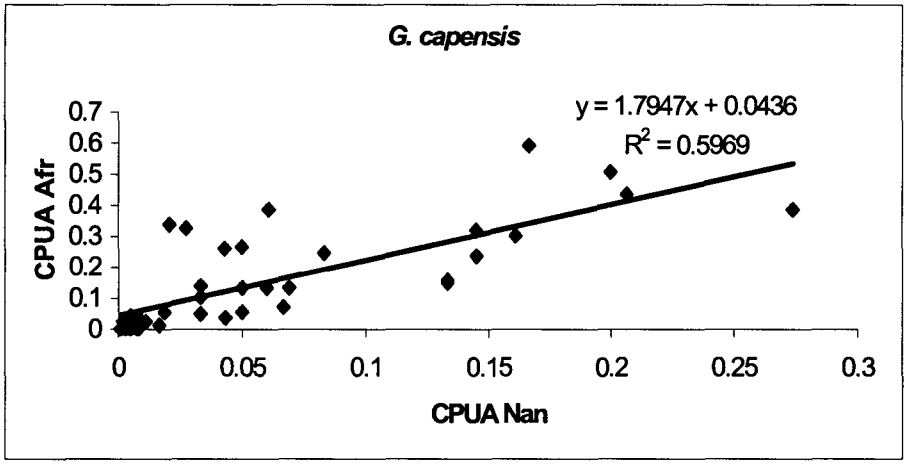


Figure 5.5: Scatterplot of *Genypterus capensis* after the removal of the outlier

Figure 5.6 presents histograms of catch per unit area (CPUA in g/m^2) of both research vessels for each species. The results are also consistent with the old trawl gear results presented by Figure 5.2 in section 5.2.1 in the sense that the data are skewed to the right and low values of CPUA were considerably more frequent than high CPUA values.

The two hake species, *Merluccius capensis* and *Merluccius paradoxus*, the two most important commercial species in South African fisheries industry, were very dominant in respect of mass within the catches, especially *Merluccius paradoxus*. One can observe from the histograms that *Merluccius paradoxus* CPUA went up to $35g/m^2$. CPUA values of $5g/m^2$ occurred most frequently.

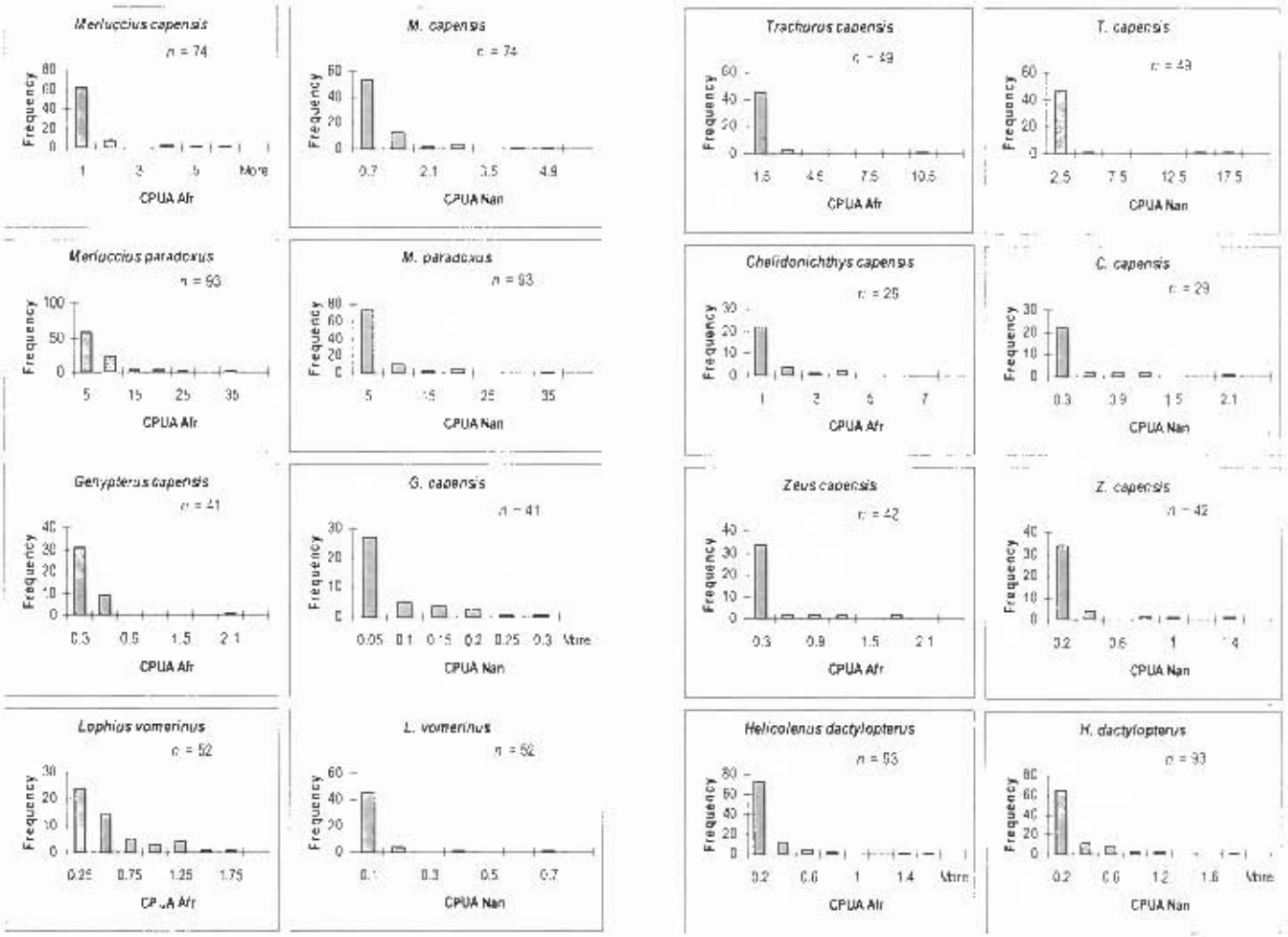


Figure 5.6: Histograms of catch per unit area (CPUA) of both research vessels, for each species

Figure 5.7 presents plots of the difference and ratio of CPUA (g/m^2) for FRS *Africana* (with new trawl gear) and RV *Dr Fridtjof Nansen* against trawl depths for each species. These plots provide a clear picture of the range of depths at which one vessel was dominant over the other for that particular species. The negative difference in CPUA implies RV *Dr Fridtjof Nansen* had higher catches than FRS *Africana*. These negative differences occurred often for *Helicolenus dactylopterus*. A similar feature emerged in the comparative trials when FRS *Africana* was using the old trawl gear.

The difference and ratio plots also exhibit the fact that FRS *Africana* was dominant, in most of the species caught, over RV *Dr Fridtjof Nansen* at almost all depths. Examining the *Genypterus capensis* CPUA difference and ratio plots, it appears that the CPUA values were similar, except for depths between 224m and 299m where there was a huge CPUA ratio favouring FRS *Africana*. The plot of *Trachurus capensis* also provides an interesting observation especially the CPUA difference of -20 g/m^2 at depth 328m.

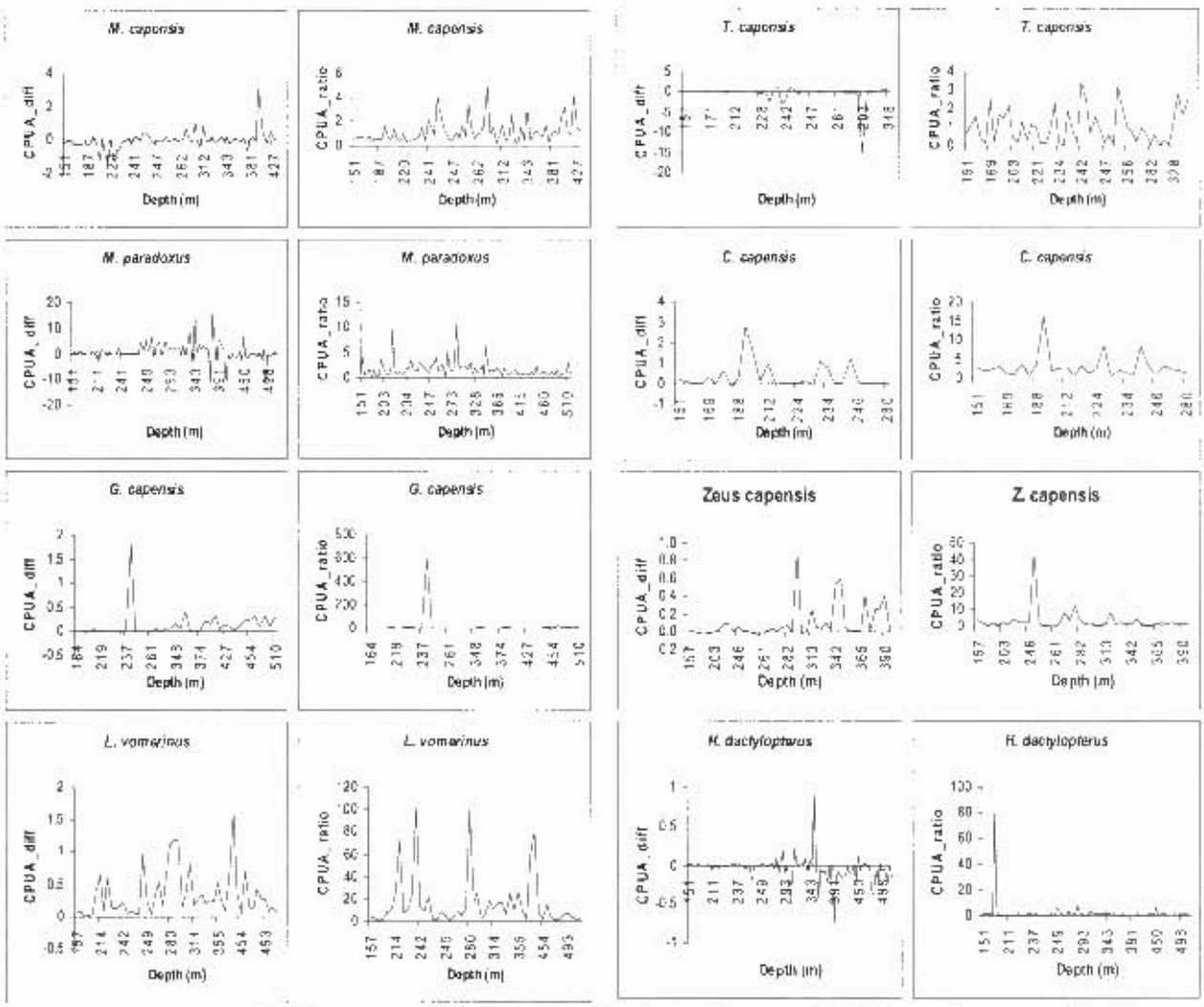


Figure 5.7: Plots of difference (left) and ratio (right) of vessel CPUA against trawl depth, for each species

As in section 5.2.1, we also present descriptive statistics (Tables 5.3 and 5.4) of the paired data from the comparative trial between FRS *Africana* (new trawl gear) and RV *Dr Fridtjof Nansen* for all 8 species under study. The statistics are similar to the statistics in section 5.2.1 with the CPUA variance being reasonably smaller than expected, especially with the high skewness and kurtosis.

Two species, *Merluccius paradoxus* and *Trachurus capensis* have the highest CPUA variances in both tables. This feature can be explained by taking a closer review of the histograms (Figure 5.6) of these two species, and noting the size of the maximum CPUA values in relation to all other values in the data.

As in section 5.2.1, the descriptive statistics in Tables 5.3 and 5.4 provide strong evidence to suggest that the species CPUA data sets deviate completely from normality, and transformation methods need to be applied before conversion factors are calculated by methods invoking Gaussian assumptions.

RV *Dr Fridtjof Nansen*								
	M.cap	***M.par***	***G.cap***	***L.vom***	***T.cap***	***C.cap***	***Z.cap***	***H.dact***
Mean	0.72	4.18	0.06	0.06	0.89	0.29	0.14	0.19
Std error	0.09	0.54	0.01	0.01	0.41	0.08	0.04	0.03
Median	0.44	2.93	0.03	0.04	0.03	0.10	0.02	0.07
Std dev	0.82	5.17	0.07	0.10	2.84	0.46	0.27	0.27
Variance	0.67	26.69	0.00	0.01	8.04	0.21	0.07	0.08
Range	4.69	31.13	0.27	0.63	15.38	1.99	1.26	1.71
Min	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00
Max	4.71	31.15	0.27	0.63	15.38	2.00	1.27	1.71
Count	74	93	41	52	49	29	42	93

Table 5.3: Descriptive statistics of CPUA data from RV *Dr Fridtjof Nansen*, for each species.

5.3.1. FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen*

5.3.1.1. Weighted Least Squares

Before we apply the weighted least squares method, firstly we determine if the data satisfy the condition that the two variables are linearly related. Then we need to test if the intercept is at the origin. Under confirmation of these two conditions we apply this method in order to reduce the variation and the skewness in the data that are submitted for estimation.

From Table 5.5, we observed that only 5 species showed some evidence of linear relationship between the two research vessels. Those species are *Merluccius capensis*, *Merluccius paradoxus*, *Chelidonichthys capensis*, *Zeus capensis*, and *Helicolenus dactylopterus*. Section 3.4.1.1 discusses two methods of weighted least squares in relation to the ordinary least squares, i.e., WLS regression line through origin (1); and WLS regression line through origin (2).

For method 1: after applying the weights: $w_i = \frac{\frac{1}{\sqrt{x_i}}}{\sum\left(\frac{1}{\sqrt{x_i}}\right)}$ to impose the following model: $\left(\frac{y_i}{\sqrt{x_i}}\right) = \beta_{01}\sqrt{x_i} + e_i$ where $\text{var}(e_i) = \sigma^2$; we obtain the following results:

Species (x vs y)	$\hat{\beta}_{01}$ (conv factor)	Std error $\left(\hat{\beta}_{01}\right)$	95% CI for $\hat{\beta}_{01}$
Merluccius capensis	1.50	0.1171	[1.2687; 1.7277]
Merluccius paradoxus	1.03	0.1030	[0.8239; 1.2277]
Chelidonichthys capensis	6.01	0.5531	[4.9231; 7.0913]
Zeus capensis	1.79	0.2088	[1.3821; 2.2007]
Helicolenus dactylopterus	0.87	0.0811	[0.7088; 1.0269]

Table 5.5: Results of WLS regression line through origin (1)

FRS *Africana*								
	M.cap	***M.par***	***G.cap***	***L.vom***	***T.cap***	***C.cap***	***Z.cap***	***H.dact***
Mean	0.69	5.04	0.19	0.40	0.50	0.66	0.24	0.15
Std error	0.11	0.60	0.05	0.05	0.21	0.18	0.07	0.02
Median	0.37	3.45	0.10	0.27	0.03	0.17	0.04	0.06
Std dev	0.94	5.75	0.30	0.37	1.49	0.98	0.44	0.24
Variance	0.88	33.11	0.09	0.14	2.21	0.96	0.19	0.06
Range	5.53	32.63	1.81	1.60	9.76	3.58	1.80	1.47
Min	0.02	0.00	0.00	0.03	0.00	0.02	0.00	0.00
Max	5.54	32.63	1.82	1.63	9.76	3.60	1.80	1.47
Count	74	93	41	52	49	29	42	93

Table 5.4: Descriptive statistics of CPUA data from FRS *Africana* (new trawl gear), for each species.

NB: *M.cap – Merluccius capensis;*
M.par – Merluccius paradoxus;
G.cap – Genypterus capensis;
L.vom – Lophius vomerinus;
T.cap – Trachurus capensis;
C.cap – Chelidonichthys capensis;
Z.cap – Zeus capensis;
H.dact – Helicolenus dactylopterus

5.3. Application of the methods

This section analyses the data from the comparative trials between the FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* and FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen* by using the methods discussed in section 3.4. We present the analyses in the order provided in section 3.4 beginning with the weighted least squares. The following notation will be used for all the methods:

x = RV *Dr Fridtjof Nansen* CPUA (g/m^2) data

y = FRS *Africana* CPUA (g/m^2) data

The last column in Table 5.5 presents confidence intervals for the slope parameter $\hat{\beta}_{01}$. These confidence intervals arise because we apply a method of deriving or describing boundaries for which there is a 95% probability that the true slope of the linear regression through origin, though unknown, lies between the lower and upper boundaries.

The width (the difference between the two confidence limits) of the above confidence intervals is small for all species except for *Chelidonichthys capensis*, with the width of 2.1682. This large confidence width for *Chelidonichthys capensis* is due to the fact that its slope parameter (conversion factor) is far from one, implying that FRS *Africana* was highly efficient in catching this species.

For method 2: after applying the weights: $w_i = \frac{\frac{1}{x_i}}{\sum\left(\frac{1}{x_i}\right)}$ to obtain the following model: $\left(\frac{y_i}{x_i}\right) = R_i = \beta_{02} + e_i$ where $\text{var}(e_i) = \sigma^2$; we obtain the following results:

Species (x vs y)	$\hat{\beta}_{02}$ (conv factor)	Std error $\left(\hat{\beta}_{02}\right)$	95% CI for $\hat{\beta}_{02}$
Merluccius capensis	1.67	0.2053	[1.2626; 2.0674]
Merluccius paradoxus	1.74	0.1815	[1.3877; 2.0991]
Chelidonichthys capensis	7.21	0.8993	[5.4436; 8.9688]
Zeus capensis	3.75	0.7609	[2.2588; 5.2414]
Helicolenus dactylopterus	1.12	0.1826	[0.7668; 1.4827]

Table 5.6: Results of WLS regression line through origin (2)

The results in Table 5.5 appear to be much better compared to the results in Table 5.6 with the low standard errors and smaller interval widths.

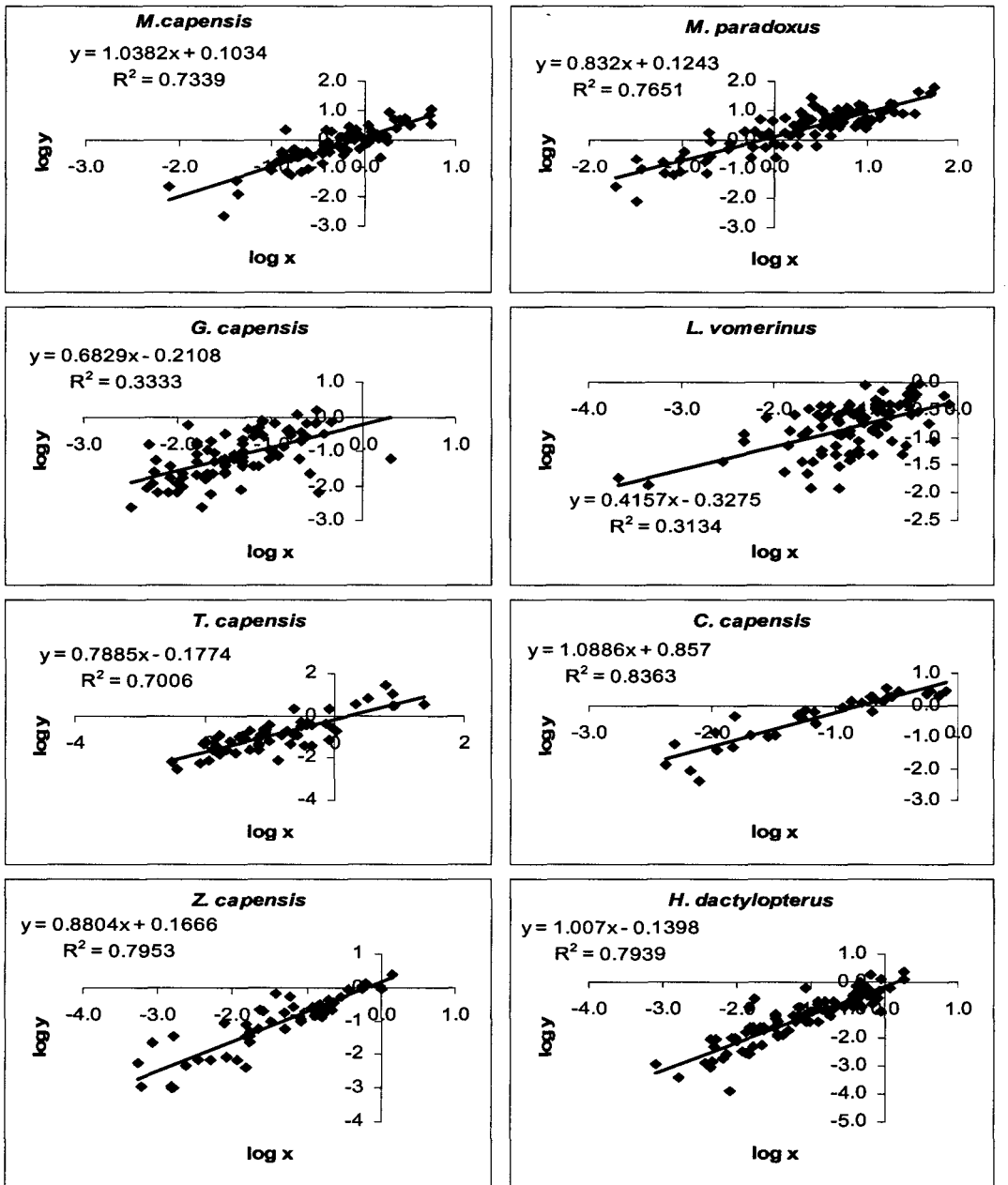


Figure 5.8: Scatterplots of log-transformed CPUA (g/m^2) values of FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen*, for each species.

We also use the Shapiro-Wilks test to investigate whether the log-transformed CPUA values show any evidence against being marginally univariate Gaussian in distribution. Table 5.7 presents the results of the Shapiro-Wilks test for normality for the log-transformed CPUA values for both research vessels.

5.3.1.2. Log transformation

From Section 3.4.1.2 we noted that $X \sim \log Normal(\mu, \sigma^2)$ if $\log X \sim N(\mu_{\ln X}, \sigma^2_{\ln X})$ and $Y \sim \log Normal(\mu, \sigma^2)$ if $\log Y \sim N(\mu_{\ln Y}, \sigma^2_{\ln Y})$. Now, we apply this method only to those species that show evidence of linear relationship in their transformed variables, in order to control for the effects of large catches.

Comparing the scatterplots in Figure 5.8 to the scatterplots in Figure 5.1, it appears that the transformation has increased the R^2 values substantially for all species except for *Zeus capensis*, and all scatterplots appear to have a plausible linear relationship between the two vessels. The transformation seems to have dealt with the outlier values that are evident in the untransformed scatterplots of *Trachurus capensis* and *Genypterus capensis*.

H_0: Data set have a Gaussian distribution

H_1: Data set does not have a Gaussian distribution

Species	Vessel	Shapiro-Wilks W	p-value
Merluccius capensis	Nansen (logx)	0.9725	0.072
	Africana (logy)	0.9556	**0.006**
Merluccius paradoxus	Nansen (logx)	0.9568	**0.002**
	Africana (logy)	0.9359	**0.000**
Genypterus capensis	Nansen (logx)	0.9867	0.554
	Africana (logy)	0.9793	0.200
Lophius vomerinus	Nansen (logx)	0.9083	**0.000**
	Africana (logy)	0.9599	**0.006**
Trachurus capensis	Nansen (logx)	0.9619	0.066
	Africana (logy)	0.9583	0.044
Chelidonichthys capensis	Nansen (logx)	0.9488	0.123
	Africana (logy)	0.8941	**0.004**
Zeus capensis	Nansen (logx)	0.9621	0.097
	Africana (logy)	0.9306	**0.005**
Helicolenus dactylopterus	Nansen (logx)	0.963	**0.007**
	Africana (logy)	0.9637	**0.008**

Table 5.7: Shapiro-Wilks normality tests for all species from both research vessels.

The bold p-values imply statistical significance. It appears that only *Genypterus capensis* and *Trachurus capensis* CPUA values had p-values not significant at 5% level, all other sets of values give evidence that they deviate from a log Gaussian distribution. Therefore the two methods discussed in section 3.4.1.2 of Chapter 3 will be applied only to *Genypterus capensis* and *Trachurus capensis*.

The inversion of the linear regression equation method discussed in Section 3.4.1.2.1, requires that if the slope parameter $\hat{\beta} \cong 1$, then the conversion factor k will be e^{α}. The results of the first method are presented in Table 5.8. Three species which did not meet the requirement of $\hat{\beta} \cong 1$ are excluded from the analysis. Under the condition $\hat{\beta} \cong 1$, the model $Y = e^{\alpha} X^{\beta}(E)$ becomes $Y = e^{\alpha} X(E)$, which implies that e^{α} can be treated as the slope parameter.

Species (x vs y)	Regression line	Is $\hat{\beta} \cong 1$?	e^{α} (Conv factor)	Std error for e^{α}	95% CI for e^{α}
Genypterus capensis	$y = 0.6829x - 0.2108$	No	-	-	-
Trachurus capensis	$y = 0.7885x - 0.1774$	No	-	-	-

Table 5.8: Results of first method in log-transformation

It appears from Table 5.8 that the method did not work well with the data, as there are no species that satisfied the condition $\hat{\beta} \cong 1$. Slope parameters for the two species, *Genypterus capensis* and *Trachurus capensis*, which were 0.6829 and 0.7885 respectively, were not approximately equal to one.

The inversion of the linear regression through origin method, which is the second method under log transformation, discussed in Section 3.4.1.2., requires that we first test the hypothesis that the intercept, from the log linear regression model $\log Y = \beta \log X + \log E$, is zero ($\alpha$ = 0). Table 5.9 presents the results of the hypothesis test (α = 0) for all species except for *Merluccius paradoxus*.

H_0: Intercept is at origin (α = 0)
H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	α	p-value
Merluccius capensis	0.103	0.014
Genypterus capensis	-0.211	0.180
Lophius vomerinus	-0.328	0.000
Trachurus capensis	-0.177	0.048
Chelidonichthys capensis	0.857	0.000
Zeus capensis	0.167	0.118
Helicolenus dactylopterus	-0.140	0.044

Table 5.9: Hypothesis test for zero intercept

Only two species (*Genypterus capensis* and *Zeus capensis*) showed evidence of zero intercept, all other species had p-values that were significant at 5% level. Applying the second method of log transformation to the two species, we obtain results in Table 5.10.

Species	$\hat{\beta}$ (conv factor)	Std error$(\hat{\beta})$	95% CI for $\hat{\beta}$
Genypterus capensis	0.52	0.071	[0.384; 0.663]
Zeus capensis	0.80	0.034	[0.728; 0.862]

Table 5.10: Results of second method in log-transformation

This method, inversion of the linear regression through origin method, produces low standard errors and confidence widths for both species in Table 5.10. However, the slope parameters are somewhat misleading. They (slope parameters) are very small compared to the results in the weighted least squares methods.

5.3.1.3. Kappenman's method

Kappeman (1992)'s method is discussed in section 3.4.1.3 of Chapter 3. In this method, Kappenman (1992) transforms the *X* variable to X^d and the *Y* variable transformed to $r^d Y^d$. This choice means that the two transformed variables (X^d and $r^d Y^d$) must have the same Gaussian distribution. Table 5.11 presents Shapiro-Wilks normality tests of the two variables (X^d and $r^d Y^d$).

H_0: Power-transformed data follows a Gaussian distribution
H_1: Power-transformed data does not follow a Gaussian distribution

Species	Vessel	*d*	Shapiro-Wilks W	p-value
Merluccius capensis	Nansen (X^d)	0.174	0.990	0.758
	Africana (r^dY^d)		0.995	0.993
Merluccius paradoxus	Nansen (X^d)	0.176	0.987	0.430
	Africana (r^dY^d)		0.982	0.173
Genypterus capensis	Nansen (X^d)	0.010	0.986	0.511
	Africana (r^dY^d)		0.980	0.214
Lophius vomerinus	Nansen (X^d)	0.261	0.993	0.894
	Africana (r^dY^d)		0.983	0.293
Trachurus capensis	Nansen (X^d)	-0.122	0.983	0.610
	Africana (r^dY^d)		0.985	0.686
Chelidonichthys capensis	Nansen (X^d)	0.219	0.954	0.171
	Africana (r^dY^d)		0.942	0.077
Zeus capensis	Nansen (X^d)	0.131	0.976	0.370
	Africana (r^dY^d)		0.968	0.177
Helicolenus dactylopterus	Nansen (X^d)	0.131	0.970	0.025
	Africana (r^dY^d)		0.993	0.866

Table 5.11: Shapiro-Wilks normality tests of variables X^d and r^dY^d for each species from both research vessels.

It appears from Table 5.11 that all species (both X^d and r^dY^d) showed evidence of normality, with the exception of *Helicolenus dactylopterus* X^d, for which the p-value for Shapiro-Wilks statistic was significant at 5% level. Table 5.12 presents results of the Kappenman (1992)'s method which were determined using the Microsoft Excel Solver routine with the Newtonian algorithm option.

Species	*d*	r	Conversion factor	95% CI
Merluccius capensis	0.174	0.772	1.301	[1.117; 1.500]
Merluccius paradoxus	0.176	0.877	1.14	[0.935; 1.313]
Genypterus capensis	0.010	0.608	1.64	[1.231; 2.227]
Lophius vomerinus	0.261	0.489	2.04	[1.602; 2.369]
Trachurus capensis	-0.122	0.925	1.08	[0.763; 1.381]
Chelidonichthys capensis	0.219	0.169	5.93	[4.693; 7.282]
Zeus capensis	0.131	0.477	2.10	[1.701; 2.674]

Table 5.12: Results of the Kappenman's method

The Kappenman (1992) method does not produce standard errors. The 95% confidence intervals were calculated by using a percentile method from 1000 bootstrap statistics of Kappenman conversion factors. The conversion factors do correspond to the results from other methods. Figure 5.9 presents 1000 bootstrap estimates of the Kappenman conversion factors. The average of the bootstrap estimates is shown by the broken line, and the lower and upper values of the 95% confidence interval shown by the solid lines. One can notice that the average of the bootstrap estimates is approximately equal to the Kappenman conversion factors presented in Table 5.12 for all species, which implies that those conversion factors can be considered unbiased.

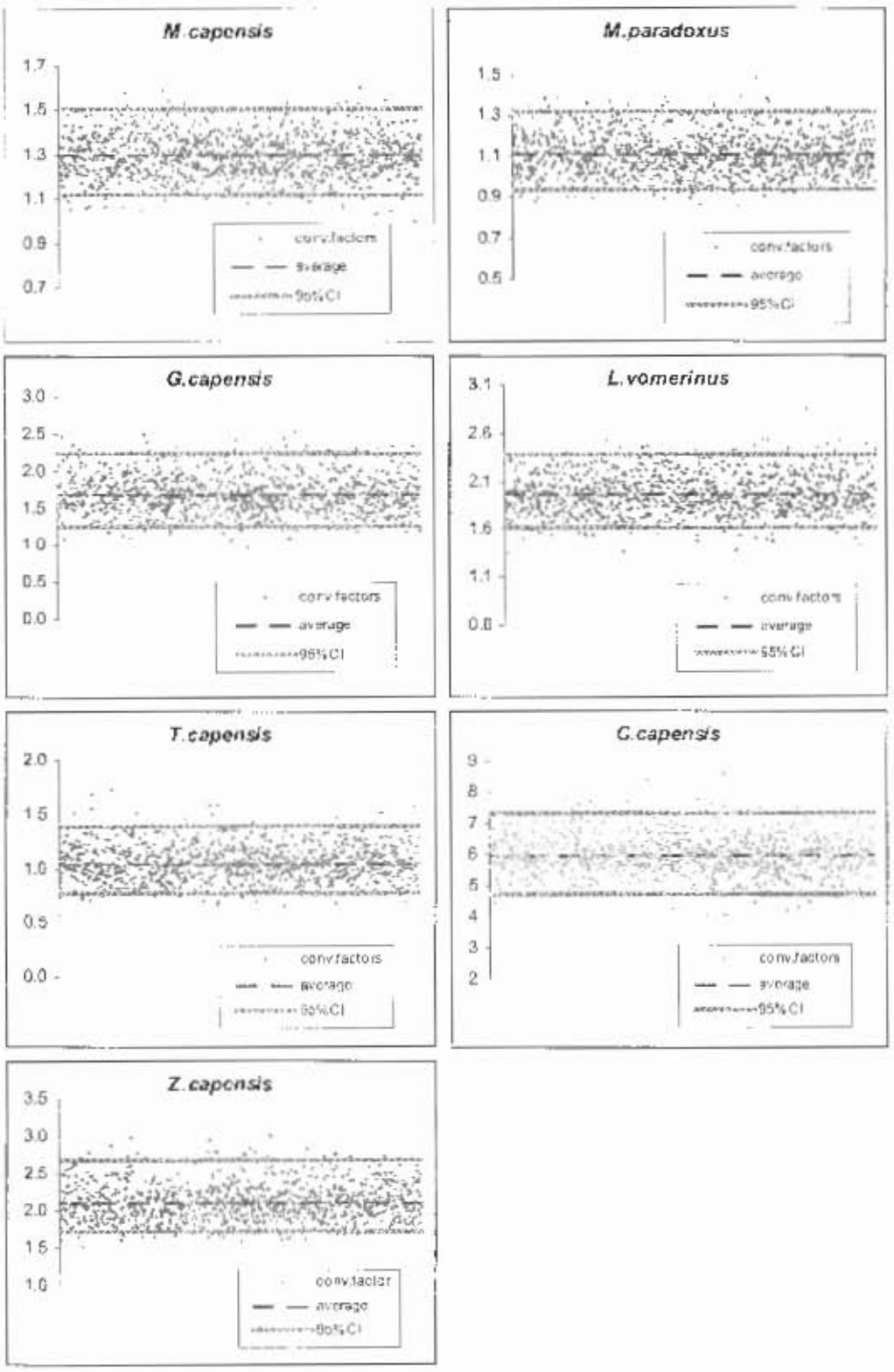


Figure 5.9: 1000 bootstrap estimates of the Kappenman conversion factors of seven species

5.3.1.4. Box-Cox Extension Method

The Box-Cox Extension method is discussed in Section 3.4.1.4. This method enables the original data to be transformed to Gaussian distribution using one of the three options: 1) transformation using a common transformation parameter for both x and y; 2) transformation using different transformation parameters for x and y; and 3) univariate transformation of x and y. After the data have been transformed, the regression analysis through the origin will be applied to the transformed data (v and w) so as to obtain the slope estimate, which then is the conversion factor.

All three options mentioned in the above paragraph will be applied to data of all suite of species. If one recalls from Chapter 3, the likelihood function was obtained and the transformation parameters λ and k were solved iteratively using Microsoft Excel Solver routine with the Newtonian algorithm option.

Option 1: Univariate transformation of x and y

The results in Table 5.13 were obtained by maximising the likelihood functions given in Section 3.4.1.4 for univariate transformation of x and y and solving for transformation parameters, λ_v and λ_w. The table presents transformation parameters λ_v and λ_w; the mean and variance of transformed variables v and w; and the ratio of the two variances ($variance\ ratio = \frac{s_w^2}{s_v^2}$).

The variances have all been stabilised by the transformation, though seemingly high for *Merluccius paradoxus*. However, the size of these variances after transformation is not an appropriate feature to examine. One may prefer to examine the ratio of the variances instead.

The variance ratios in Table 5.13 are all above 1, except for *Trachurus capensis* and *Helicolenus dactylopterus*, showing the high variation in catches from FRS *Africana* compared to the catches of RV *Dr Fridtjof Nansen*.

Species	**Trans. variables**	**Trans. par**	μ	s^2	***Variance ratio***
Merluccius capensis	v	0.154	-0.292	0.402	2.12
	w	0.177	-0.163	0.853	
Merluccius paradoxus	v	0.154	1.825	11.632	1.19
	w	0.199	2.548	13.809	
Genypterus capensis	v	-0.051	-0.132	0.003	4.33
	w	0.047	-0.202	0.013	
Lophius vomerinus	v	0.218	-0.243	0.007	3.29
	w	0.236	-0.344	0.023	
Trachurus capensis	v	-0.12	-0.257	0.059	0.95
	w	-0.123	-0.259	0.056	
Chelidonichthys capensis	v	0.109	-0.211	0.012	36.25
	w	0.294	-0.243	0.435	
Zeus capensis	v	0.096	-0.142	0.006	4.33
	w	0.17	-0.233	0.026	
Helicolenus dactylopterus	v	0.111	-0.225	0.021	0.62
	w	0.14	-0.191	0.013	

Table 5.13: Transformation results of Box-Cox Extension method Option 1

Table 5.14 presents Shapiro-Wilks normality tests to determine if the transformed variables show any evidence against marginal Gaussian distributions.

H_0: Power-transformed data set follows a Gaussian distribution

H_1: Power-transformed data set does not follow a Gaussian distribution

Species	Transformed variables	Shapiro-Wilks W	p-value
Merluccius capensis	v	0.990	0.777
	w	0.995	0.993
Merluccius paradoxus	v	0.987	0.409
	w	0.983	0.203
Genypterus capensis	v	0.988	0.653
	w	0.980	0.237
Lophius vomerinus	v	0.992	0.822
	w	0.983	0.267
Trachurus capensis	v	0.983	0.607
	w	0.985	0.687
Chelidonichthys capensis	v	0.956	0.192
	w	0.947	0.110
Zeus capensis	v	0.976	0.361
	w	0.973	0.279
Helicolenus dactylopterus	v	0.971	0.026
	w	0.993	0.881

Table 5.14: Shapiro-Wilks test results

It appears from Table 5.14 that all species (both v and w) showed evidence of normality, with the exception of *Helicolenus dactylopterus* v, for which the p-value for Shapiro-Wilks statistic was significant at 5% level. Therefore, due to deviation from normality at 5% level of significance, *Helicolenus dactylopterus* was excluded from the analysis.

The next step is to perform regression analysis through the origin, but only after the hypothesis test $\alpha = 0$ has been conducted for all species that satisfied the normality tests in Table 5.14. The results of the hypothesis tests ($\alpha = 0$) are presented in Table 5.15, which shows test statistics values t_{calc} and their associated p-values. The test statistic t-calc is calculated as described in Section 3.4.1.4 of Chapter 3. The d.f. in brackets are the degrees of freedom (n-2) used when calculating the t-test statistic.

H_0: Intercept is at origin ($\alpha = 0$)

H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	t_{calc}	p-value (d.f)
Merluccius capensis	2.738	0.008 (81)
Merluccius paradoxus	2.916	0.004 (102)
Genypterus capensis	0.605	0.547 (81)
Lophius vomerinus	0.651	0.517 (90)
Trachurus capensis	1.115	0.270 (56)
Chelidonichthys capensis	1.709	0.098 (31)
Zeus capensis	0.558	0.579 (50)

Table 5.15: Hypothesis tests for $\alpha = 0$

The null hypothesis, H_0, was rejected in only two species, namely, *Merluccius capensis* and *Merluccius paradoxus* (significant p-values at 0.05 level). The reason for rejection of H_0 for these two species could be the fact that the small values of v and w do not correlate with each other. Therefore, these two species will be excluded from further analysis of this method.

Table 5.16 presents the results of the regression line through the origin for the 5 species that satisfied the condition $\alpha = 0$.

Species	$\hat{\beta}$(conv factor)	95% CI for $\hat{\beta}$
Genypterus capensis	1.47	[1.357; 1.574]
Lophius vomerinus	1.36	[1.251; 1.461]
Trachurus capensis	0.91	[0.804; 1.015]
Chelidonichthys capensis	2.10	[1.348; 2.855]
Zeus capensis	1.70	[1.574; 1.815]

Table 5.16: Results of Box-Cox Extended Method for Option 1

Option 2: Common transformation parameter

The results in Table 5.17 were obtained by maximising the likelihood function given in Section 3.4.1.4 of Chapter 3 and solving for transformation parameter, λ. The table presents common transformation parameter λ; the mean and variance of transformed variables v and w; and correlation, r, between the two transformed variables.

Species	**Trans. variables**	**Trans. par**	μ	σ^2	r
Merluccius capensis	v	0.155	-0.292	0.402	0.86
	w		-0.176	0.854	
Merluccius paradoxus	v	0.107	1.711	11.724	0.85
	w		2.350	14.245	
Genypterus capensis	v	-0.022	-0.137	0.003	**0.58**
	w		-0.190	0.013	
Lophius vomerinus	v	0.240	-0.246	0.007	**0.55**
	w		-0.346	0.023	
Trachurus capensis	v	-0.081	-0.256	0.059	0.83
	w		-0.260	0.056	
Chelidonichthys capensis	v	0.199	-0.234	0.012	0.92
	w		-0.278	0.444	
Zeus capensis	v	0.166	-0.156	0.006	0.91
	w		-0.232	0.026	
Helicolenus dactylopterus	v	0.111	-0.225	0.021	0.89
	w		-0.185	0.013	

Table 5.17: Transformation results of Box-Cox Extension method Option 2

It is evident from Table 5.17 that the two transformed variables v and w are linearly related for all species, but that relationship is not very strong for the two species Genypterus *capensis* and *Lophius vomerinus* (the bolded *r*-values). The variances have all been stabilised by the transformation, though slightly high for *Merluccius paradoxus*. Table 5.18 presents Shapiro-Wilks normality tests to determine if the transformed variables show any evidence of Gaussian distribution.

H_0: Power-transformed data set follows a Gaussian distribution

H_1: Power-transformed data set does not follow a Gaussian distribution

Species	Transformed variables	Shapiro-Wilks W	p-value
Merluccius capensis	v	0.990	0.777
	w	0.995	0.987
Merluccius paradoxus	v	0.987	0.409
	w	0.983	0.203
Genypterus capensis	v	0.988	0.623
	w	0.978	0.160
Lophius vomerinus	v	0.993	0.890
	w	0.983	0.273
Trachurus capensis	v	0.981	0.475
	w	0.981	0.499
Chelidonichthys capensis	v	0.955	0.182
	w	0.940	0.066
Zeus capensis	v	0.974	0.302
	w	0.973	0.268
Helicolenus dactylopterus	v	0.971	0.026
	w	0.991	0.767

Table 5.18: Shapiro-Wilks test results

As in Option 1, it appears from Table 5.18 that all species (both v and w) showed evidence of normality, with the exception of *Helicolenus dactylopterus* v, for which the p-value for Shapiro-Wilks statistic was significant at 5% level. Therefore, due to deviation from normality at 5% level of significance, *Helicolenus dactylopterus* was excluded from the analysis.

As was done in Option 1, the next step is performing regression analysis through the origin, but only after the hypothesis test $\alpha = 0$ has been conducted for all species that satisfied the normality tests in Table 5.18. The results of this hypothesis tests ($\alpha = 0$) are presented in Table 5.19.

Option 3: Different transformation parameters

The results in Table 5.21 were obtained by maximising the likelihood function given in Section 3.4.1.4 of Chapter 3 and solving for transformation parameters, λ and k. The table presents transformation parameters λ and k of v and w respectively; the mean and variance of transformed variables v and w; and correlation, r, between the two transformed variables.

Species	**Transformed variables**	***Trans. par***	μ	σ^2	r
Merluccius capensis	v	0.139	-0.296	0.402	0.85
	w	0.161	-0.172	0.854	
Merluccius paradoxus	v	0.08	1.644	11.866	0.86
	w	0.128	2.395	14.069	
Genypterus capensis	v	-0.091	-0.126	0.003	**0.59**
	w	0.039	-0.201	0.013	
Lophius vomerinus	v	0.235	-0.244	0.007	**0.56**
	w	0.259	-0.352	0.023	
Trachurus capensis	v	-0.081	-0.256	0.059	0.83
	w	-0.081	-0.260	0.056	
Chelidonichthys capensis	v	0.04	-0.197	0.012	0.94
	w	0.295	-0.243	0.435	
Zeus capensis	v	0.128	-0.148	0.006	0.91
	w	0.205	-0.240	0.026	
Helicolenus dactylopterus	v	0.082	-0.220	0.021	0.89
	w	0.119	-0.187	0.013	

Table 5.21: Transformation results of Box-Cox Extension method Option 3

It is evident from Table 5.21 that the two transformed variables v and w are linearly related for all species, but that relationship is not very strong for the two species Genypterus *capensis* and *Lophius vomerinus* (the bolded r-values). The variances have all been stabilised by the transformation, though slightly high for *Merluccius paradoxus*. Table 5.22 presents Shapiro-Wilks normality tests to determine if the transformed variables show any evidence of Gaussian distribution.

H_0: Intercept is at origin ($\alpha = 0$)

H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	t_{calc}	**p-value (d.f)**
Merluccius capensis	2.580	0.012 (81)
Merluccius paradoxus	2.652	0.009 102)
Genypterus capensis	0.387	0.700 (81)
Lophius vomerinus	0.586	0.560 (90)
Trachurus capensis	1.115	0.270 (56)
Chelidonichthys capensis	1.530	0.136 (31)
Zeus capensis	0.821	0.416 (50)

Table 5.19: Hypothesis tests for $\alpha = 0$

Similar conclusions as in Option 1 can be made from the results of Table 5.19, in the sense that the null hypothesis, H_0, was rejected in only two species, namely, *Merluccius capensis* and *Merluccius paradoxus* (significant p-values at 0.05 level).

Table 5.20 presents the results of the regression line through the origin for the 5 species that satisfied the condition of zero intercept.

Species	$\hat{\beta}$**(conv factor)**	**95% CI for** $\hat{\beta}$
Genypterus capensis	1.35	[1.213; 1.488]
Lophius vomerinus	1.36	[1.259; 1.469]
Trachurus capensis	0.91	[0.81; 1.018]
Chelidonichthys capensis	1.99	[1.276; 2.70]
Zeus capensis	1.57	[1.448; 1.682]

Table 5.20: Conversion factors from Box-Cox Extended Method Option 2

H_0: Power-transformed data set follows a Gaussian distribution

H_1: Power-transformed data set does not follow a Gaussian distribution

Species	Transformed variables	Shapiro-Wilks W	p-value
Merluccius capensis	v	0.990	0.769
	w	0.995	0.990
Merluccius paradoxus	v	0.979	0.090
	w	0.976	0.060
Genypterus capensis	v	0.987	0.586
	w	0.980	0.237
Lophius vomerinus	v	0.992	0.880
	w	0.983	0.292
Trachurus capensis	v	0.981	0.474
	w	0.981	0.500
Chelidonichthys capensis	v	0.952	0.155
	w	0.947	0.110
Zeus capensis	v	0.976	0.373
	w	0.975	0.324
Helicolenus dactylopterus	v	0.970	0.023
	w	0.992	0.816

Table 5.22: Shapiro-Wilks test results

It appears from Table 5.22 that all species (both v and w) showed evidence of normality, with the exception of *Helicolenus dactylopterus* v-variable, for which the p-value for Shapiro-Wilks statistic was significant at 5% level. Therefore, due to deviation from normality at 5% level of significance, *Helicolenus dactylopterus* was excluded from the analysis.

As was done for the first two Options, regression analysis through the origin is performed, but only after the hypothesis test $\alpha = 0$ has been conducted for all species that satisfied the normality tests in Table 5.22. The results of the hypothesis tests ($\alpha = 0$) are presented in Table 5.23.

H_0: Intercept is at origin ($\alpha = 0$)

H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	t_{calc}	p-value (d.f)
Merluccius capensis	2.676	0.009 (81)
Merluccius paradoxus	3.152	0.002 (102)
Genypterus capensis	0.712	0.478 (81)
Lophius vomerinus	0.634	0.528 (90)
Trachurus capensis	1.115	0.270 (56)
Chelidonichthys capensis	1.817	0.079 (31)
Zeus capensis	0.609	0.545 (50)

Table 5.23: Hypothesis tests for $\alpha = 0$

The null hypothesis, H_0, was rejected in only two species, namely, *Merluccius capensis* and *Merluccius paradoxus*. The reason for rejection of H_0 for these two species could be the fact that the small values of v and w do not correlate with each other. Therefore, these two species will be excluded from further analysis of this method.

Table 5.24 presents the results of the regression line through the origin for the 5 species that satisfied the condition $\alpha = 0$.

Species	$\hat{\beta}$(conv factor)	95% CI for $\hat{\beta}$
Genypterus capensis	1.51	[1.363; 1.661]
Lophius vomerinus	1.39	[1.287; 1.499]
Trachurus capensis	0.91	[0.810; 1.018]
Chelidonichthys capensis	2.28	[1.504; 3.047]
Zeus capensis	1.68	[1.560; 1.793]

Table 5.24: Results of Box-Cox Extended Method for Option 3

5.3.1.5. Summary of results of FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen*

A summary of the results of all the methods applied to the comparative trials data between FRS *Africana* (old trawl gear) and RV *Dr Fridtjof Nansen* is presented in Table 5.25.

Merluccius capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	1.50	0.12	1.27	1.73	0.46
WLS2	1.67	0.21	1.26	2.07	0.80
Log1	-	-	-	-	-
Log2	-	-	-	-	-
Kappenman	1.30	-	1.09	1.50	0.41
Box-Cox1	-	-	-	-	-
Box-Cox2					
Box-Cox3					

Merluccius paradoxus

	k	s.e	95% Lower	95% Upper	CI width
	1.03	0.10	0.82	1.23	0.40
	1.74	0.18	1.39	2.10	0.71
	-				
	-	-	-	-	-
	1.14	-	0.96	1.31	0.35
	-	-	-	-	-

Genypterus capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	-	-	-	-	-
WLS2	-	-	-	-	-
Log1	-	-	-	-	-
Log2	0.52	0.07	0.38	0.66	0.28
Kappenman	1.04	-	0.79	2.17	0.98
Box-Cox1	1.47	-	1.36	1.57	0.22
Box-Cox2	1.35	-	1.27	1.49	0.23
Box-Cox3	1.51		1.36	1.66	0.30

Lophius vomerinus

	k	s.e	95% Lower	95% Upper	CI width
	-				
	-	-	-	-	-
	-	-	-	-	-
	-	-	-	-	-
	2.04	-	1.65	2.47	0.83
	1.36	-	1.25	1.46	0.21
	1.36	-	1.26	1.47	0.21
	1.29		1.29	1.50	0.21

Trachurus capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	-	-	-	-	-
WLS2	-	-	-	-	-
Log1	-	-	-	-	-
Log2					
Kappenman	1.08		0.76	1.45	0.67
Box-Cox1	0.91	-	0.80	1.02	0.21
Box-Cox2	0.91	-	0.81	1.02	0.21
Box-Cox3	0.91	-	0.81	1.02	0.21

Chelidonichthys capensis

	k	s.e	95% Lower	95% Upper	CI width
	6.01	0.55	4.92	7.09	2.17
	7.21	0.50	5.44	8.97	3.53
	-	-	-	-	-
	-	-	-	-	-
	5.93		4.55	7.53	2.98
	2.10	-	1.35	2.86	1.51
	1.99	-	1.28	2.70	1.42
	2.28	-	1.50	3.05	1.54

Zeus capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	1.79	0.21	1.38	2.20	0.82
WLS2	3.75	0.76	2.26	5.24	2.98
Log1					
Log2	0.60	0.03	0.73	0.86	0.13
Kappenman	2.12	-	1.64	2.65	1.00
Box-Cox1	1.70	-	1.56	1.85	0.28
Box-Cox2	1.57	-	1.46	1.68	0.23
Box-Cox3	1.66	-	1.56	1.79	0.23

Helicolenus dactylopterus

	k	s.e	95% Lower	95% Upper	CI width
	0.97	0.06	0.74	1.03	0.32
	1.12	0.16	0.77	1.43	0.72
	-	-	-	-	-
	-		-	-	
	-	-	-	-	-

Table 5.25: Summary of results for FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen*

For convenience:

WLS1 – refers to WLS regression line through the origin (1);

WLS2 – refers to WLS regression line through the origin (2);

Log1 – refers to Inversion of the Linear Regression equation;

Log2 – refers to Inversion of the linear Regression Equation through the Origin;
Box-Cox1 – refers to Option 1 of the Box-Cox Extended method: univariate transformation;
Box-Cox2 – refers to Option 2 of the Box-Cox Extended method: common transformation parameters; and
Box-Cox3 – refers to Option 3 of the Box-Cox Extended method: different transformation parameters.

The dashes for some methods imply that the data for the corresponding species did not meet all requirements of that method, and hence that the method cannot be applied to calculate a conversion factor or relationship for that species. The Log2 method appears problematic as it was applicable to only 2 species (*Genypterus capensis* and *Zeus capensis*), and the standard errors of the conversion factor from this method were very low compared to other methods.

The 8 methods are not directly comparable, as they are probably on different scales. Therefore, the best conversion factor can be chosen as the geometric mean of all the 8 conversion factors, and this value can then be used to convert catch from RV *Dr Fridtjof Nansen* to FRS *Africana* (old trawl gear).

Species	**Conversion factor**
Merluccius capensis	1.48
Merluccius paradoxus	1.27
Genypterus capensis	1.21
Lophius vomerinus	1.51
Trachurus capensis	0.95
Chelidonichthys capensis	3.67
Zeus capensis	1.75
Helicolenus dactylopterus	0.99

Table 5.26: Conversion factors to convert RV *Dr Fridtjof Nansen* catch to FRS *Africana* (old trawl gear) catch.

Results for calculating conversion factors to convert RV *Dr Fridtjof Nansen* catch data to FRS *Africana* (old trawl gear) are given in Table 5.26. All the conversion factors were above one, except for *Helicolenus dactylopterus*, whose conversion factor was 0.86. This feature implies that FRS *Africana* (old trawl gear) was less efficient in catching *Helicolenus dactylopterus*, than RV *Dr Fridtjof Nansen*, but more efficient in catching all other species.

Conversion factor for *Chelidonichthys capensis* was very high, even though the data were standardised into density. The reason for this high conversion factor is because there were only 33 non-zero data points for this species, and RV *Dr Fridtjof Nansen*'s catches were very low.

5.3.2. FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen*

5.3.2.1. Weighted Least Squares (WLS)

Discussion of the two methods under weighted least squares is already covered in Chapter 3. We have seen in Figure 5.4 that the scatterplots for all species except for *Genypterus capensis* and *Lophius vomerinus* showed evidence of a linear relationship between the catch per unit area (CPUA) data from the two research vessels. As the WLS method requires the linearity assumption to be satisfied before it can be applied, we will exclude the two species (*Genypterus capensis* and *Lophius vomerinus*) from this method.

For method 1: after applying the weights: $w_i = \frac{\frac{1}{\sqrt{x_i}}}{\sum\left(\frac{1}{\sqrt{x_i}}\right)}$ to impose the model:

$\left(\frac{y_i}{\sqrt{x_i}}\right) = \beta_{01}\sqrt{x_i} + e_i$ where $\text{var}(e_i) = \sigma^2$; we obtain the following results:

Species	$\hat{\beta}_{01}$ (conv factor)	Std error $\left(\hat{\beta}_{01}\right)$	95% CI for $\hat{\beta}_{01}$
Merluccius capensis	0.97	0.072	[0.826; 1.109]
Merluccius paradoxus	1.20	0.092	[1.022; 1.384]
Trachurus capensis	0.56	0.091	[0.38; 0.738]
Chelidonichthy capensis	2.31	0.433	[1.46; 3.159]
Zeus capensis	1.75	0.142	[1.468; 2.026]
Helicolenus dactylopterus	0.77	0.062	[0.645; 0.886]

Table 5.27: Results of the WLS regression line through origin (1) method

The last column in Table 5.27 presents confidence intervals for the slope parameter $\hat{\beta}_{01}$. The width of the confidence intervals is narrow for all species except for *Chelidonichthy capensis*, with the width of 1.699. The conversion factors for *Merluccius capensis*, *Trachurus capensis*, and *Helicolenus dactylopterus* are all less than one, implying that according to this WLS method, FRS *Africana* (new trawl gear) was less efficient in catching these species than RV *Dr Fridtjof Nansen*.

For method 2: after applying the weights: $w_i = \frac{\frac{1}{x_i}}{\sum\left(\frac{1}{x_i}\right)}$ to impose the following model: $\left(\frac{y_i}{x_i}\right) = R_i = \beta_{02} + e_i$ where $\text{var}(e_i) = \sigma^2$; we obtain the following results:

Species	$\hat{\beta}_{02}$ (conv factor)	Std error $\left(\hat{\beta}_{02}\right)$	95% CI for $\hat{\beta}_{02}$
Merluccius capensis	1.19	0.114	[0.969; 1.415]
Merluccius paradoxus	1.70	0.174	[1.361; 2.041]
Trachurus capensis	1.02	0.131	[0.769; 1.280]
Chelidonichthys capensis	3.00	0.597	[1.826; 4.167]
Zeus capensis	3.37	1.058	[1.302; 5.447]
Helicolenus dactylopterus	2.00	0.846	[0.338; 3.655]

Table 5.28: Results of WLS regression line through origin (2)

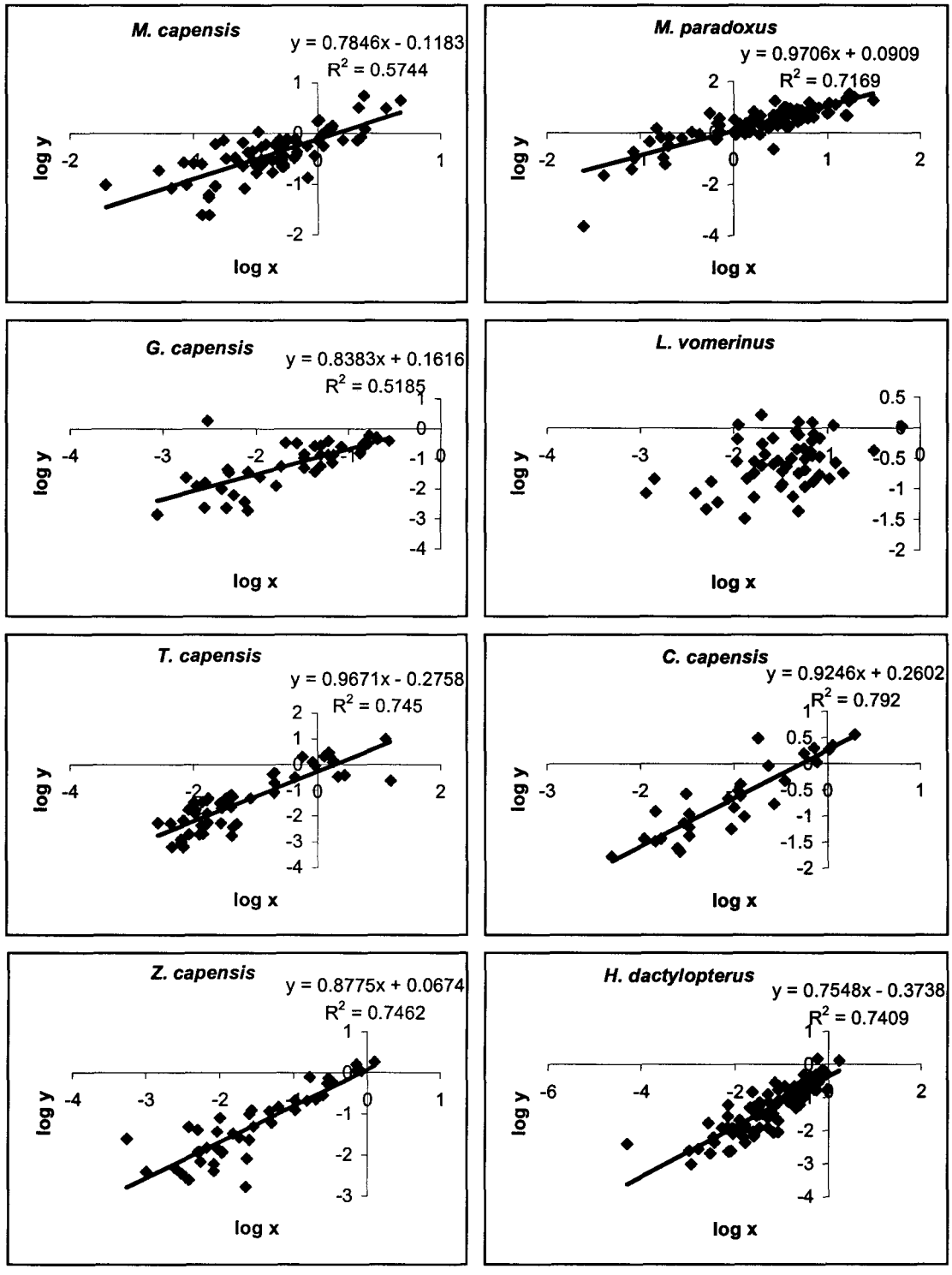


Figure 5.10: Scatterplots of log-transformed CPUA (g/m2) values of FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen*, for each species

Results in Table 5.28 appear to have low standard errors and narrow interval widths. The conversion factors are all greater than one, implying that according to this WLS method, FRS *Africana* (new trawl gear) was highly efficient in catching all species in Table 5.28 compared to RV *Dr Fridtjof Nansen*. This outcome is in stark contrast to the results in Table 5.27.

5.3.2.2. Log transformations

In section 3.4.1.2 of Chapter 3, two methods under log transformations were discussed. These were 1) Inversion of the linear regression equation and 2) Inversion of the linear regression through origin. Both these methods require that the two log-transformed variables show evidence of a linear relationship before they can be applied. Figure 5.10 presents scatterplots of the log-transformed CPUA variables for FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen* for all species under study.

Merluccius capensis, *Genypterus capensis*, *Chelidonichthys capensis*, and *Zeus capensis*, and these are the species that will be analyzed in both methods of this section.

Table 5.30 presents results of the first method under log transformation, inversion of the linear regression equation. All species were declared to have met the requirement that $\hat{\beta} \cong 1$, except only for *Merluccius capensis*, whose slope parameter value of 0.7846 was far from one. Therefore, *Merluccius capensis* was excluded.

Species (x vs y)	**Regression line**	**Is $\beta \cong 1$?**	**e^{α} (Conv factor)**	**[6]Inverted Std error for e^{α}**	**[7]Inverted 95% CI for e^{α}**
Merluccius capensis	$y = 0.785x + 0.118$	No	-	-	-
Genypterus capensis	$y = 0.838x - 0.162$	Yes	1.18	1.254	[0.747; 1.848]
Chelidonichthys capensis	y = 0.925x + 0.260	Yes	1.30	1.1183	[1.037; 1.623]
Zeus capensis	y = 0.878x + 0.067	Yes	1.07	1.1531	[0.804; 1.422]

Table 5.30: Results of inversion of the linear regression equation

It appears from Table 5.30 that the confidence intervals produced by this method were also not very wide. The standard errors of e^{α} were also reasonably lower, although not as lower as in the results for the weighted least squares methods.

The second method, inversion of the linear regression through origin model, requires that we first test the hypothesis that the intercept is zero ($\alpha = 0$) and the results are presented in Table 5.31.

[6] Inverted Std error for e^{α} = exp (std error for α)
[7] Inverted 95% CI for e^{α} = exp (95% CI for α). See Chapter 3.

All the species in the above scatterplots (Figure 5.10) appear to satisfy the requirement of linearly relationship between the two vessels, with the exception of only *Lophius vomerinus*. It is very interesting to note that before transformation, for *Genypterus capensis*, there was no evidence that the two vessels were linearly related, which is not the case for their log-transformed variables.

The log transformed variables for the species that showed linear relationship between the two vessels were then tested for univariate Gaussian using the Shapiro-Wilks test, which is a very good test statistic for normality. The results are presented in Table 5.29.

H_0: Data set have a Gaussian distribution
H_1: Data set does not have a Gaussian distribution

Species	Vessel	Shapiro-Wilks W	p-value
Merluccius capensis	Nansen (logx)	0.992	0.905
	Africana (logy)	0.981	0.316
Merluccius paradoxus	Nansen (logx)	0.844	**0.000**
	Africana (logy)	0.947	**0.001**
Genypterus capensis	Nansen (logx)	0.952	0.081
	Africana (logy)	0.945	0.047
Trachurus capensis	Nansen (logx)	0.901	**0.001**
	Africana (logy)	0.964	0.135
Chelidonichthys capensis	Nansen (logx)	0.966	0.452
	Africana (logy)	0.946	0.142
Zeus capensis	Nansen (logx)	0.967	0.257
	Africana (logy)	0.965	0.220
Helicolenus dactylopterus	Nansen (logx)	0.956	**0.003**
	Africana (logy)	0.979	0.130

Table 5.29: Shapiro-Wilks test for normality for each species

The bolded p-values in Table 5.29 implies significant difference for Ho at α = 0.05. Only four species showed evidence of Gaussian for both variables, i.e.,

H_0: Intercept is at origin ($\alpha = 0$)

H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	α	p-value
Merluccius capensis	-0.118	0.012
Genypterus capensis	0.162	0.226
Chelidonichthys capensis	0.260	0.028
Zeus capensis	0.067	0.639

Table 5.31: Hypothesis test for zero intercept

It appears from Table 5.31 that only *Genypterus capensis* and *Zeus capensis* have non-significant p-values at 5% level. Therefore, the second method under log transformation, inversion of the linear regression through origin, applied only to *Genypterus capensis* and *Zeus capensis*, produces the results in Table 5.32.

Species	$\hat{\beta}$ (conv factor)	Std error $(\hat{\beta})$	*95% CI for* $\hat{\beta}$
Genypterus capensis	0.75	0.049	[0.657; 0.848]
Zeus capensis	0.84	0.038	[0.770; 0.917]

Table 5.32: Results of inversion of the linear regression through origin method

This method, inversion of the linear regression through origin method, produces low standard errors and confidence interval widths for both species in Table 5.32, however, the slope parameters are a bit misleading. They are very small compared to the results in Table 5.30 and also in the weighted least squares methods in Section 5.3.2.1.

5.3.2.3. Kappenman's method

We've seen from Section 5.3.1.3 that for all species except for *Helicolenus dactylopterus*, the Kappenman (1992)'s transformation method was successful in

transforming the variables x and y to Gaussian distribution. Table 5.33 presents Shapiro-Wilks normality tests for the transformed FRS *Africana* $\left(X^d\right)$ and RV *Dr Fridtjof Nansen* $\left(r^dY^d\right)$ variables for all species.

H_0: Data set have a Gaussian distribution

H_1: Data set does not have a Gaussian distribution

Species	Vessel	*d*	Shapiro-Wilks W	p-value
	Nansen (X^d)		0.994	0.973
Merluccius capensis	Africana (r^dY^d)	0.050	0.994	0.364
	Nansen (X^d)		0.982	0.241
Merluccius paradoxus	Africana (r^dY^d)	0.248	0.994	0.955
	Nansen (X^d)		0.957	0.121
Genypterus capensis	Africana (r^dY^d)	0.140	0.968	0.300
	Nansen (X^d)		0.971	0.235
Lophius vomerinus	Africana (r^dY^d)	0.109	0.983	0.669
	Nansen (X^d)		0.940	**0.015**
Trachurus capensis	Africana (r^dY^d)	-0.124	0.968	0.203
	Nansen (X^d)		0.969	0.545
Chelidonichthys capensis	Africana (r^dY^d)	-0.081	0.953	0.215
	Nansen (X^d)		0.973	0.413
Zeus capensis	Africana (r^dY^d)	-0.058	0.969	0.297
	Nansen (X^d)		0.987	0.479
Helicolenus dactylopterus	Africana (r^dY^d)	0.126	0.984	0.302

Table 5.33: Shapiro-Wilks normality tests of variables X^d and r^dY^d for each species from both research vessels.

It appears from Table 5.33 that the Kappenman (1992) transformation method could not transform the RV *Dr Fridtjof Nansen* (X) variable of *Trachurus capensis* to Gaussian (p-value significant at α = 0.01). Table 5.34 presents results of the Kappenman (1992)'s method which were determined using the solver tool in Microsoft Excel.

Species	*d*	r	conversion factor	95% CI
Merluccius capensis	0.050	1.098	0.91	[0.768; 1.072]
Merluccius paradoxus	0.248	0.807	1.24	[1.081; 1.454]
Genypterus capensis	0.140	0.365	2.74	[1.978; 4.072]
Lophius vomerinus	0.109	0.127	7.87	[5.772; 10.674]
Chelidonichthys capensis	-0.081	0.456	2.19	[1.686; 2.938]
Zeus capensis	-0.058	0.548	1.82	[1.336; 2.432]
Helicolenus dactylopterus	0.126	1252	0.80	[0.681; 0.951]

Table 5.34: Results of the Kappenman's method

As in Section 5.3.1.2, the 95% confidence intervals were calculated by using a percentile method from 1000 bootstrap statistics of Kappenman conversion factors. The conversion factors do correspond to the results from other methods.

Figure 5.11 presents 1000 bootstrap estimates of the Kappenman conversion factors. The average of the bootstrap estimates is shown by the broken line, and the lower and upper values of the 95% confidence interval shown by the solid lines. One can notice that the average of the bootstrap estimates is approximately equal to the Kappenman conversion factors presented in Table 5.34 for all species, which implies that Kappenman conversion factors can be considered very reliable.

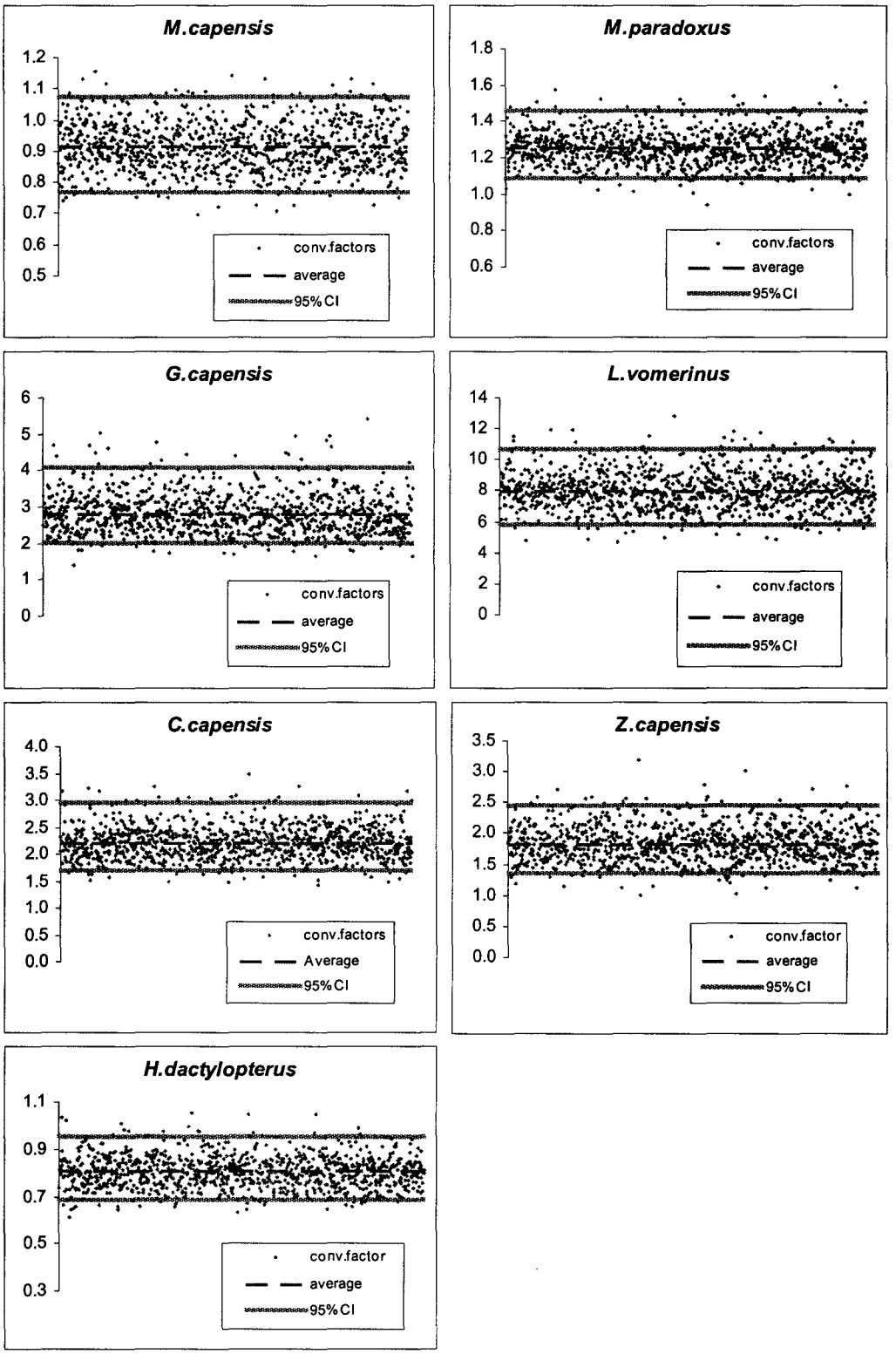


Figure 5.11: 1000 bootstrap estimates of the Kappenman conversion factors of seven species

5.3.2.4. Box-Cox Extension method

The Box-Cox Extension method is discussed in Section 3.4.1.4. This method enables the original data to be transformed to Gaussian distribution using one of the three options: 1) univariate transformation of x and y; 2) transformation using a common transformation parameter for both x and y; and 3) transformation using different transformation parameters for x and y. After the data have been transformed, the regression analysis through the origin will be applied to the transformed data (v and w) so as to obtain the slope estimate, which then is the conversion factor.

All three options mentioned in the above paragraph will be applied to data of all suite of species. If one recalls from Chapter 3, the likelihood function was obtained and the transformation parameters λ and k were solved iteratively using Microsoft Excel Solver routine with the Newtonian algorithm option.

Option 1 – univariate transformation of x and y

The results in Table 5.35 were obtained by maximising the likelihood function given in Section 3.4.1.4.1 of Chapter 3 and solving for transformation parameters, λ and k. The table presents transformation parameters (λ for v and k for w); the mean and variance of transformed variables; and the ratio of the two variances ($variance\ ratio = \frac{s_w^2}{s_v^2}$).

Species	Transformed variables	Shapiro-Wilks W	p-value
Merluccius capensis	v	0.994	0.975
	w	0.982	0.367
Merluccius paradoxus	v	0.983	0.252
	w	0.996	0.990
Genypterus capensis	v	0.957	0.122
	w	0.969	0.310
Lophius vomerinus	v	0.971	0.232
	w	0.985	0.738
Trachurus capensis	v	0.956	0.065
	w	0.971	0.261
Chelidonichthys capensis	v	0.969	0.540
	w	0.953	0.225
Zeus capensis	v	0.973	0.420
	w	0.968	0.293
Helicolenus dactylopterus	v	0.988	0.529
	w	0.984	0.333

Table 5.36: Shapiro-Wilks normality tests for variables v and w in Option 1

It appears from Table 5.36 that all species (both v and w) exhibit evidence of Gaussian distribution, showing the power of this Box-Cox method in transforming variables to exhibit Gaussian distribution.

The next step is to perform regression analysis through the origin, but only after the hypothesis test $\alpha = 0$ has been conducted for all species in Table 5.36. The results of the hypothesis tests ($\alpha = 0$) are presented in Table 5.37, which shows test statistics values t_{calc} and their associated p-values. The test statistic t-calc is calculated as described in Section 3.4.1.4 of Chapter 3. The d.f. in brackets are the degrees of freedom (n-2) used when calculating the t-test statistic.

H_0: Intercept is at origin ($\alpha = 0$)
H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	Trans. Variables	*Trans. par*	μ	s^2	*Variance ratio*
Merluccius capensis	*v*	0.063	-0.357	0.199	0.892
	w	0.039	-0.365	0.178	
Merluccius paradoxus	*v*	0.205	1.552	7.193	1.426
	w	0.277	2.455	10.257	
Genypterus capensis	*v*	0.11	-0.108	0.001	9.08
	w	0.149	-0.202	0.012	
Lophius vomerinus	*v*	0.119	-0.136	0.001	44.655
	w	0.046	-0.358	0.063	
Trachurus capensis	*v*	-0.219	-0.16	0.016	0.5
	w	-0.067	-0.115	0.008	
Chelidonichthys capensis	*v*	-0.065	-0.214	0.021	5.055
	w	-0.098	-0.327	0.108	
Zeus capensis	*v*	-0.063	-0.092	0.003	3
	w	-0.052	-0.144	0.009	
Helicolenus dactylopterus	*v*	0.152	-0.198	0.012	0.583
	w	0.079	-0.167	0.007	

Table 5.35: Transformation results of Box-Cox Extension method Option 1

It appears from Table 5.35 that the variances have all been stabilised by the transformation, though slightly high for *Merluccius paradoxus* ($\sigma_v^2 = 7.193$ and $\sigma_w^2 = 10.257$). This feature could be attributed to the high catches of *Merluccius paradoxus* in the trawls.

Table 5.36 presents Shapiro-Wilks normality tests to determine if the transformed variables exhibit any evidence of Gaussian distribution.

H_0: Power-transformed data follows a Gaussian distribution

H_1: Power-transformed data does not follow a Gaussian distribution

Species	t_{calc}	p-value (d.f)
Merluccius capensis	1.654	0.103 (72)
Merluccius paradoxus	3.059	**0.003** (91)
Genypterus capensis	0.131	0.897 (39)
Lophius vomerinus	0.005	0.996 (50)
Trachurus capensis	0.799	0.429 (47)
Chelidonichthys capensis	0.797	0.433 (27)
Zeus capensis	0.036	0.971 (40)
Helicolenus dactylopterus	1.283	0.203 (91)

Table 5.37: Hypothesis tests for $\alpha = 0$ in Option 1

One species had a significant p-value (at 0.05 level), namely, *Merluccius paradoxus*. Therefore, this species will be excluded from further analysis of this method.

Table 5.38 presents the results of the regression line through the origin for the 7 species that satisfied the condition $\alpha = 0$.

Species	$\hat{\beta}$(conv factor)	95% CI for $\hat{\beta}$
Merluccius capensis	0.84	[0.726; 0.956]
Genypterus capensis	1.90	[1.676; 2.122]
Lophius vomerinus	2.64	[2.177; 2.122]
Trachurus capensis	0.68	[0.608; 0.743]
Chelidonichthys capensis	1.68	[1.434; 1.918]
Zeus capensis	1.56	[1.400; 1.715]
Helicolenus dactylopterus	0.80	[0.765; 0.843]

Table 5.38: Conversion factors from Box-Cox Extension method Option 1

Option 2 – common transformation parameter

The results in Table 5.39 were obtained by maximising the likelihood function given in Section 3.4.1.4.1 of Chapter 3 and solving for transformation parameters, λ and k. The table presents transformation parameters (λ for v and k

for w); the mean and variance of transformed variables v and w; and correlation, r, between the two transformed variables.

Species	Transformed variables	λ	μ	s^2	r
Merluccius capensis	v		-0.357	0.201	
	w	0.137	-0.36	0.181	0.78
Merluccius paradoxus	v		1.563	7.194	
	w	0.213	2.345	10.428	0.85
Genypterus capensis	v		-0.102	0.001	
	w	0.081	-0.188	0.012	0.71
Lophius vomerinus	v		-0.128	0.001	
	w	0.085	-0.363	0.063	0.40
Trachurus capensis	v		-0.168	0.017	
	w	-0.066	-0.115	0.008	0.85
Chelidonichthys capensis	v		-0.218	0.021	
	w	-0.036	-0.328	0.109	0.89
Zeus capensis	v		-0.126	0.003	
	w	0.16	-0.177	0.011	0.92
Helicolenus dactylopterus	v		-0.199	0.012	
	w	0.156	-0.184	0.007	0.88

Table 5.39: Transformation results of Box-Cox Extension method Option 2

It appears from Table 5.39 that the two transformed variables v and w are linearly related for all species, except for *Lophius vomerinus* (*r = 0.4*). The variances have all been stabilised by the transformation, though slightly high for *Merluccius paradoxus* ($s_v^2 = 7.194$ and $s_w^2 = 10.428$).

Table 5.40 presents Shapiro-Wilks normality tests to determine if the transformed variables (only for species whose transformed variables are linearly related) show any evidence of Gaussian distribution.

H_0: Power-transformed data follow a Gaussian distribution

H_1: Power-transformed data does not follow a Gaussian distribution

Species	t_{calc}	p-value (d.f)
Merluccius capensis	1.61	0.112 (72)
Merluccius paradoxus	2.621	**0.010** (91)
Genypterus capensis	0.204	0.839 (39)
Chelidonichthys capensis	0.857	0.399 (27)
Helicolenus dactylopterus	1.726	0.088 (91)

Table 5.41: Hypothesis tests for $\alpha = 0$ in Option 2

The null hypothesis was rejected in only one species, namely, *Merluccius paradoxus*. Therefore, a conversion factor using this method will not be calculated for this species.

Table 5.42 presents the results of the regression line through the origin for the 4 species that satisfied the condition $\alpha = 0$.

Species	$\hat{\beta}(conv\,factor)$	**95% CI for** $\hat{\beta}$
Merluccius capensis	0.85	[0.737; 0.963]
Genypterus capensis	1.88	[1.645; 2.108]
Chelidonichthys capensis	1.65	[1.412; 1.894]
Helicolenus dactylopterus	0.87	[0.826; 0.909]

Table 5.42: Conversion factors from Box-Cox extension method Option 2

Option 3 – different transformation parameters

This option attempts at using the Box-Cox method to transform x and y to v and w using two different transformation parameters, but in the same likelihood function. The statistics results of this option are presented in Table 5.43 and these results were obtained by maximising the likelihood function given in Section 3.4.1.4.1 of Chapter 3 and calling the Solver tool to find solutions for transformation parameters, λ and k. The table presents transformation parameters (λ for v and k for w); the mean and variance of transformed variables v and w; and correlation, r, between the two transformed variables.

Species	Transformed variables	Shapiro-Wilks W	p-value
Merluccius capensis	v	0.991	0.864
	w	0.975	0.152
Merluccius paradoxus	v	0.983	0.257
	w	0.989	0.653
Genypterus capensis	v	0.956	0.117
	w	0.962	0.184
Trachurus capensis	v	0.925	**0.004**
	w	0.971	0.261
Chelidonichthys capensis	v	0.968	0.513
	w	0.949	0.177
Zeus capensis	v	0.923	**0.008**
	w	0.932	**0.015**
Helicolenus dactylopterus	v	0.988	0.527
	w	0.982	0.220

Table 5.40: Shapiro-Wilks normality tests for variables v and w in Option 2

It appears from Table 5.40 that 2 species should be excluded from further analysis of this method because, namely, *Trachurus capensis* (v p-value < 0.05) and *Zeus capensis* (both v and w p-values < 0.05).

The next step is to perform regression analysis through the origin, but only after the hypothesis test $\alpha = 0$ has been conducted for species that satisfied the Gaussian assumptions conducted in Table 5.40. The results of the hypothesis tests ($\alpha = 0$) are presented in Table 5.41, which shows test statistics values t_{calc} and their associated p-values. The test statistic t-calc is calculated as described in Section 3.4.1.4 of Chapter 3. The d.f. in brackets are the degrees of freedom (n-2) used when calculating the t-test statistic.

H_0: Intercept is at origin ($\alpha = 0$)

H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	Trans. variables	Trans. par	μ	σ^2	r
Merluccius capensis	v	0.163	-0.35	0.202	
	w	0.118	-0.361	0.18	0.78
Merluccius paradoxus	v	0.139	1.446	7.276	
	w	0.234	2.381	10.333	0.86
Genypterus capensis	v	0.107	-0.107	0.001	
	w	0.068	-0.186	0.012	0.71
Lophius vomerinus	v	0.11	-0.133	0.001	
	w	0.015	-0.354	0.063	0.4
Trachurus capensis	v	-0.163	-0.162	0.016	
	w	0.007	-0.121	0.008	0.87
Chelidonichthys capensis	v	-0.018	-0.22	0.021	
	w	-0.061	-0.326	0.108	0.89
Zeus capensis	v	0.144	-0.123	0.003	
	w	0.184	-0.183	0.011	0.92
Helicolenus dactylopterus	v	0.184	-0.206	0.012	
	w	0.105	-0.172	0.007	0.88

Table 5.43: Transformation results of Box-Cox Extension method Option 3

It appears from Table 5.43 that the two transformed variables v and w are linearly related for all species, except for *Lophius vomerinus* (r = *0.4*). The variances have all been stabilised by the transformation, though slightly high for *Merluccius paradoxus* ($\sigma_v^2 = 7.276$ and $\sigma_w^2 = 10.333$).

Table 5.44 presents Shapiro-Wilks normality tests to determine if the transformed variables (only for species whose transformed variables are linearly related) show any evidence of Gaussian distribution.

H_0: Power-transformed data follow a Gaussian distribution

H_1: Power-transformed data does not follow a Gaussian distribution

Species	Transformed variables	Shapiro-Wilks W	p-value
Merluccius capensis	v	0.988	0.722
	w	0.977	0.207
Merluccius paradoxus	v	0.978	0.114
	w	0.993	0.886
Genypterus capensis	v	0.957	0.122
	w	0.96	0.157
Trachurus capensis	v	0.948	**0.031**
	w	0.962	0.118
Chelidonichthys capensis	v	0.967	0.485
	w	0.951	0.199
Zeus capensis	v	0.929	**0.012**
	w	0.924	**0.008**
Helicolenus dactylopterus	v	0.986	0.458
	w	0.984	0.334

Table 5.44: Shapiro-Wilks normality tests for variables *v* and *w* in Option 3

It appears from Table 5.44 that all species (both v and w) showed evidence of normality, except for *Trachurus capensis* v (p-value for Shapiro-Wilks statistic was significant at 5% level), and *Zeus capensis* (both transformed variables v and w had p-values significant at 5% level). Therefore, due to deviation from normality at 5% level of significance, *Trachurus capensis* and *Zeus capensis* were excluded from the analysis using this method.

The next step is to perform regression analysis through the origin, but only after the hypothesis test $\alpha = 0$ has been conducted for 5 species that satisfied the Gaussian assumptions conducted in Table 5.44. The results of the hypothesis tests ($\alpha = 0$) are presented in Table 5.45, which shows test statistics values t_{calc} and their associated p-values. The test statistic t-calc is calculated as described in Section 3.4.1.4 of Chapter 3. The d.f. in brackets are the degrees of freedom (n-2) used when calculating the t-test statistic.

H_0: Intercept is at origin ($\alpha = 0$)

H_1: Intercept is not at origin ($\alpha \neq 0$)

Species	t_{calc}	p-value (d.f)
Merluccius capensis	0.184	0.097 (72)
Merluccius paradoxus	3.316	0.001 (91)
Genypterus capensis	0.271	0.788 (39)
Chelidonichthys capensis	0.867	0.394 (27)
Helicolenus dactylopterus	1.229	0.222 (91)

Table 5.45: Hypothesis tests for $\alpha = 0$ in Option 3

The null hypothesis was rejected in only one species, namely, *Merluccius paradoxus*. Therefore, this species will be excluded from further analysis of this method.

Table 5.46 presents the results of the regression line through the origin for the 4 species that satisfied the condition $\alpha = 0$.

Species	$\hat{\beta}(conv\ factor)$	95% CI for $\hat{\beta}$
Merluccius capensis	0.85	[0.736; 0.962]
Genypterus capensis	1.78	[1.558; 2.006]
Chelidonichthys capensis	1.64	[1.399; 1.877]
Helicolenus dactylopterus	0.8	[0.766; 0.840]

Table 5.46: Conversion factors from Box-Cox extension method Option 3

5.3.2.5. Summary of results of FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen*

The dashes for some methods imply that the data for that species could not meet all requirements of that method, meaning that the method could not be applied to calculate a conversion factor for that species. A similar conclusion to Section 5.3.1.5 can be said about the Log2 method. This (Log2 method) was again

applicable to only 2 species (*Genypterus capensis* and *Zeus capensis*), and the conversion factors from this method were very low compared to conversion factors from other methods.

Merluccius capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	0.97	0.07	0.83	1.11	0.28
WLS2	1.19	0.11	0.97	1.41	0.45
Log1	-	-	-	-	-
Log2	-	-	-	-	-
Kappenman	0.91	-	0.76	1.05	0.29
Box-Cox1	0.84	-	0.73	0.96	0.23
Box-Cox2	0.85	-	0.74	0.96	0.23
Box-Cox3	0.85	-	0.74	0.96	0.23

Merluccius paradoxus

	k	s.e	95% Lower	95% Upper	CI width
WLS1	1.20	0.09	1.02	1.38	0.36
WLS2	1.70	0.17	1.36	2.04	0.68
Log1	-	-	-	-	-
Log2	-	-	-	-	-
Kappenman	1.24	-	1.07	1.39	0.32
Box-Cox1	-	-	-	-	-
Box-Cox2	-	-	-	-	-
Box-Cox3	-	-	-	-	-

Genypterus capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	-	-	-	-	-
WLS2	-	-	-	-	-
Log1	1.18	1.25	0.75	1.85	1.10
Log2	0.75	0.05	0.66	0.85	0.19
Kappenman	2.74	-	1.94	3.73	1.78
Box-Cox1	1.90	-	1.68	2.12	0.45
Box-Cox2	1.88	-	1.65	2.11	0.46
Box-Cox3	1.78	-	1.56	2.01	0.45

Lophius vomerinus

	k	s.e	95% Lower	95% Upper	CI width
WLS1	-	-	-	-	-
WLS2	-	-	-	-	-
Log1	-	-	-	-	-
Log2	-	-	-	-	-
Kappenman	7.87	-	5.30	10.63	5.33
Box-Cox1	2.64	-	2.18	3.10	0.92
Box-Cox2	-	-	-	-	-
Box-Cox3	-	-	-	-	-

Trachurus capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	0.56	0.09	0.38	0.74	0.36
WLS2	1.02	0.13	0.77	1.28	0.51
Log1	-	-	-	-	-
Log2	-	-	-	-	-
Kappenman	-	-	-	-	-
Box-Cox1	0.68	-	0.608	0.743	0.135
Box-Cox2	-	-	-	-	-
Box-Cox3	-	-	-	-	-

Chelidonichthys capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	2.31	0.43	1.46	3.16	1.70
WLS2	3.00	0.60	1.83	4.17	2.34
Log1	1.30	1.12	1.04	1.62	0.59
Log2	-	-	-	-	-
Kappenman	2.19	-	1.62	2.93	1.31
Box-Cox1	1.68	-	1.43	1.92	0.48
Box-Cox2	1.65	-	1.41	1.89	0.48
Box-Cox3	1.64	-	1.40	1.88	0.48

Zeus capensis

	k	s.e	95% Lower	95% Upper	CI width
WLS1	1.75	0.14	1.47	2.03	0.56
WLS2	3.37	1.06	1.30	5.45	4.15
Log1	1.07	1.15	0.80	1.42	0.62
Log2	0.84	0.04	0.77	0.92	0.15
Kappenman	1.82	-	1.30	2.43	1.14
Box-Cox1	1.56	-	1.40	1.72	0.32
Box-Cox2	-	-	-	-	-
Box-Cox3	-	-	-	-	-

Helicolenus dactylopterus

	k	s.e	95% Lower	95% Upper	CI width
WLS1	0.77	0.06	0.64	0.89	0.24
WLS2	2.00	0.85	0.34	3.65	3.32
Log1	-	-	-	-	-
Log2	-	-	-	-	-
Kappenman	0.80	-	0.67	0.94	0.27
Box-Cox1	0.80	-	0.77	0.84	0.08
Box-Cox2	0.87	-	0.83	0.91	0.08
Box-Cox3	0.80	-	0.77	0.84	0.07

Table 5.47: Summary of results for FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen*

For convenience:

WLS1 – refers to WLS regression line through the origin (1);

WLS2 – refers to WLS regression line through the origin (2);

Log1 – refers to Inversion of the Linear Regression equation;

Log2 – refers to Inversion of the linear Regression Equation through the Origin;

Box-Cox1 – refers to Option 1 of the Box-Cox Extended method: univariate transformation;

Box-Cox2 – refers to Option 2 of the Box-Cox Extended method: common transformation parameters; and
Box-Cox3 – refers to Option 3 of the Box-Cox Extended method: different transformation parameters.

As is the case in Section 5.3.1.5, the 8 methods are not directly comparable, as they are probably not on the same scale. Therefore, the best conversion factor can be chosen as the geometric mean of all the 8 conversion factors, and this value can then be used to convert catch from RV *Dr Fridtjof Nansen* to FRS *Africana* (new trawl gear).

Species	Conversion factor
Merluccius capensis	0.93
Merluccius paradoxus	1.36
Genypterus capensis	1.58
Lophius vomerinus	4.56
Trachurus capensis	0.73
Chelidonichthys capensis	1.90
Zeus capensis	1.57
Helicolenus dactylopterus	0.94

Table 5.48: Conversion factors to convert RV *Dr Fridtjof Nansen* catch to FRS *Africana* (new trawl gear) catch.

Results for calculating conversion factors to convert RV *Dr Fridtjof Nansen* catch data to FRS *Africana* (new trawl gear) are given in Table 5.48. Conversion factors for *Merluccius capensis*, *Trachurus capensis*, and *Helicolenus dactylopterus* were below one, implying that FRS *Africana* (new trawl gear) was less efficient in catching these species than RV *Dr Fridtjof Nansen*, but more efficient in catching all other species.

The conversion factor for *Lophius vomerinus* was very high, even though the data were standardised in density. This factor is due to very low catches of RV *Dr Fridtjof Nansen*.

5.3.3. FRS *Africana* (old trawl gear) against FRS *Africana* (new trawl gear)

This section uses the results found in Sections 5.3.1 and 5.3.2. Section 5.3.1 analyses the comparative data between FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen*, and Section 5.3.2 analyses comparative data between FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen*. The methods used in this section are discussed in detail in Section 3.5 of Chapter 3.

The conversion factors from the comparative trials between FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* (Tables 5.27) and between FRS *Africana* (new trawl gear) against RV *Dr Fridtjof Nansen* (Tables 5.48) are used to calculate conversion factors to convert FRS *Africana* (old trawl gear) historical data to FRS *Africana* (new trawl gear), by the use of the following method:

$$Afr_{new} = \frac{\hat{k}_2}{\hat{k}_1} Afr_{old}$$

and the results are presented in Table 5.49

Species	Conv factor (Afr_{Old} vs Nan)	Conv factor (Afr_{New} vs Nan)	Conv factor (Afr_{Old} to Afr_{New})
Merluccius capensis	1.48	0.93	0.63
Merluccius paradoxus	1.27	1.36	1.07
Genypterus capensis	1.21	1.58	1.31
Lophius vomerinus	1.51	4.56	3.02
Trachurus capensis	0.95	0.73	0.77
Chelidonichthys capensis	3.67	1.9	0.52
Zeus capensis	1.75	1.57	0.9
Helicolenus dactylopterus	0.99	0.94	0.95

Table 5.49: Conversion factors to convert FRS *Africana* (old trawl gear) to FRS *Africana* (new trawl gear).

It is interesting to note that the conversion factors to convert FRS *Africana* (old trawl gear) catch to FRS *Africana* (new trawl gear) for *Merluccius capensis* (0.63) and *Merluccius paradoxus* (1.07) are not far from those found by the GLM

concluductory remarks on the whole project and also provides some recommendations on issues that came up during the course of the study.

analysis method in Brandão *et al* (2004), which were 0.61 for *Merluccius capensis* and 0.95 for *Merluccius paradoxus*.

It should be noted that, though the hake dataset in Brandão *et al* (2004) is the same as the one used in this study, they didn't standardise the catch (kg) into density (g/m^2) as was done in this study, thereby not accounting for the differences in catching efficiencies of the two research vessels.

It appears from Table 5.49 that the new gear of FRS *Africana* was less efficient compared to the old gear of FRS *Africana* in catching most species (*M. capensis*; *T. capensis*; *C. capensis*; *Z. capensis*; and *H. dactylopterus*), and more efficient in catching *M. paradoxus*, *G. capensis*, and *L. vomerinus*.

5.4. Conclusion

This chapter began by reporting exploratory data analysis (graphical and descriptive statistics) of the comparative trials CPUA data collected by the two research vessels. It has also shown what effects removal of zero catches from either vessel would have on the estimates of the regression coefficients. These contrasts are reflected by observing the behaviour of the r^2 value before and after the removal of the zero catches.

Another important feature of this chapter was plotting the CPUA differences and ratios of the two research vessels at different sea depths. These plots explore any depth effects that might have to be taken into account on calculation of the conversion factors.

The analytical methods discussed in Chapter 3 were then followed in calculating the conversion factors for the two research vessels. The next chapter provides

CHAPTER 6

CONCLUSION AND RECOMMENDATIONS

6.1. Introduction

As in any intercalibration study, the main aim of this study was to calculate conversion factors from comparative trawl experiments between FRS *Africana* (old and new trawl gears) against RV *Dr Fridtjof Nansen*. 10 Species were used in the study, but due to absent or insufficient data, conversion factors for *Cynoglossus zanzibarensis* (Red-spotted tonguefish) and *Chelidonichthys queketti* (Lesser Gurnard) could not be calculated. Studies at Marine and Coastal Management on Sole fish are underway to calculate conversion factors for this species.

There is an absence of recent reading material on intercalibration studies, which meant that new methods of calculating conversion factors were not available. However, in this thesis, regression analysis based methods were used. It is always possible that data, even after transformation, may not satisfy all the requirements of that particular method. For this reason more than one method was discussed in the thesis. These methods are discussed in Chapter 3 and applied to the data and the results presented in Chapter 5.

The discussion on the two Hake species, *Merluccius capensis* and *Merluccius paradoxus*, is provided in Chapter 4. These two species are the two commercially most important species in the South African fisheries industry. The observations during the collection of data suggested that, from the response of the different size classes to the trawl gear, it might well be worthwhile to ascertain if there are any size effects in the hake data that will warrant calculation of size-based and depth-based conversion factors. Frequency data for the two species were

analysed graphically and log-linear modelling performed to establish any size and depth effects that may have influenced the catch rate of the two vessels. It was therefore concluded from the rank correlation tests that though there is evidence of depth effect, it was not strong enough to warrant the calculating of conversion factors at different depths. Similarly, the size-based effect was also rejected.

The aim of the study, which is given in Chapter 1, is repeated here for ease of reference.

1. To calculate conversion factors to relate biomass estimates determined by FRS *Africana* (with old trawl gear) with biomass estimates obtained by RV *Dr Fridtjof Nansen* in 2000 and 2001 for the following species:

    - *Merluccius capensis* (Shallow-Water Cape Hake)
    - *Merluccius paradoxus* (Deep-Water Cape Hake)
    - *Genypterus capensis* (Kingklip)
    - *Cynoglossus zanzibarensis* (Red-spotted tonguefish)
    - *Lophius vomerinus* (Monkfish)

    And where extent of data permits:

    - *Trachurus capensis* (Horse Mackerel)
    - *Chelidonichthys capensis* (Cape Gurnard)
    - *Chelidonichthys queketti* (Lesser Gurnard)
    - *Zeus capensis* (Cape Dory)
    - *Helicolenus dactylopterus* (Jacopever)

2. To calculate conversion factors to relate biomass estimates determined by FRS *Africana* (with new trawl gear) since 2003 with biomass estimates obtained by RV *Dr Fridtjof Nansen* for the suite of 10 listed species in 1.

the biomass survey data collected by FRS *Africana* (old trawl gear). Finally, stock assessment[8] of the species concerned might be revised.

Some concerns were raised during the course of the study. Marine and Coastal Management (MCM) might usefully examine these concerns and investigate them further.

The unavailability of comparative trials data between FRS *Africana* (new trawl gear) and FRS *Africana* (old trawl gear) needs to be addressed. Calculating conversion factors to convert FRS *Africana* (old trawl gear) catch to FRS *Africana* (new trawl gear) catch was one of the major issues the study needed to address, but lack of data resulted in methods discussed in Section 3.5 of Chapter 3 to be used. The reliability of this method is somehow questionable as it does not produce any standard errors or confidence intervals. If possible, MCM therefore might arrange comparative trials between the two trawl gears so that reliable conversion factors can be calculated.

Studies describing the behaviour of the species in response to the trawl gears used in the study could not be found. There is a need for such studies to be undertaken, because they provide essential information to the calculation of conversion factors. Such studies could include vessel avoidance (some species react to light or noise generated by survey vessels and can initiate avoidance responses well in advance of the sampling gear (Gunderson 1993)) and gear avoidance (when fish are in contact with the sampling gear, some can escape capture by actively avoiding the trawl gear (Gunderson 1993)).

It is further recommended that comparative trials for Sole fish and *Chelidonichthys queketti* (Lesser Gurnard) should be undertaken, as these species are also very important to the South African fishery.

[8] Stock assessment is the process of collecting and analyzing biological and statistical information to determine the changes in the abundance of fishery stocks in response to fishing, and, to the extent possible, to predict future trends of stock abundance (FAO website: http://www.fao.org)

3. To combine the old gear and new gear conversion factors to relate FRS *Africana* (with old trawl gear) to FRS *Africana* (with new trawl gear) for the relevant suite of species.

4. To determine whether or not the conversion factors are independent of:
   - Size of the catch
   - Mean length of fish in catch (hake only)
   - Fishing depth

5. To determine whether or not the calibration factors are independent of various environmental parameters such as water temperature, oxygen content, etc. This aspect of the study is of lower priority and is dependent on the suitability of the available data and time constraints.

The objectives of the study were all satisfied, with the exception of the analysis of environmental factors, which was deemed to be of lower priority and therefore was not tackled in this study. Size and depth effect were only examined for the two Hake species. It was concluded that there were effects, but they were not large enough to calculate size-based and depth-based conversion factors.

6.2. Recommendations

The next step following the completion of this project would be to use the conversion factors calculated from comparative trials between FRS *Africana* (old trawl gear) against RV *Dr Fridtjof Nansen* and apply them to the biomass survey data collected by RV *Dr Fridtjof Nansen* when FRS *Africana* was not operational. Thereafter one might use the conversion factors calculated between FRS *Africana* (new trawl gear) against FRS *Africana* (old trawl gear) and apply them to

6.3. Conclusion

This study has calculated conversion factors from comparative trials between FRS *Africana* (old and new trawl gear) and RV *Dr Fridtjof Nansen*, and also calculated conversion factors to convert catch data from FRS *Africana* (old trawl gear) to catch data from FRS *Africana* (new trawl gear). Methods to calculate these conversion factors were developed using transformation rules and regression analysis.

REFERENCES

Agresti A. 1990. Categorical data analysis. John Wiley & Sons. New York.

Aitchison J. and Brown J.A.C. 1957. The lognormal distribution. Cambridge University Press, New York.

Bartlett M.S. 1947. The use of transformations. Biometrics 3: 39 – 52.

Box G.E.P. and Cox D.R. 1964. An analysis of transformations. Journal of Royal Statistics. B 26: 211-252.

Brandão A., Rademeyer R.A. and Butterworth D.S. 2004. First attempt to obtain a multiplicative bias calibration factor between the *Africana* with the old and the new gear. Marine and Coastal Management Demersal Working Group Document WG/11/04/D:H:26.

Byrne C.J. and Forrester J.R.S. 1991. Relative fishing power of *NOAA R/Vs Albatross IV* and *Delaware II*. National Marine Fisheries Service, Stock Assessment Workshop Working Paper SAW/12/P1. 8 pp (mimeo).

Casas J.M., del Rio J.L. and Gonzalez Troncoso D. 2004. Northern shrimp (*Pandulus borealis*) on *Flemish Cap* surveys 2003 and 2004. NAFO SCR Doc 04/77.

Cochran W.G. 1977. Sampling techniques. Wiley, New York.

Daniel W. 1978. Applied nonparametric statistics. Houghton Mifflin Company, Palo Alto, California, USA.

Dobson A.J. 1990. An introduction to generalized linear models. Chapman & Hall. London.

Dolby J.L. 1963. A quick method for choosing a transformation. Technometrics 5: 317 – 326.

Doubleday W.G. 1981. Manual on groundfish surveys in the Northwest Atlantic. Northwest Atlantic Fisheries Organization. Scientific Council Studies: Number 2.

Draper N.R. and Hunter W.G. 1969. Transformations: Some examples revisited. Technometrics 11: 23 – 40.

Draper N.R. and Smith H. 1966. Applied regression analysis. Wiley, New York.

Engas A., Lokkerborg S., Solda A.V. and Ona E. 1996. Comparative fishing trials for cod and haddock using commercial trawl and longline at two different stock levels. Journal of Northwest Atlantic Fisheries Science 19: 83–90.

Forrester J.R.S. (m.s.). A trawl survey conversion coefficient suitable for lognormal data. National Marine Fisheries Service, Woods Hole laboratory 17 pp (mimeo)

Gavaris S. and Brodie W.B. 1984. Results of comparative fishing between the *A.T. Cameron* and the *Wilfred Templeman* during July – August 1983. CAFSAC Res. Doc. 84/41

Gill J. 2000. Generalized linear models: A unified approach. Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-134. Thousand Oaks, CA: Sage.

Hald A. 1952. Statistical theory with engineering applications. Wiley, New York.

Heemstra P. and Heemstra E. 2004. Coastal fishes of Southern Africa. South African Institute for Aquatic Biodiversity (SAIAB) and National Inquiry Services Center (NISC).

Hollander M. and Wolfe D.A. 1973. Nonparametric statistical methods. Wiley, NewYork.

Kappenman R. 1992. Robust estimation of the ratio scale parameters from positive random variables. Communications in statistics, Theory and Methods 21:2983 – 2996.

Koller P. and Smith S.J. 1983. Preliminary analysis of *A.T. Cameron – Lady Hammond* comparative fishing experiments 1979 – 1981. CAFSAC Res. Doc. 83/59

Knoke D. and Burke P.J. 1980. Log-linear models. Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-020. Beverly Hills and London.

Kowalski C.J. 1970. The performance of some rough tests for bivariate normality before and after coordinate transformations to normality. Technometrics 12: 517 – 544.

Miller J.C. and Miller J.N. 1984. Statistics for analytical chemistry. Wiley, New York.

Neave H.R. and Worthington P.L. 1988. Distribution free tests. Unwin Hyman Ltd, London.

Newby J., Gooday P. and Elliston L. 2004. Structural adjustment in Australian fisheries, ABARE eReport 04.17 Prepared for the Fisheries Resources Research Fund, Canberra, November.

Pelletier D. 1998. Intercalibration of research survey vessels in fisheries: a review and an application. Canadian Journal of Fisheries and Aquatic Sciences 55: 2672 – 2690.

Punt A.E. and Leslie R.W. 1991. Estimates of some biological parameters for the Cape hake off the South African west coast. South African Journal of Marine Science. 10: 271-284.

Quinn G.P. and Keough M.J. 2002. Experimental design and data analysis for biologists. Cambridge University Press, United Kingdom.

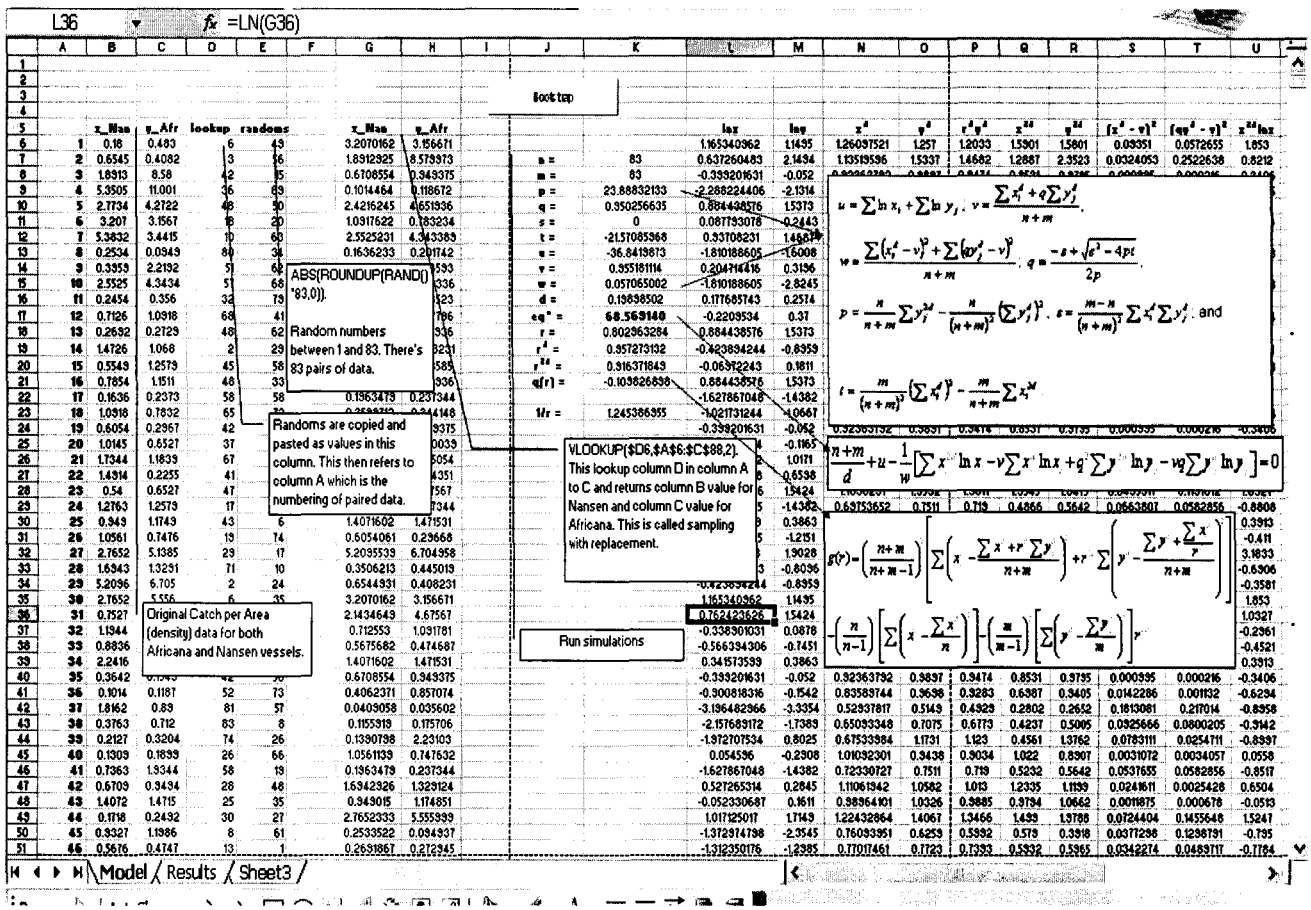
Smith M.M. and Heemstra P.C. 1986. Smith's sea fishes. JLB Smith Institute of ichthyology, New York.

Sissenwine M.P. and Bowman E.W. 1978. An analysis of some factors affecting the catchability of fish by bottom trawls. ICNAF Research Bulletin 13: 81 – 87.

Sprent P. 1989. Applied nonparametric statistical methods. Chapman and Hall, New York.

APPENDICES

Appendix A1: Kappenman method



Appendix A2: The Visual Basic code to run bootstrapping Macros for the Kappenman method

```
Sub MCap_kappenman_oldgear()
'
' MCap_kappenman_oldgear Macro
' Macro recorded by Luyanda Antony
```

Theil H. 1950. A rank invariant method of linear and polynomial regression analysis, I, II, III. Koninklijke Nederlandsche Akadamie van Wetenschappen. A, 53, 386 – 392, 521 – 525, 1397 – 1412.

Vazquez A. 2002. Catchability comparison between *Lofoten* and *Campelen* gears. NAFO SCR Doc. 02/74.

Von Szalay P.G. and Brown E. 2001. Trawl comparisons of fishing power differences and their applicability to National Marine Fisheries Service and Alaska Department of Fish and Game Trawl Survey Gear. Alaska Fishery Research Bulletin 8(2) 85 – 95.

Walsh S.J. and Orr D. 1998. Conversion factors for *Yellowtail Flounder* survey indices derived from comparative trials between the *Engel 145 Otter Trawl* and the *Campelen 1800 Shrimp Trawl*. NAFO SCR Doc. 98/60.

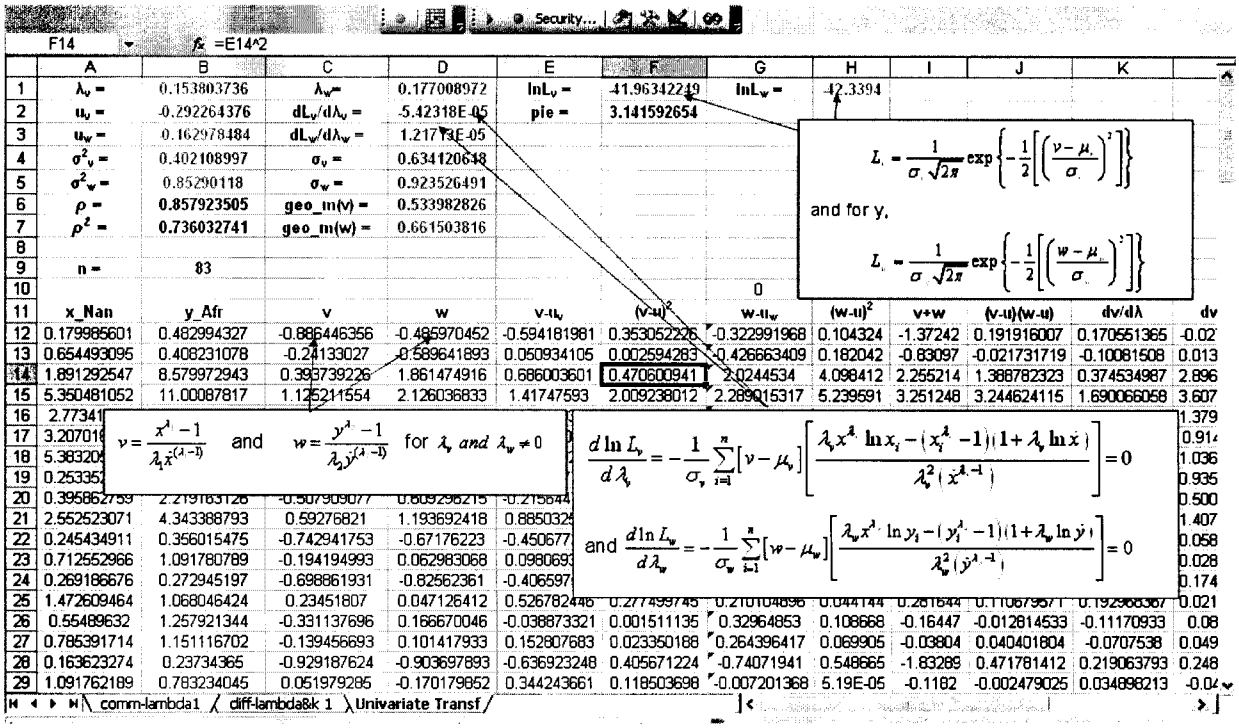
Warren G.W. 1996. Report on the comparative fishing trial between the *Gadus Atlantica* and *Teleost*. NAFO SCR Doc 96/28.

Wilderbuer T.K., Kappenman R.F. and Gunderson D.R. 1998. Analysis of fishing power correction factor estimates from a trawl comparison experiment. North American Journal of Fisheries Management 18: 11 – 18.

Zar. J.H. 1984. Biostatistical analysis. 2nd Edition. Prentice Hall International, New Jersey.

```
 For i = 1 To 1000
' turn on randoms
   Calculate

' copy randoms and paste as values in column D and solve for parameters
   Range("E6").Select
   Range(Selection, Selection.End(xlDown)).Select
   Selection.Copy
   Range("D6").Select
   Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
      :=False, Transpose:=False
   SolverSolve (True)

 ' copy Kappenman conversion factor and paste as value in Results worksheet
   Range("K23").Select
   Selection.Copy
   Sheets("Results").Select
   Range("B" & i + 6).Select
   Selection.PasteSpecial Paste:=xlPasteValues, Operation:=xlNone, SkipBlanks _
      :=False, Transpose:=False

 ' activate worksheet Model
   Sheets("Model").Select

 Next i
End Sub

Private Sub CommandButton1_Click()
 Module1.MCap_kappenman_oldgear

End Sub
```

Appendix B1: Box-Cox Extended Method: Option 1 – Univariate transformation parameter for *Merluccius capensis*



F14 =E14^2

	A	B	C	D	E	F	G	H	I	J	K	
1	λ_v =	0.153803736	λ_w=	0.177008972	lnL$_v$ =	-41.96342249	lnL$_w$ =	-42.3394				
2	u_v =	-0.292264376	dL$_v$/dλ_v =	-5.42318E-05	pie =	3.141592654						
3	u_w =	0.162978484	dL$_w$/dλ_w =	1.21773E-05								
4	σ^2_v =	0.402108997	σ_v =	0.634120648								
5	σ^2_w =	0.85290118	σ_w =	0.923526491								
6	ρ =	0.857923505	geo_m(v) =	0.533982826								
7	ρ^2 =	0.736032741	geo_m(w) =	0.661503816								
8												
9	n =	83										
10							0					
11	x_Nan	y_Afr	v	w	v-u$_v$	(v-u)^2	w-u$_w$	(w-u)^2	v+w	(v-u)(w-u)	dv/dλ	dv
12	0.179985601	0.482994327	-0.886446356	-0.485970452	-0.594181981	0.353052226	-0.322991968	0.104324	-1.37242	0.191916007	0.170551365	-0.02
13	0.654493095	0.408231078	-0.24133027	-0.589641893	0.050934105	0.002594283	-0.426663409	0.182042	-0.83097	-0.021731719	-0.10081508	0.013
14	1.891292547	8.579972943	0.393739226	1.861474916	0.686003601	0.470600941	2.0244534	4.098412	2.255214	1.388782323	0.374534987	2.896
15	5.350481052	11.00087817	1.125211554	2.126036833	1.41747593	2.009238012	2.289015317	5.239591	3.251248	3.244624115	1.690066058	3.607
16	2.77341											1.379
17	3.20701											0.91
18	5.38320											1.036
19	0.25335											0.935
20	0.39586											0.500
21	2.552523071	4.343388793	0.59276821	1.193692418	0.8850325							1.407
22	0.245434911	0.356015475	-0.742941753	-0.67176223	-0.450677							0.058
23	0.712552966	1.091780789	-0.194194993	0.062983068	0.098069							0.028
24	0.269186676	0.272945197	-0.698861931	-0.82562361	-0.406597							0.174
25	1.472609464	1.068046424	0.23451807	0.047126412	0.526782446	0.277499745	0.210104896	0.044144	0.281644	0.110679571	0.192968367	0.021
26	0.55489632	1.257921344	-0.331137696	0.166670046	-0.038873321	0.001511135	0.32964853	0.108668	-0.16447	-0.012814533	-0.11170933	0.08
27	0.785391714	1.151116702	-0.139456693	0.101417933	0.152807683	0.023350188	0.264396417	0.069905	-0.03804	0.040401804	-0.0707538	0.049
28	0.163623274	0.23734365	-0.929187624	-0.903697893	-0.636923248	0.405671224	-0.74071941	0.548665	-1.83289	0.471781412	0.219063793	0.248
29	1.091762189	0.783234045	0.051979285	-0.170179852	0.344243661	0.118503698	-0.007201368	5.19E-05	-0.1182	-0.002479025	0.034898213	-0.0

comm-lambda1 / diff-lambda&k 1 / Univariate Transf /

$$L_v = \frac{1}{\sigma_v\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\left(\frac{v-\mu_v}{\sigma_v}\right)^2\right]\right\}$$

and for y,

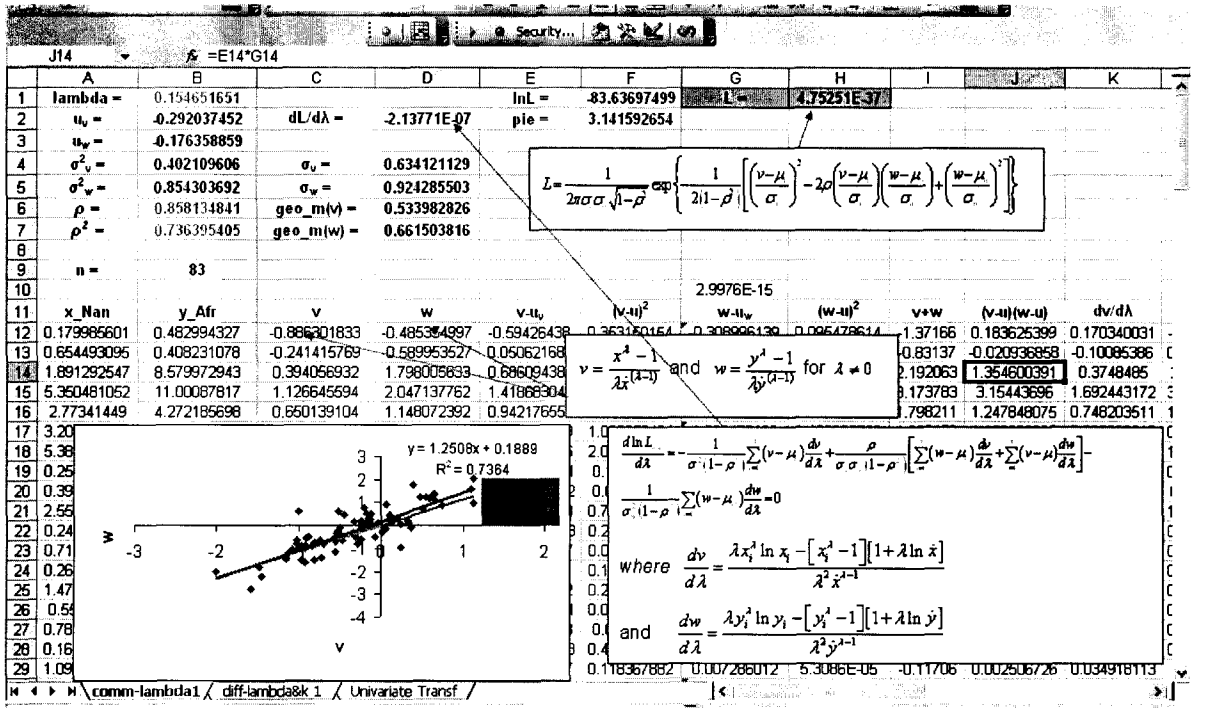
$$L_w = \frac{1}{\sigma_w\sqrt{2\pi}}\exp\left\{-\frac{1}{2}\left[\left(\frac{w-\mu_w}{\sigma_w}\right)^2\right]\right\}$$

$$v = \frac{x^{\lambda_v}-1}{\lambda_v \dot{x}^{(\lambda_v-1)}} \quad \text{and} \quad w = \frac{y^{\lambda_w}-1}{\lambda_w \dot{y}^{(\lambda_w-1)}} \quad \text{for } \lambda_v \text{ and } \lambda_w \neq 0$$

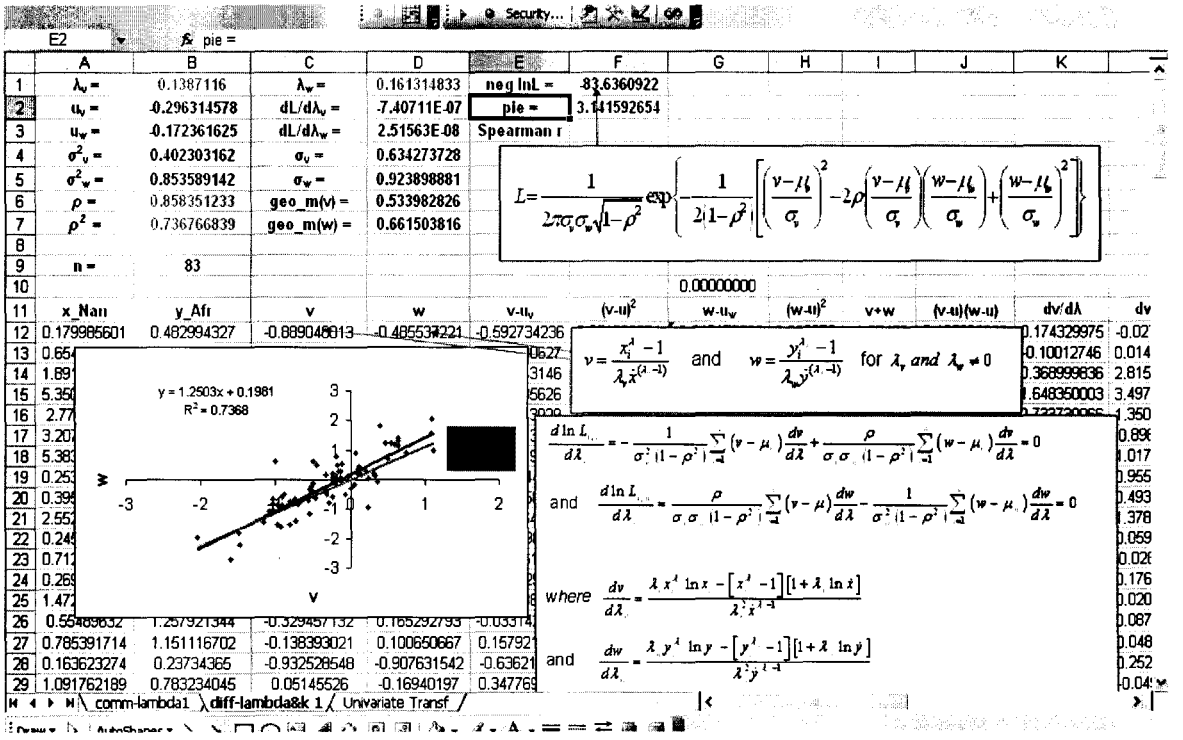
$$\frac{d\ln L_v}{d\lambda_v} = -\frac{1}{\sigma_v}\sum_{i=1}^{n}\left[v-\mu_v\right]\left[\frac{\lambda_v x^{\lambda_v}\ln x_i - (x_i^{\lambda_v}-1)(1+\lambda_v \ln \dot{x})}{\lambda_v^2(\dot{x}^{\lambda_v-1})}\right] = 0$$

and $$\frac{d\ln L_w}{d\lambda_w} = -\frac{1}{\sigma_w}\sum_{i=1}^{n}\left[w-\mu_w\right]\left[\frac{\lambda_w x^{\lambda_w}\ln y_i - (y_i^{\lambda_w}-1)(1+\lambda_w \ln \dot{y})}{\lambda_w^2(\dot{y}^{\lambda_w-1})}\right] = 0$$

Appendix B2: Box-Cox Extended Method: Option 2 – common transformation parameter for *Merluccius capensis*



Appendix B3: Box-Cox Extended Method: Option 3 – Different transformation parameter for *Merluccius capensis*



	A	B	C	D	E	F	G	H	I	J	K
1	λ_v =	0.1387116	λ_w =	0.161314833	neg lnL =	-83.6360922					
2	u_v =	-0.296314578	$dL/d\lambda_v$ =	-7.40711E-07	pie =	3.141592654					
3	u_w =	-0.172361625	$dL/d\lambda_w$ =	2.51563E-08	Spearman r						
4	σ^2_v =	0.402303162	σ_v =	0.634273728							
5	σ^2_w =	0.853589142	σ_w =	0.923898881							
6	ρ =	0.858351233	geo_m(v) =	0.533982826							
7	ρ^2 =	0.736766839	geo_m(w) =	0.661503816							
8											
9	n =	83									
10							0.00000000				
11	x_Nam	y_Afr	v	w	v-u_v	$(v-u)^2$	w-u_w	$(w-u)^2$	v+w	(v-u)(w-u)	dv/dλ	dv
12	0.179985601	0.482994327	-0.889048813	-0.485534221	-0.592734236						0.174329975	-0.02
13	0.654				0627						-0.10012746	0.014
14	1.891				3146						0.368999836	2.815
15	5.350				5626						0.648350003	3.497
16	2.77				3009						0.727230066	1.350
17	3.207											0.898
18	5.383											1.017
19	0.253											0.955
20	0.399											0.493
21	2.552											1.378
22	0.245											0.059
23	0.711											0.028
24	0.269											0.176
25	1.472											0.020
26	0.55469632	1.257921344	-0.329457132	0.165292793	-0.03314							0.087
27	0.785391714	1.151116702	-0.138393021	0.100650667	0.15792							0.048
28	0.163623274	0.23734365	-0.932528548	-0.907631542	-0.63621							0.252
29	1.091762189	0.783234045	0.05145526	-0.16940197	0.347769							-0.04

$$L=\frac{1}{2\pi\sigma_v\sigma_w\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{v-\mu_v}{\sigma_v}\right)^2-2\rho\left(\frac{v-\mu_v}{\sigma_v}\right)\left(\frac{w-\mu_w}{\sigma_w}\right)+\left(\frac{w-\mu_w}{\sigma_w}\right)^2\right]\right\}$$

$$v=\frac{x_i^{\lambda_v}-1}{\lambda_v\dot{x}^{(\lambda_v-1)}} \quad \text{and} \quad w=\frac{y_i^{\lambda_w}-1}{\lambda_w\dot{y}^{(\lambda_w-1)}} \quad \text{for } \lambda_v \text{ and } \lambda_w \neq 0$$

$$\frac{d\ln L_{\lambda_v}}{d\lambda_v}=-\frac{1}{\sigma_v^2(1-\rho^2)}\sum_{i=1}^{n}(v-\mu_v)\frac{dv}{d\lambda_v}+\frac{\rho}{\sigma_v\sigma_w(1-\rho^2)}\sum_{i=1}^{n}(w-\mu_w)\frac{dv}{d\lambda_v}=0$$

$$\text{and} \quad \frac{d\ln L_{\lambda_w}}{d\lambda_w}=\frac{\rho}{\sigma_v\sigma_w(1-\rho^2)}\sum_{i=1}^{n}(v-\mu_v)\frac{dw}{d\lambda_w}-\frac{1}{\sigma_w^2(1-\rho^2)}\sum_{i=1}^{n}(w-\mu_w)\frac{dw}{d\lambda_w}=0$$

$$\text{where} \quad \frac{dv}{d\lambda_v}=\frac{\lambda_v x^{\lambda_v}\ln x-\left[x^{\lambda_v}-1\right]\left[1+\lambda_v\ln\dot{x}\right]}{\lambda_v^2\dot{x}^{\lambda_v-1}}$$

$$\text{and} \quad \frac{dw}{d\lambda_w}=\frac{\lambda_w y^{\lambda_w}\ln y-\left[y^{\lambda_w}-1\right]\left[1+\lambda_w\ln\dot{y}\right]}{\lambda_w^2\dot{y}^{\lambda_w-1}}$$