

Use of commercial vessels for hake surveys, calibration versus gear replication for achieving short term continuity.

Prepared by

**OLRAC-SPS
Silvermine House
Steenberg Office Park
Tokai 7945**

On behalf of

**SADSTIA
The South African Deep Sea
Trawling Industry Association**

November 2013

Summary

This document addresses the use of un-calibrated surveys carried out with standardised fishing gear in stock assessments and management advice in the short term. The international panel from the workshop on fish stock assessment held in 2012 at UCT was strongly of the opinion that the short term use of such data should be contingent on the execution of inter-calibration studies. There are however still divergent views on this topic in South Africa. There is a view that vessel length, power and noise levels affect the catchability properties of the gear, necessitating calibrations whenever a different vessel is used. An opposing view suggests that were that the case then the associated economic advantages would have had an impact on trawler design, since even a small advantage would represent a significant economic incentive. Representatives of the South African offshore trawling industry aver that this economic incentive has not featured in trawler design. Indeed, a short search for trawling optimisation turns up information about energy usage and net design, but not vessel design in relation to optimal catchability.

The literature reports the results of numerous inter-calibration experiments but only a small selection of these are relevant to the use of common standard gear on different vessels. For two relevant papers cited here, one did not find a significant vessel effect while the other did. The latter reports on a large number of species – the vessel effect for cod is marginally significant. Statistical methods

employed to test for the presence of a vessel effect are varied, and most examined do not correct for the impact of multiple comparisons on probability levels. This most likely leads to an overstatement of the extent of vessel effects.

A statistical approach to interpreting the data from inter-calibration experiments leads to the conclusion that if the relative vessel effect is less than 10%, then experiments with a scientifically acceptable Type I (false positive) and II (false negative) error rate require unrealistically large sample sizes, equivalent to sample sizes which are an order of magnitude or more times the research effort for a typical demersal survey.

Use of an uninformative prior on the size of the relative vessel effect is also problematic. If the relative vessel effect is less than 10%, and if one assumes that a pragmatic 100 paired tows is the most that is likely to be funded, then the benefits of the experiment are very little. The impact of such data on stock assessments is to introduce unwarranted variance and uncertainty into the calculations.

A promising route to avoid these problems is to use an informative prior for the relative vessel effect, for example a mean of 0 and a variance of $5\%^2$ (0.05^2 for $\Delta \ln q$).

There is a paper in the literature (Walsh, 2006) that reports that deck geometry can influence the geometry of the net while fishing at depth. In general the geometry of the trawl is something that can be monitored with net, door and footrope contact sensors. Furthermore, the standard correction to the catch, i.e. dividing it by the swept area (which includes wing spread), intentionally corrects for geometry impacts, by converting the catch to a measure of fish density. An analysis of gear geometry for two surveys conducted with the Africana using the new gear is presented here, proposing the use of structural equation models to analyse the data. That study rejects the null hypothesis that gear geometry has an impact on the calculated fishing density.

Lastly it is suggested that an analysis of the size structure of the catch by the Andromeda during the 2013 survey may throw light on the existence of a vessel effect, making use of comparisons with the Nansen data for summer 2013, and the Africana data for other surveys. The rationale behind this is that catchability differences are likely to be associated with size selectivity differences.

The following proposals are made here:

1. Use of an informative prior on the vessel effect when using commercial vessels with standard fishing gear
2. Use of the OMP framework to explore the benefits of inter-calibration experiments for a number of scenarios (to be agreed) for the future use of different vessels for surveys.
3. Monitoring of gear geometry in all survey circumstances, and retrospective analysis of these data to test for (a) differences between vessels and (b) possible impacts on fishing density.
4. Analysis of catch size structure data from the Andromeda when fishing with standard survey fishing gear to throw light on the present and scale of possible relative vessel effects.

The basic sentiment behind this discussion document is that the trawling industry would like to find a means to make use of demersal surveys run using commercial vessels in the short term (while operating standardised fishing gear), without this being reliant on costly inter-calibration experiments.

Background

Recent years have seen a reduction in the availability of the South African research vessel, the *Africana*, for demersal surveys. An alternative research vessel, the RV Dr Fridtjof Nansen, has been used for carrying out some of the routine survey work. For some survey slots no alternative survey vessel was available, and the survey was not carried out. In January / February / March of 2013 a commercial vessel, the *Andromeda* (*Andromeda III*, 32.511 metres, 30 crew, 521.2 GT, 2000 year of construction), and the research vessel the *Nansen* were both used for the routine summer survey. The gear used on the *Andromeda III* was one of the new gear nets used on the *Africana* (see Anon, 2013 and Appendix B). Only the data for the RV Dr Fridtjof Nansen was however used to provide input data for the OMP calculations that generate the TAC.

In late 2012 at the International Fisheries Stock Assessment Workshop held at the University of Cape Town, some discussion about the usability of survey data from commercial vessels was held. A position that was explored was the possibility of treating a commercial vessel using the standard (new) survey gear used by the *Africana* as equivalent to a survey by the *Africana* with the new gear. The view of the international panel was that in order to use such data in the short term, calibrations of the commercial vessel against the previous research vessel, the *Africana*, was necessary (Smith et al, 2012): “*However the new data point should not be used in the OMP until a sufficient time series is built up or a calibration with Africana is completed*” (quoting from the panel report of the international workshop. The discussions cited, inter alia, Helser et al (2004) to substantiate a view that vessel effects need to be estimated in any application of survey data based on the deployment of multiple vessels, even when fishing gear has been standardised. The meeting was of the view that in the long term a survey time series could be established which provides scientific value in its own right, without calibration. Making the short term use of survey data collected using commercial vessels contingent on inter-calibration studies presents a number of challenges.

There are considerable costs associated with running the commercial vessel to achieve the calibrations, and the availability of the *Africana* for these calibrations is uncertain. Also uncertain is whether the same commercial survey vessel could be made available routinely. Thus, if the use of commercial vessel(s) becomes routine, the issue and costs of inter-calibration experiments arises. There is a view that the strict replication and standardisation of the survey gear avoids the need for inter-calibrations, that there is no difference in catchability under these circumstances, and that use of a likelihood term based on the data from calibration experiments would reduce the value of the commercial vessel data by introducing variance into stock assessment calculations. Opposing views to this draw on the argument that the length, power and noise spectrum for the commercial vessel differs from the *Africana* and that this will give rise to catchability differences which cannot be ignored. Yet other opinions (various meetings, discussions, pers. comms.) are for example that the general recommendation for calibration is inconsistent given a number of uncalibrated (and undocumented) changes to gear, vessel, crew and skippers on the *Africana* over the years.

Results in the literature

Cadigan et al (2006) examined paired trawl experiments to estimate the relative difference between the catchabilities of the *William Templeman* and the *Alfred Needler* when both vessels fished the *Campelen 1800* shrimp trawl, which is the standard survey trawl used by the Newfoundland Region of Fisheries and Oceans Canada. Their results do not demonstrate statistically significant differences in catchability between the two vessels. However they include the following comment in their paper:

“Vessel differences in survey catchabilities, even when using the same gear, are common (e.g. Cotter 2001, Pelletier 1998, Wilderbuer and Kappenman 1998). Vessel differences in trawl geometry and swept area using the same fishing gear were observed by McCallum and Walsh (2002).”

Closer examination shows that the comparison provided by Cotter (2001) which is closest to comparing the performance of different vessels using the same gear is confounded because of the use of different quarters of the year. Other candidate comparisons with potential to reveal the existence of a vessel effect when standard fishing gear is used are compromised because different regions are involved in the comparison. Cotter’s (2001) work is not based on paired trawl experiments.

Wilderbuer et al (1998), referred to by Cadigan et al (2006), reports on the results of experiments using a common vessel and two different nets. The nets had exactly the same design and rigging but were constructed out of different mesh materials, viz. polyethylene versus nylon. Results in that paper are thus not relevant to whether it is possible to view the use of different vessels with a common standard gear as a standard experimental type, even given their use of a paired trawl experimental design.

The commentary on page 255 of Hesler et al (2004) suggests that the vessel effects in that study were confounded by gear differences between vessels. On first reading these results do not seem to be relevant to the role of vessel distinct from gear.

Pelletier (1998) reports analyses based on data from paired trawl experiments, using a standard gear type but different vessels. This work does present evidence of statistically significant catchability differences between the two vessels in the experiment, for certain of the species in a multispecies trawl fishery. In the analysis of the data it is assumed that the abundance of fish available to the two gear types was the same – this is a common assumption for all work of this type. The analyses do not appear to have corrected the probability level threshold for testing the null hypothesis of no change for multiple comparisons. As a result the number of species for which statistically significant results are reported is unduly inflated. This omission seems quite common in studies of this nature. Also of note is that the results in one region were not significant for cod (the Celtic Sea), and in another region (the North Sea) they are only marginally significant at the 10% level, and only for GADU MOR <50 but not for GADU MOR >50. Cod is mentioned here because it is assumed to be the most ‘hake-like’ species in Pelletier’s (1998) study. This study reports results for two areas, the North Sea and the Celtic Sea. The two studies also disagree on MELA AEG (significant in one area not the other), MICR KIT (significant in one area not the other), SCYL CAN (significant in one area not the other). These three, together with cod, are the only 4 common species between the two areas, and there is disagreement between the two areas for all of them (insofar as in one area the relative vessel catchability is statistically significant, but not in the other area).

Vessel differences in trawl geometry and swept area using the same fishing gear are reported by McCallum and Walsh (2002). This they ascribe to a different geometry of the gallows blocks for one of the three vessels compared, although they do suggest that many of the differences could be due to the use of different fishing depths in the comparisons.

Statistical considerations

There seems to be consensus that the design of inter-calibration experiments should be based on paired trawls. What is an acceptable sample size for a paired trawl calibration study? Work reported in Brandao et al (2004) uses about 100 paired trawls, i.e. 100 for each of two vessels. This is

equivalent to double the number of tows in a typical South African demersal survey. Statistical considerations for experimental designs involves fixing the Type I and Type II error rates. A rule of thumb is to use a Type I error rate of $\alpha = 10\%$ and a Type II error rate of $\beta = 0.20$, i.e. the power of the test is $1 - \beta = 80\%$.

For *Merluccius paradoxus*, Brandao et al (2004) report a standard error for the estimated relative log of catchability ($\Delta \ln q$) between the old and new Africana gear of about 0.0827.

*(The experiments involved in the work reported on by Brandao et al (2004) compared the old gear + Africana and the Nansen in one experiment and the new gear + Africana and the Nansen in another experiment. The standard error for the new to old Africana $\Delta \ln q$ is taken as the standard error cited by Brandao et al for *M. paradoxus* of 0.117 divided by $2^{1/2}$, since the new and old African gear were only comparable via their mutual comparison to the Nansen. This gives a value of 0.0827 as the standard error of an estimate of $\Delta \ln q$ for a paired trawl experiment for *M. paradoxus*).*

The estimate of $\Delta \ln q$ between old and new Africana gear obtained by Brandao et al (2004) is 0.053. For such an experimental result, one would not reject the null hypothesis of $\Delta \ln q = 0.00$ at $\alpha = 10\%$. If 0.053 was the true value of $\Delta \ln q$ (i.e. H_A), then for $\alpha = 10\%$, for a paired t-test, the power of the test is 17%. In order to achieve a power of 80% under this H_A , the sample size given by n , the number of paired tows, would have to be increased to $n=1510$ tows for each vessel in the pair, or 3020 tows in total. In terms of number of tows, this is equivalent to 30 typical South African demersal surveys. (these results were produced using the IBM SPSS Sample Power software).

However, as the 'true' value of $\Delta \ln q$ is increased, the number of tows to achieve a power of 80% at $\alpha = 10\%$ declines (see Table 1).

If $\Delta \ln q$ is 0.20 or greater, then fewer than 110 paired tows would achieve a power of 80% at $\alpha = 10\%$. If there is a $\Delta \ln q$ of 0.10 or less, then more than 430 paired tows are required to achieve a power of 80% at $\alpha = 10\%$.

An alternative to the approach discussed above is the approach which has been used in South Africa for incorporating the data from inter-calibration work into stock assessments. This has been to use a posterior distribution (posterior to the calibration experiment) for $\Delta \ln q$ in the likelihood function for the stock assessments. The prior (prior to the calibration experiment) for $\Delta \ln q$ is implicitly uninformative, i.e. $\Delta \ln q \sim U(-\infty, +\infty)$. There is however no clear guidance for the sample size that should be used, and the following problems arise:

1. For an approach that, for example, assumes that $\Delta \ln q$ is zero if the null hypothesis cannot be rejected, an inordinate amount of inter-calibration work needs to be done to reach this point (reject/accept H_0) when $\Delta \ln q$ is less than 0.10 and $\alpha=10\%$, $1-\beta = 80\%$.
2. For an approach that includes the findings of the mean and the variance from a calibration study via a likelihood term, a large amount of unnecessary variance is introduced into stock assessments when $\Delta \ln q$ is small.

To address these problems, one possibility is the use of an informative prior for $\Delta \ln q$, for example a normal with mean zero and variance 0.05^2 . Table 2 illustrates how much the posterior would shift from the prior at a realistic 100 paired tows.

Pragmatic considerations

These statistical considerations need to be combined with a more explicit assessment of the value for management of different approaches to the use of data from a commercial vessel using standard gear.

The OMP framework provides a mechanism to explore (quantify risks and benefits) a number of specific scenarios going forward, involving the use of the following:

- Africana fishing old gear
- Africana fishing new gear
- Dr Fridtjof Nansen fishing its own gear as in the past
- A commercial vessel using Africana new gear.
- Inter-calibration work involving some plausible combinations of the above, vessel availability dependent.
- Data from the commercial vessel surveys assuming $\Delta \ln q = 0$, but with added precautionary adjustments to address the associated risks. In this scenario avoidance of the costs of inter-calibration studies could, for example, be exchanged for the cost of a smaller average TAC.

A number of scenarios could be mapped out and explored. Because of the high cost of inter-calibration work, specific focus should fall on the benefits of inter-calibration studies. Clearly the choice of a prior for $\Delta \ln q$ would be very influential. The work that is proposed above relates to the following recommendation in the international panel's report from the international workshop held in 2012 (see Smith et al, 2012):

*“Whatever surveys are conducted in January 2013, the Panel **strongly recommends** that a comprehensive plan be developed as soon as possible to ensure the ongoing provision of scientific survey data for hake (and other demersal species). Such a plan should factor in contingencies concerning the future availability of survey vessels (both industry and government vessels), and the need for calibration between vessels. Priority should be given to ensuring the continuity of comparable survey indices. If future surveys are to be conducted using industry vessels, consideration should be given to the use of multiple vessels and the estimation of between-vessel efficiency. This is particularly important if it is likely that the same industry vessel will not be used for each survey.”*

Gear geometry: The literature reports evidence of changes in gear geometry between different vessels using the same net. This can of course be monitored, and should be, as a possible driver of $\Delta \ln q \neq 0$. If data are already available, then a retrospective analysis should be carried out comparing the net geometry on the Andromeda with data from the Africana, correcting for depth effects. A preliminary analysis based on data from the Africana only (summarised in Appendix B) suggests that the standard practice of correcting the survey catch rate by the wing spread eliminates gear geometry impacts from consideration of the survey results.

Changes in catch size structure: It is very likely that if there are catchability differences between vessels using the same gear then this will be manifested as differences in the catch size structure. The size structure data from the Andromeda survey in early 2013 should therefore be analysed to see what evidence there is for this, in comparison with the catch size structure for the results from the Dr Fridtjof Nansen's survey over the same period.

Conclusions and Proposals

This document is intended to stimulate a further discussion about the role of demersal surveys using commercial vessels. The review of the literature suggests that in some cases vessel effects as distinct from gear effects are reported. Other studies which compared vessels using the same gear have not found evidence for a vessel effect. There is some indication however that some of the conclusions in

the literature that vessel is a significant factor may not withstand a more rigorous statistical assessment. One example appears to be that multiple comparison corrections to the overall false positive rate have not been implemented when large numbers of species are considered in the same experiment.

The following situation thus arises. If there is no substantial vessel effect, then it will prove unmanageably expensive to verify this using catch statistics. And if the experiment is reduced in scale to financially manageable proportions, then it introduces additional uncertainty into the stock assessments. Clearly, if a sizeable catchability difference does exist, then calibrations are necessary.

It seems unavoidable that some judgement needs to be exercised as to the likely scale of relative vessel effects, e.g. Africana cf. commercial vessel. A number of factors could be used to inform this judgement. These include, for example:

- A closer look at the catch statistics from the Andromeda survey run in early 2013 (size structure)
- A closer look at gear geometry data for the Africana in previous surveys and for the Andromeda survey run in early 2013
- A critical review of the evidence in the literature for vessel effects (where a common gear was used).
- Any other professional judgement / evidence / informed opinion that can be brought to bear.

The outcome of this process would most likely be a statement about the possible scale of relative vessel effects, which could be viewed as an informative prior. Given this, it would help to go further and assess the benefits to management of inter-calibration studies. One way to do this would be to assess the benefit of calibration studies using the same performance statistics that are used for evaluating candidate OMPs for resource management. A broader suite of possibilities could be explored, involving future likely combinations of:

- Africana fishing old gear
- Africana fishing new gear
- Dr Fridtjof Nansen fishing its own gear as in the past
- A commercial vessel using Africana new gear.
- Inter-calibration work involving some plausible combinations of the above, vessel availability dependent.

References

Anon, 2013. German trawl design. FISHERIES/2013/AUG/SWG-DEM/18.

Brandao, A., Rademeyer, R.A., and D.S. Butterworth. 2004. First attempt to obtain a multiplicative bias calibration factor between the Africana with the old and the new gear. WG/11/04/D:H:26.

Cotter, A.J.R. 2001. Intercalibration of North Sea International Bottom Trawl Surveys by fitting year-class curves. ICES J. Mar. Sci. 58: 622-632.

Fairweather, T., Rademeyer, R., and D. Butterworth. 2013. Abundance Estimates for Hake – Nansen vs Africana. FISHERIES/2013/OCT/SWG-DEM/54.

- Helser, T.E., Punt, A.E., and R.D. Methot. 2004. A generalized linear mixed model analysis of multi-vessel fishery resource survey. *Fisheries Research*. 70: 251 – 264.
- McCallum, B., and Walsh, S.J. 2002. An update on the performance of the Campelen 1800 during bottom trawl surveys in NAFO Subareas 2 and 3 in 2001. NAFO SCR Doc.02/32:16p.
- Pelletier, D. 1998. Intercalibration of research survey vessels in fisheries: a review and an application. *Can. J. Fish. Aquat. Sci.* 55: 2672-2690.
- Smith, A.D.M., Fernandez, C., Ortiz, M. and A.E. Punt. 2012. International Review Panel Report for the 2012 International Fisheries Stock Assessment Workshop. 26 – 30 November 2012, UCT.
- Walsh, S.J. 2006. Effect of vessels and crew on geometry and performance of Campelen 1800 shrimp trawl during comparative fishing in the Newfoundland region in the spring and fall of 2005. DFO Can. Sci. Advis. Sec. Res. Doc. To appear.
- Wilderbuer, T.K., and Kappenman, R.F. 1998. Analysis of Fishing Power Correction Factor Estimates from a Trawl Comparison Experiment. *N. Am. J. Fish. Manag.* 18: 11-18.

Table 1. The relative logarithm of the catchability between the Africana and the Nansen, for the African using the old gear. n is the number of paired tows required to reject the null hypothesis when the $\Delta \ln q$ is as given on the left at $\alpha = 10\%$ and $1-\beta = 80\%$.

$\Delta \ln q$	$n =$ number of paired tows
0.053	1510
0.100	430
0.200	110
0.300	50
0.400	30

Table 2. The mean value of $\Delta \ln q$ when a prior of $\Delta \ln q \sim N(0,0.05^2)$ is used, and the experimental result is as given in the first column, for an experimental variance on the estimate of $\Delta \ln q$ of 0.0827^2 .

$\Delta \ln q$ experiment, $n = 100$ paired tows, S.E. = 0.0827	Mean of $\Delta \ln q$ in posterior
0.053	0.014
0.100	0.027
0.200	0.054
0.300	0.080
0.400	0.107

Appendix A: Research ship Africana.

BACKGROUND

Africana, was built and commissioned in 1982 for the then Sea Fisheries Research Institute. She has achieved close to 40 000 scientific stations. The Africana is based in Cape Town and owned by the Department of Agriculture, Forestry and Fisheries.

VESSEL PARTICULARS:

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 T
 Nett tonnage: 617.78 T
 Length: 69.886 m
 Breadth: 15,25 m
 Horsepower: 1790 kW (2400 BHP)
 Complement: 52 comprising 34 crew and 18 scientific/other staff

PROPULSION:

Single screw, Diesel Electric
 3 x APE Allen Diesel Engines each developing 1044 kW at 600 rpm
 3 x 750 kW Siemens DC Generators supplying Siemens Motors of 1790 kW at 174 rpm.
 Bunker capacity 532 tonnes

ELECTRICAL POWER:

2 x 550 kW Siemens Alternators driven by 2 of the 3 Main Engines.
 Auxiliary power by 437 kW Deutz Diesel Engine driving 350 kW Siemens Alternator.

DECK MACHINERY:

Hydraulic Brattvag Windlass
 2 Pesci PI700 Cranes 0.5 ton
 Petrel Knuckle Boom Crane SWL 2.8 tonnes at 10 meters
 4000m x 11.88mm Conductor Cable
 Petrel Stern Towing Winch 3 tons at 60m per minute
 Clarke Chapman Main Trawl Winch: 2 x 25 ton finger capstans, 2 storage drums each with 5000m x 26 mm Warp
 Petrel Net Drum Winch 7.5m – 6.2 tons
 Petrel Auxiliary Trawl Winch 2 x 10 tons
 Petrel Fishing/Mooring Winches 2 x 10 tons
 Elac Net Sonde Winch 2500mm Cable, constant tension
 Clarke Chapman Hydrographic Winch Single Finger Capstan with 1 storage drum
 5000m x 9.5mm Conductor Cable and 1 storage drum 6000m x 6mm SWR.
 Lebus Large Towing Winch with 1 drum 2500m x 12.5mm Conductor Cable and 1 drum 2500m x 12.5 Conductor Cable.
 Lebus Vertical Plankton Winch with 1 drum 1000m x 7.2 Conductor Cable
 Lebus Small Towing Winch with 1 drum 1000m x 9.5mm Conductor Cable

ACOUSTICAL CHARACTERISTICS:

Diesel electric configuration
 Isolation of reciprocating machinery and special treatment of hull and propeller, permits acoustics surveys to be conducted at speeds of up to 12 knots.

OTHER FEATURES:

Passive Stabiliser Tanks 2 x 60 ton
 Retractable Bow Propeller, 360° rotation (Azimuth Thruster)
 Hospital, two bed, fully equipped with Dispensary

REPLACEMENT PROGRAMME

Within next five years

Appendix B:

Do survey gear dimensions during trawling affect hake catch rate?

By
OLRAC-SPS
Silvermine House
Steenberg Office Park
Tokai 7945

16 August 2011

Background

Cruise reports for Voyages 270 and 273 contain information about the dimensions of the new trawl gear during the execution of surveys. A preliminary analysis of the data for Voyage 273 suggested that there was a correlation between the total hake catch and wing spread. This raised some concern since it would demonstrate a link between precisely how the gear is 'flown' and the results of the survey. While the trawling industry believe that this is the case, there has as yet been no clear statistical demonstration of this.

The preliminary result alluded to above was the motivation for a slightly more extensive examination using more sophisticated multivariate techniques.

Methods and Results

Data for both Voyages 270 and 273 were digitised and subjected to a group comparison analysis using SEM (structural equation modelling) software. The available variables included:

1. Nr : Survey trawl number
2. Stn : Trawl location code
3. Date
4. Time
5. GMT
6. Dpth ; Trawl average depth
7. Lat
8. Long
9. Dur : Duration of the trawl
10. Head : average trawl heading
11. Spd : average trawl speed
12. Dist : distance covered by trawl
13. Door_sprd : door spread (metres)
14. Wing_sprd : wing spread (metres)
15. Vert_sprd : net vertical opening spread (metres)

- 16. SST : Sea surface temperature
- 17. Wind_spd : wind speed
- 18. Wind_dir
- 19. Total_Catch : all species total catch
- 20. HKC_catch : catch of Merluccius capensis
- 21. HKP_catch : catch of Merluccius paradoxus
- 22. HakeTot : catch of all hake = HKC_catch + HKP_catch

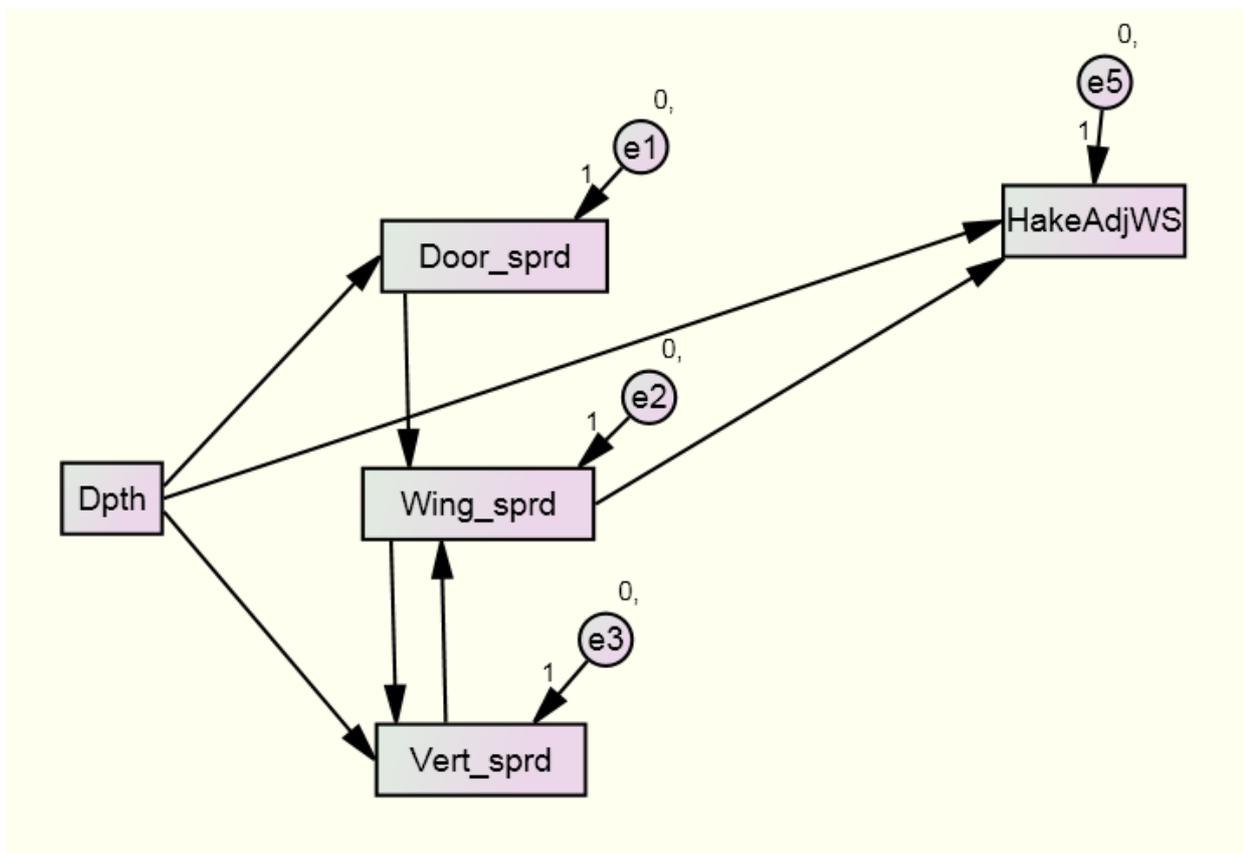
In order to carry out the statistical analyses a proxy for hake density was calculated, HakeAdjWS, where:

$$\text{HakeAdjWS} = \text{HakeTot} / \text{Dist} / \text{Wing_sprd}.$$

The relationship between the following variables was used for further statistical analyses:

- Dpth
- Door_sprd
- Wing_sprd
- Vert_sprd
- HakeAdjWS

A structural equation model (SEM) was constructed to explore the relationship between these five variables. The basic idea behind SEMs is to obtain a model guided by theory which provides an adequate match between the observed variance-covariance matrix (which could include means as moments as well) and the variance-covariance matrix implied by the model. The following path diagram was found to provide a reasonable statistical fit (measured by the χ^2 statistics):



The interpretation attached to this model is that:

- trawl depth has a causal effect on the net dimensions of door spread, wing spread and vertical spread.
- door spread has a causal effect on wing spread and wing spread and vertical spread affect each other.
- there is also a direct affect of trawl depth on the density of hake.

This model has 8 degrees of freedom and $\chi^2 = 8.5$. The associated probability level is 0.383. This version of the model uses the net dimension data for voyages 270 and 273 as separate groups. In its general form separate parameter values are fitted for each parameter for each group. In its present form, for each group, the parameters are:

Regression weights:

1. Door_sprd <--- Dpth
2. Wing_sprd <--- Door_sprd
3. Vert_sprd <--- Dpth
4. HakeAdjWS <--- Dpth
5. Vert_sprd <--- Wing_sprd
6. Wing_sprd <--- Vert_sprd

Means:

7. Dpth

Intercepts:

8. Door_sprd
9. Wing_sprd
10. Vert_sprd
11. HakeAdjWS

Variances:

12. Dpth
13. e1
14. e3
15. e2
16. e5

There are 20 sample moments (5 sample means, 5 sample variances and 10 sample covariances). There are thus $20 - 16 = 4$ degrees of freedom for each voyage, 8 in total considering both voyages.

The model as presented in Fig. 1 posits that the direct link Wing_sprd ---> HakeAdjWS is equal to zero. If this link were non-zero it implies that wing spread could influence the estimate of hake density which would impart bias to the survey. This hypothesis can be tested using a nested model approach, i.e. the difference in χ^2 between the model with and without the effect, and the difference in the number of degrees of freedom ($\Delta df = 2, 1$ for each voyage). It comes down to comparing the χ^2 value with and without the additional effect present. The χ^2 with the effect included is 6.6, which is 1.9 units less than the model which assumes the effect is 0. The p-value associated with this

difference is > 0.05 , i.e. the difference is not statistically significant. We are therefore unable to reject the null hypothesis that the effect is zero.

Conclusion

Based on the data at hand, there is no evidence that the gear dimensions are impacting on the estimate of hake density, over and above the correction which is being made routinely, to divide hake catch by the wing spread and distance trawled to calculate a proxy for hake population density.

Appendix C: Gear design, new gear, demersal surveys in South Africa (from Anon, 2013).

