INTERNATIONAL REVIEW PANEL REPORT FOR THE 2019 INTERNATIONAL FISHERIES STOCK ASSESSMENT WORKSHOP
2 -5 December 2019, UCT

[David J. Die[1], André E Punt[2]*, Ralph Tiedemann[3], Robin Waples[2,4], Michael J. Wilberg[5]]

Summary of general issues

The Panel recognised the very high quality of the research presented at the 2019 International Fisheries Stock Assessment Review Workshop. This included research on stocks of southern African hakes, sardine and rock lobster. New information and research related to an experimental spatial fishing closure near African penguin colonies was also presented.

This report starts with observations from the Panel on some general issues for the species and analysis programmes reviewed, and then focuses on answering the questions posed by DEA Research, providing a more detailed technical review where necessary, and finally recommending further work concerning each topic. The recommendations are annotated by their priorities (H, M, L). Much of this report reflects responses to the questions. For ease of reading, answers to the questions that also have research components are indicated by an asterisk.

Hake

The focus of the Panel related to evaluating the information from new genetics studies, and the implications for stock-structure hypotheses for use in assessments and future OMP evaluations. The results of the hake assessment will feed into the MSC recertification and the Panel deliberated on how hake fit into the meta-population structure descriptions by the MSC. The new microsatellite studies supported the existing two-stock hypothesis for *M. capensis* and provided evidence as to seasonal shifts in population boundaries. For *M. paradoxus*, the new genetics information did not change the 2014 Panel's view that a single stock is the most plausible hypothesis. The Panel refined the set of models to be used for assessment and OMP evaluations for *M. paradoxus*. The models identified by the Panel require data from Namibia, and the Panel **strongly encourages** data sharing and ideally collaborative development of assessments and management strategies between Namibian and South African scientists.

Sardine

The Panel evaluated several supporting analyses for long- and short-term projections (and hence management advice) for sardine. The currently depressed state of the resource requires new approaches to making catch limit recommendations (e.g., how to conduct short-term projections given that "Exceptional Circumstances" provisions have been invoked, and the agreed OMP has consequently had to be set aside) and introduces new challenges, such as how best to use the (apparently anomalous) length-frequency distributions from the 2018 survey in analyses. The Panel emphasizes the importance of continued monitoring of the population, and the application of methods for analysing survey data to provide biomass estimates that are more robust to the patchy distributions expected for a population at low abundance.

West Coast rock lobster

The Panel evaluated questions in relation to how to develop projections for rock lobster in the context of a stock that is severely depleted, and did not rebuild towards previous management

1 Department of Marine Ecosystems and Society, Rosenstiel School of Marine and Atmospheric Sciences, Miami USA

2 School of Aquatic and Fishery Sciences, University of Washington, USA (participated remotely by Skype to the extent possible)

3 Unit of Evolutionary Biology/Systematic Zoology, University of Potsdam, Germany

4 Northwest Fisheries Science Center, NOAA, Seattle, USA

5 Chesapeake Biological Laboratory, University of Maryland, Center for Environmental Science, Solomons, MD, USA

targets, in spite of having been managed under agreed OMPs with such objectives because of lower recruitment than anticipated and a substantial increase in poaching. Recommendations are made on how estimates of recruitment might be improved through a change in the definition of recruitment, which should also simplify the assessment model. This change may also lead to reductions in uncertainty of estimates of recruitment and will reduce the time lag between the assessment year and the most recent value of estimated recruitment.

Penguins
The closure experiment to provide information on the effects of fishing near colonies on penguin population parameters continues to provide information and increases the power to detect effects of fishing (if they exist). In principle, use of data for individual penguins is to be preferred over aggregated data, particularly when there are covariates that pertain to individual penguins. However, using individual penguin data without appropriate accounting for the statistical problems that can arise due to common factors jointly impacting the responses of penguins may violate assumptions regarding statistical independence among individual penguins, and can lead to parameter estimates, including the effects of fishing on penguins, that appear more precise than is actually the case. The Panel reviewed models for analysing individual penguin data and provided recommendations for model exploration and simulation evaluation.

The Panel notes that few of the scientists at the 2019 workshop appear to be fully aware of the details of the models (and more importantly their underlying biological assumptions). The Panel suggests that a "simple persons' guide" be developed to explain these assumptions to broaden the group of scientists capable of discussing these models meaningfully. Biologists should also be more involved in identifying potential covariates for use in models. The Panel appreciated results provided by "two groups", but noted that poor communication between these groups probably reduced the value of the work conducted, and compromised the comparability of the analyses, resulting in additional requests for analyses in short time.

The Panel reiterates the 2014 Panel's note that "cessation of fishing around the islands by itself is unlikely to be sufficient for the penguin population to recover", and also reiterates the **recommendation** from that Panel that "appropriate authorities work together to identify goals for both the pelagic fishery and penguin recovery, to develop and implement a comprehensive research program that aims to identify the core reasons for the reduction in penguin population numbers, and identify any potential mitigation measures".

General issues
The Panel was requested to comment on the appropriate fraction of population growth to be allocated to resource recovery (instead of to the fishery) for very depleted populations (west rock lobster and sardine). Some jurisdictions have very clear rules regarding rebuilding strategies for depleted populations (e.g., the USA), while other jurisdictions (e.g., Australia and New Zealand) have agreed harvest policy documents that outline general principles that relate to depleted populations. The key decisions when developing a rebuilding plan are how much of the expected population growth should be allocated to rebuilding and how should account be taken of uncertainty. These are policy decisions, but the Panel **strongly advises** that the ability for scientists to support decision making will enhanced if a general national policy is developed that includes both of these considerations. The more guidance managers provide on the actions to be taken when an OMP has to be supplanted because of Exceptional Circumstances, the better. Such guidance should be agreed upon not only in advance of the implementation of a new OMP, but also in the current Exceptional Circumstances, which have necessitated annual best assessment-based recommendations for the immediate future.

The Panel process was mostly effective: 1) the terms of reference were clear and specific, 2) the documents provided were informative and mostly sufficient to inform on the topics to be addressed by the Panel, 3) all participants were given a reasonable opportunity to express their opinions, and 4) the organizers and local scientists were quick to provide additional information required by the Panel. There were, however, some challenges to the process,

including 1) the amount of information required to be reviewed prior to the meeting, 2) the number of topics to be addressed and 3) the specificity of some of the questions that could only be addressed by a small subset of the Panel members, and 4) the narrowness of the methodological questions related to the penguin analyses. Organizers of the review, and particularly session chairs, did everything in their power to help the Panel to overcome such challenges, and for that the Panel thanks them. In the future, however, consideration should be given to reduce the number of topics to be reviewed by the Panel.

B. Hake

B.1 In 2014, the Panel considered a single paradoxus stock in SA and Namibia the most plausible stock structure hypothesis for paradoxus (see MARAM/IWS/2019/Hake/P1). Does more recently available information indicate a need for a change in that view, and if so to what?

No, a single panmictic population in South Africa and Namibia is still the most plausible hypothesis for *M. paradoxus*. However, as was the case in 2014, there are some scenarios involving more than one stock that cannot be rejected by the current genetic data due to lack of statistical power. An earlier figure (reproduced in Figure 2 of MARAM/IWS/2019/Hake/BG7) depicting the parameter space consistent with cryptic stock structure has been updated to reflect new information in MARAM/IWS/2019/Hake/P2 (Figure 1). Appendix 2b describes a technical evaluation of the new genetic information.

B.2 The MSC ask that metapopulations be categorised in terms of Table G2 in the excerpt below from their Standard; which option does the Panel consider provides the most appropriate categorisation of the SA-Namibian hake complex?

- *M. capensis*: Category B with one northern local population (Namibia) and one southern population (South Africa), as detailed in Figure 2.
- *M. paradoxus*: The Panel notes that categories A and D are equivalent from a genetics point of view. Spatially-explicit spawning/length data (presented in MARAM/IWS/2019/Hake/P6) indicate that the major spawning area is mostly in northwestern South Africa, but also straddling into southwestern Namibia (Figure 2), from which some of the recruits go north (Namibian waters), while the remainder stay in South African waters. Under the understanding that category D essentially comprises a single panmictic (meta-) population, within which local stocks may be identified for demographic and/or management reasons, this species could be categorized as D. Categorization solely based on currently available genetic data suggests categorization as A, although current genetic data do not have sufficient power to rule out the existence of more stock structure (see response above to question B.1). In summary, the definitions of the categories were unclear and conflicted in some parts, so that the Panel was unable to unambiguously assign *M. paradoxus* to a specific category.

B.3 What are the priorities for future genetics data collection and analysis (including analysis of existing samples)?

The Panel encourages developing and utilizing the genomic resources, i.e., informative Single Nucleotide Polymorphisms (SNPs). The double digest restriction site associated DNR (ddRAD) approach presented in MARAM/IWS/2019/Hake/P3 is generally suitable. However, future analyses should: 1) substantially increase the absolute number of sequence reads per typed specimen; 2) type SNPs such that the data can be assigned to individuals (i.e., by individual-specific labelling); and 3) ensure that sampling areas of interest are representatively covered. The sampling should specifically include the spawning area (as depicted in MARAM/IWS/2019/Hake/P6) and samples to the north and south of it (respectively in Namibian and South African waters). Within the spawning area, spatially denser sampling is **recommended** to test for any small-scale heterogeneity, i.e., the northernmost (southernmost) spawners exhibiting genetic affinity to the specimens migrating to Namibia (South African) waters. The suggested individual-based genomic approach is less prone to biases due to

differences in sample quality, as low quality samples can be identified *post hoc* and eventually dismissed. The use of samples from previous collections facilitates analysis of utilization of already existing samples.

B.4 How should the various SA-Namibian hake stock structure alternatives shown schematically in Hake/P4 be amended, given more recent information, for initiating further assessment analyses?
See response to question B.5 below (and Figure 3).

B.5 MARAM/IWS/2019/Hake/P4 proposed, and the 2014 Panel concurred [MARAM/IWS/2019/Hake/P1], that the first (base case) joint assessment attempt be based on a fleets-as-areas approach, which implicitly assumes complete "re-mixing" of stocks each year so as to distribute themselves in an unchanged way from year to year. Should this remain the first step. And if not, what else?
There should be two basic model structures for *M. paradoxus*, which is the species of the greatest current conservation concern, given that it appears that the stock(s) of *M. capensis* fished in South Africa are not shared with Namibia and those stocks are most probably at a high proportion of their unfished level(s).

- Base case: A single (homogenous) stock found from northern Namibia to the south coast (Figure 3-left). This model will need to include multiple fleets, some of which will be found in Namibia and others in South Africa. These fleets should be modelled using the fleets-as-area approach, which eliminates the need for multiple spatial strata.
- A two component (nominally "South Africa" and "Namibia") model, with the two components overlapping south of Luderitz to northern South Africa (the "spawning area"). This model should also be based on a fleets-as-areas approach. The specifications of this model will need to be completed and agreed before coding begins, and the Panel makes the following comments and suggestions:
  - Consider the case in which density dependence for each stock is a weighted sum of the spawning biomass by component (e.g., MARAM/IWS/2019/Hake/P7), and in which density-dependence is common to the two components. In the latter case the model needs to calculate total recruitment from the stock-recruitment relationship, estimate annual recruitment deviations and estimate a set of parameters for the proportion of total recruitment contributing to each component.
  - It should be feasible to estimate the proportion of total recruitment to each component given survey data that are comparable for the regions in the model.
  - The parameters (*a* and *b*) determining the weighting of spawning biomass (section II of MARAM/IWS/2019/Hake/P7) are likely inestimable, so sensitivity will likely need to be examined to alternative values.
  - The only movement is that of pre-recruits, and this occurs after density-dependence.

B.6 The base case multiple stock model would assume no permanent interchange between any two stocks of the same species, and furthermore that the spatial distribution (in relative terms) of any one stock at the start of each year was time invariant (as is implicit in the fleets-as-areas approach).
  a) **Is there a need to consider the possibility of permanent interchange between different hake stocks of the same species; if so, how might the rate of that interchange be best estimated from (e.g.) genetic data (or alternatively specified in some way)?**
  b) **If there is a need to assume some "inertia" in a stock's recovery of its original spatial distribution following spatially different fishing mortality levels exerted the previous year, how best would this be modelled, and how might one estimate how quickly the distribution would be expected to revert to its**

original pattern?

The Panel notes that the assumption of "no permanent interchange between any two stocks of the same species" – although implemented in the base case model – is biologically unrealistic, as such complete reproductive isolation would generally qualify the respective strata as different species (not stocks or populations), or at least incipient species. A more realistic assumption would be that separate populations of the same species exchange genes/migrants at a rate sufficient to keep them parts of the same species, but low enough to lead to demographic independence.

While there is no need in the base case model to consider permanent exchange of individuals among stocks, this could be a sensitivity (robustness) test for the two-component model. For *M. paradoxus*, inference from F_{ST} and effective population size values (Figure 1) could be used to develop exchange rate hypotheses, while for *M. capensis*, estimates of interchange rates could be derived from F_{ST} values or population assignment (e.g., Figures 3 and 5 in MARAM/IWS/2019/Hake/P2 and information on seasonal movements in Figure 3 in MARAM/IWS/2019/Hake/P3).

In relation to b) the two model structures outlined in the response to question B.5 are extremes in regard to interchange (perfect interchange to potentially no interchange after larval dispersal). The second model structure implicitly includes inertia with respect to the redistribution of fish following fishing.

B.6 Other recommendations

B.6.1 (*) The success of the proposed *M. paradoxus* assessment will depend on the information available for Namibia. The ideal case is for the two models to be fitted to data for Namibia and South Africa and the Panel **strongly encourages** data sharing and ideally collaborative development of assessments and management strategies between Namibia and South African scientists. SADSTIA (2019) outlines a proposal regarding data needed from Namibia. The minimum amount of data needed to implement requisite models is catches by species north and south of Luderitz over time, although assumptions will need to be made regarding the selectivity patterns (which include both gear selectivity and availability) for the fisheries off Namibia if these are the only data provided.

B.6.2 (*) Whether the totally homogeneous model is viable can be examined using the residuals for the survey indices; should there be patterns in these residuals (assuming that survey selectivity is time-invariant – except due to changes of gear), this would indicate a lack of homogeneity.

B.6.3 (*) The evaluation of the impacts of fishing in Namibia on the South African fishery for *M. paradoxus* will require making hypotheses about future catches of *M. paradoxus* off Namibia.

B.6.4 (M) The assumption that there is no exchange of post-recruits in the two-component model may lead to unrealistic results for the Namibian component (e.g., exceptionally high fishing mortality on old animals or an extreme "dome" to selectivity patterns). This may imply the need for post-recruit migration to be included in the two-component model.

B.6.5. (M) Explore the potential utility of close-kin genetic analyses (focusing on parent-offspring pairs and/or siblings) to provide insights into the stock structure of *M. paradoxus* (e.g., Økland et al. 2009). Evaluations should start by estimating the sampling effort that would be required to produce useful information because close-kin matches will be very rare in a population as large as $\sim 10^{8}$ adults..

C. West Coast Rock Lobster

C.1 Have the Panel's 2018 recommendations about the analysis of the compliance data to estimate poaching trends been appropriately addressed?

The approach in MARAM/IWS/2019/WCRL/P1 follows the recommendation of the 2018 Panel and improves it by applying a truncated negative-binomial distribution for the new database and by estimating separate overdispersion parameters for the 'old' and 'new' databases. Thus, the 2018 Panel recommendation has been appropriately addressed.

C.2 Is it worth pursuing the $CPUE \propto q\sqrt{B}$ relationship for commercial CPUE further in updated assessments?

The CPUE of male lobsters above the minimum size (hereafter, "legal male lobsters") in the fishery should be compared with an index of legal male lobsters developed from the fishery-independent surveys (FIMS). This comparison will help determine whether it is appropriate to continue with the hypothesis $CPUE \propto qB$ or it is worth considering an alternative hypothesis such as $CPUE \propto q\sqrt{B}$. The relationship between CPUE and biomass should be modelled as $CPUE \propto qB^{\alpha}$ and not just $CPUE \propto q\sqrt{B}$ if an alternative to $CPUE \propto qB$ needs to be considered. One issue with estimating the α parameter is that there are multiple super areas with independently fitted models. Two approaches may be attempted: 1) estimate α separately for each super area, and 2) specify α at a common value across super areas. Given likely differences in the spatial distribution of the fishery and the lobsters among the super areas, it is theoretically possible for α to differ among super areas, which supports option 1. Option 2 is easier to implement, but choosing among values of α is more difficult.

C.3 Is the method now used to estimate (particularly recent) and projected recruitment satisfactory; if not what alternatives are suggested?

The current method to determine recruitment in the forecast period is to assume that it is a stochastic process with mean obtained from historical estimates of recruitment and deviations about the mean also estimated from historical recruitment estimates. The method requires choosing a period over which such means and deviations are calculated, presently from 1970+ to the last year with reliable estimates of recruitment. MARAM/IWS/2019/WCRL/P2 raises the concerns: 1) should the most recent estimates of recruitment should be used in the determination of projections because of their high uncertainty, and 2) whether the possibility that the stock is close to the point when recruitment will be declining as a result of the low levels of spawning stock biomass needs to be considered.

Recruitment is currently defined as the number of very small (1-15 mm CL) lobster settling into the population, but there is no information in the current data that directly informs the abundance of such lobsters. The estimation of recruitment relies on the backward reconstruction of numbers-at-size from size-classes that are present in the data (e.g., size classes observed in the FIMS data set), on the basis of assumptions about growth rates (modelled as a moulting process that combines the intermoult periods and moult increments) and natural mortality. The model correctly accounts for yearly changes in growth rates, but assumes stationary mortality-at-size and no mortality associated with moulting. It is well known that natural mortality of crustaceans is highest during moulting and shortly thereafter.

Growth rates have been estimated to have been decreasing over time so lobsters would take longer to grow from the settlement size to the sizes observed in the FIMS samples and the fishery. Moreover, natural mortality will also increase as growth declines and lobsters grow more slowly because natural mortality is assumed to be greater the smaller the lobster. This implies an increased cumulative mortality between settlement and the sizes that are first seen in the data. This increased mortality translates into larger numbers of recruits required to sustain any given harvest. This increase in natural mortality of lobsters prior to joining the harvestable portion of the population may partially explain why recruitment seems to be increasing while stock biomass continues to decline.

There are two ways to address this concern:

- define recruits in relation to the size of the smaller lobsters observed in the FIMS data set; or
- define natural mortality in different ways: for example, $M = f(\text{size, duration of intermolt period, moulting process})$.

Neither of these options are likely to completely eliminate the potential problems associated with changes in growth rate. However, the extent of the strength of the effect of growth rate on recruitment estimates can be investigated through implementation of these changes.

C.4 Recent recruitment is poorly estimated; would approaches other than that used currently be preferable?
The Panel identified two alternatives:
- redefine the size of recruits in the model (see response to question C.3) - if recruits are redefined as much larger lobsters than is currently the case, the lag between observations and recruitment estimates could be reduced considerably, and depend less on the assumptions about natural mortality of very small lobsters, consequently allowing more reliable and timely estimates of recruitment; and
- conduct retrospective analyses on recruitment estimates (regardless of how recruits are defined) to understand whether recent estimates are biased - retrospective results could help determine the lag that needs to be considered before recruitment estimates are deemed to be reliable).

C.5 Should projections include taking account of stock-recruitment effects (and in the manner suggested)?
Short-term projections (<5 years) can be made on the basis of estimates of recent recruitment without considering stock-recruitment relationships. The most recent recruitment estimate/s should be used only if they are estimated with reasonable confidence. The choice of how many estimates of past recruitment should be used for short-term projections is generally a subjective choice. This should be related to the number of years to be projected (the greater the number of years to be projected, the greater the number of historical estimates to be used).

Longer-term projections should use a stock-recruitment relationship, if available, and especially if there is an indication that recent spawning stock status is very low. It is challenging to predict to what a redefinition of recruitment size would lead in terms of stock-recruitment relationships. For some lobster populations, however, it has been hypothesized that the main density-dependence processes are associated with competition for hiding spaces, which protect lobsters from predation. This hypothesis assumes that such processes act upon a wide range of lobster sizes, and are related to the availability of hiding spaces. Consequently, the stock-recruitment relationship may change depending on how recruits are defined.

C.6 How might the basis used to provide TAC recommendations be improved?
See Section F of the report for responses related to the basis for providing short-term TAC recommendations.

C.7 Other recommendations
C.7.1 (H) Projections of the rock lobster biomass should show projected catches, the proportion of population growth allocated to resource recovery versus harvest, and the probability of not achieving the target reference point.

C.7.2 (M) The analysis of the compliance data to estimate poaching trends should consider treating the year-effects as random rather than fixed effects.

C.7.3. (M) Estimates of equilibrium yield-per-recruit for each year and area during the period 1970-present should be calculated to demonstrate the effects of changes in lobster growth.

Estimates should incorporate the growth curve and appropriate selectivity operating in each year and area.

C.7.4 (M) Survey approaches have been developed to estimate illegal activities such as poaching. One such method is Randomized Response Techniques (e.g., St. John et al., 2010). This type of approach asks a survey respondent to answer the question truthfully based on a random event such as a flip of a coin. This means that one cannot tell if an individual participated in an illegal activity, but the aggregate results can be used to estimate the "inappropriate" behaviour.

C.7.5 (M) Many decades of high harvest rates on large male lobsters may have caused an evolutionary response in the population toward a lower growth rate and maturation at a smaller size (fishery-induced evolution; FIE); this possibility could be explored. In addition, the empirical data for changes in size at maturity of female rock lobsters should be examined. Marine ectotherms such as rock lobsters that have indeterminant growth and increasing fecundity with age are particularly prone to FIE, which has been empirically documented in several marine fish species (Barson et al., 2015; Enberg et al., 2012; Heino et al., 2012, 2013). Even if direct selection occurs primarily on males, females can also be expected to show reduced growth rates as well because of sharing of genes through reproduction. If FIE affects growth of young lobsters, it may lead to a reduction in survival of recruits (because they take longer to grow to larger sizes), and may thus cause a decline in the productivity of the population.

D. Sardine
D.1 Short term projections, particularly to November 2020
D.1.1 What it the best way to account for the under-representation of large sardine in the November 2018 survey length frequency? What is the best way to determine if a commercial or survey length frequency should be down weighted (in particular, for example, the 2019 length frequencies)? If a survey length frequency is biased or unrepresentative, what is the best length frequency (e.g. weighted survey and commercial) i) to apply to the interval-specific densities used in calculating the survey estimate of biomass and ii) to use as data to condition the assessment model. If a commercial length frequency is biased or unrepresentative, what is the best length frequency for conditioning the assessment model?
The approach for developing length-frequency distributions for the November 2018 survey in MARAM/IWS/2019/Sardine/P3 involves pooling the survey and commercial length-frequencies. However, while there must have been some large sardine when the survey took place, there is no basis to select a weighting for the two types of length-frequencies. In addition, there is no way to account for spatial structure when conducting this pooling. Moreover, differences between survey and commercial length-frequencies have been observed in the past, meaning that a decision may need to be made whether to pool survey and commercial length-frequency data in the future. The survey estimates should be based on the observed survey frequencies (at least in the short-term), noting that biased or unrepresentative length frequencies are likely a consequence of relatively few trawls containing sardines, and often very small sample sizes when sardines are caught, caused by a small population biomass.

D.1.2 What is the best way to model recruitment in the *immediate future*[1] for both the west and south components of South African sardine?
Short-term projections should not be based on the entire time-series of estimates of recruitment and the most recent 5-10 years of recruitments seems appropriate (Wiedenmann and Jensen, 2019). However, the Panel is concerned that the November 2017 estimate of recruitment is likely quite imprecise (CV of recruitment of about 130% compared to an average CVs for the 2013-16 recruitments of about 40%). Appropriate ways to address this concern are to drop the most-recent estimate when generating future recruitment, or to compute an inverse-variance

weighted average (and SD) of the five most-recent estimates of recruitment and then to generate future recruitments from the resulting distribution. The Panel expressed a preference for the first of these options because it would be simpler, and because there was a clear difference in (the greater) precision of recruitment estimates for the previous four years relative to the most recent year.

D.1.3 For the highly depleted sardine resource under Exceptional Circumstances provisions, the decision on TAC and TAB recommendations amounts to the choice of what fraction of any projected population growth should remain in the resource to improve future recovery prospects, and what should be taken as a harvest, noting the high dependency of 1-year ahead projections for this short-lived species on the assumed incoming recruitment. What are typical splits used in these circumstances for similar fisheries elsewhere?
See Section F of the report for responses related to the basis for providing short-term TAC recommendations.

D.2 Long-term analyses / General Best Practice
D.2.1. As D.1.1, but with a longer-term focus.
The potentially biased or unrepresentative length frequency for November 2018 is likely a consequence of a small population biomass, which may continue for some years. The Panel advocates using a model-based smoother approach (i.e., develop a two-dimensional density surface for length-frequency) to make inferences regarding survey length-frequency spatially to make use of any spatial correlation structure in the population length-frequency. Woillez et al. (2016) provide one example of an approach for such modelling, while the general vector autoregressive state space modeling platform VAST (Thorson and Barnett, 2017) could also be used. There may be value in the future of using spatial models to predict species mixes in the survey based on the survey catches.

D.2.2. How might one best model future recruitment to both the west and south components of South African sardine?
The Panel supports the current approach (MARAM/IWS/2019/Sardine/P3) of using a stock-recruitment relationship and sampling residuals non-parametrically for the west coast and using the 2-stage model for the south coast.

D.2.3. What is the best way to account for the apparent conflict in assumptions about the shape of the growth curve and range of t_0 between the fit to the data (particularly the parasite prevalence-at-length) and 'conventional understanding' that peak spawning/recruitment occurs between September and January for the west component.
It is not uncommon for the growth not to follow the von Bertalanffy growth curve shape, and alternative curves (e.g., Richards) might provide a better fit. In addition, two-stage growth curves are implemented in some assessment packages (e.g., Stock Synthesis), where a different functional form is assumed for young ages. Alternatively, the L_∞ or κ parameters could be treated as estimable (even though variation among years in t_0 is more plausible biologically).

D.3. Other recommendations
D.3.1 (H) Projections of the sardine biomass should show catches, the proportion of population growth allocated to resource recovery (instead of harvest), and the probability of not achieving an interim rebuilding abundance target.

D.3.2. (M) Consider basing long-term projections of recruitment on a relationship between recruitment and the environment (if such a relationship can be found and downscaled climate forecasts are available). This could be used as a sensitivity or as a robustness trial in OMP evaluations.

D.3.3. (M) Consider automating the process of assigning trawls to intervals, for example using the approach adopted by NOAA Southwest Fisheries Science Center for its acoustic surveys of anchovy, sardine, mackerel and herring.

D.3.4. (M) The problem that the maturity-at-age relationship does not asymptote to 1 is expected given that maturity is specified as a function of length and there is variation in length-at-age. However, this should not be an analysis concern because spawning biomass, while reported in tonnes, is treated as a relative index for management purposes.

D.3.5. (M) Consider a three- rather two-parameter logistic selectivity function for survey selectivity to improve the fits to the survey length-frequency data.

E. Penguins
E.1 Do the estimation models of PENG/P4 based on fits to responses for individual penguins produce negatively biased estimates of the standard error of the parameter related to the impact on penguins of fishing in the neighbourhood of island colonies?
The types of models in MARAM/IWS/2019/PENG/P4 are capable of providing estimates of parameters (including the effect of closures on penguin population parameters) that are negatively biased if covariates common to individuals are ignored (i.e., pseudo replication). Random effects models are used to account for such "latent" covariates in designed experiments. However, in the natural experiments such as the closure experiment, it is a working hypothesis that including random effects chosen using model selection methods will appropriately account for the pseudo-replication. This is a working hypothesis because it can never be guaranteed that including random effects will fully address pseudo-replication when the "true" sampling design is not known (i.e., how individuals are selected). Results presented to the Workshop suggest that estimates of closure parameters using models fitted to aggregated and individual data had similar standard errors (at least for the effects of closure parameter for the South Coast colonies evaluated), but they may differ for other population parameters. It is also difficult to compare the precision of the standard error estimates because the models do not always have the same covariates. The Panel makes the following **recommendations**:
- Given the nature of the experiment, use of individual data is to be preferred. However, this is only the case if an appropriate random effects structure is chosen.
- Model selection methods should be applied to select an appropriate random effects structure. The penalty for model complexity for complex random effects models needs to be carefully considered (e.g., Spiegelhalter et al., 2002).
- Use of a Bayesian fitting process combined with the Widely Applicable Information Criterion (WAIC; Watanabe, 2010) comparisons should be used to check model selection.
- The set of covariates to consider in the analyses should be identified by the relevant DEA working group. This requires an understanding of how the individual data are collected.
- Best practices for fitting mixed effects models (e.g., Zuur et al., 2009) should be followed (if this is not the case already). This should include standard residual analysis as well as residual analyses that are tailored to the problem at hand (e.g., temporal, spatial or within-season plots of residuals).

E.2 Has adequate adjustment been made for the non-independence of data in the individual-penguin-based estimation models of PENG/P4?
See response to question E.1.

E.3 What can be concluded from the simulation studies presented for various estimation models for the parameter estimating the impact on penguins of fishing in the

neighbourhood of islands; are further analyses of this type needed; if so, please provide suggestions as to their specification in general terms?

The simulations conducted to date have been useful in focusing attention on the potential issues associated with the difficulty to correctly address pseudo-replication. A key uncertainty at present is lack of validation of the working hypothesis that a "good" random effect is adequate even when it is not the "correct" random effect. This should be explored further by constructing a simulation experiment with multiple possible random effects (e.g., year, month within year, individual if multiple measurements occur on the same individual) and including testing of the model selection process. To the extent possible, the simulation should be tailored to reality (e.g., effect sizes that match those in the data, sample sizes that match the data, appropriate error variances, etc.). The simulation study should also further examine lack of balance (e.g., missing individual covariates, years, months, etc.) which could impact the performance of methods based on aggregated or individual data.

E.4 Particularly in the context of responses to the preceding questions, comment on the reliability of results presented for the impact on penguins of fishing in the neighbourhood of the South Coast colonies at Bird and St Croix islands.

The Panel did not address this issue. The Panel **recommends** shifting the debate from estimation methods to the consequences of the estimates. Thus, it is important to conduct projections of penguin population dynamics given the estimates from the current studies (particularly chick survival) and measures of their uncertainty.

F. The basis used to provide TAC recommendations be improved for depleted populations?

F.1 General policy considerations

The Panel is not aware of cases where short-term decisions are needed for depleted stocks, with advice based on projections of an assessment model, for the jurisdictions in which they work. Rather, management advice for depleted stocks in those jurisdictions is based on pre-agreed policy. For example, within the US, rebuilding plans are needed for all stocks designated to be overfished (biomass less than 50% of that corresponding to MSY), with timelines for rebuilding based on mean generation time and the time to rebuild to a target biomass in the absence of future removals. Furthermore, many of the management plans in the US specify that the fraction of population growth allocated to the fishery should decrease as population biomass decreases. The Panel offers the following general **advice**:

- The proportion of the resource growth assigned to recovery, rather than to the fishery, should be higher the more depleted the population.
- Projections should be undertaken, following which a decision should be made on how much of the population growth projected to allocate to the resource, with that decision based on median projections.
- The resultant amount calculated to be allocated to the fishery should be reduced, based on the extent of uncertainty.
- Uncertainty can be computed using simulations that report probabilities of recovery and simulation intervals, but these depend on model assumptions. Thus, adoption of a pragmatic approach, such as reducing the allowable catch by 20% (the default in the Northeast US), is likely be more robust.

F.2 Case-specific considerations

The general advice in F.1 needs to be modified by case-specific considerations, including:

- Longer-lived species should be managed more conservatively than shorter-lived species because of their longer expected recovery times, and the consequential increased probability of depensatory effects coming into play.
- There may be transient dynamics effects which impact estimated medium- to longer-term rates of recovery (e.g. a large recruitment may lead to a high predicted short-term rate of increase which then reduces as more average recruitments enter the population).

- Some species have dependent predator populations, such that lower rates of recovery could have detrimental impacts on the (generally much longer-lived) predator populations.

References

Barson N.J., Aykanat T., Hindar K., Baranski M., Bolstad G.H., et al., 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in Atlantic salmon. *Nature* 528: 405–408.

Enberg, K., Jørgensen, C., Dunlop, E.S., Varpe, Ø., Boukal, D.S., Baulier, L., Eliassen, S. and M. Heino, M., 2012. Fishing-induced evolution of growth: concepts, mechanisms and the empirical evidence. *Marine Ecology* 33(:1-25.

Heino, M., Baulier, L., Boukal, D.S., Ernande, B., Johnston, F.D., Mollet, F.M., Pardoe, H., Therkildsen, N.O., Uusi-Heikkilä, S., Vainikka, A. and R. Arlinghaus. 2013. Can fisheries-induced evolution shift reference points for fisheries management*? ICES Journal of Marine Science* 70L707-721.

Økland, J.M., Haaland, Ø.A. and Skaug, H.J., 2009. A method for defining management units based on genetically determined close relatives. ICES Journal of Marine Science, 67:551-558.

SADSTIA. 2019. A request for data related to the Namibia hake resource, submitted in terms of MCS rules. Letter dates 7 August 2019.

Spiegelhalter D.J., Best N.G., Carlin B.P., and A. van der Linde. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society* 64:583–616.

St. John, F.A.V., Edwards-Jones, H. Gibbons, J.M. and J.P.G. Jones. 2010. Testing novel methods for assessing rule breaking in conservation. *Biological Conservation* 143: 1025–1030.

Strømme, T., Lipinski, M.R. and P. Kainge 2016. Life cycle of hake and likely management implications. *Reviews in Fish Biology and Fisheries*. 26: 235–248.

Thorson, J.T. and L.A. Barnett. 2017. Comparing estimates of abundance trends and distribution shifts using single-and multispecies models of fishes and biogenic habitat. *ICES Journal of Marine Science* 74: 1311–1321.

Watanabe, S. 2010. Asymptotic Equivalence of Bayes Cross Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research*. 11: 3571–3594.

Wiedenmann, J. and O.P. Jensen. 2019. Could recent overfishing of New England groundfish have been prevented? A retrospective evaluation of alternative management strategies. *Canadian Journal of Fisheries and Aquatic Sciences* 76: 1006-1018.

Woillez, M., Walline, P.D., Ianelli, J.N., Dorn, M.W., Wilson, C.D. and A.E. Punt, A. 2016. Evaluating total uncertainty for biomass- and abundance-at-age estimates from eastern Bering Sea walleye pollock acoustic-trawl surveys. *ICES Journal of Marine Science* 73: 2208-2226.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16: 97–159.

Zuur A.F., Ieno E.N., Walker N.J., Saveliev A.A., and G.M. Smith. 2009 Mixed Effects Models and 1134 Extensions in Ecology with R Springer, New York
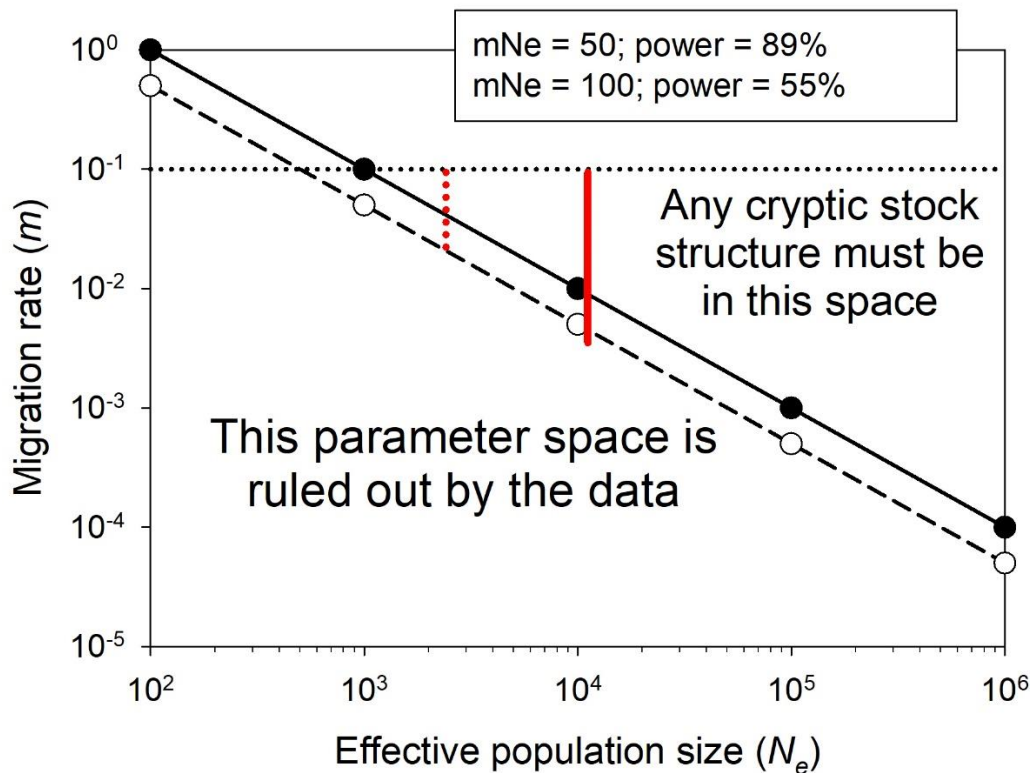
Figure 1. The parameter space that is and is not compatible with more than one stock of *M. paradoxus* [Modified from the 2014 Panel Report]. The diagonal lines plot combinations of effective population size (*Ne*) and migration rate (*m*) that are expected to produce *Fst* values of 0.005 (*mNe* = 50) and 0.0025 (*mNe* = 100). The values from the power analysis come from Henriques *et al.* (unpublished data). The parameter space below the lines can be ruled out as implausible with specified probabilities based on genetic data. The parameter space consistent with multiple stocks is further constrained if one assumes that separate stocks must exchange migrants at a rate below a certain threshold (in this case $m = 0.1 = 10\%$ per generation). Finally, an estimate of N_e from MARAM/IWS/2019/Hake/P2 (not available for review by the 2014 Panel) further constrains the parameter space consistent with multiple stocks. The solid red line is the point estimate $(1-2 \times 10^4)$ and the dotted red line is the lower bound of the 95% confidence interval $(2-3 \times 10^3)$; the upper bound is ∞. Some caveats about the above relationships between *m*, *Ne*, and power need to be appreciated: i) they are based on a widely-used but simplistic relationship between *mNe* and *Fst* [E(*Fst*) $\approx 1/(1+4mNe)$] developed by Wright (1931); ii) the relationships shown above are probably qualitatively robust, but caution should be taken in quantitative applications; iii) Wright's relationship assumes that an equilibrium has been reached between the homogenizing effects of migration (*m*) and divergence due to genetic drift (indexed by *Ne*), which might not be the case; and iv) the underlying model assumes discrete generations - translating between annual and generational migration rates needs more research.
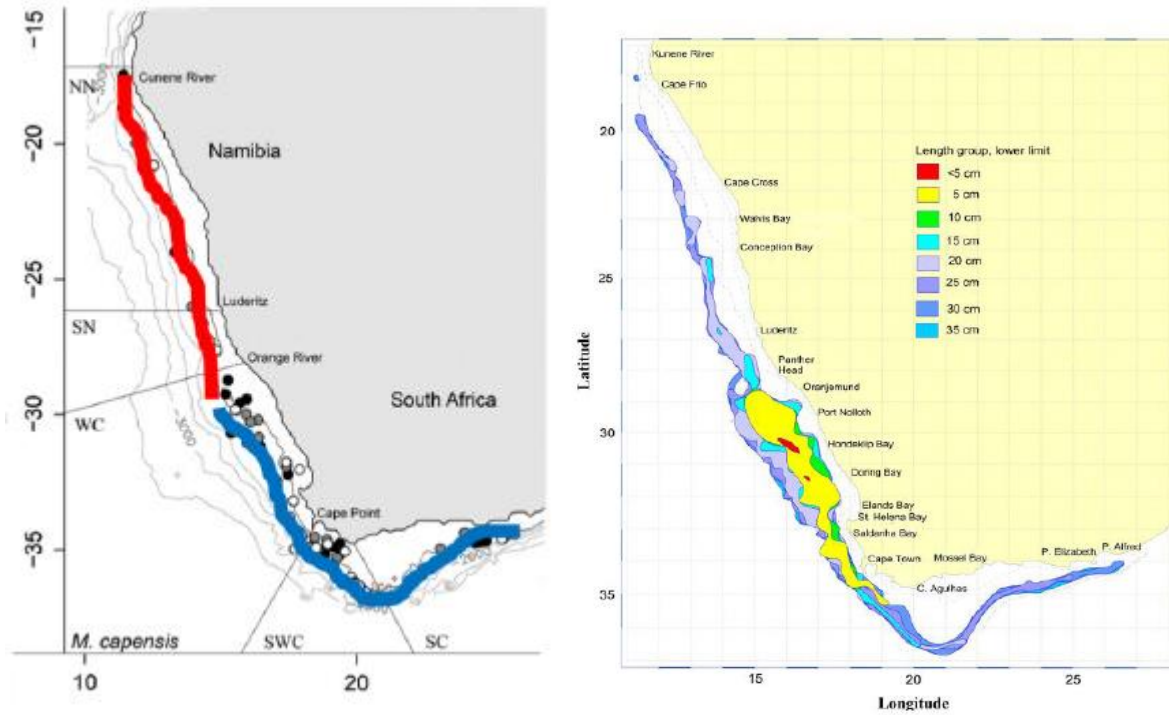
Figure 2. Map of *M. capensis* stock structure (left) from microsatellite data in MARAM/IWS/2019/Hake/P2 and locations of small *M. paradoxus,* which are thought to reflect the major area of successful spawning from Strømme et al (2015). In the left panel, red denotes the northern population, and blue denotes the southern population.
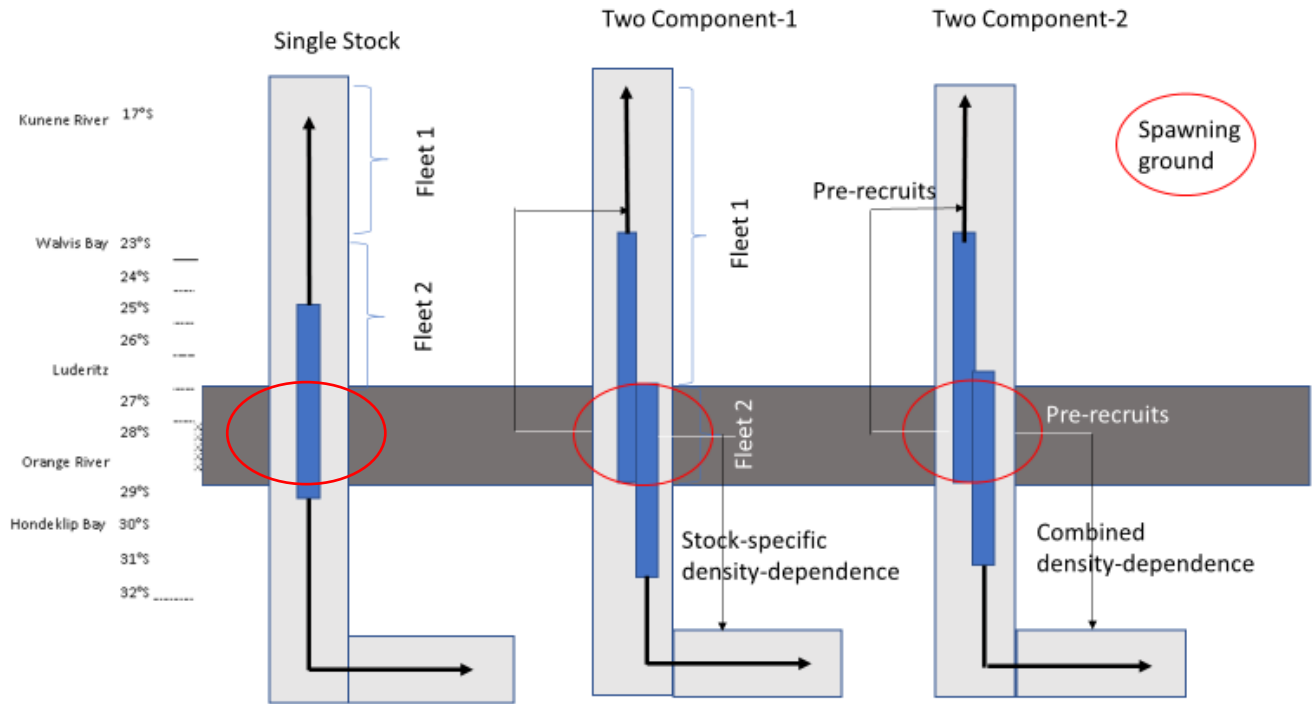
Figure 3. Schematic illustration of the revised *M. paradoxus* stock structure hypotheses. The spawning area is the horizontal shaded area. Blue bars indicate the nursery area and arrows indicate directions of movement. The difference between the Two Component-1 and Two Component-2 stock structure hypotheses is how density dependence is represented in the stock-recruitment relationship, with density dependence applied at the population level for hypothesis Two Component-1 and at the combined population level in hypothesis Two Component-2.

Appendix 1. Rock lobster stock and recruitment

It is important to understand how the process of stock recruitment is modelled in the historical period and the forecasts. It is clear that mature lobster do not move across areas, thus spawning can be characterized at the regional level.

$R_{y,a}$ is the recruitment in year y and area a

$S_{y,a}$ is the spawning biomass in year y and area a which generated the $R_{y,a}$.

Assuming an appropriate lag between spawning and recruitment even if that is not indicated in the subscripts, there are several alternative models for stock-recruitment relationships that could be considered:

A) Independent population models for each area (assuming there is no larval connectivity between areas):

$$R_{y,a} = \frac{S_{y,a}}{\alpha_a + \beta_a S_{y,a}}$$

B) A single population model for all areas combined (full mixing of larvae):

$$R_y = \frac{\sum_a S_y}{\alpha + \beta \sum_a S_y}$$

C1) Population models for each area linked through a common stock-recruitment relationship. The recruitment process is such that density dependence operates at global level although a constant proportion of the available recruits arrive at each area:

$$R_{y,a} = p_a \frac{\sum_a S_{y,a}}{\alpha + \beta \sum_a S_{y,a}}$$

where p_a is the proportion of recruits that recruit in each area

C2) Population models for each area linked through stock-recruitment relationships. The recruitment processes are such that density dependence operates at area level, but a constant proportion of the available recruits arrive at each area:

$$R_{y,a} = p_a \sum_a \frac{S_{y,a}}{\alpha_a + \beta_a S_{y,a}}$$

Appendix 2. hakes
2A. Evidence of spawning and hake migration
There is evidence of spawning of *M. paradoxus* in the southern coast and western coast of South Africa and the southern coast of Namibia, south of Luderitz. However, small *M. paradoxus* are detected in large numbers only off the west of South Africa, but not off the south coast of South Africa nor on the southern coast of Namibia. This is consistent with the hypotheses of Jensen that the majority of the successful recruitment of small *M. paradoxus* is to the west coast of South Africa. These recruits would then spread to occupy the entire distribution of the stock, some going north, some south and east, and some staying off the west coast. A portion of those northern and southern migrants would return to the main spawning grounds as they become larger and older. This hypothesis is consistent with the genetic results suggesting a single population of *M. paradoxus*.

An alternative hypothesis is that there is that although there is initial dispersal of small fish towards the north and south of the main spawning area, there is no natal homing of large fish back to the main spawning grounds. Additionally, spawning products always drift north so that spawners in southern Namibian waters would never contribute to recruitment off South Africa. The lack of large and old fish off Namibia and southern South Africa could be explained by greater fishing mortality of adult fish in those areas. Such a possibility would imply that areas north of Luderitz currently are acting as a population sink for part of the South African stock, but that the stock in Namibian waters would not affect productivity of the South African stock.

2B. Consideration of new SNP data for hakes
The Panel welcomed the genomic approach to improve hake stock structure resolution. While the data on Single Nucleotide Polymorphisms (SNPs) presented in MARAM/IWS/2019/Hake/P3 demonstrate the general feasibility of a ddRAD based genomic survey for hake, yielding over 50,000 putatively informative SNPs, the Panel concluded that the resulting data are too premature to be informative in current stock structure assessments. This conclusion was based on two major lines of evidence: 1) methods for generating the data; and 2) plausibility of the resulting data.

Regarding methodology, the primary issue is that DNA from multiple individuals was pooled before sequencing. Sequence coverage (about 54×10^6 reads for over 52,000 SNPs) translates into an average of about 1,000 reads/SNP in total for several hundred typed specimens, which produces an average of only five or fewer sequences per specimen and SNP. Given the inherent stochasticity in several analytical steps of the ddRAD method, and the low coverage of the analysis presented in MARAM/IWS/2019/Hake/P3, it is likely that most SNPs were assessed in only a (potentially small) fraction of the samples. Unfortunately, as samples were pooled prior to analysis, the actual number of specimens contributing to estimated allele frequencies for any locus is not known. Furthermore, if DNA quality was not equal across the pooled samples, uneven sample contribution to the SNP data may have been exacerbated. Taken together, these effects could have reduced the number of individuals that effectively contributed to the resulting data, such that the effective sample size was smaller than the number of individuals used. This in turn would lead to underestimation of the contribution of sampling error and an upward bias in the estimate of genetic differences between populations.

Regarding plausibility, two general themes are often found when moving from analysis of a small number of microsatellites to large numbers of SNPs.
- For a given locus, the maximum possible value of F_{ST} is constrained by genetic diversity within populations, as measured by heterozygosity (H): $F_{ST} < 1-H$ (Hedrick, 2005). For highly variable loci such microsatellites, the expected value of H is often 0.8-0.9 in marine fish, whereas 0.5 is the maximum expected H for diallelic loci like SNPs. Therefore, when the same samples are analyzed for both marker types, one generally expects to find a higher mean F_{ST} for SNPs than for microsatellites.
- When large numbers of SNPs are assayed, it is often possible to find some that are 'outliers' in the sense that they have much higher F_{ST} than the background level

associated with most other loci. A common interpretation is that these outlier loci are influenced by different selection pressures in different populations, acting either on the typed loci themselves or on other linked loci.

Results reported in MARAM/IWS/2019/Hake/P3 for *M. capensis* do not conform to either of these expectations. First, mean F_{ST} values between the two putative *M. capensis* stocks did not increase with the new SNP data; instead, they declined about 3-4 fold, from ~0.15 to ~0.04. Secondly, the clear signal of high differentiation across areas (Namibia versus South Africa) and weak differentiation within areas disappeared; instead, the new SNP data showed approximately the same level of differentiation for all pairwise comparisons. In addition, the Manhattan plots provided by the authors of MARAM/IWS/2019/Hake/P3 showed that high-F_{ST} loci are distributed widely across the genome, rather than being concentrated in one or a few locations. New SNP results for *M. paradoxus* are qualitatively the same as for *M. capensis*: mean F_{ST} values are similar for all pairwise comparisons, and high-F_{ST} SNPs are widely distributed across the genome.

The authors of MARAM/IWS/2019/Hake/P3 suggest that the pattern of relatively high levels of genetic differentiation in *M. paradoxus,* without any geographic signal, could be explained by stochastic survivability and reproductive success among cohorts – a process that has been referred to as "chaotic genetic patchiness" (e.g., Johnson and Black, 1982; Broquet et al., 2013). The Panel does not doubt that such processes could occur in one or both hake species, but it does not consider this a plausible explanation for the high F_{ST} values in *M. paradoxus* SNP data (F_{ST} ~ 0.05) compared to microsatellites (F_{ST} close to 0). A good approximation to the magnitude of the increase in F_{ST} due to chaotic genetic patchiness is $1/(2N_b)$, where N_b is the effective number of parents producing each sample (Waples, 1998). To produce $F_{ST} = 0.05$ by genetic patchiness thus would require that $N_b = 10$ – that is, that each of the samples analyzed represents the offspring of just 10 parents. Given that each of the samples included individuals from several different age classes, and that the estimated number of *M. paradoxus* adults is $> 10^8$, the Panel did not consider this explanation credible. Instead, it seems more likely that the new SNP data for both species are influenced by artefacts not related to biological processes in hake.

References

Broquet, T., Viard, F. and J.M. Yearsley. 2013. Genetic drift and collective dispersal can result in chaotic genetic patchiness. *Evolution* 67: 1660-1675.

Hedrick, P.W. 2005. A standardized genetic differentiation measure. *Evolution*, 59:1633-1638.

Johnson, M. S., and R. Black, R. (1982). Chaotic genetic patchiness in an inter-tidal limpet, Siphonaria sp. *Marine Biology*, *70*(2), 157–164.

Waples, R.S. 1998. Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity* 89:438-450.

Appendix 3. Documents and presentations considered by the panel are listed in document MARAM/IWS/2019/General 3.

This list is not duplicated here because it is available to be downloaded from http://www.maram.uct.ac.za/maram/workshops/2019.