SOME NOTES ON THE STRUCTURE OF AN OPERATIONAL MODEL IN THE IOTC CONTEXT

Alejandro Anganuzzi

IOTC Secretariat

ABSTRACT

Possible applications of operating models in the context of the IOTC scientific process are discussed. The two main uses in other fora have been the evaluation of management procedures or as a benchmark to assess the performance of various methods of analysis, primarily stock assessment procedures.

Considering the current stage of development of management advice in IOTC, it is suggested that the testing of assessment technologies will take priority in the near future. Considering this perspective, various desirable features of the operating model are discussed.

PRELUDE

The Scientific Committee in its last meeting in December 2001 recommended that the Secretariat prepare a report on the possibility of developing an operating model for IOTC and to explore the possibility of taking advantage of the experience of other fisheries bodies. These notes are intended as a starting point to discuss some of the features of operating models and their use in the context of IOTC.

BACKGROUND: THE ORIGINAL CONCEPT OF THE OPERATING MODEL

As far as we are aware, the term 'operating model' was originally mentioned in Linhart and Zucchini (1986) as a way to assess approximation and estimation errors involved in a model formulation and as a general approach to select between competing models [1]. The concept and the term were applied during the evaluation of management procedures in the International Whaling Commission, ICES, IATTC, ICCAT, and evaluation of assessment methods in ICES and NRC, and other fora. This year, CCSBT started on the development of an operating model to test management procedures.

The basic idea behind the application of an operating model is simple: simulate a virtual fishery which will generate data similar to those available to scientists and managers. Then use those artificially generated data to assess the efficacy of your procedures in providing a proper picture of the status of the resource and the adequacy of the decision rules that translate into management decisions. Therefore, although the term operating model refers only to this 'virtual reality', the 'operating model approach' (OMA) should also include a protocol for testing and measuring performance.

There is a useful distinction to make between applications of OMA directed towards testing:

1. Methods of data analyses, in particular the assessment procedures (including, for example, data analyses leading to indices of abundance or catch-at-age matrices or stock status indicators), and

2. Management procedures (including decision rules and management actions arising from these decision rules).

The two uses are obviously related but there are some differences that are relevant when we think about the most likely application in the case of IOTC. Given that the Commission has not yet even considered the possibility of adopting formal management procedures, the use listed under point 2.- does not seem like a pressing issue. On the other hand, during the meetings of the species Working Groups there is often insufficient time to fully understand the properties of (new or conventional) methods to the specific patterns of data of the resources involved. The usefulness of even limited simulations in isolating important questions was evident during the meeting of the WP on Methods in 2001 and led to the recommendation of further exploring the possibility of having a standard OM to deal with the testing according to point 1 above.

Note that, in any case, these two uses are not exclusive but complementary. The development of an OM oriented to the testing of analytical methods can be easily extended to

[1] Linhart and Zucchini define the operating model as the nearest representation of the 'true situation' which is possible to construct by means of a probability model, as opposed to the approximating model which is the model 'used to fit the data' usually with a simpler structure that the operating model.

evaluation of management strategies when the latter is required in the future. Also, a comprehensive evaluation of an analytical method should, in principle, include consideration of its final influence on the current management advice. For example, if the current advice is purely based on the perceptions about the trend of some index of abundance, the management advice might be robust to a lack of precision in the index or even a mild violation of the linearity between catch and effort. It has been often the case that the simpler methods of analysis could perform on a par with complex assessment methods under certain management scenarios, particularly in data-poor situations.

For example, even if an assessment method performs poorly most of the time, its results could be robust to grave mistakes in the data so that the overall performance of the assessment method-management strategy combination is acceptable. However, as we will discuss later, the difference in emphasis on which aspect is the main objective for the testing has some consequences on the structure of the implementation.

ELEMENTS OF AN OPERATING MODEL

The fundamental objective of the application of an operating model is to test the robustness of either a method for analysis or a management procedure to departures from the basic assumptions incorporated in the models or to major uncertainties in the dynamics of the resource and the data generating procedures.

In many cases, the implementation is cast to let different procedures compete for the assessment (see for example IWC or NRC) but the fundamental goal is to better understand the behaviour of the model in the face of uncertainties. Irrespective of the primary objectives, the use of an operating model entitles the same basic steps:

Repeat a large number of times (replicas):

For each year:

1. Simulate the dynamics of the population.
2. Simulate the catch process.
3. Simulate the generation of data for the purpose of the assessment.
4. Conduct an assessment using the data generated in step 3.

If the objective is to assess a management procedure:

5. Apply a decision rule on the basis of the assessment conducted in step 4.
6. Apply a management action as specified by the decision rule in step 5.

Repeat step 1 through 6 for a number of years.

In each replica, compile the information necessary to compute an indicator of the performance of the method or management strategy for that particular scenario.

It is essential that, in each of these components, uncertainties related to the intrinsic stochasticity of the dynamics of the fishery (the term **'process error'** is often associated with these components) are simulated. There will be also uncertainties related to the inability to obtain precise measures of the states of the system (or **'measurement error'**). Examples of the latter include errors in the catches, errors in estimating the age composition of the catch, sampling error associated with a particular index of abundance, etc.). In some cases, also considered is the **implementation error** that arises from an imperfect application of a management action.

The steps described above represents a particular realization of the history of the fishery, a time series whose properties depend on the specific values taken by the stochastic components of the whole process. If we repeat those steps, the values obtained will be different due to both process and measurement errors. Therefore, in the application of the operating model approach, the process described above is repeated a large number of times to obtain different replicas. If the level of process and observation errors that are represented in the operating model are adequate compared to the real situation, then the average (or any other statistic) performance over the replicas of an assessment technique or management procedure would be a fair representation of what we might expect the future.

As mentioned above, in order to assess performance of assessment methods or management strategies we need some measure (metric) of how well the results match the (known) reality of the OM or how far we are from the management goals. In the case of management strategy evaluation, these **performance indicators** are often simply a measure of how often in the replicas for a particular scenario, an undesirable event happens (e.g. in how many replicas the spawning stock falls below 20% of the virgin stock). In testing methods of analyses, the performance indicators usually measure how far the results of the method are from the 'virtual reality'.

Random variability (be process or measurement error) is not the only source of problems to be evaluated. Often, the robustness of the procedure to specific violations of assumptions implicit in the assessment technique needs to be tested. For example, what are the consequences of assuming linearity in the relation between an index of abundance and biomass of a portion of the population. An important question always in this example is how robust our conclusions and decisions are if the true relationship is non-linear.

These are often called **robustness trials**, or in some implementations they are considered as different operating models, although often the distinction is merely semantic. The important principle is that these trials should be conducted and then contrasted with 'base case' scenarios to assess the performance of the procedures under plausible scenarios.

Robustness trials take center stage when the objective of the OM approach is to learn about the properties of the methods in place. In this case, it is often more informative to look for those conditions that the cause the procedure to 'break down' and start giving answers that are unacceptably biased or uncertain. The results of the robustness trials could then be used to give a list of the conditions under which we can expect and acceptable performance (or in the words of some colleagues 'to map the area of applicability').

SOME PRACTICAL CONSIDERATIONS

Level of detail of the operating model. In practice, often it is not necessary to have a very detailed internal representation of the fishery in the OM if we are able anticipate how these details would translate into the parameters of aggregated processes. For example, for a given population, a plausible hypothesis could be that the spawning stock is divided into two components with little mixing and different productivities. One way of representing this in an operating model is having effectively two 'sub-populations', each with its own stock-recruitment parameters. However, the effect of overfishing one of these two stocks could be equally well explored if we assume a stock-recruitment relationship whose parameters are dependent on the stock size itself. Such a model could be used to represent processes such as depensation at low-biomass levels.

However, for the sake of communication, especially with non-technical audience, it is sometimes better to explain certain features in term of familiar quantities. For example, an explicit representation of the fleet dynamics could be used to better understand and communicate the effects of observable features such as aggregation of the fleet or exchange of information among fishermen. However, this presupposes often a long process of tuning a large of parameters so that the model exhibits all the observable properties of the real fishery. Often, interactions between parameters determine that the model would behave in an unexpected way under certain circumstances.

Therefore, *there is a balance between the level of detail in the representation of the processes in the operating model and the difficulties that the modeler will face when tuning the model to obtain a reasonable behavior*. A classic example is when trying to model fish movement as a response of certain environmental parameters. Most often, it is not necessary to discover why fish move to and stay in a certain area. While this question has a scientific interest of its own, it is often not fundamental in the context of an operating model. From the management point of view, the relevant question is not why the fish move to a particular area, but how many fish move in and out of an the area and how long do they stay in a particular area.

Conditioning of the operating model. It is obvious that a desirable feature of the operating model is that, at least, it reflects adequately the range of uncertainties in the population dynamics. Only in that case the robustness of the assessment techniques and management procedures can be adequately tested. When the representation of the operating model is aggregated at the same level of the data available, then a process known as **conditioning or tuning** could be used to estimate parameter values.

Some implementations of operating models describe the underlying processes in such a disaggregated way (e.g. with detailed spatial structure) that many parameter combinations can produce a behavior observed in reality. The problem here lies in finding which of those plausible sets of parameter values is more adequate.

If the operating model has been formulated at a level of detail similar to that of the assessment, the model can be 'tuned' or 'conditioned' to the historical data in a process often indistinguishable from the actual conduct of the assessment. After this conditioning is complete, the model (with the parameters estimated) can be used to project the future dynamics, although complications arise when the future scenarios are very different from the historical situation.

In any case, *for the purposes of testing and understanding the properties of methods it is not crucial that the OM behaves exactly as the real fishery,* only that certain patterns of data can be generated. These patterns can be as extreme as necessary to understand where the method under testing ceases to work acceptably.

WHERE TO GO FROM HERE

The WPM-2001 produced a list of desirable features for an operating model:

- The model should be age, size, time and space-structured.
- It should allow for multiple fleets with different and variable selectivity.
- It should allow for testing trends and variability in catchability, including effects due to targeting.
- It should allow for the distribution of fishing effort and catchability to change spatially and temporally.
- The spatial structure should be sufficient to allow testing for effects due to concentration of effort and density-dependent population responses in habitat use, but it does not necessarily need to be a highly realistic representation of the Indian Ocean.
- The spatial component should allow for movement between areas that contains both a random and directed components. This may need to vary with age and season to account for spawning behaviour, etc.
- The model should allow for flexibility in modelling of growth. It should allow for both density-dependent and size-based models to be incorporated. However, this is not seen as an initial high priority and the initial modelling of growth would be done using a standard VBG formulation.
- Recruitment should be flexibly modelled to include a range of possible stock recruitment relationships with a random component. It should also allow for a range of different hypotheses for how recruitment is spatially distributed, including the possibility of localized stock/recruitment relationships.
- The observational data produced by the model needs to include realistic levels of variability and potential biases.

The WPM also noted that the development of the operating model should be seen as an evolving process, with initial priority being given to those features that have been identified as most problematic in Indian Ocean Tuna assessments (e.g. multiple fleets with different selectivities and different temporal effort trends; non-random and localized concentration of fishing effort).

The WPM also recommended that the computer implementation of the model be designed in an object-oriented framework to allow increasing complexity to be added progressively. An additional advantage of such an implementation is that different components of the OM can be developed independently by interested scientists (even in different computer languages) provided that certain conventions are followed. This offers a possible way to proceed in the development of the OM as, currently; the staff levels and commitments of the Secretariat do not allow it to develop the full model. A small task group could coordinate with the Secretariat how to partition the tasks and monitor the progress in the development.