

Proposed procedure of selecting agreeable Management Procedure and
results of feasibility experiment.

Sachiko Tsuji

National Research Institute of Far Seas Fisheries
Fisheries Research Agency

Summary: Feasibility experiment about the procedure to assign overall rating among different MPs by examining stock and quota trajectories of individual run was conducted with four participants. The procedure seems to provide reasonable rating scores reflecting integrated judgment quite easily.

要約: 個々の資源、漁獲枠の変化を調べることで、異なる MP に対して全体的な評価を行う手法の実現可能性を 4 人の被験者の協力で実施した。この手法で、容易に全体的な判断に基づく妥当な数値評価ができるように思われる。

Background:

The Third Meeting of Management Procedures Workshop selected four candidates of Management Procedures (MP). Those four were selected to represent a range of characteristics of MPs developed by Members. Next step is to choose the final one and this process requires vigorous participation from both industries and managers. Though scientists accustomed to view and compare various aspects using statistics, those statistics is not easy and some misleading for non-technical people. Especially, by showing median and 10 percentiles of stock and quota trends that seem quite smooth, people tend to think all expected trajectories to be much smoother than reality. It is extremely important for managers and industries to realize that what expected to happen is only one of those realizations, and that they must select such MP to be acceptable in any of those circumstances. For this purpose, we found it much easier to work with multiple examples than explaining using summary statistics. This documents proposes to examine a certain number of randomly selected runs for final selection of MPs and reports on results and observations obtained from quick feasibility experiment.

Feasibility Experiment:

Feasibility experiment was conducted with four participants using the file showing 2000 runs of quota and stock trajectories by four selected MP under 1.1 tuning provided Advisory Panel at the end of MPWS3. The participants were not scientists and varied in understandings on the process of developing Management Procedures. The participants were asked to examine every 50 runs and record run ID and preferred order or preferred levels of 4 MPs for each run. Whether scoring by rating preferable orders or preferred level as ranks was left to examinee's choice. Three chose scoring with orders and one chose scoring with ranks. All participants worked independently. Because of ambiguity of protocol, participants chose different set of runs. Author also did both ordering and level rankings independently for 160 runs selected by either

one of examinee, independently.

Examination of 40 runs took about one hour if rating principles were already settled, and required some level of patience. This number is not adequate to regenerate an intended occurrence of scenarios (Table 1). Despite of that, accumulated scores by one person for four sets of runs shows reasonable consistency (Table 2). Difference in scoring whether using orders or ranks does not seem to lead different results.

Table 3 shows results of four participants. In spite of lack of explanations, all participants do not seem to have a difficulty to complete assigned task. Careful examination of details results suggests all following their individual judging criteria consistently for 40 runs selected. It seems rating is based on overall consideration of various issues. Priority and emphasis of judging criteria seem to differ each other. Still overall ratings shows similar pattern.

Other observations and comments obtained are as follows:

- Showing both historical and future stock trends in one plate is preferable. Stock level to be maintained cannot be judged without historical trajectories of stock in conjunction with future trends.
- MSY level should also be added from the same reason. The need to show historical catch trends is less critical since a current catch level is known.
- Accumulated results are different from the first impression when seeing a summarized figure and few cases without serious thinking.
- Even with different judging criteria, the ratings are frequently obtained.
- There are number of runs where none of MPs working satisfactory. Some runs are easy to judge.
- Forty plates might be too many to examine. Patterns of MP responses and stock trajectories can be classified into some smaller number. Examination of those typical patterns may be adequate if expected probabilities by patterns presented together.

Table 1. Occurrences of six scenarios for sets of runs that four participants (A to D) worked

	S1	S2	S3	S4	S5	S6
A:	0.2	0.3	0.125	0.025	0.25	0.1
B:	0.15	0.425	0.05	0.025	0.275	0.075
C:	0.05	0.325	0.175	0.075	0.3	0.075
D:	0.15	0.3	0.15	0.075	0.2	0.125
2000 runs	0.12	0.36	0.12	0.08	0.24	0.08

Table 2. Comparison of relative rating points by one person using different sets of runs. A to D correspond to sets of data selected by participants A-D. Scores are standardized with the highest score within each runs set.

	MP1	MP2	MP3	MP4
When rating with order:				
A:	1.00	0.80	0.82	0.20
B:	1.00	0.81	0.77	0.15
C:	1.00	0.70	0.80	0.25
D:	1.00	0.69	0.58	0.16
When rating with ranks:				
A:	1.00	0.72	0.72	0.56
B:	1.00	0.81	0.75	0.57
C:	1.00	0.87	0.82	0.60
D:	1.00	0.82	0.73	0.55

Table 3. Comparison of rating results by different participants. Scores are standardized with the highest rated score by participants.

	MP1	MP2	MP3	MP4
A (order):	0.95	0.51	1.00	0.43
B (order):	1.00	0.89	0.92	0.39
C (order):	0.61	0.20	1.00	0.45
D (rank):	1.00	0.67	0.68	0.32