 Collect. Vol. Sci. Pap. ICCAT, 77(2): 434-440 (2020)

WHAT DO CURRENT RESULTS USING THE PACKAGE INDICATE REGARDING WHICH UNCERTAINTY AXES "MATTER" REGARDING CMP PERFORMANCE, AND WHAT ARE THE NEXT STEPS NEEDED IN THE ABFT MSE PROCESS

D. S. Butterworth and R. A. Rademeyer[1]

SUMMARY

An approach is put forward, using the results from the CMPs applied to the 96 Operating Models (OMs) of the interim grid in Butterworth and Rademeyer (SCRS/2020/075), to assess the extent to which CMP performance is impacted ("matters") for the various uncertainty axes currently included in this grid. For each uncertainty axis, the range of median Br30 values for the factors along that axis across the OMs for a full cross of the factors across all the other uncertainty axes is considered. Results indicate that some uncertainty axes do "matter" much less than others, rendering associated inferences possible before any decision on desired CMP tuning with respect to final abundance targets need be reached. However, for reliable results, developers first need to refine their CMPs further to improve the robustness of their performances so as to be closer to that which might be possible for their final forms. A sequential set of steps for taking the MSE process further is suggested, which distinguishes the roles of assessing whether OMs "matter" and according them plausibility weightings of some form.

RÉSUMÉ

Une approche est proposée, utilisant les résultats des CMP appliqués aux 96 modèles opérationnels (OM) de la grille provisoire de Butterworth et Rademeyer (SCRS/2020/075), pour évaluer dans quelle mesure les performances des CMP sont affectées (« comptent ») pour les différents axes d'incertitude actuellement inclus dans cette grille. Pour chaque axe d'incertitude, la gamme des valeurs médianes de Br30 pour les facteurs le long de cet axe parmi les OM pour un croisement complet des facteurs à travers tous les autres axes d'incertitude est examinée. Les résultats indiquent que certains axes d'incertitude « comptent » beaucoup moins que d'autres, ce qui rend possible les inférences associées avant qu'une décision sur le calibrage souhaité des CMP en ce qui concerne les objectifs d'abondance finale ne soit prise. Toutefois, pour obtenir des résultats fiables, les développeurs doivent d'abord affiner davantage leurs CMP pour améliorer la robustesse de leurs performances afin de se rapprocher de ce qui pourrait être possible de leurs formes finales. Un ensemble séquentiel d'étapes pour faire progresser le processus de MSE est suggéré, incluant les rôles d'évaluation des OM « qui comptent » et leur attribution de pondérations de plausibilité sous une forme ou une autre.

RESUMEN

Se presenta un enfoque, usando los resultados de los CMP aplicados a los 96 modelos operativos (OM) de la matriz provisional de Butterworth y Rademeyer (SCRS/2020/075) para evaluar la medida en que los diversos ejes de incertidumbre incluidos en esta matriz impactan («importan») en el desempeño de los CMP. Para cada eje de incertidumbre, se considera el rango de los valores de la mediana de Br30 para los factores a lo largo de dicho eje entre los OM para un cruce completo de los factores entre los demás ejes de incertidumbre. Los resultados indican que algunos ejes de incertidumbre «importan» mucho menos que otros, lo que hace posible las inferencias asociadas antes de que deba tomarse cualquier decisión sobre la calibración deseada de los CMP respecto a los objetivos de abundancia final. Sin embargo, para lograr resultados fiables, los desarrolladores primero deben mejorar más sus CMP para aumentar la robustez de sus desempeños con el fin de que estén lo más cerca posible de lo que podrían ser sus formas

[1] Marine Resource Assessment and Management Group (MARAM), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch 7701, South Africa

finales. Se sugiere un conjunto secuencial de pasos para avanzar en el proceso de MSE, que distingue los papeles para evaluar si los OM «importan» y atribuyéndoles ponderaciones de plausibilidad de alguna forma.

KEYWORDS

Management Strategy Evaluation, Candidate Management Procedure, Operating Model grid, Atlantic bluefin tuna, performance, plausibility

Assessing which Uncertainty Axes "matter"

Approach

This document uses the results for three of the CMPs considered in Butterworth and Rademeyer (2020) to suggest how to assess the extent to which CMP performance is impacted ("matters") for the various uncertainty axes currently included in the interim grid. These three are the C=Cur (continue current catches) and the 05-05 and 075-075 feedback-control CMPs, with both deterministic and stochastic outputs provided for the last two. Results are presented for the Cav30 and Br30 performance statistics in the same format as in Table 2 of Butterworth and Rademeyer (2020), which reported summaries of the distributions of outputs for these performance statistics across a set of OMs (there the full interim grid of 96 OMs).

The focus here is on the Br30 statistic, which provides spawning stock biomass relative to dynamic Bmsy after 30 years, and for which the Commission's primary objective is attainment of Bmsy (i.e. Br30=1). Results are reported for each of the OM factors for each uncertainty axis of the initial grid, i.e. for each such factor for one uncertainty axis, the summary statistics reported here apply to the distribution across a full cross of the factors for the other uncertainty axes. The maximum difference in the medians for Br30 across the factors for each uncertainty axis is then taken to be a measure of the extent to which that axis influences CMP performance ("matters").

Results and Discussion

Table 1 lists the results described above for the C=cur and the two feedback-control (05-05 and 075-075) CMPs, with stochastic as well as deterministic results included for the last two. Table 2 provides a summary of the results for the maximum difference in the medians for Br30 across the factors for each uncertainty axis in a manner that contrasts behaviour for the C=cur with that the feedback-control CMPs for both the eastern and western origin stocks. Figure 1 shows these median Br30 maximum differences in histogram form for each uncertainty axis for each CMP application, and for the eastern and western origin stocks.

Two important features of these results are immediately evident from **Figure 1**:

i) The biggest of the maximum differences are reduced by the feedback-control CMPs compared to the "poor" C=cur CMP, i.e. feedback-control assists in providing more robust performance in terms of stock status targets/recovery.

ii) Some uncertainty axes "matter" less (indeed substantially less) than others.

Table 2 and **Figure 1** indicate that in terms of inferences about "mattering", it is of little importance whether deterministic or stochastic results are considered, or what tuning level is adopted for the CMP (i.e. whether a more or less intensive level of harvesting, corresponding respectively to lower or higher final abundances). However, the "quality" of the CMP can impact results, because results from a "poor" CMP can hide the fact that an uncertainty axis matters through subsuming its impact in results for another axis for which the CMP provides inadequately robust performance. Note here how the "poor" C=cur CMP hides the fact that the L-H uncertainty axis does indeed "matter" for the western stock.

In moving forward, these results suggest that two important inferences can be drawn:

- Decisions about which uncertainty axes "matter" in terms of CMP performance can be made before any decision on desired CMP tuning with respect to final abundance targets need be reached.

- The approach above can provide a basis for assessing which uncertainty axes "matter", and hence also a basis for perhaps deleting some of these in the current interim grid and also considering replacing them with others. However, before this can be done reliably, the CMPs must be refined further by their developers to improve the robustness of their performances to closer to that which might be possible for their eventual final forms.

Next steps in the ABFT MSE process

Note: Though "decision" structures may differ, the sequence of steps set out below is effectively a summary of that which has been followed in the SCs (and their sub-committees) of other RFMOs which have adopted MPs, such as the IWC, CCSBT and NAFO (though some took some short cuts!).

1) Developers refine their CMPs by use of the present interim grid together with the more important robustness trials.

   Objective: To determine which of the interim grid axes "matter" most in terms of the relative extent to which they impact key performance statistics. Note that the inferences drawn above from the results discussed there indicate that:
   - Final inferences **do** require the application of realistic CMPs.
   - Final inferences **do not** require prior agreement on tuning levels/recovery targets.
   - At this stage of the process very few performance statistics need to be considered.

2) The BFT WG agrees (by virtual meeting – possibly the one replacing the physical meeting originally scheduled for July) to possible modification of the axes to be included in the interim grid, and designates which are the most important robustness trials to be considered.

3) Developers (through the MSE technical group) consider and propose a small set (probably three) of interim "recovery" targets to which to "development tune" their CMPs for the purpose of convenient and comparable presentation of initial results.
   - These would need to span the range likely to be of interest to the Commission.
   - They might (for ease of implementation) most readily be defined in terms of the median value in a stochastic implementation of one "centrally performing" OM in the interim grid of a biomass performance statistic such as Br30.

4) The BFT WG agrees (by virtual meeting) to this selection (though note that in other fora this has been delegated to the equivalent of the developers/MSE technical group, since the outcome is needed only to facilitate presentation of results, and does not constitute a final decision on tuning for implementation in any way).

5) Developers refine their CMPs further, now using the development tunings in 3), and taking account also of a fuller range of performance statistics.
   - Based on these results, developers (through the MSE technical group) also suggest a smaller set of performance statistics which capture the key differences in performance (independent of the main catch-recovery trade-off reflected by the different development tuning choices) amongst their CMPs. (Note that in practice, many performance statistics are so highly positively correlated that their further consideration is not helpful in distinguishing qualitative differences in CMP performance.)

6) Considering the results from developers from the process above, the BFT WG (possibly at its September meeting) proceeds to discuss the following topics and to agree with what combination of the possibilities for each it wishes to proceed in some iterative manner, which will in turn be linked to further runs of the CMPs (note that in this process, prior consultation of BFT WG members with their principals will be helpful to inform deliberations):

- Agreement to possibly modified values for interim development tuning targets.
- Agreement to a possibly modified "smaller set" of performance statistics (note that in many past actual cases, the ultimate selection of an MP has involved consideration of performance for at most only two performance statistics).
- Agreement on the final grid (or reference set) of OMs – note that the constituent uncertainty axes are unlikely to change at this time, but the values of the factors on these axes might change from the initial "extremes" to somewhat more central options – e.g. in some instances, a "central" OM (or very few OMs) from the grid have been chosen as a basis for the primary final comparison of the performances of different CMPs and presentation of their results. What decisions are made at this stage will depend also on decisions on the matters raised in the bullet below.
- Agreement on a system for assignment of plausibilities/plausibility weightings to different OMs, and on how to utilise these in developing recommendations to stakeholder groups (see section 7 of the report of the February 2020 meeting of the ICCAT Bluefin Tuna MSE Technical Group).
- Culling the CMPs surviving to this stage of the process to a very few (maybe two or at most three) for which to present results to stakeholders, and refine further on the basis of their feedback
- Agreement on the range of tunings for which to present such results (these could be identical to the development tunings – they are NOT final – advice on a final tuning range for which to present results would come from iterative interaction with stakeholders and ultimately be provided by the Commission).

7) At this stage of the process, the BFT WG presents results for a first set of options through the SCRS to the Commission for their response and specification of an iterative interactive process of interaction with stakeholders to lead to a final proposal of options for an MP to be made to the Commission.

Reference

Butterworth DS and Rademeyer RA. 2020. Can the wide range of resource behaviours evident across the ABFT MSE interim grid be "tamed" by the feedback control provided by a CMP? SCRS/2020/075.

Table 1: Summary statistics (median, standard deviation and minimum (for Br30)) of the distributions of key performance statistics AvC30 and Br30 across each of the OM factors for each uncertainty axis of the initial grid for the three CMPs considered, i.e. for each such factor for one uncertainty axis, the summary statistics apply to the distribution across a full cross of the factors for the other uncertainty axes. For C=cur CMP, only the deterministic results are shown, while for the "05-05" and "075-075" CMPs, both the deterministic and stochastic results are shown. Note that AvC30 refers to the catch from the East or West area, whereas Br30 refers to the eastern or western origin stock. The maximum difference in the medians for Br30 across the factors for each uncertainty axis is also shown as a measure of the extent to which that axis influences CMP performance ("matters").

a) C=cur (Deterministic)

C=cur (Det)	EAST AvC30 Median	EAST AvC30 SD	EAST Br30 Median	EAST Br30 SD	EAST Br30min	EAST Max diff. in Br30 Median	WEST AvC30 Median	WEST AvC30 SD	WEST Br30 Median	WEST Br30 SD	WEST Br30min	WEST Max diff. in Br30 Median
L	36.00	7.11	1.39	0.95	0.00	0.48	2.35	0.38	0.99	0.76	0.00	0.07
H	36.00	4.48	1.88	1.14	0.00		2.35	0.22	1.06	0.65	0.00	
--	36.00	7.90	1.49	1.04	0.00	0.58	2.35	0.45	0.58	0.65	0.00	0.73
-+	36.00	2.96	1.84	1.08	0.00		2.35	0.18	0.51	0.58	0.00	
+-	36.00	7.70	1.34	1.05	0.00		2.35	0.36	1.25	0.81	0.00	
++	36.00	1.19	1.91	1.12	0.00		2.35	0.00	1.16	0.66	0.06	
MixI	36.00	6.55	1.53	1.08	0.00	0.07	2.35	0.32	1.10	0.72	0.00	0.29
MixII	36.00	5.52	1.61	1.08	0.00		2.35	0.32	0.81	0.71	0.00	
A	36.00	5.36	1.53	1.05	0.00	0.15	2.35	0.28	1.04	0.70	0.00	0.01
B	36.00	6.58	1.68	1.10	0.00		2.35	0.34	1.05	0.72	0.00	
R1	36.00	0.00	2.50	0.35	1.64	2.50	2.35	0.00	1.68	0.31	1.04	1.68
R2	29.17	7.64	0.00	0.19	0.00		2.13	0.43	0.00	0.24	0.00	
R3	36.00	0.00	1.61	0.41	0.56		2.35	0.00	1.06	0.40	0.13	

bi) "05-05" (Deterministic)

05-05 (Det)	EAST AvC30 Median	EAST AvC30 SD	EAST Br30 Median	EAST Br30 SD	EAST Br30min	EAST Max diff. in Br30 Median	WEST AvC30 Median	WEST AvC30 SD	WEST Br30 Median	WEST Br30 SD	WEST Br30min	WEST Max diff. in Br30 Median
L	23.71	7.23	1.98	0.51	1.06	0.59	1.47	0.68	1.02	0.75	0.38	0.49
H	23.55	5.72	2.57	0.67	1.24		1.04	0.61	1.51	0.50	0.75	
--	25.00	7.07	1.98	0.62	1.06		2.14	0.71	1.03	0.40	0.43	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	

bii) "05-05" (Stochastic)

05-05 (Stoch)	EAST AvC30 Median	EAST AvC30 SD	EAST Br30 Median	EAST Br30 SD	EAST Br30min	EAST Max diff. in Br30 Median	WEST AvC30 Median	WEST AvC30 SD	WEST Br30 Median	WEST Br30 SD	WEST Br30min	WEST Max diff. in Br30 Median
L	23.71	7.23	1.98	0.51	1.06	0.59	1.47	0.68	1.02	0.75	0.38	0.49
H	23.55	5.72	2.57	0.67	1.24		1.04	0.61	1.51	0.50	0.75	
--	25.00	7.07	1.98	0.62	1.06		2.14	0.71	1.03	0.40	0.43	
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	

ci) "075-075" (Deterministic)

075-075 (Det)	EAST AvC30 Median	EAST AvC30 SD	EAST Br30 Median	EAST Br30 SD	EAST Br30min	EAST Max diff. in Br30 Median	WEST AvC30 Median	WEST AvC30 SD	WEST Br30 Median	WEST Br30 SD	WEST Br30min	WEST Max diff. in Br30 Median
L	32.34	10.25	1.50	0.50	0.73	0.56	1.92	0.77	0.74	0.73	0.11	0.52
H	32.48	8.34	2.06	0.64	0.92		1.34	0.75	1.26	0.52	0.36	
--	33.97	10.11	1.57	0.59	0.73	0.62	2.61	0.87	0.69	0.37	0.13	0.99
-+	31.42	7.99	2.00	0.58	1.08		2.61	0.85	0.60	0.33	0.11	
+-	34.80	11.25	1.50	0.59	0.81		1.30	0.29	1.59	0.56	0.65	
++	31.27	7.68	2.12	0.60	1.15		1.65	0.33	1.59	0.49	0.62	
MixI	32.76	9.62	1.73	0.64	0.73	0.01	1.84	0.86	1.14	0.65	0.14	0.08
MixII	32.31	9.27	1.74	0.63	0.88		1.83	0.72	1.06	0.64	0.11	
A	32.90	9.10	1.74	0.57	0.99	0.01	1.89	0.57	1.16	0.61	0.18	0.23
B	31.97	9.58	1.73	0.68	0.73		1.48	0.97	0.93	0.68	0.11	
R1	38.08	3.83	2.40	0.44	1.69	1.20	2.21	0.89	1.68	0.65	0.57	0.90
R2	18.33	1.14	1.21	0.27	0.73		1.33	0.32	0.78	0.38	0.39	
R3	32.48	2.73	1.83	0.47	0.95		2.21	0.75	1.11	0.65	0.11	

cii) "075-075" (Stochastic)

075-075 (Stoch)	EAST AvC30 Median	EAST AvC30 SD	EAST Br30 Median	EAST Br30 SD	EAST Br30min	EAST Max diff. in Br30 Median	WEST AvC30 Median	WEST AvC30 SD	WEST Br30 Median	WEST Br30 SD	WEST Br30min	WEST Max diff. in Br30 Median
L	33.13	12.01	1.30	0.57	0.28	0.57	1.91	0.85	0.59	0.71	0.03	0.52
H	33.60	10.03	1.87	0.72	0.61		1.41	0.84	1.11	0.52	0.21	
--	34.02	12.04	1.37	0.69	0.28	0.63	2.63	0.99	0.61	0.37	0.03	0.94
-+	32.30	9.39	1.81	0.63	0.78		2.66	0.97	0.47	0.32	0.07	
+-	36.09	13.35	1.28	0.69	0.42		1.33	0.25	1.38	0.64	0.33	
++	32.16	9.14	1.90	0.68	0.81		1.65	0.32	1.41	0.51	0.41	
MixI	33.50	11.42	1.59	0.72	0.28	0.05	1.61	0.95	1.00	0.66	0.09	0.07
MixII	33.48	10.91	1.53	0.70	0.50		1.78	0.76	0.93	0.64	0.03	
A	34.10	10.99	1.57	0.67	0.48	0.04	1.88	0.60	1.00	0.61	0.03	0.20
B	32.92	11.11	1.61	0.74	0.28		1.33	1.05	0.79	0.68	0.06	
R1	39.39	4.00	2.28	0.44	1.53	1.39	2.33	0.89	1.51	0.64	0.45	0.86
R2	15.97	1.66	0.89	0.31	0.28		1.28	0.21	0.64	0.40	0.13	
R3	33.49	3.24	1.60	0.47	0.75		2.30	0.79	0.84	0.60	0.03	

Table 2: Maximum difference between Br30 medians across each factor for each uncertainty axis for the eastern and western origin stocks for the three CMPs. For C=cur, only the deterministic results are shown, while for "05-05" and "075-075" the average and range (min, max) across both and for both the deterministic and stochastic results are shown.

		C=cur	05-05 and 075-075 deterministic and stochastic	
		Det.	av.	Range
EAST	L-H	0.48	0.58	(0.56; 0.59)
	Scale	0.58	0.61	(0.60; 0.63)
	Mixing	0.07	0.02	(0.01; 0.05)
	A-B	0.15	0.04	(0.01; 0.05)
	Rec	2.50	1.29	(1.20; 1.39)
WEST	L-H	0.07	0.50	(0.49; 0.52)
	Scale	0.73	0.97	(0.94; 0.99)
	Mixing	0.29	0.05	(0.02; 0.08)
	A-B	0.01	0.15	(0.08; 0.23)
	Rec	1.68	0.98	(0.86; 1.08)

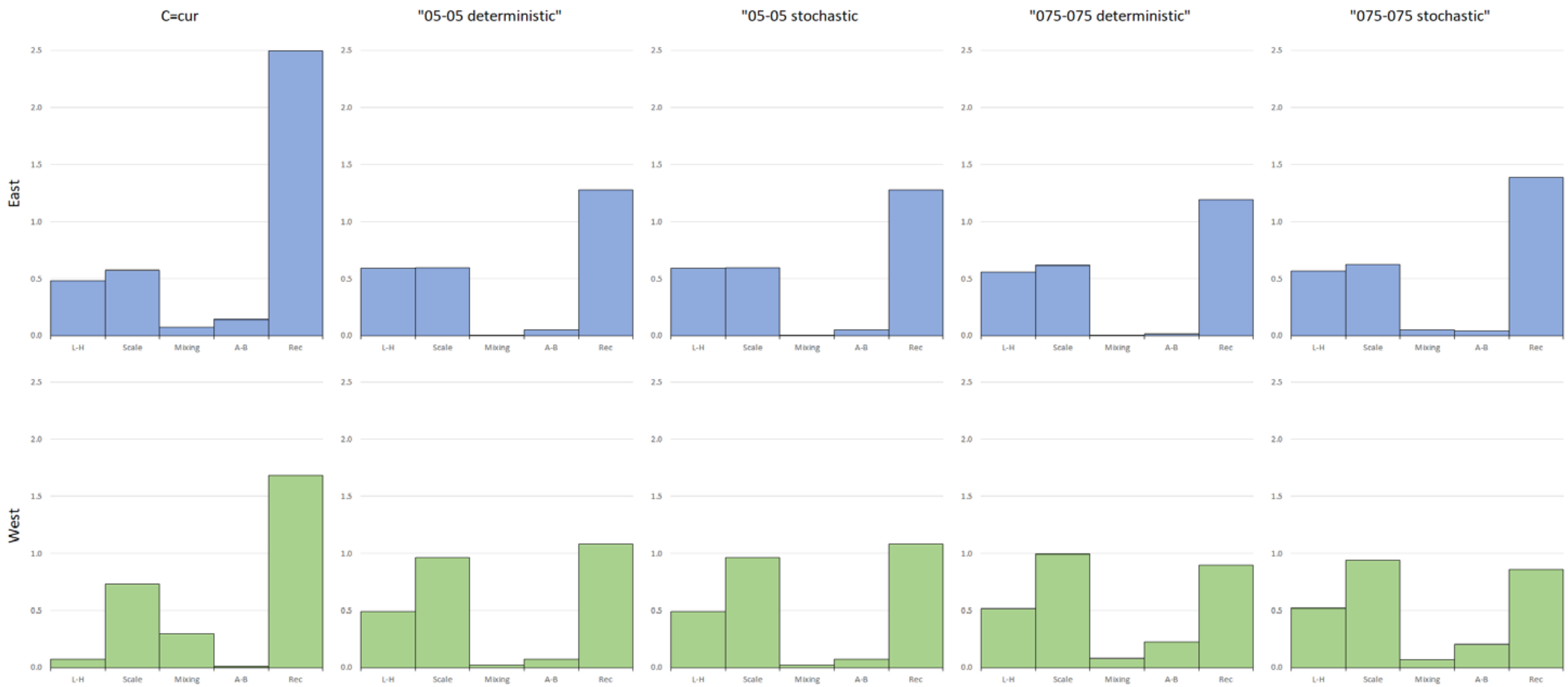

Figure 1: Bar plots of the maximum differences between Br30 medians across each factor for each uncertainty axis for the eastern and western origin stocks for the three CMPs. For C=cur, only deterministic results are shown, while for “05-05” and “075-075”, both the deterministic and stochastic results are plotted.