MSE Convergence dianostics WKMSE2

S. Cerviño

February 4, 2019

Contents


```
source('convTools.R')
```

Background and aim

Management Strategy Evaluation (MSE) are statistical simulation experiments designed to test competing management measures (e.g. harvest control rules) in an uncertain world. MSE are usually designed as Monte-Carlo simulations and then are limited by the finite number of iterations causing incomplete statistical convergence. Although recent computing advances have allowed improved efficiency in the simulation process, the number of replications (or iterations) to be performed is still an issue in MSE. A low value can drive to inaccurate advice and high values can be computational expensive.

Figure 1. Management Strategy Evaluation scheme.

Convergence is a desirable property in random variable when we make statistical simulations. Convergence may be defined as the approach of an infinite series to a finite limit. Translated to MSE simulations we could say that "true" value of performance measures are just achieved with infinitum iterations. These performance measures include two sources of errors: **Statistical Error**, coming from error model (observation, process, implementation), and **Simulation Error**, coming from sampling size simulation, i.e. number of iterations. Simulation error is reduced when iterations increases and disappear totally with infinitum iterations.

A fast and incomplete review of publications on MSE can show the number of iterations ranging from 100 to 10 000. However there is not a usual practice in these reports to find a explanatioin anbout the number of iterations used neither evaluating the impact of limited iterations in MSE simulations. 30 MSE developed in ICES since 2013 (Colm Presentation). Most of them use 1000 iterations although this can be different. De Oliveira presentation on number of replications analisis (500 to 10000 each 500) to decide 1000 they are enough.

There is a lack of agreed criteria to set the number of iterations. Punt et al.(2016) says that "The number of simulations for each trial should be selected to ensure that the percentiles of the distributions on which performance statistics are based can be calculated with the precision required for the decisions to follow"

ICES MSE (2013) did some recomendations to the number of iterations and this work was adressed focussing on analitycal estimation of probability errors. The probability calculated on the basis of n *iterations* will be within the interval $prob \pm 1.96 * \sqrt{\frac{prob*(1-prob)}{nIter}}$ in aproximately 95% of the cases. This was the base to analize errors around 3 different risk scenarios as follow:

- Prob1.- average probability that is below, where the average (of the annual probabilities) is taken across years.
- Prob2.- probability that is below at least once during years.

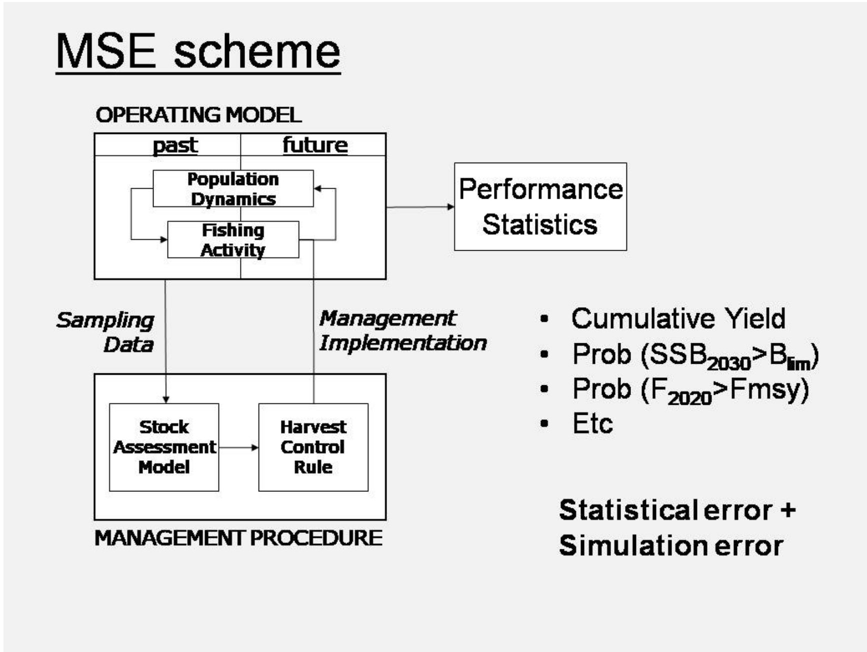

Figure 1: ...

- Prob3.- maximum probability that SSB is below Blim, where the maximum (of the annual probabilities) is taken over years.

Current ICES approach to considering a management plan/strategy precautionary is based on the probability of the spawning stock biomass (SSB) for the stocks concerned being below the precautionary reference point B lim. For medium to long-live stocks a management plan/strategy is precautionary if the maximum probability that SSB is below B lim is les than 5%, where the maximum (of the annual probabilities) is taken over all years in the plan.

Computing time is allways limiting our simulations. When is then CONVERGENCE achieved? The common sense says that to achieve convergence the number of iterations should be those that do not alter the MSE conclusions. For instance, whether the aim of MSE simulations is to identify if a HCR performs better than other (comparing some selected statistical measures for different OMs), then a correct number of iterations, i.e. convergence, should reduce to a low level the probability of saying, for instance, that HCR1 gives higher Pr(SSB>Blim) than HCR2 when in fact this is false or that HCR3 is precautionary when in fact is not.

In this exercise we present some simulations to evaluate the impact of the number of iterations with the aim of understand better the factors affecting convergence, the impact of limited iterations and to provide some recommendations to set a sound amount of iterations in MSE.

Illustrative example

Just to avoid confusion in terminology, here a **Simulation** is a stochastic performance that combines a OM with a MP producing one or more performance measures. This stochastic simulation is made of n **iterations** or **replications**. **Repetition** is used here to reffer to a new performance of a simulation with a different random seed.

Let's suppose that Operating Model 1 (OM1) under Harvest Control Rule 1 (HCR1) gives a SSB in 2030 with a median 25000 tonnes and CV 0.5 after ∞ replications. In the real world this density can be aproximated with a limmited number of replications. Next plot compares the true density with the density obtained with 100 and 1000 iterations for 4 simulations.

```r
set.seed(1234)
logsd <- sqrt(log(0.5^2+1))
logm <- log(25000)
SSB30 <- rlnorm(1000000, logm, logsd )
```

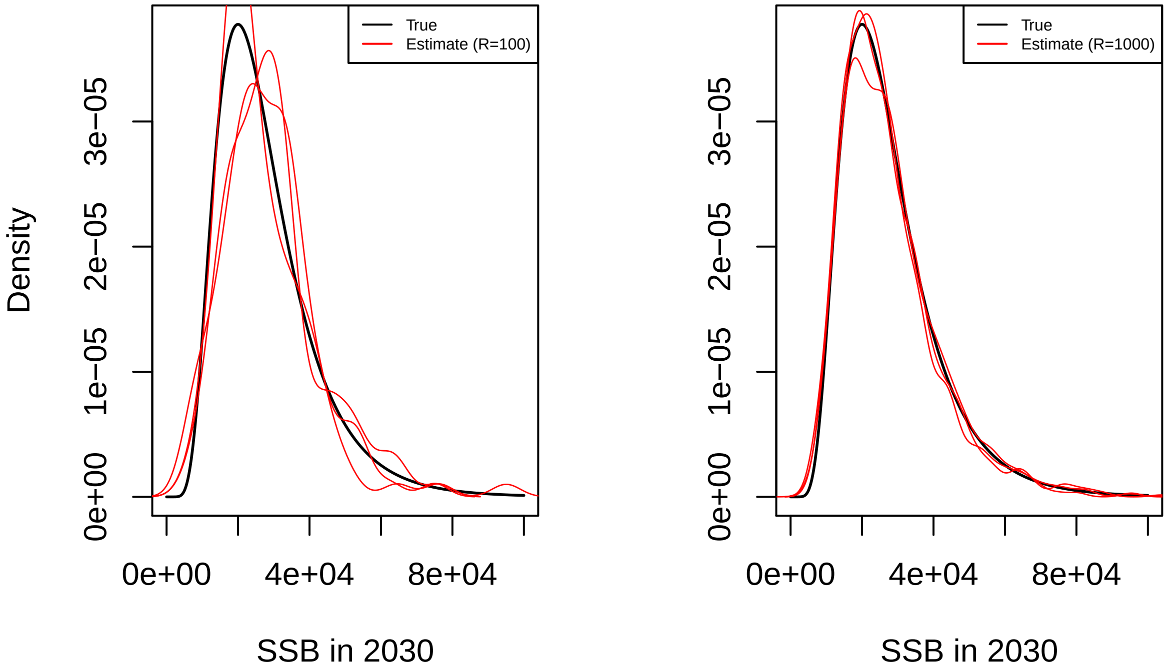

Fig 2. True SSB in 2030, i.e. simulation error is zero; compared with 3 samples of 100 replications (left) and 1000 replicates (right)

Convergence and cumulative plots

In MSE we are more interested in statistical properties such as mean, median, percentiLes or probability of being above or below a given level. Let's see how a limited number of iterations affects convergence and SE for these statistics. A simple way to do it is though cumulative plots. It requires a limited number of samples ("iterations or replications") taken from the true distribution.

```r
# Data
  dat <- SSB30[1:5000]
  cumLag <- 10
  Blim <- qlnorm(0.05, logm, logsd, lower.tail = TRUE)
  grd <- seq(cumLag, length(dat), by=cumLag)
  statsDF <- cumPerc(dat, percs=c(0.01, 0.05, 0.5, 0.95, 0.99), cumLag)
  statsDF$mean <- cumMean(dat, cumLag)
```

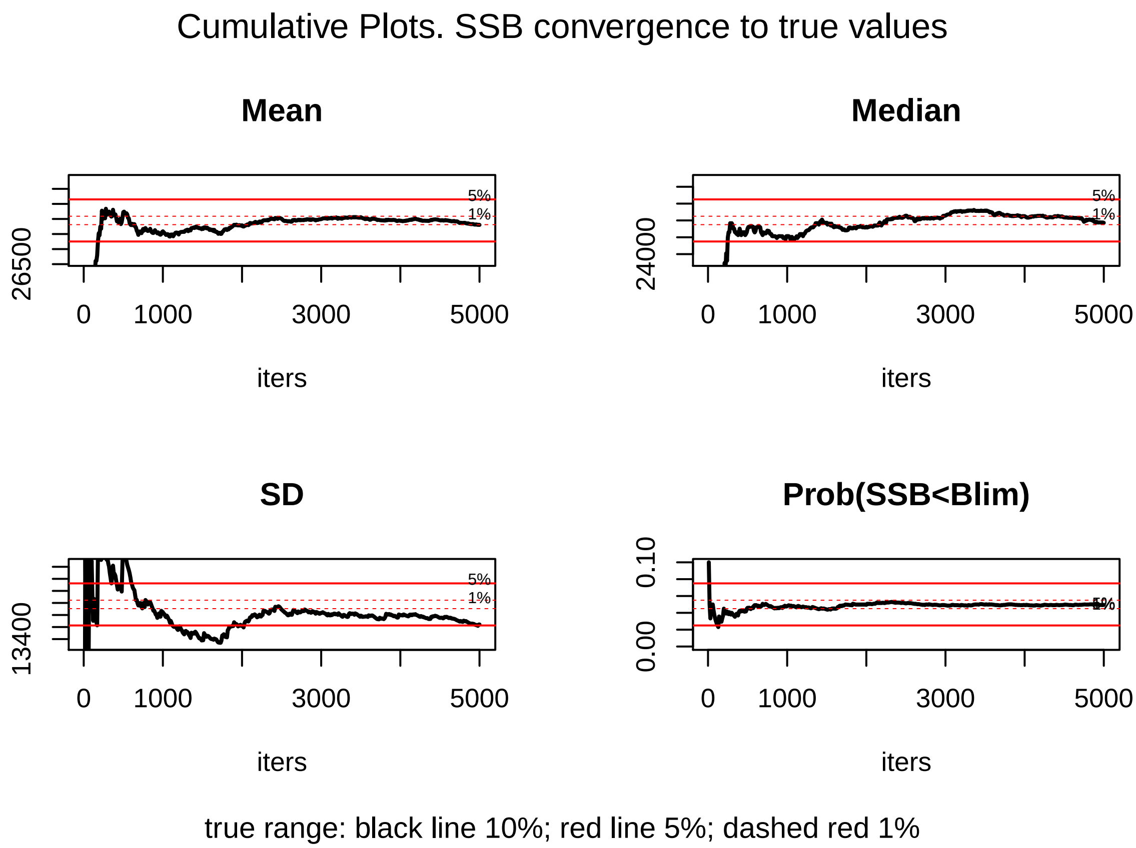

Fig. 3. Convergence trends for some pèrformance statistics of SSB.

Different statistics converge at different rate towards true figures. The plot set as references proportional ranges around the true figures. Obviously when running MSEs true figures are unknown. However similar plots with ranges around last figure can help to have a rough idea about the current level of convergence.

How different statistics converge at different speed is clear with the cumulative percentiles. Percentiles near the tail converge slower than central figures such as median. Upper tail converge even slower in log normal distributions.

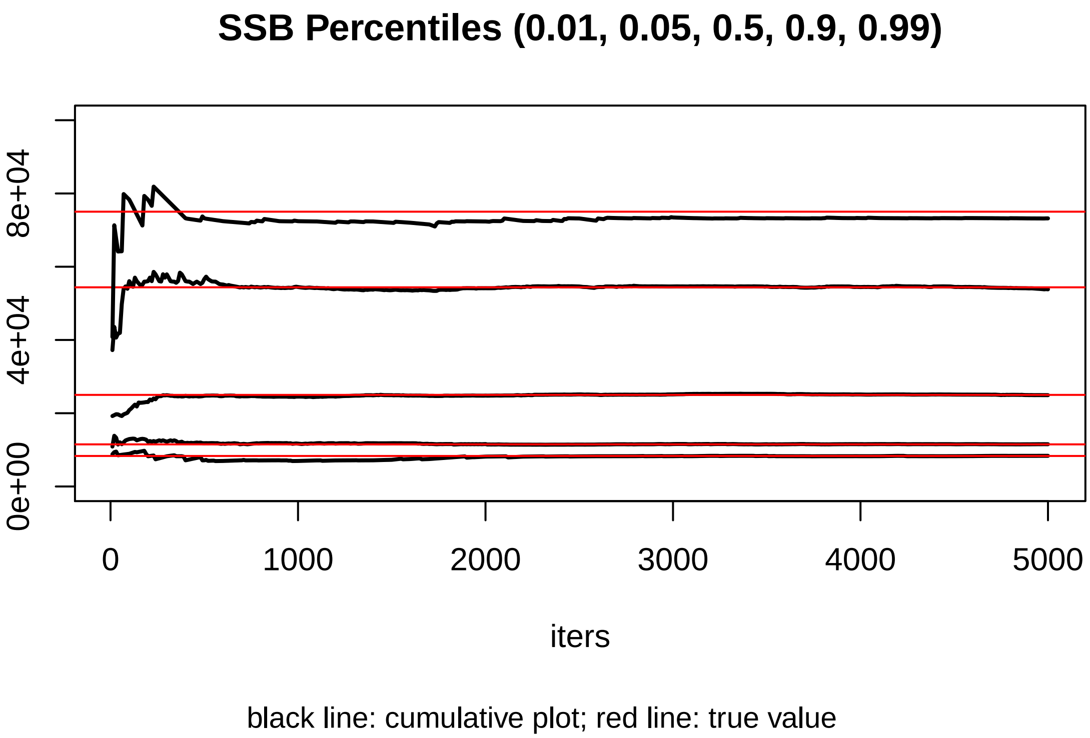

Fig. 4. Convergence trends for SSB 2030 percentiles.

The plot shows how different percentiles converge to true figures. However this convergence looks quite fast for the lower percentiles (0.5, 0.1 and 0.01). The 0.5 percentile cumulative line can be compared with the median convergence in the previous plot. Both are the same data with different scale for the y axis. This highliths the importance of set correctly the y axis ranges for this cumulative plots to be usefull to check convergence.

Simulation Error with repeated simulations

Good help for Latex equations in wikibooks

MSE simulation studies are under sampling error in a similar way than any research based in a finite sample from a broader population. This implies than when a simulation is repeated with a different seed different results are obtained. Here we call this between-simulation variability Simulation Error (SE), wich is defined as the standard deviation of a MSE performance measure (Koelher et al., 2009), taken across repetitions of the simulations with the same amount of replications.

$$SE(\phi_{it}) = \sqrt{Var(\phi_{it})}$$

where ϕ ia a MSE performance measure of interest and ϕ_{it} denote the estimate of ϕ from a simulation with it iterations. In a similar way, the Simulation Coeficient of Variance (SCV) is $SCV(\phi_{it}) = \frac{SE(\phi_{it})}{mean(\phi_{it})}$. Obviously $SE(SSB_{\infty}) = 0$. Convergence then is defined as the limit of SE when the number replications (it) tend to ∞

$$\lim_{it \to \infty} SE = 0$$

Simulation error is reduced when iterations increases and disappear totally with infinitum iterations. Since this is not possible a criteria to decide how many iterations to perform is needed. A simple approach is repeating the simulation with a different random seed and check you get the same results. But this can be done in a more conclusive way. The idea is to perform a big number of iterations and check than performance measures are stable at a level that do not compromises your conclusions. For example, with a MSE simulation with 5000 iterations, first, estimate the Pr[SSByear<Blim] for the first 10, 20 30, ... 5000 iterations and then, plot these probabilities against iterations number. Looking at this plot we can see how stability is achieved as the number of iterations increase. We can decide the number of iterations needed to achieved the required stability. If the number of scenarios is big, this exercise should be done at least with some representative scenarios.

Repeted cumulative plots can also serve to estimate simulation errors.

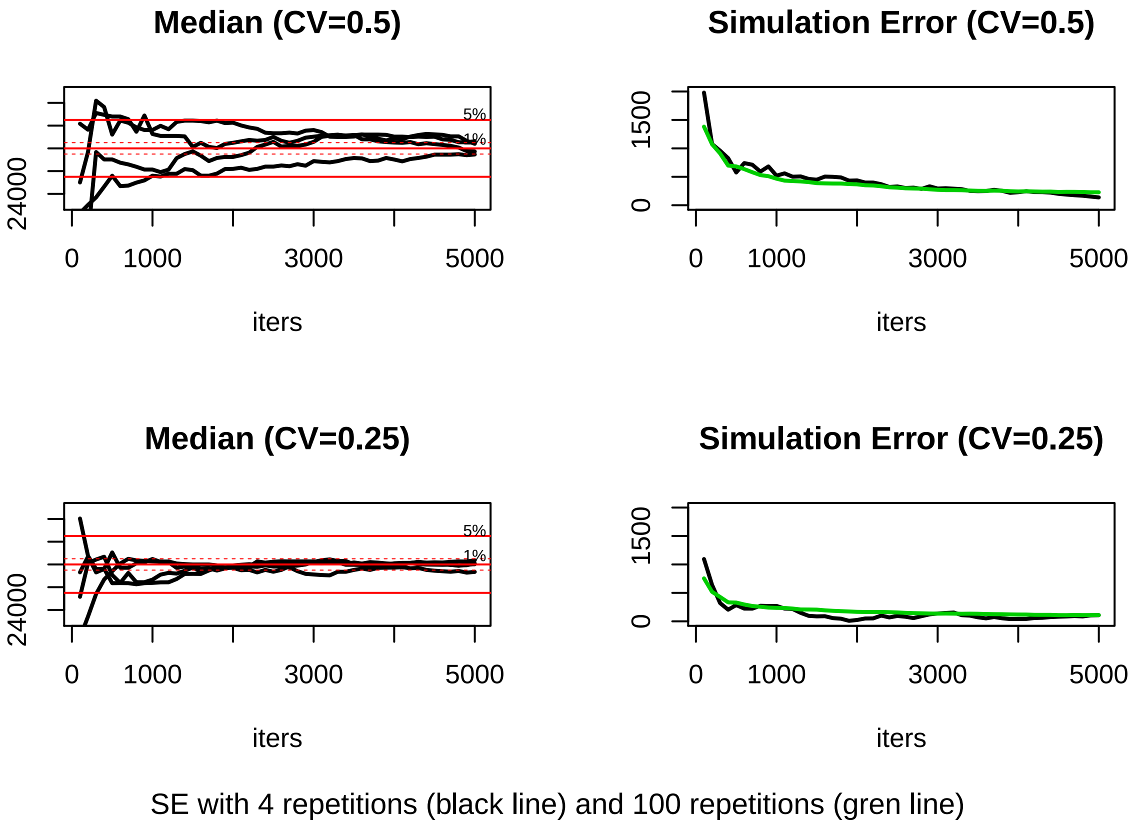

Fig. 5. Convergence trends for four simulations of the same distribution with a maximun of 5000 iterations (left) and, in the right, the Simulation Error estimated with the 4 simulations (black) or with 100 (green). Upper plots with SSB CV = 0.5 and lower plots with SSB CV=0.25

Repeating the simulation with a different random number (in case it was fixed in the script) wound allow to check whether the different statistical mesures converge to the same amount, or whether the diferrences are acceptable or not and more iterations would be needed. Furthermore, repeating the simulation also allows to have a measures of the Simualtion Error. However, to have an accurate measure we need a bigger amount of simulations something that is not ussually feasible. A way to check whether the number of repetitions is enough is looking at the SE trend, that should look like a negative exponential function zero without noise.

Fig 5 also shows that variability of the performance measure (SSB in this case) contributes to convergence sped. Performance statistis estimeted from a SSB with a CV=0.5 converge slower than other with a CV=0.25. SE with cv=0.5 is about two times SE with CV=0.25.

Precautionary approach convergence

ICES approach to considering a management plan/strategy precautionary is based on the probability of the spawning stock biomass (SSB) for the stocks concerned being below the precautionary reference point B lim.

For medium to long-live stocks a management plan/strategy is precautionary if the maximum probability that SSB is below B lim is less than 5%, where the maximum (of the annual probabilities) is taken over all years in the plan (i.e. short as well as long terms). This criteria coincides with the "Prob2" criteria described in ICES MSE (2013) section 3.2. and the analisys performed there still applies. This precautionary criteria can be biased under slow number of iterations increasing the risk of setting a plan/strategy as non-precautionary when in fact it is. This risk increases also with high number of years considered. Time independency also increasese this risk.

Preliminary conclusions and recommendations

- There is not a definite rule to set the number of iterations. However some explicit rationale about the number of iterations chosen should be presented in ME reports since this number can affect the MSE conclusions.
- The variability of the variable, the performance statistic or the proximity to reference points affect the convergence speed.
- Visual diagnostics such as cumulative plots can help as a final check to convergence. Repeated simulations can help to identify critical convergence levels. However we can not spend more time on diagnostics than in MSE when computing time is issue. -An alternative can be to identify one or two critical scenarios where convergence is an issue, explore in depth this scenario to set the minimum iterations needed. And apply these to all the scenarios.
- Precautionary criteria is a critical issue with limited iteration. Pay special attention to convergence!

References

ICES. 2013. Report of the Workshop on Guidelines for Management Strategy Evaluations (WKGMSE) , 21 - 23 January 2013, ICES HQ, Copenhagen, Denmark. Koehler, E., Brown, E., & Haneuse, S. J.-P. A. (2009). On the Assessment of Monte Carlo Error in Simulation-Based Statistical Analyses. The American Statistician, 63, 155-162.