Noname manuscript No.
(will be inserted by the editor)

Sensitivity to MDA steps and ensemble size

Geir Evensen

the date of receipt and acceptance should be inserted later

1 Sensitivity of ESMDA to the number of steps

We have included a sensitivity experiment to examine the convergence properties of the ESMDA algorithm. We expect that the accuracy of the solution will improve with the number of steps, until a certain level where there is nothing more to gain. Evensen (2018) examined the convergence of ESMDA for a simple nonlinear scalar case and obtained minimal improvement after 16-32 steps. Fig. 1 presents the posterior solution for deaths and hospitalized when using ESMDA with 1, 2, 4, 8, 16, 32, 64, and 128 steps, and in Fig. 2 we show the corresponding estimates of $R(t)$. From visual inspection, it is hard to justify more than 16 steps. We decided to use 32 steps in all the simulations presented in this manuscript to ensure convergence. All these cases used 5000 model realizations to reduce the impact of sampling errors (see next section).

2 Sensitivity of ESMDA to the size of the ensemble

Finally, we have examined the convergence of ESMDA concerning the ensemble size. ESMDA being a Monte Carlo algorithm, means that we can always improve the solution by increasing the ensemble size. However, we need to decide on a tradeoff between ensemble size and the number of ESMDA steps due to limitations on computing power. While the number of ESMDA steps impacts the actual convergence of the algorithm towards the correct solution, the ensemble size impacts the precision of the statistical estimate of the final solution. We find it most important to first converge to the correct physical solution and, then, use an as large as possible ensemble to reduce the sampling errors. In Figs. 3 and 4 we show the results using different ensemble sizes and two different random seeds. First of all, these plots demonstrate the robustness of ensemble-based assimilation methods. Even using 100 realizations, the posterior predictions are consistent with the data and very similar to the cases with larger ensemble sizes. There is a visual dependency on the random seed, which is seen more clearly in the estimated $R(t)$. When using a 1000 realizations, there is still a significant difference in the estimated $R(t)$. For the predictions, it is hard to note any dependency on the seed, or difference from the case with 5000 realizations. With 5000 and 10000 realizations, we can hardly see any difference in the parameter estimates or predictions. Thus, we use 5000 realizations in all the simulations in the paper.

Address(es) of author(s) should be given

References

Evensen G (2018) Analysis of iterative ensemble smoothers for solving inverse problems. Computat Geosci 22(3):pp. 885–908, DOI 10.1007/s10596-018-9731-y, URL https://doi.org/10.1007/s10596-018-9731-y

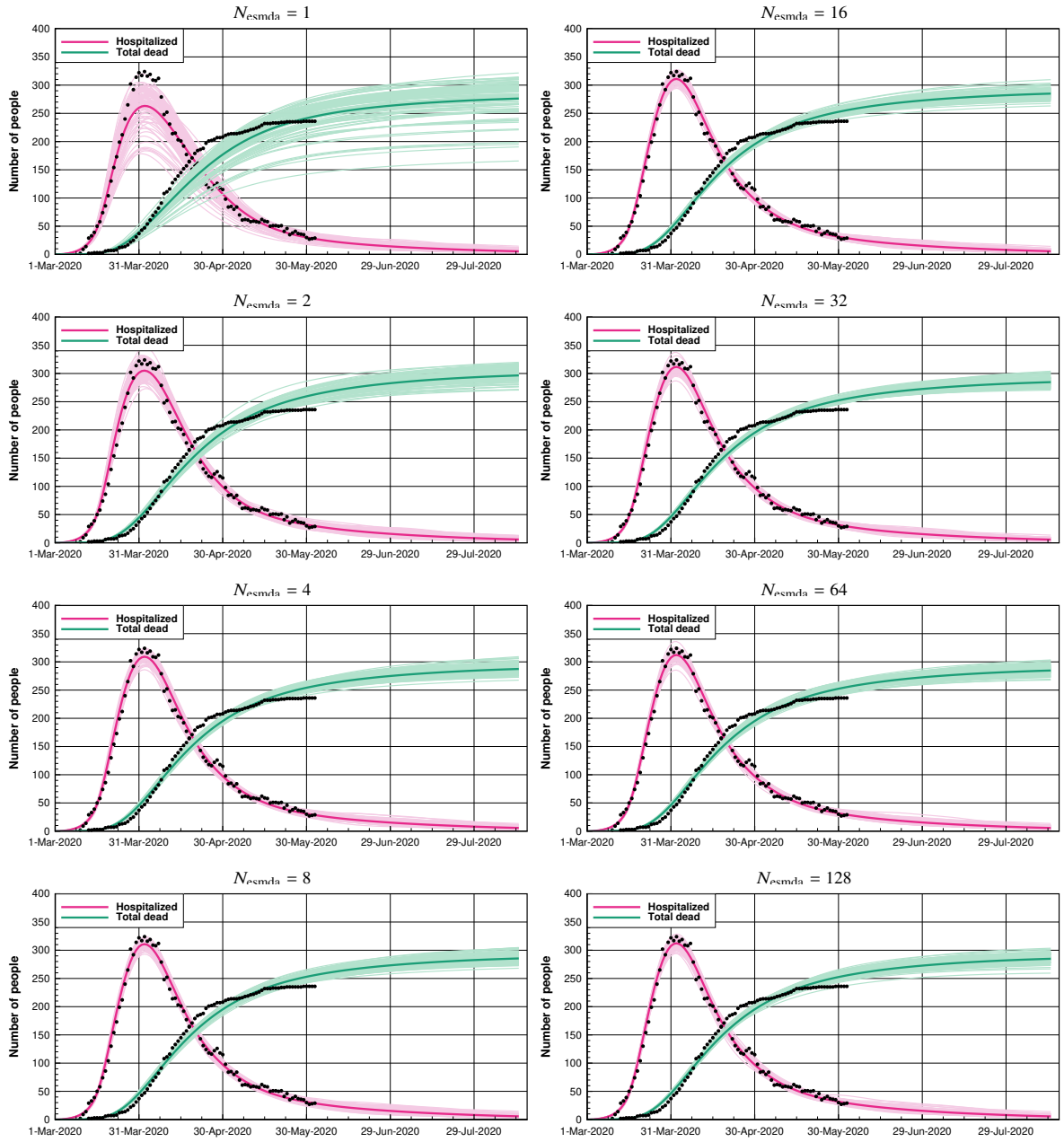

Fig. 1 Norway (convergence of ESMDA using different numbers of MDA steps): The plots show the number of deaths and hospitalized with 1, 2, 4, 8, 16, 32, 64, and 128 steps.

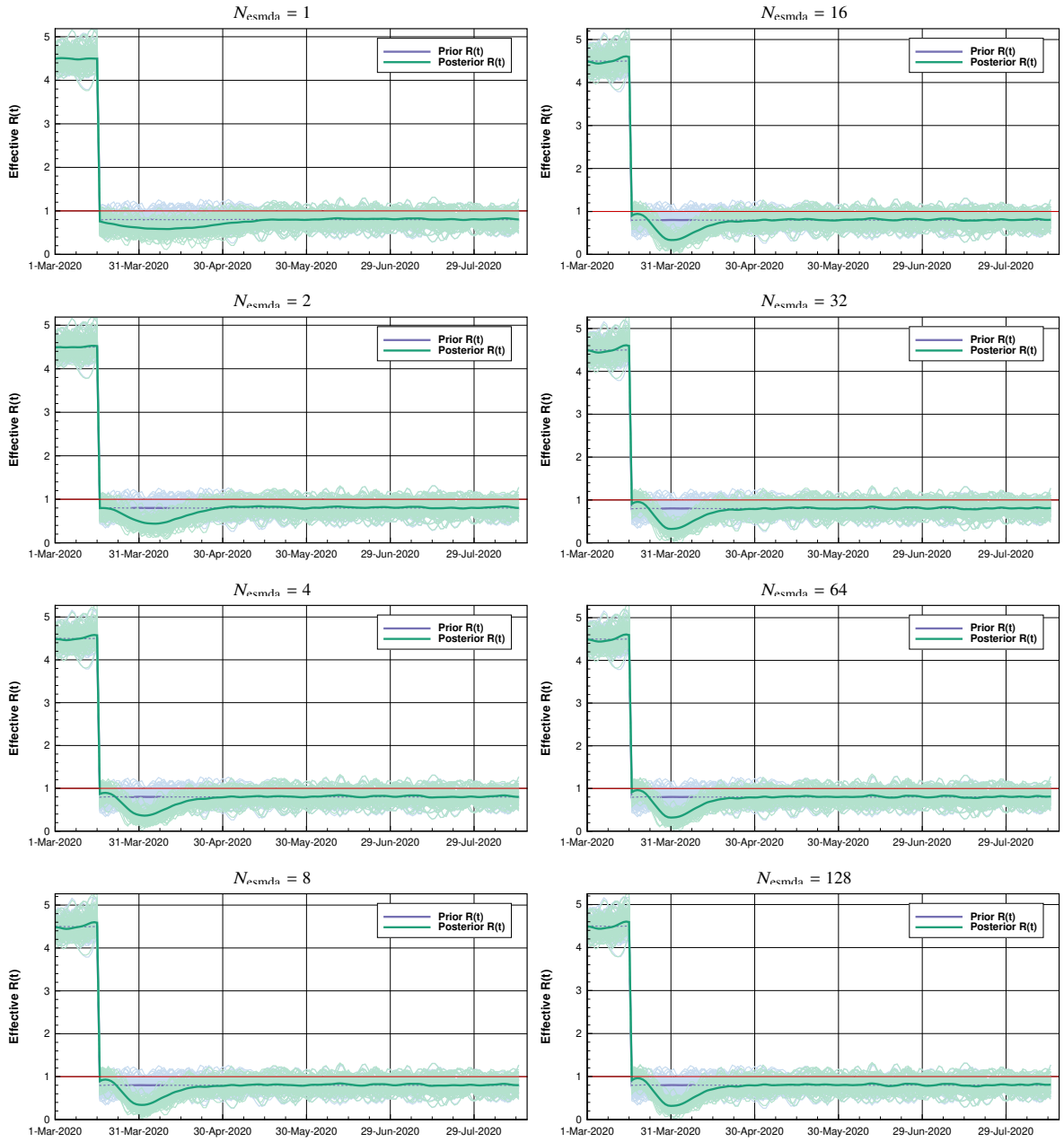

Fig. 2 Norway (convergence of ESMDA using different numbers of MDA steps): The plots show the estimate of $R(t)$ using 1, 2, 4, 8, 16, 32, 64, and 128 MDA steps.

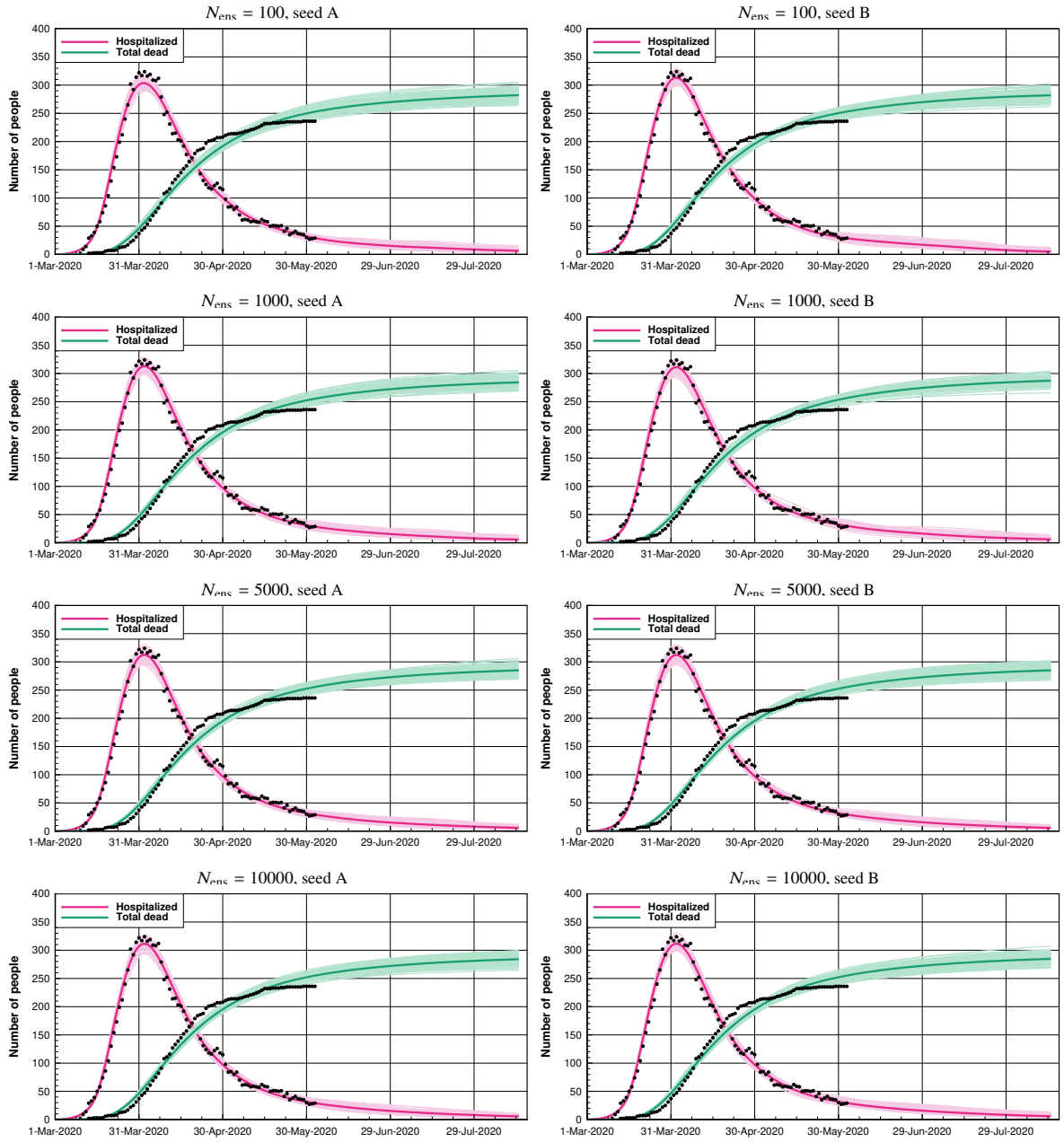

Fig. 3 Norway (convergence of ESMDA with ensemble size): The plots show the number of deaths and hospitalized using different ensemble sizes.

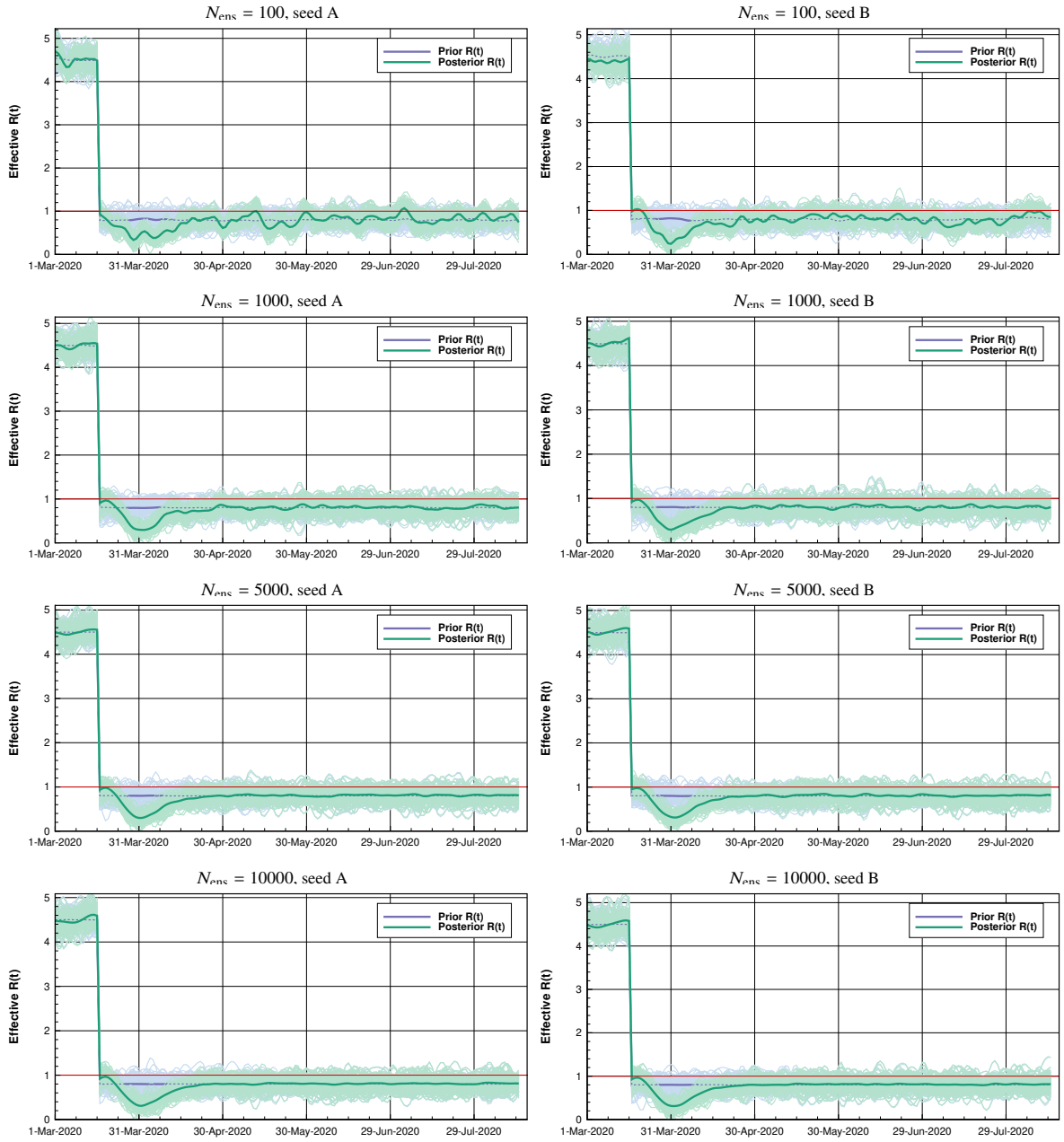

Fig. 4 Norway (convergence of ESMDA with ensemble size): The plots show the estimated $R(t)$ using different ensemble sizes.