



Why NRMSE is not completely reliable for forecast/hindcast model test performances

Lorenzo Mentaschi (1), Giovanni Besio (1), Andrea Mazzino (1), and Federico Cassola (2)

(1) DICCA, Università di Genova, Via Montallegro 1, 16145 Genova, Italy, (2) DIFI, Università di Genova, Via Dodecanneso 33, 16146 Genova, Italy

We investigate the reliability of the statistical error indicator NRMSE (Normalized Root Mean Square Error) as an index of the performances of numerical simulation for wave forecasting. This widespread indicator, also known as Scatter Index, is defined as

$$\text{NRMSE} = \frac{\sum (S_i - O_i)^2}{\sum O_i^2} \quad (1)$$

where O_i are observed values and S_i are simulated values. A small value of NRMSE identifies a numerical simulation in good agreement with the field observations.

In our analysis we show that NRMSE can be orthogonally decomposed in two contributions related to the average bias and to the scatter of simulated values around the observed values, respectively. Results from numerical simulations show how these contributions appear statistically dependent on each other because positive or negative amplifications of the simulation average generally involve a corresponding amplification in the scatter of simulated values. An almost linear dependency between bias and scatter can be found and it implies that minimum NRMSE simulation and unbiased simulation do not coincide, since minimum NRMSE requires a minimum squared sum of its bias and scatter contributions. This finding suggests that a lower value of NRMSE is not always associated to the best results, and that this indicator is not always reliable if used to find the best simulation. This flaw of NRMSE indicator is already known in literature, and some authors proposed the usage of corrected indicators to overcome it (e.g. Hanna and Heinolds, 1985[1]).

References

- [1] Hanna, S.R. and Heinold, D.W. 1985 Development and Application of a Simple Method for Evaluating Air Quality, API Pub. No. 4409, Washington, D.C.