

Supplementary Information for *Disentangling Density and Geometry in Weather Regime Dimensions using Stochastic Twins*

Paul Platzter^{1,2*}, Bertrand Chapron^{1,2} and Gabriele Messori^{3, 4}

¹Laboratoire d’Océanographie Physique et Spatiale, Univ.
Brest/Ifremer/CNRS/IRD, Plouzané, F-29280, France.

²Odyssey, Inria/IMT/CNRS, Plouzané, F-29280, France.

³Department of Earth Sciences, Uppsala University, Uppsala, 752 36,
Sweden.

⁴Department of Meteorology, Stockholm University, Stockholm, 106 91,
Sweden.

*Corresponding author(s). E-mail(s): paul.platzer@ifremer.fr;

This supplementary information is attached to the main paper entitled “Disentangling Density and Geometry in Weather Regime Dimensions using Stochastic Twins”.

In Fig. S1 are compared, on the one hand, the time-autocorrelation functions of the projections of the three leading empirical orthogonal functions of the 10 days running-average of ERA5 500 hPa geopotential height anomaly (full lines), and on the other hand, the three first coordinates of 4 out of the 20 stochastic twins used in this study (dotted lines). This mild visual agreement was used as a basis for setting the time-step of the stochastic twin integration scheme.

In Fig. S2 is shown the average of weather regime indices (WRI) computed over life-cycles defined as exceedances of the threshold $WRI_{i,1}$ for at least 5 consecutive days and centered on the time of maximum WRI, for each regime index, and averaged over all 20 stochastic twins. This plot highlights that no information on the asymmetry was included in the stochastic twins. For instance, it is known that there are preferred transitions between regimes [1], which results in asymmetry around rescaled time in Fig. 2 of the main paper. Also, the curves are more peaked around $t = 0$ than in Fig. 2 of the main paper, highlighting the erratic nature of the stochastic twins compared to ERA5, also witnessed in Fig. 4 of the main paper.

Fig. S3 shows the dependence of the correlation coefficient between density-based local dimension (from stochastic twin analogs) and full local dimension variations (from ERA5 analogs) to the number of analogs K . This correlation coefficient is the one printed in text over Fig. 5 in the main paper. The fact that the correlation coefficient is a growing function of K is expected since for low values of K the estimation of local dimension is more noisy : estimating a tail parameter from fewer data allows to decrease bias but at the cost of increasing variance. Although there is a strong dependence on K , the fact that the correlation coefficients are always non-negligible indicates that a large fraction of variations of local dimension estimates are density-based.

Fig. S4 plots the local dimension estimated from whole fields of smoothed z500 anomaly versus the local dimension computed from stochastic twin analogs of *projected* smoothed z500 anomaly. The results are almost identical to the ones of Fig. 5 of the main paper. Similarly, Fig. S5 and S7 show that the results in terms of local dimensions over weather-regime life cycles and peak index are insensitive to the fact that local dimensions are computed on projected rather than full ERA5 fields.

Fig. S6 and S8 repeat the operations of the main paper but using only stochastic twins (both as targets and analogs). For Fig. S8, all peak WRIs from the 20 stochastic twins are used, which is why the plot is more dense than Fig. 7 of main paper. In the case of Fig. S6 and S8, the results differ significantly from the case where ERA5 data is used as target. In particular, the decrease of estimated local dimension is much weaker when using stochastic data alone for NAO-, while it is stronger in absolute value for Atlantic Ridge (it is not in relative value, see the similar discussion of Fig. 6 in main paper). However, the observed effect of decreased dimension around peak WRI is still witnessed using stochastic twin data alone, confirming the density-based nature of the decrease of estimated local dimension near peak weather regime indices.

For more details, the reader is referred to the main paper “Disentangling Density and Geometry in Weather Regime Dimensions using Stochastic Twins”.

References

- [1] Kondrashov, D., Ide, K., Ghil, M.: Weather regimes and preferred transition paths in a three-level quasigeostrophic model. *Journal of the atmospheric sciences* **61**(5), 568–587 (2004)

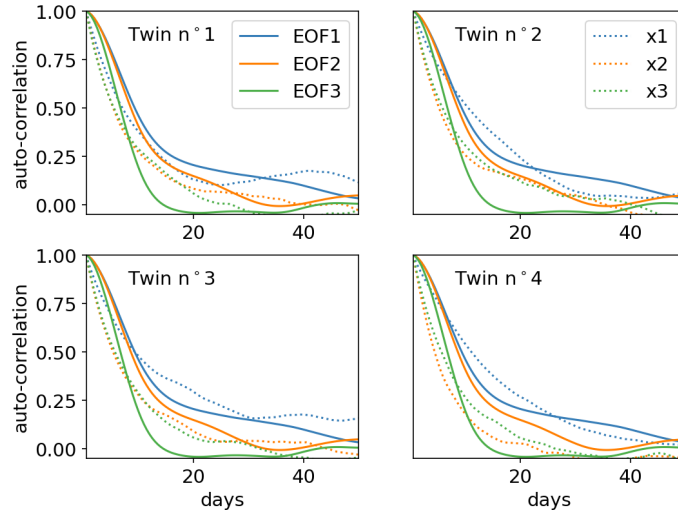


Fig. S1 Auto-correlation functions for EOF₁₂₃, compared to auto-correlation functions of 4 stochastic twins.

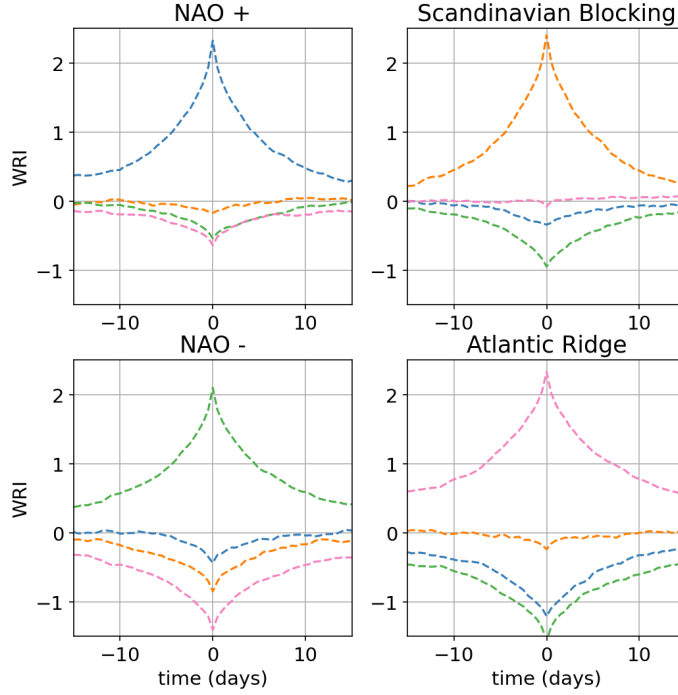


Fig. S2 Weather regime index (WRI) averaged around time of maximum WRI, for life-cycles of at least 5 days during which WRI > 1.

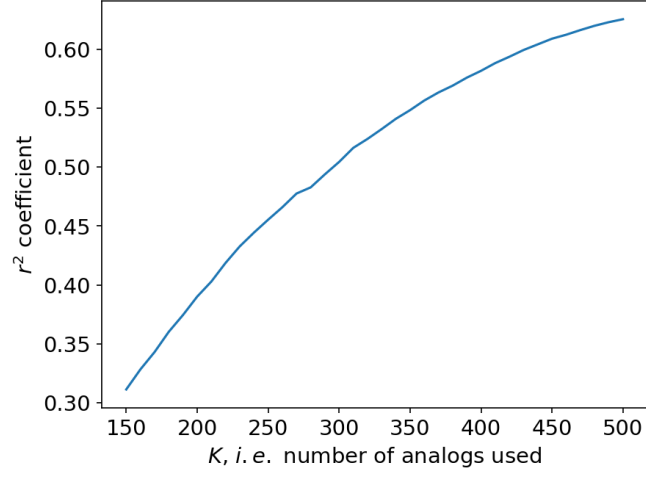


Fig. S3 Sensitivity of the correlation coefficient between density-based local dimension (from stochastic twin analogs) and full local dimension variations (from ERA5 analogs) to the number of analogs K . The coefficient of 0.60 shown in Fig. 5 of main paper corresponds to $K = 450$ analogs.

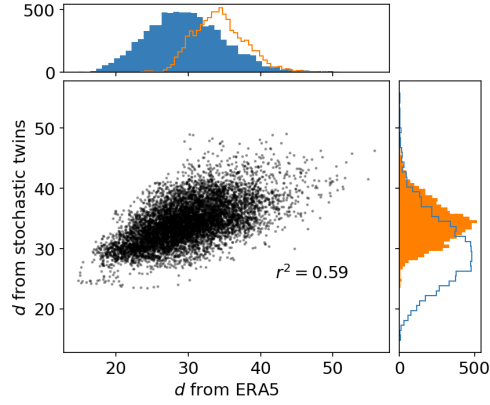


Fig. S4 Same as Fig. 5 of main paper, but the local dimension from ERA5 alone are computed based on the whole fields of smoothed z500 anomaly rather than on the projection onto the leading 65 EOFs.

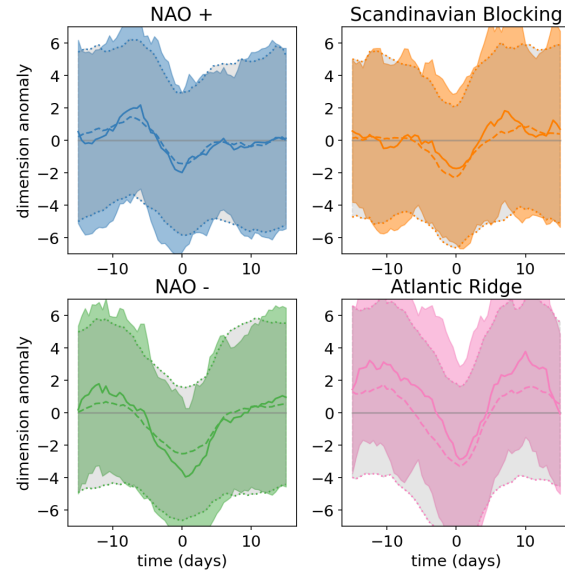


Fig. S5 Same as Fig. 6 of main paper, but the local dimension anomaly from ERA5 analogs are computed based on the whole fields of smoothed z500 anomaly rather than on the projection onto the leading 65 EOFs.

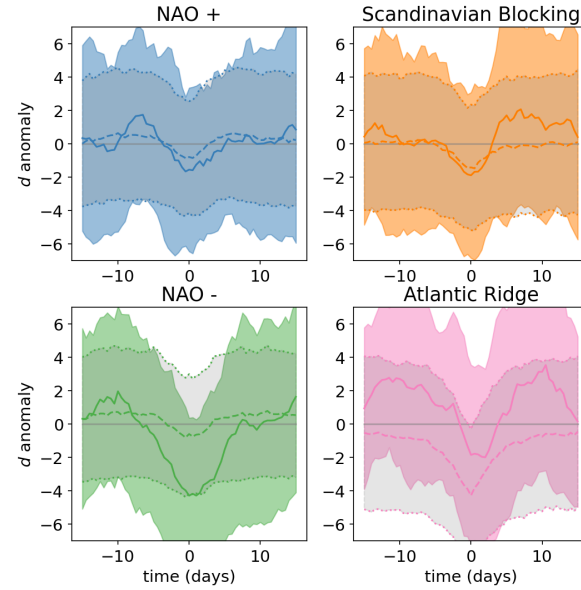


Fig. S6 Same as Fig. 6 of main paper, but here stochastic twins are used both as analogs and as targets to build the dashed curves.

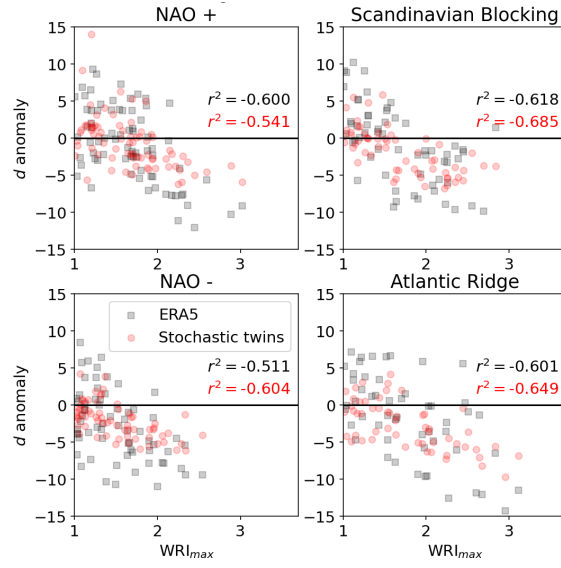


Fig. S7 Same as Fig. 7 of main paper, but the local dimension anomaly from ERA5 analogs are computed based on the whole fields of smoothed z500 anomaly rather than on the projection onto the leading 65 EOFs.

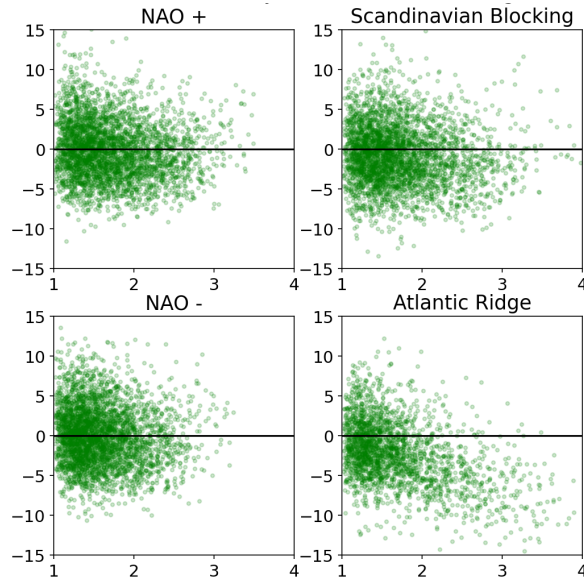


Fig. S8 Same as Fig. 7 of main paper, but using only stochastic twin data.