Decomposing weather maps into interpretable patterns using Latent Dirichlet Allocation

Lucas Fery[1,2], Bérengère Dubrulle[3], Bérengère Podvin[4], Flavio Pons[1], and Davide Faranda[1,5,6]

[1] Laboratoire des Sciences du Climat et de l'Environnement, CEA Saclay l'Orme des Merisiers, UMR 8212 CEA-CNRS-UVSQ, Université Paris-Saclay & IPSL, 91191, Gif-sur-Yvette, France
[2] Department of Physics, Ecole Normale Supérieure de Lyon, 69364, Lyon, France
[3] SPEC, CEA, CNRS, Université Paris-Saclay, F-91191 CEA Saclay, Gif-sur-Yvette, France
[4] LISN, CNRS, Université Paris-Saclay, 91405, Orsay, France
[5] London Mathematical Laboratory, 8 Margravine Gardens, London, W6 8RH, UK
[6] LMD/IPSL, Ecole Normale Superieure, PSL research University, 75005, Paris, France
`lucas.fery@lsce.ipsl.fr`

The chaotic dynamics of mid-latitude atmospheric circulation is often described with a few number of weather regimes. These are typical field configurations featuring a combination of several synoptic objects (cyclones and anticyclones). Such intrication makes it arduous to quantify recurrence and intensity of climate extremes and to get insight into their genesis and dynamical properties, as extreme events are often associated with specific synoptic structures. Here we apply Latent Dirichlet Allocation (LDA) [1], which has been used for topic modeling in natural language processing (NLP), to build a weather dictionary: we define daily maps of a gridded target observable as documents, and the grid-points composing the map as semantic words. LDA then provides a representation of weather maps in terms of a combination of spatial patterns named motifs, which are latent variables inferred from the set of snapshots. For atmospheric data (daily sea-level pressure anomalies), we find that motifs correspond to pure synoptic objects (cyclones and anticyclones), that can be seen as building blocks of weather regimes. We show that LDA weights provide a natural way to characterize the impact of climate change on the recurrence of patterns associated with extreme events.

References

1. D. M. Blei, A. Y. Ng, & M. I. Jordan, Latent dirichlet allocation, *Journal of machine Learning research*, **3**, 993-1022, (2003).
2. L. Fery, B. Dubrulle, B. Podvin, F. Pons & D. Faranda, Learning a weather dictionary of atmospheric patterns using Latent Dirichlet Allocation, *Geophysical Research Letters*, **49**, e2021GL096184. https://doi.org/10.1029/2021GL096184 (2022).

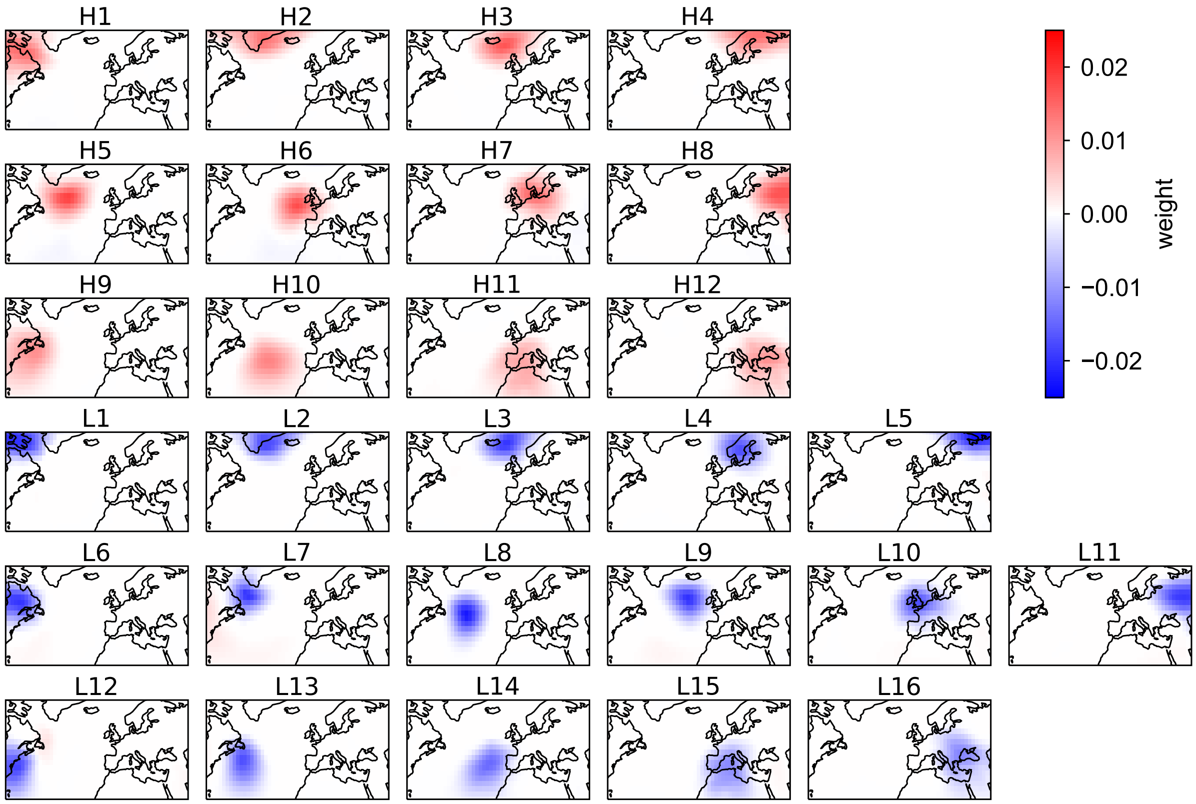

Figure 1. The 28 motifs identified by LDA by using as documents the daily sea-level pressure anomaly maps over North-Atlantic from 1948 to 2018 from NCEP reanalysis. Positive anomalies and negative anomalies are respectively represented in red and blue and are grouped accordingly in high pressure motifs (H) and low pressure motifs (L). The motifs are then roughly ordered by their geographic location from West to East and North to South. LDA outputs normalized weights, so the color intensity should be interpreted as a normalized anomaly amplitude.