Article

Unsupervised detection of large-scale weather patterns in the northern hemisphere via Markov State Modelling: from blockings to teleconnections

Check for updates

Sebastian Springer[1,6], Alessandro Laio[1,2,6] ✉, Vera Melinda Galfi [3,6] & Valerio Lucarini [4,5,6] ✉

Detecting recurrent weather patterns and understanding the transitions between such regimes are key to advancing our knowledge of the low-frequency variability of the atmosphere and have important implications in terms of weather and climate-related risks. We adopt an analysis pipeline inspired by Markov State Modelling and detect in an unsupervised manner the dominant winter mid-latitude Northern Hemisphere weather patterns in the Atlantic and Pacific sectors. The daily 500 hPa geopotential height fields are first classified in about 200 microstates. The weather dynamics are then represented on the basis of these microstates and the slowest decaying modes are identified from the spectral properties of the transition probability matrix. These modes are defined on the basis of the nonlinear dynamical processes of the system and not as tentative metastable states, as often done in Markov state analysis. When focusing on a shifting longitudinal window of 60°, we find that the longitude-dependent estimate of the longest relaxation time is smaller where stronger baroclinic activity is found. In the Atlantic and Pacific sectors slow relaxation processes are mainly related to transitions between blocked regimes and zonal flow. We also find strong evidence of a dynamical regime associated with the simultaneous Atlantic-Pacific blocking. When the analysis is performed on a broader geographical region of the Atlantic sector, we discover that the slowest relaxation modes of the system are associated with transitions between dynamical regimes that resemble teleconnection patterns like the North Atlantic Oscillation and weather regimes like the Scandinavian and Greenland blocking, yet have a much stronger dynamical foundation than classical methods based e.g. on EOF analysis. Our method clarifies that, as a result of the lack of a time-scale separation in the atmospheric variability of the mid-latitudes, there is no clear-cut way to represent the atmospheric dynamics in terms of few, well-defined modes of variability. The approach proposed here can be seamlessly applied across different regions of the globe for detecting regional modes of variability, and has a great potential for intercomparing climate models and for assessing the impact of climate change on the low-frequency variability of the atmosphere.

Low-frequency variability in meteorology refers to a broad spectrum of atmospheric processes occurring over time scales ranging from approximately one week to about a month. Despite extensive research efforts, a comprehensive understanding of its underlying nature remains an ongoing challenge[1,2]. From a practical perspective, achieving efficient medium and extended range intra-seasonal forecasts, beyond the horizon of deterministic predictability of typically 7–10 days in the mid-latitudes, poses significant difficulties. Furthermore, the climatic impacts of low-frequency

A full list of affiliations appears at the end of the paper. ✉e-mail: laio@sissa.it; v.lucarini@leicester.ac.uk

1

phenomena are substantial[3], also in terms of modelling changing climatic conditions[4]. Long-time scales are typically entangled with large spatial scales[1,5,6]. Large-scale weather configurations in the mid-latitudes hold significant importance due to their impact on weather conditions over vast regions. Understanding low-frequency variability, in space and time, is vital as it has implications for weather forecasting, climate studies, and disaster preparedness.

One of the critical aspects of the mid-latitude low-frequency variability is the presence of transitions between two distinct flow regimes: blockings and zonal flow[7–10]. Blockings are synoptic features that are characterized by persistent high-pressure systems, leading to a disruption of the typical west-to-east flow of the jet stream[11–13]. Blocking events typically manifest in either the Atlantic or Pacific sectors and, more rarely, in both concurrently[14]. The lifetime of blocking events can span several days to a few weeks, resulting in the emergence of extreme and prolonged weather anomalies with significant regional implications[15]. Depending on geographic location, season, and pre-existing conditions, blockings can induce diverse weather phenomena such as heat waves, cold spells, extensive droughts leading to wildfires, and floods[16–23]. Predicting the onset and the decay of blockings is rather elusive even for modern weather prediction systems[22,24,25], and, similarly, modern climate models face difficulties in achieving a correct representation of their statistics[26–28]. Indeed, large uncertainties exist regarding the response of the statistics of blocking events to climate change[29], even if recent studies indicate the possibility of an overall increase in their size[30] and intensity[31], with a possible reduction in their persistence in the European sector[32]. Nonetheless, as discussed in the most recent IPCC report, results are not entirely conclusive in terms of the impact of climate change on blockings, as there is at best medium confidence in *a projected decrease in the frequency of atmospheric blocking over Greenland and the North Pacific for the late* 21st *century as compared to the reference climatology*[33]. These difficulties possibly reflect the fact that the blockings are associated with atypical and anomalously unstable conditions of the atmosphere[34–36], leading to serious fundamental implications in terms of overall model accuracy in describing their dynamics and statistical properties[37].

Accurate identification of blockings is usually achieved through various indices based on the geopotential height[38] or potential vorticity (PV) fields[39]. More recently, identification of blockings has been performed using statistical indicators relying on Empirical Orthogonal Functions (EOFs)[40] as well as multidimensional methodologies[41]; see ref. [42] for a recent survey on the topic.

The low-frequency variability of the northern hemisphere mid-latitudes features also atmospheric configurations having spatial and temporal scales larger than blockings, exhibiting links between weather phenomena affecting far-away regions on Earth. These are called teleconnection patterns, and are responsible for driving the dynamics of the atmosphere at planetary scale and for facilitating the coupling between atmospheric and oceanic processes. Particularly relevant are the so-called Pacific-North American teleconnection pattern (PNA)[43] and the North Atlantic Oscillation (NAO)[44]. The PNA pattern covers a substantial portion of the North Pacific and North America and exerts influence on synoptic activity and the subtropical jet stream over the North Pacific, leading to climate impacts over extensive regions of North America. The NAO is a dipolar pattern of mean sea level pressure over the North Atlantic, extending from the subtropical (Azores high) to sub-Arctic latitudes (Icelandic low). It significantly affects the variability of westerly winds over eastern north Atlantic and Western Europe, making it a crucial factor in shaping Europe's winter climate[45]. Whereas the PNA and the NAO are considered to be large-scale but still regional teleconnections patterns, hemispheric teleconnections have been detected as well, such as the Northern Annular Mode (NAM, named also Arctic Oscillation[46]) and the circumglobal teleconnection pattern[47]. These have nontrivial connections with both the NAO and the PNA[14,48]. See ref. [2,49–51] for a detailed discussion of the dynamics and detection methods of teleconnection patterns and their nontrivial link with blockings.

Besides the above-mentioned teleconnection patterns or main modes of variability, other recurring circulation patterns have been shown to be useful for describing in a coarse-grained sense the weather evolution over Europe and North America. These go under the umbrella of the so-called weather regimes (see for example[52–54]). They are more persistent and larger than synoptic scale weather systems, and are often associated with surface extremes over vast regions[55–58]. In many cases, the dominant feature of North-Atlantic weather regimes is a persistent ridge or blocking (over the North-Atlantic, Greenland, central Europe or Scandinavia), but some of them are predominantly cyclonic regimes[54,59].

A rather popular way, among others, to describe in a coarse-grained sense the evolution of the weather in the mid-latitudes amounts to constructing via data-driven methods finite state Markov models, where each state is a statistical cluster associated with a weather pattern, and then investigating the properties of the resulting Markov chain[34,60–63]. Recently, more advanced machine learning-based methods have been used to detect recurrent mid-latitude atmospheric regimes in a simple model of the atmosphere[64]. In this work, we wish to advance this perspective and propose a procedure to detect in an unsupervised manner anomalously persistent and recurrent weather patterns. The procedure follows the strategy of Markov State modelling (MSM), an approach originally introduced for studying metastability in molecular systems[65,66]. The key idea is first classifying all the data - e.g. the molecular configurations, or, in the context of this work, the daily averaged 500 hPa geopotential height (Z500) fields - in a finite set of *microstates*. Each microstate contains rather similar data. One then empirically estimates the transition probability, at a fixed time lag, between the microstates. The eigenvalues of this transition matrix allow estimating the relaxation times of the system, in other words how fast the system is changing. If a gap in the spectrum exists, say after the nth eigenvalue, one can represent the dynamics as a Markov process between n states, which can be considered metastable on the time scale defined by the $(n+1)^{th}$ eigenvalue. Such states are identified by the sign pattern of the eigenvectors associated to the *slow* eigenvalues.

A key difference to standard k-means clustering analysis is that in our approach we use clusters (microstates) as a basis to represent dynamics, rather than as potential metastable configurations based on the analysis of the realised configurations of the flow. We choose the number of clusters by using a criterion that is not typical in ordinary clustering approaches: microstates should be as many as possible to provide a suitable partition of the data space. In this approach, low-frequency modes emerge from the spectral analysis of the transition probability matrix among the microstates and are vectors in the microstate space, not clusters of Z500 patterns.

This procedure, if applied out-of-the-box to the analysis of weather data, provides poor results for two reasons. First, the Z500 patterns in the northern hemisphere, as a result of atmospheric turbulence, are extremely variable. Since the number of available observations in our dataset is of the order of a few thousands - see details below in Sect. 4 - the set of Z500 patterns cannot be divided in microstates that are at the same time sufficiently populated, and that contain sufficiently similar patterns. Indeed it is known since long time that true atmospheric analogues are extremely rare[67] (yet they can be useful[68]). In order to address the first problem we here propose to perform MSM restricted to a longitude window. If one considers the Z500 patterns in a longitudinal window of, e.g., 60°, it will be more likely to find coherent patterns which are similar in the window, even if they can in principle differ outside it. We will nonetheless show below that, in agreement with the idea of atmospheric teleconnections mentioned above, coherence between the patterns is found also outside the window used for clustering purposes. In this manner, we are able to classify the Z500 patterns in microstates which, at the same time, include a sufficient number of observations, and include Z500 patterns that are qualitatively similar. Such microstates will be used as a basis to represent the dynamics, as in standard MSM. One can use this scheme to perform an analysis by using a window of fixed width, which moves on the longitudinal axis, covering in this manner the whole globe.

A second problem is that a genuine dimensional reduction using MSM can be performed only if one observes a significant gap between consecutive relaxation times, this being able to define a cutoff. Such a gap, as shown in

Fig. 1 | Kinetic Analysis. Kinetic analysis using a moving window with kernel width $\Phi = 30°$. **a** First three distinct relaxation times (days) for the analyzed windows. **b** Maximum relative increment for the first ten consecutive relaxation times.
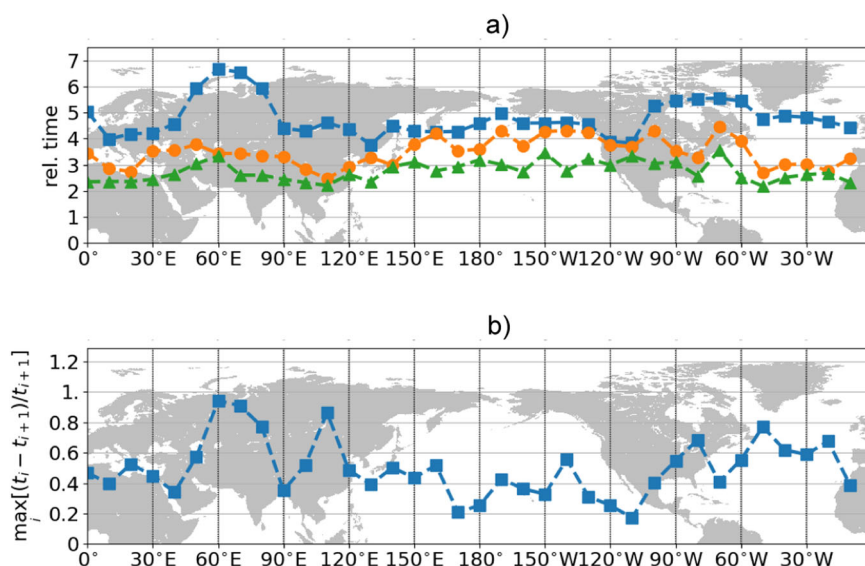
Supplementary Fig. 4 included in the Inf., is typically not present. This, strictly speaking, implies that weather dynamics cannot be unequivocally described by a Markov chain between a small number of states associated with different weather patterns. Nevertheless, all the relaxation times are associated with an eigenvector in the space of the microstates. Though these Z500 patterns may not meet the criteria for genuine metastable states in the MSM lexicon, they represent Z500 configurations that tend to be persistent and, therefore, meaningful for a coarse-grained description of weather dynamics. In this work, we exploit the approach's ability to summarize robustly localized temporal variabilities to analyze these patterns in specific geographic locations and show that the patterns are akin to persistent atmospheric states reported in the literature, such as blocking, zonal flow, and weather regimes. Moreover, the approach could also be used to study and compare dynamics predictability in different models and climate data reanalysis by inspecting their spectral properties, as an increase in the spectral gap should indicate reduced turbulence that should lead to increased predictability. This last aspect has not been elaborated further in this work as we leave it as future work.

Results

In what follows, we investigate the characteristic time scales Markov chains constructed according to the procedure detailed in the previous section, which amounts to the so-called kinetic analysis of the system[69,70]. Figure 1 shows the main findings obtained from our dataset of 1950–2022 December-January-February (DJF) daily averaged Z500 fields for the latitudinal band [32°N, 72°N], see the Section Methods for details.

In Fig. 1a we show the first three largest relaxation times of the MSM as a function of the longitude ϕ_0, for a kernel width of $\Phi = 30°$. The largest relaxation time (blue) fluctuates longitudinally between approximately 4 and 7 days, with higher values near the centre of the Eurasian and American continental masses. Instead, lower values are found over the oceans, near the Pacific west coast, and over central Europe. Such a geographical pattern follows to a good degree of approximation the patterns of synoptic variability[71], with smaller relaxation times being associated with more intense synoptic disturbances. This fits with the concept of baroclinic instability as being the catalyser of transitions between competing states of the low-frequency variability of the atmosphere[7,9]. The results are basically unchanged if one chooses a longer time lag, see Supplementary Fig. 3.

In Fig. 1b, we show the maximum relative gap between two successive relaxation times at the different longitudes for the first ten relaxation times. The gap is at most of order one, which implies that the ith relaxation time, t_i is at most twice larger than the $(i+1)$th. The lack of any significant gap in the

relaxation times implies that atmospheric dynamics cannot be meaningfully described as a Markov process between a small number of metastable states. Hence, there is always a certain degree of subjectivity in performing a cutoff in order to define a discrete, reduced-order model. In other words, the relaxation to the steady state involves a mixture of processes at different time scales which are all interlaced. This observation is not surprising: indeed it confirms the lack of temporal (and spatial) scale separation that is very apparent from spectral analyses of the mid-latitude atmospheric variability[1,5,6].

Alternation between blockings and zonal flow

In the following we describe in detail the eigenvectors associated with the two slowest relaxation times for windows centered around two longitudes: $0°W$, for the Atlantic region, and $180°W$, for the Pacific region, choosing in the first instance a window width of $60°$. The eigenvectors are visualized by plotting the average Z500, t2m, and precip fields of each of the two microstates, which have been chosen according to the rationale discussed in the Section Methods. The average is computed over all the fields belonging to a given microstate.

If the system possessed a gap in the relaxation times, each of the microstates above could be rigorously interpreted as configurations succinctly describing Markov states. Hence, the dynamics could be described at a coarse-grained level as a trapping around a microstate, a rapid transition, a trapping around another microstate, and so on[70]. Since we lack such a time scale separation, the interpretation above is valid only at a qualitative level.

Nonetheless, the eigenvectors provide, by definition, the best possible description of the transition matrix T in a span of dimension 3 (the span includes also the eigenvector that is associated to the unitary eigenvalue, which corresponds to the stationary state). Hence, while the patterns associated to these eigenvectors cannot be considered genuine metastable states, they do reflect dominant modes of variability associated with persistent configurations. Indeed, as we will see, they bear similarities to persistent atmospheric states previously described in the literature, such as blockings, zonal flow patterns, and larger-scale weather regimes associated with teleconnections.

Results for the Atlantic and Pacific sectors are presented in Fig. 2 and Fig. 3, respectively. Both figures have dedicated sections on the top for the first subdominant eigenvector and on the bottom for the second subdominant one. The arrows included in each section illustrate the fact that the two depicted microstates are, in the sense described above, counterposed and that the variability associated to each eigenvector is related to the transitions between the corresponding microstates. In each section, the graph in the bottom right corner shows the percentage of blocking days estimated using the Tibaldi-Molteni (TM) index[38] as a function of the
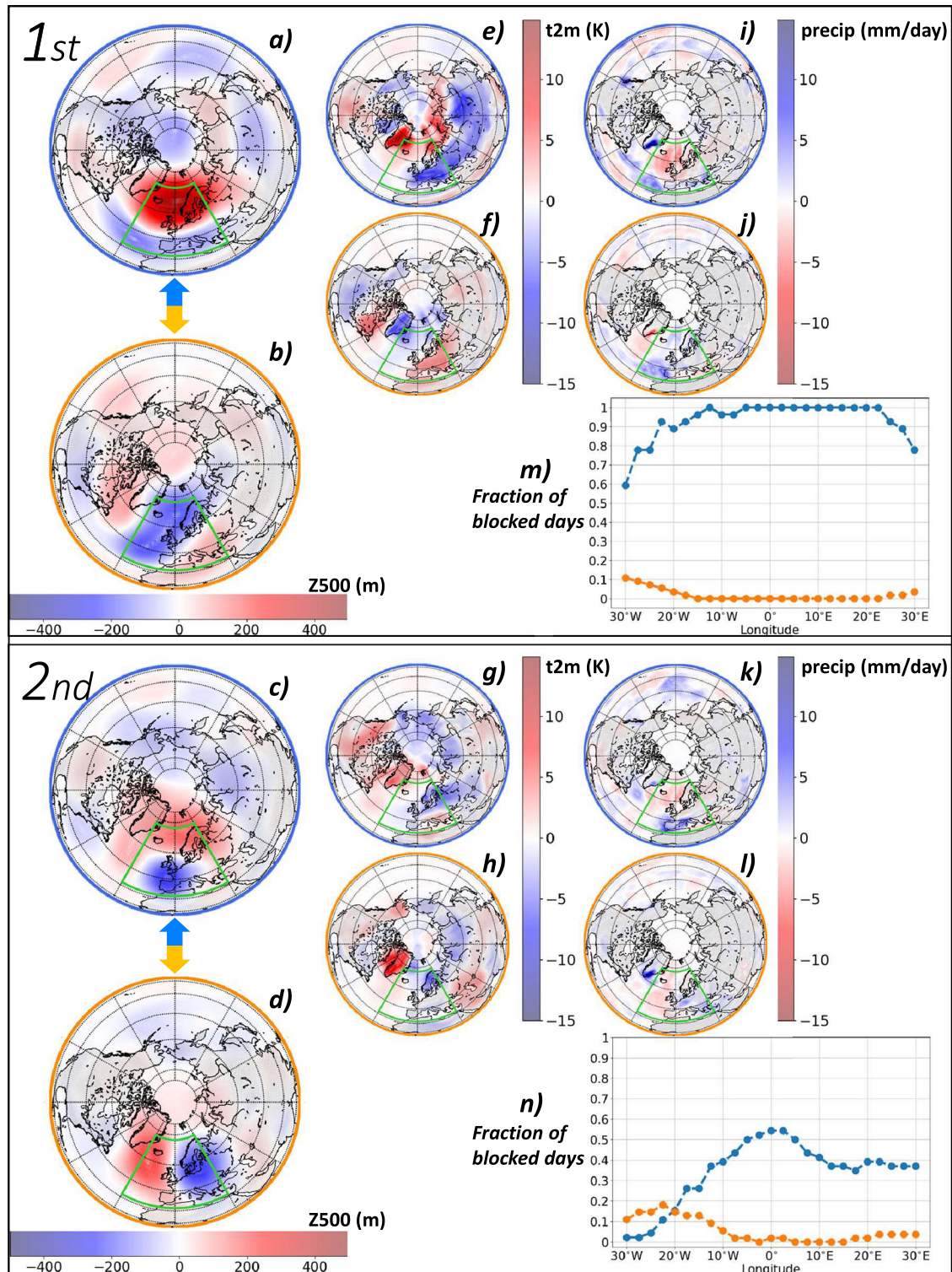
Fig. 2 | Modes of Variability of the Atlantic Sector. Analysis performed by considering Z500 data from the Atlantic sector ($[30°W, 30°E] \times [32°N, 72°N]$). Counterposed microstates mean field Z500 anomalies for the slowest relaxation mode of the system ($t_1 = 5.1$ *days*), and the second slowest relaxation mode of the system ($t_2 = 3.5$ *days*). The green window indicates the area considered for clustering the data. Slowest relaxation model, panel **a**: Z500 [m] anomaly for the microstate featuring more blocked days. Panels **e** and **i**: same as **a**, for the t2m [K] and precip [mm/day] anomalies, respectively. Panels **b**, **f**, and **j**: same as **a**, **e**, and **i**, respectively, for the microstate with fewer blocked days. Panel **m**: fraction of blocked days for each microstate at the longitudes used to cluster the data; compare the orange vs. blue color code. Panels **c**, **d**, **g**, **h**, **k**, **l**, and **n**: same as **a**, **b**, **e**, **f**, **i**, **j**, and **m**, respectively, for the second slowest relaxation mode. The full fields, given by the sum of the climatological and the anomaly patterns shown here, are portrayed in Supplementary Figs. 13a, b.
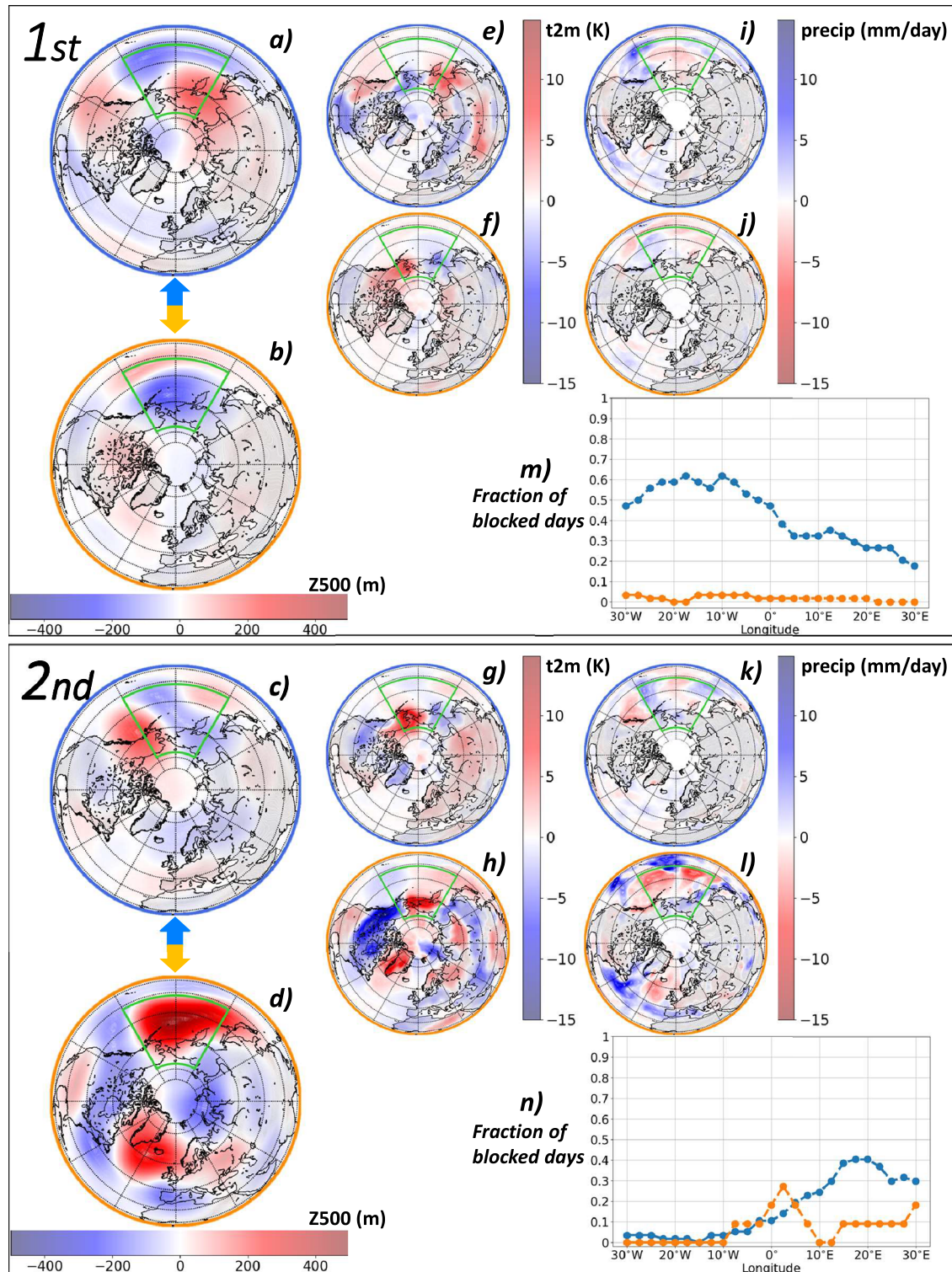
Fig. 3 | Modes of Variability of the Pacific Sector. Same as Fig. 2, panel by panel, but with the analysis performed considering Z500 [m] data from the Pacific sector ([150°W, 150°E] × [32°N, 72°N]). The two slowest relaxation modes correspond to $t_1 = 4.8$ *days* and $t_2 = 3.6$ *days*, respectively. Full fields, given by the sum of the climatological and of anomaly patterns shown here, are portrayed in Supplementary Figs. 13c, d.

longitude for the two opposing microstates obtained through the pipeline procedure. Specifically, the TM index is first computed for each longitude for all the fields belonging to the microstate. If blocked conditions are realized at the longitude of reference and in the two neighbouring longitudes and persist for at least three days, the longitude is defined as blocked and assigned unitary value. Secondly, the percentage is obtained for each latitude

by averaging over all the fields in the microstate. Note that the patterns associated with higher (lower) occurrence of blockings are indicated by a blue (orange) band around the equator. We stick to this convention for the rest of the paper and in the Supp. Inf.

The graphs concerning the first eigenvector show that the two dominant weather patterns are blocked and non-blocked atmospheric states for
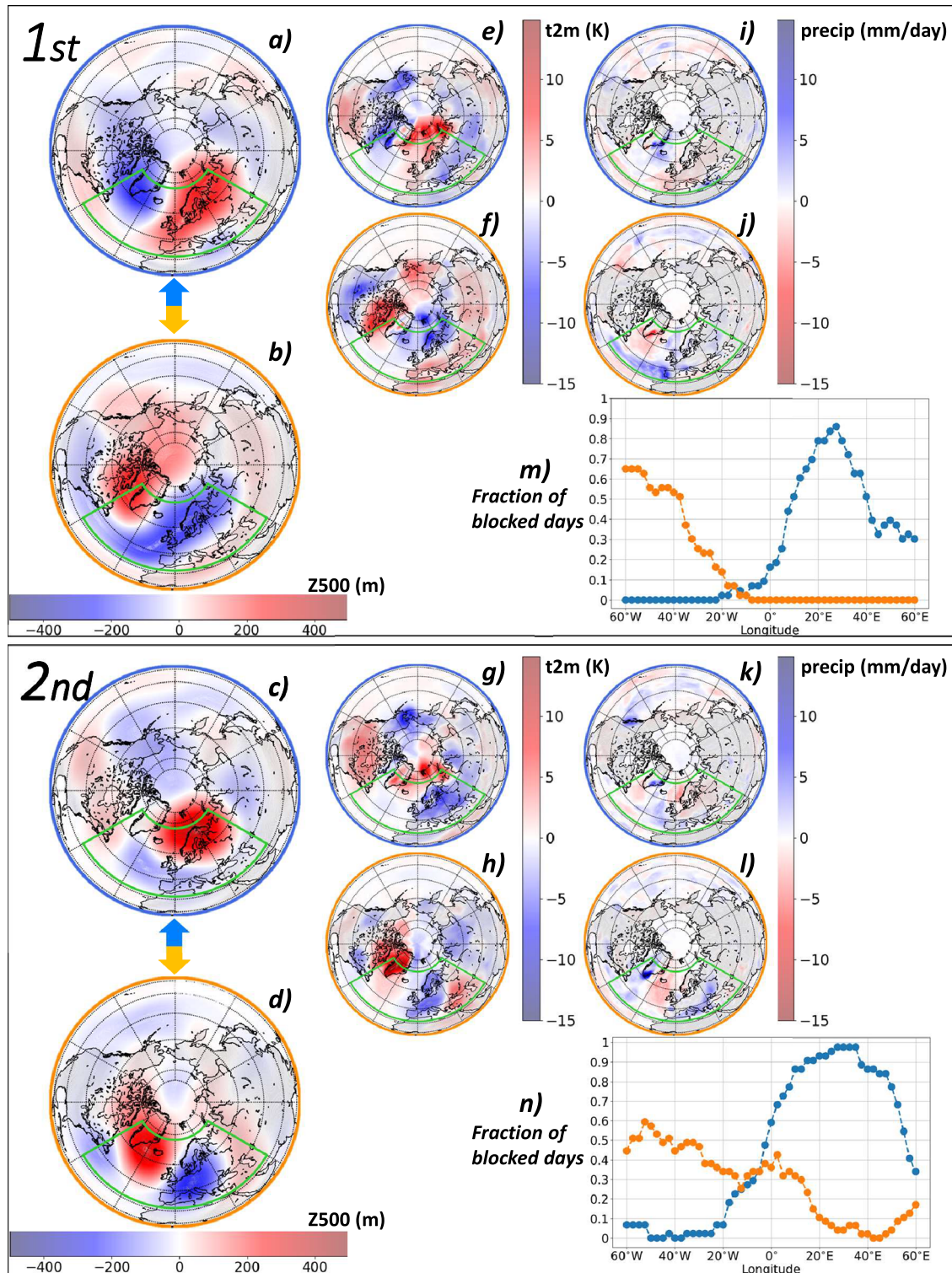
Fig. 4 | Modes of Variability of the Extended Atlantic Sector. Same as Fig. 2, panel by panel, but with the analysis performed considering Z500 data for the extended Atlantic sector ($[60°W, 60°E] \times [32°N, 72°N]$). The two slowest relaxation modes correspond to $t_1 = 3.7$ *days* and $t_2 = 3.5$ *days*, respectively. Full fields, given by the sum of the climatological and of the anomaly patterns shown here, are portrayed in Supplementary Fig. 14a, b.

both the Atlantic (Fig. 2a, b) and the Pacific sectors (Fig. 3a, b). While the dichotomy is strikingly accurate for the Atlantic sector (see Fig. 2m), over the Pacific the percentage of blocked days achieved by the state we can synoptically associate with blockings is somewhat lower (see Fig. 3m). Here, the core of the blockings is geographically shifted towards the western part of the selected region, near Siberia, in agreement with[72,73]. We also notice that

the (predominantly) zonal flow pattern over the Atlantic associated with the blockings-free microstate features the well-known meridionally tilted shape of the North-Atlantic jet stream[74]. The relaxation time of the first eigenvectors for the Atlantic and Pacific sectors (4.8 and 3.7 days, respectively) matches well the time scales of mid-latitude synoptic variability associated with baroclinic disturbances, which are responsible for the transitions

between blocked and zonal states[7,9]. Looking at the temperature and precipitation anomalies in the Atlantic sector composites (Fig. 2e, f, i, j), we find confirmation of the impact of blocking events on such meteorological fields, leading to surface extremes[22]. In what follows we focus on the impacts over land. In agreement with the original observations by Rex[75]: blockings in the Atlantic sector are associated with a dipole of warm anomalies in northern vs. cold anomalies in central and southern regions, whereas dry anomalies prevail almost everywhere, except in the southernmost regions. Note also the wet anomaly in South Greenland, which is presumably due to diverted storm tracks. On the other hand, zonal flow is characterized by moist air masses moving towards Europe, resulting in stronger precipitation. The enhanced eastward advection also leads to warmer conditions in the central and southern regions.

Changes in the advection due to the occurrence of blockings are important for determining anomalies in the surface fields in the Pacific sector too (Figs. 3e,f,,i,j). Zonal conditions are associated with strong cold anomalies in the western sector, as a result of enhanced advection from the cold continental mass, whereas the eastern sector features warm anomalies; compare with[76,77]. The occurrence of blockings leads almost everywhere within the region of interest to reduced precipitation, as a result of prevailing anticyclonic conditions. Notably, one finds increased precipitation near the western North American coast[78].

In order to illustrate the representativity of most extreme microstates shown in Figs. 2 and 3, we show in Supplementary Figs. 6 and 7 the Z500 anomaly fields corresponding to two additional pairs of microstates of the slowest relaxation mode for the Atlantic and Pacific sectors, respectively. These pairs are just slightly less far apart from each other in the complex plane of the components of the corresponding eigenvector than those depicted in Figs. 2 and 3. For both sectors, there is a close alignments among the three pairs. In all three cases, the pairs clearly describe transitions between blocked states and zonal flow. Nonetheless, one finds differences (to be expected) among variants of blocking and zonal flow configurations. Blocking can either substantially suppress the zonal flow over the midlatitudes (e.g. Supplementary Fig. 6 states 57, 170, Supplementary Fig. 7 state 80) or lead merely to its southward displacement. This is a characteristic of high-latitude blocking (Supplementary Fig. 6 state 89, Supplementary Fig. 7 state 7). The zonal flow exhibits a substantial variability as well, especially over the Atlantic region, where the jet can be either more tilted than usually affecting primarily the British Isles and Scandinavia (Supplementary Fig 6 state 41), or have a pronounced zonal orientation affecting mainly the Mediterranean region (linked to a deep trough over Europe, Supplementary Fig. 6 state 96).

Transitions between different blocking patterns
In the Atlantic sector (Fig. 2, lower panel) the Z500 maps related to the second eigenvector contain features associated with transitions between different blocking structures, whereby one microstate features a high-latitude structure (Fig. 2c) and the other one a smaller-scale positive anomaly centered in the Atlantic ocean (Fig. 2d). Quite interestingly, one extreme microstate of the second eigenvector computed for the Pacific sector (Fig. 3d) is associated with the simultaneous presence of a western North Pacific block and of a Greenland block. The fact that a specific and rather relevant microstate captures such a configuration is quite interesting. In ref. 79 it was shown that simultaneous Atlantic-Pacific blocking occur more frequently than what would be the case if the two sectoral blockings were wholly independent. In ref. 37 it was shown that such a pattern is associated with a rather special - and anomalously unstable - configuration of the large-scale flow. Additionally refs. 79, and [80] connected - yet not in the strongest terms - the simultaneous Atlantic-Pacific blockings to the negative phase of the NAM. Also in our analysis such a connection does not appear to be very strong, even if the corresponding Z500 anomaly pattern shows a disrupted polar vortex, which is the main feature of the negative NAM phase. A recent study[81] analysing co-occurring North American and Euro-Atlantic weather regimes identified patterns of simultaneous blocking in the two oceanic regions as well. While we find stronger signature of the simultaneous

blockings when targeting our analysis to the Pacific sector, a weaker global teleconnection signal emerges also when targeting the analysis to the region around Greenland, see Supplementary Figs. 15 and 16, pointing to an asymmetry in the link between the two regions. The fact that our method individuates in an unsupervised manner simultaneous Atlantic-Pacific blockings - targeting only one of the two sectors - further reinforces the idea of a physical link between the Pacific and Atlantic patterns[14,37].

Considering the remarkable nature of the simultaneous Atlantic-Pacific blocking pattern, it is worth looking at the three pairs of most extreme microstates of the second eigenvector for the Pacific sector (Supplementary Fig. 8). Indeed the simultaneous Atlantic-Pacific blocking pattern is very different compared to the other microstates associated with the occurrence of blockings in the Pacific sector. The atypical character of the simultaneous blocking is underpinned by the large distance w.r.t. the second and third most extreme eigenvalues (Supplementary Fig. 8, bottom right panel).

Exploring teleconnection patterns
The analysis of the 60° longitudinal windows captures rather well that the alternation between blocked and zonal flow states is the dominant feature of atmospheric dynamics on that scale, but also points to the relevance of considering different blocking structures and of global teleconnections. In order to be able to better explore these latter aspects, we increase the longitudinal window size to 120°. Here, we restrict ourselves to the extended Atlantic region defined by $[60°W, 60°E] \times [32°N, 72°N]$. Our main findings are portrayed in Fig. 4, which is structured along the same lines as Fig. 2.

Figure 4 shows that the Z500 fields of the two extreme microstates that are associated with the first eigenvector feature very large structures with anomalies covering the whole hemisphere, pointing to the transition between positive and negative NAO phases (NAO+ and NAO-, corresponding to Figs. 4a and 4b, respectively). This is supported by the temperature and precipitation anomaly composites, shown in Figs. 4e, f, i, j; compare with ref. 59. NAO+ is related to a strongly tilted jet and, consequently, a strong ridge over Europe leading to warm and dry anomalies over the western and northern part of the continent. Instead, NAO- is related to an anomalously zonal and southerly shifted jet, which brings abundant precipitation over the Mediterranean region, while cold arctic air masses can reach northern and central Europe. Note that the relaxation time associated with the first eigenvector - 3.7 *days* - is in good agreement with the decorrelation time - about 5 *days* - of the NAO index time series[82]. Note also that the Z500 fields of the two extreme microstates are in close agreement on a global scale with the first two EOFs of the DJF Z500 fields for the whole $20°N-90°N$ sector (not shown), thus supporting the large-scale nature of the modes detected through our method.

The Z500 anomaly fields related to the second eigenvector show similar structures, but rather than affecting the whole hemisphere, they are more confined to the North Atlantic and European regions. The dominant feature of one of the extreme microstates is a strong, meridionally restricted high Z500 anomaly over Northern Europe, representing a typical Scandinavian blocking pattern (Fig.4c). The opposite extreme microstate shows a dipole with a strong blocking over Greenland and Iceland, and a trough over continental Europe (panel d). Thus, the variability associated with the second eigenvector is related to transitions between Scandinavian blocking and Greenland blocking (or North-Atlantic ridge). Both circulation anomalies lead to anomalously cold air over most of the European continent, however, the origin of the cold air differs. The Scandinavian blocking favours cold air advection from Siberia, whereas the deep trough next to the Greenland blocking, allows Arctic airmasses to reach all the way south to the Mediterranean. This transition between circulation patterns connected to anomalously cold conditions brings prolonged periods of cold winter weather over large parts of the continent. Both circulation structures are related to mainly wet anomalies over the Mediterranean, suggesting an anomalously southern location of the jet stream.

To obtain a clearer picture of how precisely the extreme microstates are defined, we show in Supplementary Fig. 9 the Z500 anomaly fields

corresponding to the three most distant pairs of microstates of the first eigenvector for the 120° window in the Atlantic sector. We find that the NAO- pattern (Supplementary Fig. 9b) is close to two patterns (Supplementary Figs. 9d,f) that describe the occurrence of Greenland blocking. This supports previous findings that reveal a close link between NAO- and Greenland blocking and clarifies the fundamental difficulty in distinguishing between these patterns[41,83,84]. Additionally, we find that the NAO+ pattern (Supplementary Fig. 9a) is extremely close to two patterns (Supplementary Figs. 9c, e) depicting the Scandinavian blocking. This is an agreement with recent studies suggesting that these two patterns are closely dynamically associated[84,85]. We remark that this connection does not mean that the NAO is in a positive phase whenever Scandinavian blocking occurs, but points instead towards the dynamical similarity between the two weather patterns that would be classified as different weather configurations based on a clustering analysis or on the NAO index. Finally, Supplementary Fig. 9d reveals a clear connection between the first and second eigenvector; compared with Fig. 4b.

Conclusions

We have analysed the low-frequency variability of the Northern Hemisphere winter mid-latitude atmosphere via MSM. Compared to similar studies in the literature, we do not choose the number of potential metastable states a-priori. Instead, we follow a two-step procedure that allows us to distillate the main feature of the dynamics of the system, going beyond a diagnostic analysis. Following a methodology originally devised for studying molecular dynamical processes, we first partition the phase space in a number of microstates that should be as high as made possible by the available data. In our case, this number ends up being $O(100)$, and is the only free parameter of the procedure. We then construct the Markov matrix describing the transition between the microstates; and, using standard methods, end up identifying the invariant measure corresponding to the stationary probability as well as the relaxation modes, which can be ordered according to their corresponding relaxation time. Those with the longest relaxation times correspond to the slowest decaying modes. Seeking truly global modes is unfeasible because the data is insufficient. Hence, we resort to studying regional features by performing our analysis on longitudinal windows of 60° and 120°, respectively. We have the freedom to shift seamlessly our domain, but we mostly focus on the Atlantic and Pacific sectors, considering the high frequency of winter blockings in these two regions[29].

For the 60° longitudinal window, our results show that, in both the Atlantic and the Pacific sectors, the slowest modes are associated with transitions between blocking states and zonal flow, with the expected associated Z500, surface air temperature, and precipitation anomalies. Additionally, we find strong evidence of a mode of variability - almost a statistical outlier yet clearly relevant for the dynamics - strongly associated with the simultaneous occurrence of blockings in both sectors; see Fig. 3, Supplementary Figs. 15 and 16. This reinforces the conclusions of previous studies indicating the presence of a physical link between the two blockings[14] and of a rather special dynamical configuration associated with such a regime[37]. Many aspects of simultaneous Atlantic-Pacific blockings are not well understood, pointing toward the necessity of future research efforts.

Considering a 120° longitudinal window for the Atlantic sector leads as well to the detection of blockings and zonal flows as important features. Nonetheless, the slowest transitions are in this case related to large-scale dynamics, associated with teleconnections and the transition between different blocking structures. We find that the slowest and second slowest relaxation modes are associated with the transition between patterns that are closely reminiscent of the NAO+ and NAO- states, and of the Scandinavian blocking and Greenland blocking, respectively. A preliminary analysis of the 120° longitudinal window centered over the Pacific is encouraging as our method allows us to identify large-scale patterns that resemble both phases of the much-studied Pacific modes of atmospheric variability, namely the Pacific North America teleconnection pattern[86] and the North Pacific Oscillation/West Pacific teleconnection pattern[87], see Supplementary Figs. 11a–d. Clearly, this aspect deserves further investigation on its own.

The patterns we find are to some degree similar to familiar patterns obtained with other methods, such as EOF, or weather regimes based on clustering analysis, which usually leads to considering 4-8 clusters, see e.g.[59,81,88]. However, due to the large number of microstates in our analysis, our patterns resemble more accurately observed daily fields because less information is filtered out. Furthermore, they are dynamically more meaningful, being defined based on slow modes of variability, as mentioned above. Given the dynamical nature and the high degree of granularity of these patterns, they might be promising candidates to support medium-range weather prediction, along the lines of the so-called weather types[89].

Furthermore, our analysis underlines the complexity of atmospheric dynamics by revealing dynamical connections between different weather regimes. Our results also confirm the difficulty in separating clearly some of these weather patterns, as in the case of the Greenland blocking and NAO-[41] as well as Scandinavian blocking and NAO+. In ref. 83 it is shown that the NAO is strongly related to variations in the occurrence of high latitude blocking, and can be interpreted from a blocking perspective.

An important outcome of our analysis – and certainly related to the similarity of different weather patterns discussed above – is that we do not find a significant gap between relaxation times of the various decaying modes. Hence, we cannot claim that a handful of weather patterns unequivocally describe the relevant large-scale atmospheric dynamics. This underlines the irreducible complexity of atmospheric dynamics of the mid-latitudes, confirming that there is always a certain degree of arbitrariness when selecting dominant modes and excluding the rest. The lack of such a gap, which is apparent when performing a spectral analysis of the mid-latitude atmospheric variability[1,5,6], can also be seen as the reason behind the well-studied presence of deviations from exponential behaviour of the autocorrelation of the NAO index for long time lags[82,90].

In this work, we have chosen to focus on the winter season considering that blockings are more prevalent in winter compared to summer[29]. However, our method is able to detect slow modes of variability in summer as well. We have performed a preliminary analysis on the 1950-2022 summer seasons (June-July-August, JJA) for the Euro-Atlantic sector, where the longitudinal window has been shifted eastwards by 30° in order to take into account that the peak of blocking occurrence is at $\approx 30°E$[91] When considering the 60° longitudinal window, one of the identified dominant patterns of variability captures the occurrence of blockings in the region and is very similar to what was shown in previous investigations[91]. When considering the extended 120° longitudinal window, the two dominant patterns of variability have a close resemblance over Eurasia to the double-ridge and double-trough patterns presented in ref. 92. See Supplementary Fig. 12 for details.

Even if we discuss in detail only the results in two specific, much-studied regions of the Northern Hemisphere, the methodology discussed in this contribution can be adapted seamlessly for analysing dominant weather patterns in other areas of the planet that has been much less extensively studied according to this angle, as in the case of the tropics[93,94], or the mid-latitudes[95–98] and polar regions[99] of the Southern Hemisphere. Our method has the potential of revealing new and previously unknown weather patterns that are relevant in such regions, thus potentially advancing our knowledge of regional climate and facilitating the evaluation of the impacts of climate variability on human and environmental welfare.

Additionally, our approach has a great potential for comparing climate models, testing their realism, and investigating the impact of climate change on the large-scale variability of the atmosphere. In this regard, we plan to take advantage of the available data of the climate models intercomparison project[100]. The availability of much longer time series, as a result of a large ensemble climate model simulation strategy[101] might allow for extending this analysis in such a way that global atmospheric patterns can be directly studied, without the need to resort to longitudinal windows as done here. This might provide a key advancement for understanding the link between synoptic scale variability and global teleconnections and be particularly useful when trying to address the dynamic reasons behind the occurrence of persistent extremes like heatwaves and cold spells[23,102–105].

Methods

Meteorological data and definition of the microstates

Our study focuses on the winter low-frequency variability in the mid-latitudes of the Northern Hemisphere. Hence, following ref.[6], we examined the 1950–2022 DJF daily averaged Z500 field extracted from the 2.5° resolution NCEP Renalysis 1[106] for the latitudinal band [32°N, 72°N]. To ensure that the studied signal meets the basic requirements needed to be able to adopt the analysis method we intend to use, namely that it is possible to define an equilibrium measure, we have taken a necessary preprocessing step of eliminating trends and seasonal variations from the data. In particular, we have detrended and deseasonalized the data by subtracting the DJF average for each year separately before removing the seasonal cycle constructed by computing the average for each day over all years in a pointwise manner, namely, separately for each grid point. On these preprocessed data, we performed Markov state modelling (MSM). In short, MSM involves (1) dividing the data into a large set of discrete states called *microstates*. The data belonging to a microstate are assumed to be similar (see below for a detailed discussion on this point). One then (2) estimates the transition probability, in a given time, between the microstates. The dynamics of the system is (3) described on the basis of the eigenvectors of the transition probability matrix, with the eigenvalues providing information on the relaxation times.

More specifically, let's denote the data points by X_t, where $t = 1,…,N$ labels the different data. In the case of this work, X_t is the Z500 pattern, recorded on a geographical region defined below on a latitude/longitude grid, and t labels the day in which a specific pattern has been recorded. The data are then divided in microstates using a K-means clustering algorithm[107]. Specifically, one chooses a number K of microstates, and, by K-means, one finds, for the Z500 pattern X_t an integer cluster label $c_t \in \{1,…,K\}$.

To decide the value of K one should consider two different factors: (i) the minimum number of elements in each microstate should be large enough for a reliable estimation of the transition probability matrix and (ii) the Z500 patterns assigned to a microstate should be (qualitatively) similar.

Condition (i) is pretty simple to satisfy: since the transition probability in MSM is typically sparse, it is sufficient to choose K in such a way that each cluster is visited a sufficient number of times. Since the number of available days is approximately 6400, we opted for 180 microstates, which implies that each microstate will contain, on average, ~35 fields, each associated with one day, with values ranging between ~10 and ~70. The viability of this choice is demonstrated by dedicated experiments on artificial data generated by the state-of-the-art MPI-ESM-LR version 1.2 Earth system model[108] (Supplementary Figs. 2 and 4). This is among the best-performing ESMs belonging to the CMIP6 class[109] and, additionally, its skill in representing the atmospheric variability of the Northern Hemisphere mid-latitudes compares rather well with its higher resolution version[110]. We have considered the same dataset - *piControl* run featuring pre-industrial conditions in terms of atmospheric composition and land-use[111] - used in refs. 23,104 to investigate heatwaves and cold spells associated with blockings. In the Supp. Inf. we show - see Supplementary Figs. 4a, b - that the characteristic time scales of the leading modes are basically unchanged if one performs the analysis on 72 vs. 1000 winters choosing 180 microstates. Similarly, small changes in the leading characteristic time scales are found if one chooses 180 vs. 360 microstates to analyse the 1000 winters dataset. Moreover, Supplementary Fig. 5 shows that the actual patterns of the two most extreme microstates associated with the first eigenvector are robust with respect to the number of years used in their evaluation (72 vs. 1000 years).

Satisfying condition (ii) turned out to be much more difficult. If one applies K-means clustering with K of order 100 on the Z500 patterns across all longitudes in the latitudinal band of interest, one finds that patterns which are qualitatively different are assigned to the same cluster. As an example, it is possible to observe a specific and recognizable pattern in the Atlantic sector (say an omega-block) and, simultaneously, a variety of different patterns in the Pacific sector (not shown). This observation prompted us to develop a variant of MSM, specifically designed to keep into account the fact that patterns can be strongly correlated for spatial scales corresponding to zonal wavenumbers up to ≈ 3 but - in general - only weakly correlated on a truly global scale[6]. In particular, to weight appropriately local correlations we multiply the patterns by a Gaussian kernel, centered on a specific longitude value ϕ_0, with standard deviation Φ of order of some tenths of degrees (see below). Therefore, the distance between patterns X and X' is given by

$$d(X, X'|\phi_0) = \sum_{\phi,\lambda} a_{\phi,\lambda} \left(X(\phi,\lambda) - X'(\phi,\lambda) \right)^2 e^{-\frac{(\phi-\phi_0)^2}{2\Phi^2}} \tag{1}$$

where $a_{\phi,\lambda}$ is the area of the grid patch of latitude λ and longitude ϕ, and the sum over λ runs between 32°N and 72°N. Clearly, changing the latitudinal band defining the mid-latitudes has an impact on the final outcome of the procedure. Nonetheless, we have verified that the results discussed below are robust with respect to small changes in the spatial domain of references, see Supplementary Figs. 9 and 10 for an example.

Choosing the kernel width is a crucial decision as it determines the scale of climatic events we focus on. We present the results for two kernel widths $\Phi = 30°$ and $\Phi = 60°$, qualitatively corresponding to windows spanning a longitude of 60°, 120°. The former window is appropriate to detect features such as blockings occurring within the target region. The latter window is better suited to capture larger teleconnections and possibly identify global patterns of weather variability; see results below. We remark that the choice of the centre of the window is extremely relevant because the prevalence of blockings, larger scale teleconnection, and the characteristics of the jet stream in the mid-latitudes are strongly dependent on the local geographical and orographical features[2,10,20,29]. Furthermore, this selection of the effective window size is also roughly consistent with the advection length scale: taking ≈ 10 m s^{-1} as reference value for a westerly wind at 500 hPa in the region of interest[112], one obtains an advective length scale over 5 days of $\approx 4.3 \times 10^3$ km, which corresponds to about 50° longitude in the midlatitudes.

We remark that in this approach the K microstates should not be considered in any manner metastable states, but just a basis to represent the dynamics. Metastability, if present, emerges from the spectral properties of the transition probability between these microstates. This leads us to detailing the next steps in our data analysis protocol.

Transition Probability Matrix

After dividing the data into microstates, we estimate the $K \times K$ transition probability matrix T of the process from the data. Each row element of the matrix represents the index of the *departure* microstate, while each column element represents the index of the *arrival* microstate. The element $T_{i,j}$ is the estimated probability of moving from the ith microstate to the jth microstate after the time lag τ, namely $T_{i,j} = P_r(X_t + \tau = j|X_t = i)$. By construction, all the matrix elements are non-negative and the rows of the transition probability matrix sum to 1, namely T is a stochastic matrix[113]. In our specific case, we choose $\tau = 1$ day. In the Supp. Inf. we show that the slowest relaxation times remain very similar if one chooses $\tau = 2$ days, see Supplementary Fig. 3. T is estimated using only consecutive days, separately on each winter, and averaged over all the available winters in our dataset.

Also following the standard MSM pipeline, we then perform an eigendecomposition of the transition probability matrix T, finding its eigenvectors v_i and corresponding eigenvalues λ_i, so that $Tv_i = \lambda_j v_i$. Since T is a stochastic matrix, its first eigenvalue equals one, while the other eigenvalues are complex numbers with norms smaller than one. The components of the eigenvector related to the first eigenvalue - which is unique in the non-degenerate case, as found here - are proportional to the long-term probability of each microstate, thus defining the so-called invariant measure. The eigenvectors of T linked to the eigenvalues with the largest norms below 1 correspond to the slow decaying modes of the system. In the event that an eigenvalue is complex, its conjugate will also be found in the spectrum. This implies that any perturbation to the stationary state described by the ith eigenvector will decay to zero with an exponentially damped sinusoid. The time constant of the exponential, denoted by t_i, and the frequency of the

oscillation, denoted by ω_i, are given respectively by:

$$
\begin{aligned}
t_i &= \frac{-\tau}{\log(|\lambda_i|)} \\
\omega_i &= \frac{2\pi}{\left|\tan^{-1}\left(\frac{Im(\lambda_i)}{Re(\lambda_i)}\right)\right|}
\end{aligned}
\tag{2}
$$

In our case the complex eigenvalues correspond to modes where the oscillatory behaviour is much slower than the decaying one - say with timescale of one month vs. one week, so that at all practical purposes the phenomenology of relaxation is dominated by exponential damping.

It is challenging to find an efficient way to portray an eigenvector given by a combination of microstates, each representing a Z500 pattern. In the case of the subdominant eigenvectors, which correspond to $|\lambda_i| < 1$, we proceed as follows. We consider the two microstates that are as far apart as possible in the complex plane defined by the component of the eigenvector. Indeed, these two microstates represent the dominant components of the anomaly associated with the eigenvector. and they appear always in counterphase. In this sense, the two microstates can be seen in opposition and the eigenvenctor can be interpreted as describing an oscillation between these two microstates, which can be seen as weather regimes. We then visualise the Z500 field associated with each microstate by plotting the average of the Z500 fields of all the days assigned to the microstate. The same procedure can then be repeated for other meteorological fields of interest, see below. See the Supp. Inf. for further details and clarifying examples.

To check whether the relaxation dynamics have a clear signal on variables that are not explicitly controlled in the MSM procedure, we also visualize the average 2-meter air temperature field (t2m) and the precipitation rate field (precip). To have a clearer picture and show large-scale patterns, we always visualise these fields over the whole northern hemisphere, highlighting the region used for deriving the MSM by a green frame. In algorithm 1, we provide a concise overview of the whole analysis pipeline. Clearly, changing the latitudinal band defining the mid-latitudes has an impact on the final outcome of the procedure. Nonetheless, we have verified that our results are robust with respect to small changes in the spatial domain of references, see Supplementary Figs. 9 and 10 for an example.

Algorithm 1. Pipeline summary

Input: Daily average reanalysis at 500 hPa for winter months of the Northern Hemisphere

Output: Partition of the data in microstates, and label of microstates associated with the slowest mode of the system.

1 Select the window central longitude ϕ_0 and the standard deviation Φ of the Gaussian kernel;

Select the number of microstates K;

begin

2 Partition the data into K microstates by the K-means clustering algorithm;

Compute the transition probability matrix T with time lag $\tau = 1$;

Eigendecomposition of T;

Check for a possible spectral gap;

Select the first k eigenvectors of interest;

for $j = 1$ to k **do**

3 Find the indices of the two most extreme components of the j-th eigenvector (Euclidean distance on both Real and Imaginary parts) and assign them to $index_A[j]$ and $index_B[j]$

4 **end**

5 **Return** data-set partition, $index_A[:]$, $index_B[:]$

6 **end**

Data availability

Code availability

References

1. Speranza, A. Deterministic and statistical properties of the westerlies. *Paleogeophysics* **121**, 511–562 (1983).

2. Hannachi, A., Straus, D. M., Franzke, C. L. E., Corti, S. & Woollings, T. Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere. *Rev. Geophys.* **55**, 199–234 (2017).

3. Ghil, M. & Robertson, A. W. "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? *Proc. Natl Acad. Sci.* **99**, 2493–2500 (2002).

4. Masato, G., Hoskins, B. J. & Woollings, T. Winter and summer northern hemisphere blocking in CMIP5 Models. *J. Clim.* **26**, 7044–7059 (2013).

5. Fraedrich, K. & Böttger, H. A wavenumber-frequency analysis of the 500 mb Geopotential at 50° N. *J. Atmos. Sci.* **35**, 745–750 (1978).

6. Dell'Aquila, A., Lucarini, V., Ruti, P. M. & Calmanti, S. Hayashi spectra of the northern hemisphere mid-latitude atmospheric variability in the NCEP–NCAR and ECMWF reanalyses. *Clim. Dyn.* **25**, 639–652 (2005).

7. Benzi, R., Malguzzi, P., Speranza, A. & Sutera, A. The statistical properties of general atmospheric circulation: observational evidence and a minimal theory of bimodality. *Q. J. R. Meteorol. Soc.* **112**, 661–674 (1986).

8. Mo, K. C. & Ghil, M. Statistics and dynamics of persistent anomalies. *J. Atmos. Sci.* **44**, 877–902 (1987).

9. Ruti, P.M., Lucarini, V., Dell'Aquila, A., Calmanti, S. & Speranza, A. Does the subtropical jet catalyze the midlatitude atmospheric regimes? Geophys. Res. Lett. **33** https://doi.org/10.1029/2005GL024620 (2006).

10. Woollings, T., Hannachi, A. & Hoskins, B. Variability of the North Atlantic eddy-driven jet stream. *Q. J. R. Meteorol. Soc.* **136**, 856–868 (2010).

11. Rex, D. F. Blocking action in the middle troposphere and its effect upon regional climate. *Tellus* **2**, 196–211 (1950).

12. Rossby, C.-G. On the dynamics of certain types of blocking waves. *J. Chin. Geophys. Soc.* **2**, 1–13 (1951).

13. Hoskins, B.J. Theories of blocking. In: Seminar on the Nature and Prediction of Extra Tropical Weather Systems. 7-11 September 1987, vol. II, pp. pages1–10. ECMWF, Shinfield Park, Reading (1987)

14. Woollings, T. & Hoskins, B. Simultaneous atlantic-pacific blocking and the northern annular mode. *Q. J. R. Meteorol. Soc.* **134**, 1635–1646 (2008).

15. Tibaldi, S. & Molteni, F. Atmospheric Blocking in Observation and Models. Oxford University Press (2018). https://doi.org/10.1093/acrefore/9780190228620.013.611.

16. Dole, R. et al. Was there a basis for anticipating the 2010 Russian heat wave? Geophysical Research Letters **38** (6) https://doi.org/10.1029/2010GL046582 (2011).

17. Xoplaki, E. et al. Large-Scale Atmospheric Circulation Driving Extreme Climate Events in the Mediterranean and its Related Impacts. In: Lionello, P. (ed.) The Climate of the Mediterranean Region, pp. 347–417. Elsevier, Oxford (2012). https://doi.org/10.1016/B978-0-12-416042-2.00006-9.

18. Lau, W. K. M. & Kim, K.-M. The 2010 Pakistan flood and Russian heat wave: teleconnection of hydrometeorological extremes. *J. Hydrometeorol.* **13**, 392–403 (2012).

19. Buehler, T., Raible, C. C. & Stocker, T. F. The relationship of winter season North Atlantic blocking frequencies to extreme cold or dry spells in the ERA-40. *Tellus A* **63**, 212–222 (2011).

20. Hoskins, B. & Woollings, T. Persistent extratropical regimes and climate extremes. *Curr. Clim. Change Rep.* **1**, 115–124 (2015).

21. Gálfi, V. M., Lucarini, V. & Wouters, J. A large deviation theory-based analysis of heat waves and cold spells in a simplified model of the general circulation of the atmosphere. *J. Stat. Mech.: Theory Exp.* **2019**, 033404 (2019).

22. Kautz, L.-A. et al. Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review. *Weather Clim. Dyn.* **3**, 305–336 (2022).

23. Lucarini, V., Galfi, V. M., Riboldi, J. & Messori, G. Typicality of the 2021 Western North America summer heatwave. *Environ. Res. Lett.* **18**, 015004 (2023).

24. Ferranti, L., Corti, S. & Janousek, M. Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Q. J. R. Meteorol. Soc.* **141**, 916–924 (2015).

25. Lupo, A. R. Atmospheric blocking events: a review. *Ann. N. Y. Acad. Sci.* **1504**, 5–24 (2021).

26. Lucarini, V., Calmanti, S., Dell'Aquila, A., Ruti, P. M. & Speranza, A. Intercomparison of the northern hemisphere winter mid-latitude atmospheric variability of the IPCC models. *Clim. Dyn.* **28**, 829–848 (2007).

27. Di Biagio, V., Calmanti, S., Dell'Aquila, A. & Ruti, P. M. Northern Hemisphere winter midlatitude atmospheric variability in CMIP5 models. *Geophys. Res. Lett.* **41**, 1277–1282 (2014).

28. Davini, P. & D'Andrea, F. Northern hemisphere atmospheric blocking representation in global climate models: twenty years of improvements? *J. Clim.* **29**, 8823–8840 (2016).

29. Woollings, T. et al. Blocking and its response to climate change. *Curr. Clim. Change Rep.* **4**, 287–300 (2018).

30. Nabizadeh, E., Hassanzadeh, P., Yang, D. & Barnes, E. A. Size of the atmospheric blocking events: scaling law and response to climate change. *Geophys. Res. Lett.* **46**, 13488–13499 (2019).

31. Steinfeld, D., Sprenger, M., Beyerle, U. & Pfahl, S. Response of moist and dry processes in atmospheric blocking to climate change. *Environ. Res. Lett.* **17**, 084020 (2022).

32. Dorrington, J., Strommen, K., Fabiano, F. & Molteni, F. CMIP6 models trend toward less persistent European blocking regimes in a warming climate. Geophys. Res. Lett. **49**(24) https://doi.org/10.1029/2022GL100811 (2022).

33. Seneviratne, S.I. et al (eds.) Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, pp. 1513–1766. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA https://doi.org/10.1017/9781009157896.013 (2021).

34. Vannitsem, S. Toward a phase-space cartography of the short- and medium-range predictability of weather regimes. *Tellus A: Dyn. Meteorol. Oceanogr.* **53**, 56–73 (2001).

35. Schubert, S. & Lucarini, V. Dynamical analysis of blocking events: spatial and temporal fluctuations of covariant Lyapunov vectors. *Q. J. R. Meteorol. Soc.* **142**, 2143–2158 (2016).

36. Faranda, D., Messori, G. & Yiou, P. Dynamical proxies of North Atlantic predictability and extremes. *Sci. Rep.* **7**, 41278 (2017).

37. Lucarini, V. & Gritsun, A. A new mathematical framework for atmospheric blocking events. Climate Dynamics **52** https://doi.org/10.1007/s00382-019-05018-2 (2020).

38. Tibaldi, S. & Molteni, F. On the operational predictability of blocking. *Tellus A* **42**, 343–365 (1990).

39. Pelly, J. L. & Hoskins, B. J. A new perspective on blocking. *J. Atmos. Sci.* **60**, 743–755 (2003).

40. Barriopedro, D., García-Herrera, R., Lupo, A. R. & Hernández, E. A climatology of northern hemisphere blocking. *J. Clim.* **19**, 1042–1063 (2006).

41. Davini, P., Cagnazzo, C., Gualdi, S. & Navarra, A. Bidimensional diagnostics, variability, and trends of Northern hemisphere blocking. *J. Clim.* **25**, 6496–6509 (2012).

42. Pinheiro, M. C., Ullrich, P. A. & Grotjahn, R. Atmospheric blocking and intercomparison of objective detection methods: flow field characteristics. *Clim. Dyn.* **53**, 4189–4216 (2019).

43. Wallace, J. M. & Gutzler, D. S. Teleconnections in the Geopotential Height Field during the Northern Hemisphere Winter. *Month. Weather Rev.* **109**, 784–812 (1981).

44. Hurrell, J. W. Decadal trends in the north Atlantic oscillation: Regional temperatures and precipitation. *Science* **269**, 676–679 (1995).

45. Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M. In: Hurrell, J.W., Kushnir, Y., Ottersen, G., Visbeck, M. (eds.) An Overview of the North Atlantic Oscillation, pp. 1–35. American Geophysical Union (AGU), https://doi.org/10.1029/134GM01 (2003).

46. Thompson, D. W. J. & Wallace, J. M. The Arctic oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.* **25**, 1297–1300 (1998).

47. Branstator, G. Circumglobal teleconnections, the jet stream waveguide, and the North Atlantic Oscillation. *J. Clim.* **15**, 1893–1910 (2002).

48. Harnik, N., Messori, G., Caballero, R. & Feldstein, S. B. The Circumglobal North American wave pattern and its relation to cold events in eastern North America. *Geophys. Res. Lett.* **43**, 11015–11023 (2016).

49. Croci-Maspoli, M., Schwierz, C. & Davies, H. C. Atmospheric blocking: space-time links to the NAO and PNA. *Clim. Dyn.* **29**, 713–725 (2007).

50. Athanasiadis, P. J., Wallace, J. M. & Wettstein, J. J. Patterns of wintertime jet stream variability and their relation to the storm tracks. *J. Atmos. Sci.* **67**, 1361–1381 (2010).

51. Feldstein, S.B. & Franzke, C.L.E. In: Franzke, C.L.E., O'Kane, T.J. (eds.) Atmospheric Teleconnection Patterns, pp. 54–104. Cambridge University Press, Cambridge (2017)

52. Baur, F., Hess, P. & Nagel, H. Kalender der großwetterlagen europas 1881–1939. Technical report, DWD: Bad Homburg (1944)

53. Vautard, R. Multiple weather regimes over the North Atlantic: Analysis of precursors and successors. *Month. Weather Rev.* **118**, 2056–2081 (1990).

54. Madonna, E., Li, C., Grams, C. M. & Woollings, T. The link between eddy-driven jet variability and weather regimes in the North Atlantic-European sector. *Q. J. R. Meteorol. Soc.* **143**, 2960–2972 (2017).

55. Franzke, C.L.E. Persistent regimes and extreme events of the North Atlantic atmospheric circulation. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. **371** https://doi.org/10.1098/rsta.2011.0471 (2013).

56. De Luca, P. et al. Past and projected weather pattern persistence with associated multi-hazards in the British Isles. Atmosphere **10**(10) https://doi.org/10.3390/atmos10100577 (2019).

57. Madonna, E., Battisti, D. S., Li, C. & White, R. H. Reconstructing winter climate anomalies in the Euro-Atlantic sector using circulation patterns. *Weather Clim. Dyn.* **2**, 777–794 (2021).

58. Galfi, V.M. & Messori, G. Persistent anomalies of the North Atlantic jet stream and associated surface extremes over Europe. Environ. Res. Lett. **18** https://doi.org/10.1088/1748-9326/acaedf (2023).

59. Fabiano, F., Christensen, H. M., Strommen, K., Athanasiadis, P., Baker, A., Schiemann, R. & Corti, S. Euro-Atlantic weather Regimes in the PRIMAVERA coupled climate simulations: impact of resolution and mean state biases on model performance. *Clim. Dyn.* **54**, 5031–5048 (2020).

60. Franzke, C., Crommelin, D., Fischer, A. & Majda, A. J. A hidden markov model perspective on regimes and metastability in atmospheric flows. *J. Clim.* **21**, 1740–1757 (2008).

61. Kwasniok, F. Enhanced regime predictability in atmospheric low-order models due to stochastic forcing. *Philos. Trans. R. Soc. A: Math., Phys. Eng. Sci.* **372**, 20130286 (2014).

62. Tantet, A., Burgt, F. R. & Dijkstra, H. A. An early warning indicator for atmospheric blocking events using transfer operators. *Chaos: Interdiscip. J. Nonlinear Sci.* **25**, 036406 (2015).

63. Detring, C., Müller, A., Schielicke, L., Névir, P. & Rust, H. W. Occurrence and transition probabilities of omega and high-over-low blocking in the Euro-Atlantic region. *Weather Clim. Dyn.* **2**, 927–952 (2021).

64. Mukhin, D., Hannachi, A., Braun, T. & Marwan, N. Revealing recurrent regimes of mid-latitude atmospheric variability using novel machine learning method. *Chaos: Interdiscip. J. Nonlinear Sci.* **32**, 113105 (2022).

65. Prinz, J.-H. et al. Markov models of molecular kinetics: Generation and validation. *J. Chem. Phys.* **134**, 174105 (2011).

66. Husic, B. E. & Pande, V. S. Markov state models: from an art to a science. *J. Am. Chem. Soc.* **140**, 2386–2396 (2018).

67. Lorenz, E. N. Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.* **26**, 636–646 (1969).

68. Faranda, D. et al. A climate-change attribution retrospective of some impactful weather extremes of 2021. *Weather Clim. Dyn.* **3**, 1311–1340 (2022).

69. Pande, V. S., Beauchamp, K. & Bowman, G. R. Everything you wanted to know about Markov State Models but were afraid to ask. *Methods* **52**, 99–105 (2010).

70. Glielmo, A. et al. Unsupervised learning methods for molecular simulation data. *Chem. Rev.* **121**, 9722–9758 (2021).

71. Hoskins, B. J. & Hodges, K. I. The annual cycle of northern hemisphere storm tracks. part i: seasons. *J. Clim.* **32**, 1743–1760 (2019).

72. Gao, M., Yang, S. & Li, T. Assessments on simulation of pacific blocking frequency during boreal winter in cmip6 climate models. *Dyn. Atmos. Oceans* **100**, 101333 (2022).

73. Gao, M., Yang, S. & Li, T. The spatio-temporal variation of pacific blocking frequency within winter months and its relationship with surface air temperature. Atmosphere **11**(9) https://doi.org/10.3390/atmos11090960 (2020).

74. Woollings, T. & Blackburn, M. The North Atlantic jet stream under climate change and its relation to the NAO and EA patterns. *J. Clim.* **25**, 886–902 (2012).

75. Rex, D. F. The effect of Atlantic blocking action upon European Climate. *Tellus* **3**, 100–112 (1951).

76. Breeden, M. L., Hoover, B. T., Newman, M. & Vimont, D. J. Optimal North Pacific blocking precursors and their deterministic subseasonal evolution during Boreal winter. *Monthly Weather Rev.* **148**, 739–761 (2020).

77. Nabizadeh, E., Lubis, S. W. & Hassanzadeh, P. The 3D structure of Northern Hemisphere blocking events: climatology, role of moisture, and response to climate change. *J. Clim.* **34**, 9837–9860 (2021).

78. Carrera, M. L., Higgins, R. W. & Kousky, V. E. Downstream weather impacts associated with atmospheric blocking over the Northeast Pacific. *J. Clim.* **17**, 4823–4839 (2004).

79. Woollings, T. & Hoskins, B. Simultaneous Atlantic-Pacific blocking and the Northern Annular Mode. *Q. J. R. Meteorol. Soc.* **134**, 1635–1646 (2008).

80. Thompson, D. W. J. & Wallace, J. M. Regional climate impacts of the Northern hemisphere annular mode. *Science* **293**, 85–89 (2001).

81. Messori, G. & Dorrington, J. A joint perspective on North American and Euro-Atlantic weather regimes. *Geophys. Res. Lett.* **50**, 2023–104696 (2023).

82. Önskog, T., Franzke, C. L. E. & Hannachi, A. Predictability and Non-Gaussian characteristics of the North Atlantic oscillation. *J. Clim.* **31**, 537–554 (2018).

83. Woollings, T. J., Hoskins, B., Blackburn, M. & Berrisford, P. A new Rossby wave-breaking interpretation of the North Atlantic Oscillation. *J. Atmos. Sci.* **65**, 609–626 (2008).

84. Li, C. & Zhang, Q. An observed connection between wintertime temperature anomalies over Northwest China and weather regime transitions in North Atlantic. *J. Meteorol. Res.* **29**, 201–203 (2015).

85. Vihma, T. et al. Effects of the tropospheric large-scale circulation on European winter temperatures during the period of amplified Arctic warming. *Int. J. Climatol.* **40**, 509–529 (2020).

86. Linkin, M. E. & Nigam, S. The North Pacific Oscillation-West Pacific teleconnection pattern: mature-phase structure and winter impacts. *J. Clim.* **21**, 1979–1997 (2008).

87. Aru, H., Chen, S. & Chen, W. Change in the variability in the Western Pacific pattern during boreal winter: roles of tropical Pacific sea surface temperature anomalies and North Pacific storm track activity. *Clim. Dyn.* **58**, 2451–2468 (2022).

88. Hochman, A., Messori, G., Quinting, J.F., Pinto, J.G. & Grams, C.M. Do Atlantic-European weather regimes physically exist? *Geophys. Res. Lett.* **48** https://doi.org/10.1029/2021GL095574 (2021).

89. Neal, R., Fereday, D., Crocker, R. & Comer, R. E. A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorol. Appl.* **23**, 389–400 (2016).

90. Feldstein, S. B. The timescale, power spectra, and climate noise properties of teleconnection patterns. *J. Clim.* **13**, 4430–4440 (2000).

91. Drouard, M. & Woollings, T. Contrasting mechanisms of summer blocking over Western Eurasia. *Geophys. Res. Lett.* **45**, 12040–12048 (2018).

92. Yang, D. & Wang, L. The summertime circulation types over Eurasia and their connections with the North Atlantic oscillation modulated by North Atlantic SST. Atmosphere **13**(12) https://doi.org/10.3390/atmos13122093 (2022).

93. Pope, M., Jakob, C. & Reeder, M. J. Regimes of the North Australian wet season. *J. Clim.* **22**, 6699–6715 (2009).

94. Hassim, M. E. E. & Timbal, B. Observed rainfall trends over Singapore and the maritime continent from the perspective of regional-scale weather regimes. *J. Appl. Meteorol. Climatol.* **58**, 365–384 (2019).

95. Solman, S. A. & Menéndez, C. G. Weather regimes in the South American sector and neighbouring oceans during winter. *Clim. Dyn.* **21**, 91–104 (2003).

96. Wilson, L., Manton, M. J. & Siems, S. T. Relationship between rainfall and weather regimes in south-eastern Queensland, Australia. *Int. J. Climatol.* **33**, 979–991 (2013).

97. Arizmendi, F., Trinchin, R. & Barreiro, M. Weather regimes in subtropical South America and their impacts over Uruguay. *Int. J. Climatol.* **42**, 9253–9270 (2022).

98. Loikith, P. C. et al. A climatology of daily synoptic circulation patterns and associated surface meteorology over southern South America. *Clim. Dyn.* **53**, 4019–4035 (2019).

99. Pohl, B. et al. Relationship between weather regimes and atmospheric rivers in East Antarctica. *J. Geophys. Res.: Atmos.* **126**, 2021–035294 (2021).

100. Eyring, V. et al. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 10539–10583 (2016).

101. Maher, N., Milinski, S. & Ludwig, R. Large ensemble climate model simulations: introduction, overview, and future prospects for utilising multiple types of large ensemble. *Earth Syst. Dyn.* **12**, 401–418 (2021).

102. Kornhuber, K. et al. Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environ. Res. Lett.* **14**, 054002 (2019).

103. Kornhuber, K. & Tamarin-Brodsky, T. Future changes in Northern Hemisphere summer weather persistence linked to projected Arctic warming. *Geophys. Res. Lett.* **48**, 2020–091603 (2021).

104. Galfi, V. M. & Lucarini, V. Fingerprinting heatwaves and cold spells and assessing their response to climate change using large deviation theory. *Phys. Rev. Lett.* **127**, 058701 (2021).

105. Ragone, F. & Bouchet, F. Rare event algorithm study of extreme warm summers and heatwaves over Europe. *Geophys. Res. Lett.* **48**, 2020–091197 (2021).

106. Kalnay, E. et al. The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Am. Meteorol. Soc.* **77**, 437–472 (1996).

107. Lloyd, S. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **28**, 129–137 (1982).
108. Mauritsen, T. et al. Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO2. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).
109. Bock, L. et al. Quantifying progress across different CMIP phases with the ESMValTool. *J. Geophys. Res.: Atmos.* **125**, 2019–032321 (2020).
110. Müller, W. A. et al. A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *J. Adv. Model. Earth Syst.* **10**, 1383–1413 (2018).
111. Wieners, K.-H. et al. MPI-M MPI-ESM1.2-LR model output prepared for CMIP6 CMIP esm-piControl. Earth System Grid Federation https://doi.org/10.22033/ESGF/CMIP6.6553 (2019).
112. Whittleston, D., McColl, K. A. & Entekhabi, D. Multimodel future projections of wintertime North Atlantic and North Pacific tropospheric jets: A Bayesian analysis. *J. Clim.* **31**, 2533–2545 (2018).
113. Andrilli, S. & Hecker, D. Chapter 8 - additional applications. In: Andrilli, S., Hecker, D. (eds.) Elementary Linear Algebra (Fifth Edition), Fifth edition edn., pp. 513–605. Academic Press, Boston (2016). https://doi.org/10.1016/B978-0-12-800853-9.00008-6.

Acknowledgements

Author contributions

V.L. and A.L. conceived the study. S.S. performed the data analysis and produced the figures. V.M.G. led the analysis of the obtained meteorological fields. All authors discussed the results and contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41612-024-00659-5.

Correspondence and requests for materials should be addressed to Alessandro Laio or Valerio Lucarini.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

¹Physics Department, International School for Advanced Studies (SISSA), Via Bonomea 265, Trieste 34136, Italy. ²International Center of Theoretical Physics (ICTP), Viale Miramare 11, Trieste 34151, Italy. ³Institute for Environmental Studies, Vrije Universiteit Amsterdam, De Boelelaan 1105, Amsterdam 081 HV, the Netherlands. ⁴School of Computing and Mathematical Sciences, University of Leicester, University Road, Leicester LE17RH, UK. ⁵Department of Mathematics and Statistics & Centre for the Mathematics of Planet Earth, University of Reading, Whiteknights Campus, Reading RG66AX, UK. ⁶These authors contributed equally: Sebastian Springer, Alessandro Laio, Vera Melinda Galfi, Valerio Lucarini. ✉e-mail: laio@sissa.it; v.lucarini@leicester.ac.uk